Investment Sophistication and Wealth Inequality

Ehsan Azarmsa*

September 2019

Abstract

I study how differences in investment sophistication, the ability to identify profitable investment opportunities, contribute to wealth inequality and its dynamics. I analyze a financial market with a continuum of investors heterogeneously informed about the distribution of future returns. The stationary distribution of wealth shares features a thick right tail populated by the best-informed investors. Wealth inequality increases with the emergence of ever-costlier information production technologies, but the expected excess return achieved through such technologies is bounded in the long run. Empirically, wealthier households indeed tend to have more precise beliefs, and the gap in their beliefs can rationalize inequality under reasonable parameters. My findings suggest that information inequality can be a first-order contributor to top wealth inequality. Subsidizing financial education and leveling the playing field regarding information can help reduce inequality.

1 Introduction

Wealth inequality has been rising for decades (Piketty (2014), Saez and Zucman (2016)). Particularly, the finance industry has been at the center of accusations, eminently manifested in the “Occupy Wall Street” movement, since the industry deemed to represent the

---

*I am extremely grateful to my committee chair Pietro Veronesi and members Lin William Cong, Lars Peter Hansen, Lubos Pastor and Constantine Yannelis for detailed feedback and suggestions. I also thank Mohammad Akbarpour, Ben Brooks, George Constantinides, Wenxin Du, Piotr Dworczak, Niels Gormsen, Gary Gorton, Zhiguo He, John Heaton, Paymon Khorrami, Jian Li, Elliot Lipnowski, Yueran Ma, Negar Matooarian, Roger Myerson, Stefan Nagel, Doron Ravid, Marzena Rostek, Xavier Vives, Alex Zentefis, Anthony Lee Zhang, Eric Zwick and conference and seminar participants at Economic Dynamics and Financial Markets 2018-19, Crossing Disciplinary Boundaries 2018, Hand Economics Forum 2019 and SITE Banks and Financial Frictions 2019 for helpful comments and discussions. Author’s contact: azarmsa@uchicago.edu
speculators, not the innovators, of the economy. In fact, the financial industry professionals, mostly consisting of fund managers, now comprise about 20% of the Forbes 400 list of US richest individuals and families, five-times more than early ’80s (Kaplan and Rauh (2013)). Markedly, this increase overlapped with the recent developments in information technology, which have yielded investors more accurate asset valuations (Bai, Philippon, and Savov (2016)). This paper posits a model explaining how the recent technological advancements have contributed to the rise of both wealth inequality and that of highly sophisticated investors, such as fund managers, in the pool of super-rich.

Specifically, I develop a dynamic model of a financial market in which the investors have heterogeneous access to some information resources about the distribution of future returns. Therefore, endowed with different information, the investors trade assets, hold different portfolios, and consequently reap different returns on their savings, which translates into a change in the wealth distribution. In this paper, I primarily study the stationary (long-run) distribution of wealth shares in the economy. In short, the distribution features a thick right tail populated by the most informed investors, best-representing hedge fund managers, known for their sophistication and advantage in accessing and processing the relevant information.

Generally speaking, no matter how complicated the information environment is, it affects the top wealth inequality only through the extent of asymmetric information between the most informed investors and the “representative investor”, who takes a wealth-weighted average belief of all investors. The intuition is that the top inequality is directly related to the most informed investors’ expected excess return, with respect to the representative investor. As their informational advantage shrinks, so does their expected excess return, which leads to less return heterogeneity, and hence, less inequality. This result yields a heuristic rule for understanding the distributional impact of a change in the households’
information: If the change makes the wealthiest relatively more informed about the available investment opportunities compared to an average investor, then it also leads to more wealth inequality. For instance, the development of personalized financial advisors, such as robo-advisors, aiming to liberate the access to financial planning and information, are effective in reducing the inequality since they are directly targeting the less-wealthy segment of the population, reducing the extent of asymmetric information among the investors.

In addition to the distribution, my model also generates predictions about the composition of the super-rich. As we focus on a smaller and smaller pool of the top wealthiest, only the most informed investors remain and the others gradually wash out from the pool. In short, sophistication fully crowds out luck in the long-run: almost no lucky investor with a low sophistication makes to the pool of super-rich. This result explains the notable representation of fund managers in the Forbes 400 list (Kaplan and Rauh (2013)).

The presence of asymmetric information among investors might be best manifested in the hedge funds’ access to highly expensive and exclusive information sources, such as satellite images\textsuperscript{1} and private polls\textsuperscript{2}. To obtain an edge in financial markets, they search for novel, but potentially expensive, ways to produce information, resembling an arms race for generating $\alpha$. Informing the consequences, the model implies wealth inequality generally increases with the emergence of costlier information production technologies. The intuition being that such costlier technologies would be affordable for a smaller fraction of the investors, increasing the extent of asymmetric information between the adopters (informed investors) and the representative investor.

Even though the adopters, as the new informed investors, populate the tail in the new equilibrium, their superior performance diminishes over time. More specifically, their ex-

\textsuperscript{1}“Stock Picks From Space” The Atlantic, May 2019
\textsuperscript{2}“The Brexit Short: How Hedge Funds Used Private Polls to Make Millions” Bloomberg, June 2018
pected return converges to a bounded level, independent of the technology’s informativeness. The intuition is as the adopters become wealthier, they push the equilibrium prices more strongly and make them more informative, blunting their information edge, and consequently, reducing their expected returns. Therefore, such technologies only boost the adopters’ short-term expected return, consistent with the recent mediocre performance of the hedge fund Industry.\textsuperscript{3}

Regarding the comparative statics, I study the distributional impact of an increase in the precision of available information, say, obtained by the recent IT developments. First, the more precise the available information, the more informative the equilibrium prices, reducing both equity premium and asymmetric information among the investors. However, more strikingly, it might also increase wealth inequality. The intuition is that the investors, facing less uncertainty, react more aggressively to any signal not perfectly incorporated in the prices. Such aggressive behaviors magnify portfolio heterogeneity, and consequently, the return heterogeneity, increasing the inequality. Therefore, depending on whether this channel, the stronger reaction to beliefs, dominates the opposite channel, the belief convergence, the availability of more precise information might increase the inequality.

In addition to the precision, the allocation of information, namely who has access to and comprehends which source of information, impacts the wealth distribution. The concentration of information among a few increases inequality and do so even more when combined with the information becoming more precise. This result indicates that the policies tackling information inequality, such as subsidizing higher education or improving the transparency of financial markets, are both effective and crucial in reducing wealth inequality.

Along with the information environment, the financial markets have evolved tremendously

\textsuperscript{3}“Hedge Fund Trends For 2019: Industry Has Finally Hit Its Saturation Point” Yahoo Finance, January 2019
in their size, liquidity, connectedness, and complexity over the past decades. The inequality increases with the liquidity of the markets and the magnitude of non-fundamental noises. More generally, the model implies the informativeness of equilibrium prices plays a crucial role in reducing the inequality; hence, any other change in the market structure that hampers the information dissemination through the prices also increases the inequality.

Dynamic models of financial markets with asymmetrically informed investors are notoriously intractable. Nonetheless, my model delivers highly tractable dynamics for wealth distribution. I tackle the common issues and obtain tractability by using logarithmic preferences instead of CARA preferences. Assuming logarithmic preferences, I prove the wealth shares move with the investors’ probability assignments to the realized states, rendering all the intermediary steps, such as solving for the optimal portfolios or the equilibrium prices and asset allocations, unnecessary, once the dynamics are derived.

Moreover, according to the no-trade theorem of Milgrom and Stokey (1982), the presence of some “noise trades” is necessary to facilitate the trades among a group of traders with homogeneous preferences, but heterogeneous information about the asset values. Hence, to fully leverage the feature mentioned above, I introduce some noise traders whom I exogenously specify their beliefs, instead of their asset demands. Therefore, the noise traders in my model are rational, in the sense that they are Bayesian and dynamically maximize their expected utility. However, they have misspecified beliefs about the underlying data generating process. Particularly, they are assumed to base their predictions about the asset returns on a wrong signal. That said, the wrong signal, operating through the noise traders’ decisions, works as a non-fundamental noise in the prices, obscuring the information of more informed investors from the less informed ones.\(^4\)

\(^4\)In addition, the other studies mostly use CARA preferences, which disconnects the optimal portfolio from the wealth level, causing the price informativeness to be independent of the wealth distribution, which is an undesirable feature for the question at hand. Here, by using logarithmic preferences, the prices become
After building this model, several questions naturally follow. Do the wealthier actually have more precise beliefs? If so, is the gap in their precision large enough to significantly contribute to wealth inequality? In Section 2, I empirically find a significant difference in the macroeconomic beliefs, confirming the earlier empirical finding of Das, Kuhnen, and Nagel (2017). Then, leveraging the wealth dynamics derived here, I estimate the impact of the belief heterogeneity on wealth inequality. The estimates indicate a relatively large level of wealth inequality, suggesting the belief gap is large enough to contribute significantly to US wealth inequality. Finally, I discuss the role of bequest motive and security space in Section 7.

The rest of the paper is organized as follows: In Section 2, I present two motivating facts. Section 3 lays out the model and Section 4 defines and characterizes the equilibrium. I find and analyze the wealth dynamics in Section 5 and discuss the stationary distribution. Section 6 studies the case of endogenous information acquisition. Section 7 contains the empirical analysis and some theoretical extensions of the model. Section 8 concludes.

Literature

With the recent surge in wealth inequality, its causes and consequences are currently at the center of political and academic debates. The earlier studies focused more on the mechanisms relating wealth inequality to income inequality (Aiyagari (1994), Castaneda, Diaz-Gimenez, and Rios-Rull (2003)). Nonetheless, they achieved limited success, especially in explaining the top inequality (Fagereng, Guiso, Malacrino, and Pistaferri (2016)). Alternatively, a new strand of theoretical and empirical studies have examined the role of more informative as the wealth becomes more concentrated in the hand of most informed investors. Wilkinson, Pickett, and Cato (2009); Wilkinson and Pickett (2018) argue inequality creates social problems such as illiteracy, crime, and poor health. Pastor and Veronesi (2018) discuss its role in the recent uprise of populism.

6
return heterogeneities in wealth inequality. Supporting this channel, Bach, Calvet, and Sodini (2016) and Fagereng, Guiso, Malacrino, and Pistaferri (2016) document a vast heterogeneity in returns in Sweden and Norway respectively. Benhabib, Bisin, and Zhu (2011) and Benhabib, Bisin, and Luo (2015, 2019) demonstrate, theoretically and quantitatively, the importance of return heterogeneity to match the empirical wealth distribution. However, they do not discuss the root causes of return heterogeneity.

Several recent empirical studies suggest the answer lies in the portfolio heterogeneity across households. Garbinti, Goupille-Lebret, and Piketty (2017), Kuhn, Schularick, and Steins (2017) and Hubner, Krusell, and Smith Jr (2018) show that the portfolio heterogeneity and asset price movements are essential to explain the dynamics of wealth distribution. Fagereng, Guiso, Malacrino, and Pistaferri (2016) are the first to demonstrate the presence of substantial return heterogeneity even within asset classes, highlighting the role of investment sophistication, which I primarily study here.\(^6\)

A large body of empirical studies also lends support to the importance of investment sophistication in generating a higher risk-adjusted return. Grinblatt, Keloharju, and Linnainmaa (2011, 2012), Korniotis and Kumar (2013), Barber, Lee, Liu, and Odean (2014), Gargano and Rossi (2018), Clark, Lusardi, and Mitchell (2017) provide evidence that high-IQ, skilled and attentive investors hold portfolios with a higher Sharpe ratio.\(^7\) Frazzini, Kabiller, and Pedersen (2018) demonstrate Warren Buffet has generated a significantly positive alpha over Fama and French three factors and most of his wealth accrued by his stock-picking ability. Bartscher, Kuhn, and Schularick (2018) and Girshina (2019) emphasize the role

---

\(^6\)In fact, they show that the scale effects, the wealthier having access to more lucrative investment opportunities, and heterogeneity in risk preferences do not account for the whole cross-section of returns across households. Additionally, they find the individual fixed effects (after controlling for the observable characteristics and risk-taking behavior), interpreted as “investment sophistication”, substantially increases the explanatory power.

\(^7\)Further on the role of experience, Kempf, Manconi, and Spalt (2017) show fund managers’ performance in a sector increases with his or her investment experience in that sector.
of education in portfolio choice and generating higher returns on risky assets. Massa and Simonov (2006) show investors invest in stocks closely related to their non-financial income and provide evidence that they rationally prioritize familiarity over income hedging in their portfolio decisions. My model sheds light on the distributional implication of the heterogeneities identified in the studies above.

Apart from the investment sophistication, the returns are heterogeneous due to the differences in investment opportunity sets, risk preferences (Veronesi (2018)) and entrepreneurial skills (Cagetti and De Nardi (2006)). Pástor and Veronesi (2016) study a model of occupational choice with agents heterogeneous in their entrepreneurial skills and risk-aversion. Guvenen (2009) examines the role of market segmentation. However, none of these studies justify the over-representation of highly sophisticated and talented investors among the wealthiest, documented by (Kaplan and Rauh (2013)). Moreover, Fagereng, Guiso, Malacrino, and Pistaferri (2016) rigorously analyze the role of sophistication in the cross-section of returns and show it is essential to explain the vast heterogeneity in returns.

My paper also contributes to the literature of belief heterogeneity by characterizing the wealth dynamics induced by any arbitrary belief distribution. The belief heterogeneity, especially regarding the business cycle variables, is extensively documented (e.g., Mankiw, Reis, and Wolfers (2003), Vissing-Jorgensen (2003) Souleles (2004), Puri and Robinson (2007)). D’Acunto, Hoang, Paloviita, and Weber (2019) and Das, Kuhnen, and Nagel (2017) empirically show high IQ investors and the ones with a higher socioeconomic status, respectively, have more precise macroeconomic beliefs. Giglio, Maggiori, Stroebel, and Utkus (2019) also find a tight connection between investors’ beliefs and their portfolio choice. These findings suggest economic beliefs directly impact the cross-section of households’ average returns, and subsequently, the wealth distribution. The setup presented here can be employed to study the distributional implication of such belief heterogeneities.
This paper is also related to nascent literature studying the role of sophistication in wealth inequality. In addition to Fagereng, Guiso, Malacrino, and Pistaferri (2016), An, Bian, Lou, and Shi (2019) and Campbell, Ramadorai, and Ranish (2018) provide evidence on how financial markets redistribute from the less to more sophisticated investors. Among the theoretical studies, Lei (2019) studies endogenous learning of the profitability of privately owned assets. Peress (2003) and Kacperczyk, Nosal, and Stevens (2018) find that wealth and capital income inequality increase with the emergence of new costly signals (costly to obtain or process), as they favor wealthier and more sophisticated investors. Alas, having finitely many periods, their models only inform the short-term consequences of such technological changes. Particularly, their models do not account for the fact that the price informativeness indeed increases with the informed investors gaining more wealth share, leading to a smaller information gap, and hence, return heterogeneity in the long-run. Accounting for this negative feedback effect, I find a sufficiently large increase in the information cost (or the emergence of a sufficiently expensive signal) increases wealth inequality. Nonetheless, the relationship might not be monotone. Overall, to the best of my knowledge, my paper is the first to discuss the long-run distributional impact of such technological changes. Furthermore, it puts forward empirical predictions regarding the composition of the wealthiest and distribution of expected returns, absent in the earlier studies.

2 Motivating Facts

The top wealth shares have increased substantially over the past decades (Saez and Zucman (2016); Smith, Zidar, and Zwick (2019)). However, along with the distribution, the composition of the wealthiest has drastically changed. For instance, Kaplan and Rauh
(2013) find that the number of fund managers in the Forbes 400 list has increased close to five times between 1982 and 2011. More strikingly, the number of hedge fund managers have increased from 2 in 1982 to 30 in 2011, indicating a fifteen-time increase (Figure 1). Notably, a salient distinguishing feature of the hedge fund managers is their direct access to the state-of-the-art technologies for information production and processing. Linking asymmetric information to wealth distribution, my model explains the presence of highly informed investors among the wealthiest. It further implies that the presence increases with the emergence of new and expensive information production technologies, justifying the empirical pattern in Figure 1.

![Figure 1: Top 0.1% wealth shares in the US, 1980-2016.](image)

The second observation is that the wealthier have more accurate beliefs about business cycles, which in turn improve their saving decisions. In order to demonstrate this, I use a sample of 218559 individuals asked about their macroeconomic expectation in Michigan Survey Data, between 1978q1-2018q1. ⁸ Since the data does not provide the households’

---

⁸The survey includes several questions, but I am particularly interested in the respondents’ expectation
net worth, I use “investment amount in stock market” and “income” as proxy variables. Therefore, to see the relationship between wealth and beliefs, I divide the respondents into 20 groups, separately based on each proxy, and look at the fraction of respondents who expressed a high chance for having a good business condition in the year ahead. I compare this fraction with the realized growth over the next four quarters. The growth rate is considered to be high if it exceeds the historical mean, and it is considered to be low otherwise. Figures 2 and 3 show the average forecast errors for every income and investment level group respectively.

Figures 2 and 3 exhibit a negative relationship between the forecast errors in economic growth and the two proxies of wealth. Both figures show there is a substantial belief heterogeneity across the respondents and the wealthier individuals tend to have more accurate beliefs. Therefore, they are more likely to obtain a higher expected risk-adjusted return on their savings.

3 A Dynamic Model of Financial Markets

Consider an economy with discrete periods, i.e. \( t = 1, 2, \ldots \). It is an endowment economy: There is a Lucas tree that pays out \( d_t \) units of consumption good at the beginning of each period. About 1-year-ahead economic growth, captured by variable “BUS12”. More specifically, it asks “Now turning to business conditions in the country as a whole, do you think that during the next 12 months we’ll have good times financially, or bad times, or what?” The data also includes information about the household income and their investment in stock market. I only consider the respondents without any qualification. For more information about the survey, see Das, Kuhnen, and Nagel (2017).

To construct the forecast errors, I compare the respondents’ answer with the realized growth in the following year (including the survey’s quarter). I only use respondents that have a clear good or bad expectation, where they constitute more than 75% of the total respondents. Moreover, I consider an annual growth rate above its historical mean as “high”, and “low” otherwise. Then, I calculate and report the average value of \( (\text{respond}_{i,t} - I_{g_t+g_{t+1}+g_{t+2}+g_{t+3}\geq 2.4})^2 \) for each of twenty half-deciles. \( \text{respond}_{i,t} \) is a binary variable indicating whether the respondent has a positive economic view. The result is qualitatively robust to all choices of the threshold between 1.5% to 2.5%.
Figure 2: Prediction error of the growth rate for different levels of total investment in the publicly listed stocks. The predictions are binary and the realized growth rates are divided into “high” or “low” growth rate, with a cutoff equal to the annual rate of 2.4%. The figure shows $\mathbb{E}[(\text{respondent}_{i,t} - \mathbb{1}_{g_t+g_{t+1}+g_{t+2}+g_{t+3} \geq 2.4})^2]$, where $i$ is the respondent indicator and $t$ is the quarter indicator and $g_t$ is the quarter-to-quarter growth at $t$.

In (1), $\tilde{z}_t^N$ and $\tilde{z}_t^I$ are the exogenous signals based on which the agents make their prediction about the next period’s growth rate $g_{t+1}$. Following the literature, the variables with $\sim$ are random variables. Note that, in this specification, the only signal informative about the future growth rate is $\tilde{z}_t^I$. However, as I elaborate later, some “Naive” investors have misspecified beliefs and use signal $\tilde{z}_t^N$ for their predictions. $q^j, j \in \{g, h\}$ is the conditional
Figure 3: Prediction error of the growth rate for different levels of income. The predictions are binary and the realized growth rates are divided into “high” or “low” growth rate, with a cutoff equal to the annual rate of 2%. The figure shows $E[(\text{respond}_{i,t} - \mathbb{I}_{g_t+g_{t+1}+g_{t+2}+g_{t+3} \geq 2.4})^2]$, where $i$ is the respondent indicator and $t$ is the quarter indicator and $g_t$ is the quarter-to-quarter growth at $t$.

probability of $\tilde{g}_{t+1} = g^h$ when $g_t = g^i$, without knowing the realizations $z^I_t$ and $z^N_t$. For tractibility reasons, I also assume the unconditional probability of having a high or low growth rate is the same. It implies $q^h = 1 - q^l$.

**Agents.** The economy is populated with a unit measure of investors, with time-separable and logarithmic preferences over their consumption. All investors are Bayesian, that is they form and update their beliefs based on Bayes’ rule.

There are two types of investors in the economy. First, there is a set of atomistic investors with beliefs consistent with the underlying data generating process. For the rest of the paper, I refer to this set of investors as “sophisticated investors”. I denote this set by $S$ and they constitute a total measure of $1 - b$, for some $b \in (0, 1)$. The sophisticated investors are further divided into finer types $i \in S \equiv [0, 1]$. The types indicate the sophistication level of each investor in the sense that the agents with a higher $i$ observe the realization of the informative signal $z^I_t$ more frequently, as discussed next.
In each period, a random fraction of the sophisticated investors observe \( z^I_t \) and become “informed”. In particular, at the beginning of period \( t \), an i.i.d random variable \( \tilde{\lambda}_t \in (0, 1) \) is drawn with cumulative distribution \( F(\cdot) \), and once \( \lambda_t \) is realized, all agents with type \( i \geq \lambda_t \) perfectly observe \( z^I_t \). The rest of the sophisticated investors do not receive any exogenous signal. Hence, they need to make Bayesian inference about \( z^I_t \) based on the equilibrium prices. Furthermore, I denote the posterior of the informed and uninformed investors (after potentially observing the signal realization and equilibrium prices) by \( q^I_t = P^I_t(\tilde{g}_{t+1} = g^h) \) and \( q^U_t = P^U_t(\tilde{g}_{t+1} = g^h) \) respectively, where \( P^J_t, J \in \{U, I\} \), indicates the probability measure that an \( J \)-type investor assigns to the future events at time \( t \).

Therefore, the sophisticated investors are heterogeneous only in the unconditional probability that they receive the informative signal \( \tilde{z}^I_t \) and become informed. However, after the realization of \( \lambda_t \), all informed investors observe the same signal and there is no asymmetric information among the informed investors.\(^{10}\) Here, the information structure is governed by cumulative distribution \( F(\cdot) \).

The primary goal here is to study the evolution of wealth distribution among the sophisticated investors. However, to motivate trade among the sophisticated investors, we need to have an exogenous and stochastic supply of liquidity (Milgrom and Stokey (1982)). Here, the provision of exogenous liquidity is done by a set of “Naive” investors, who constitute a measure of \( b < 1 \) in this economy. The key assumption about the naive investors is that they falsely believe the growth rate follows the following Markov process, instead of the actual one described in (1):

\[
P^N(\tilde{g}_{t+1} = g^h | \tilde{g}_t = g^j, z^I_t, z^N_t) = q^j + z^N_t \quad j \in \{h, l\}
\]

\(^{10}\)The assumption is made to simplify the exposition and avoid unnecessary complications in dealing with different information sets.
Note that the only different between naive and sophisticated investors is that the naive investors falsely believe that signal $\tilde{z}_t^N$, not $\tilde{z}_t^I$, is informative about the next period’s growth rate. All naive investors perfectly observe $z_t^N$ at the beginning of each period. As a result, they do not update their beliefs based on the equilibrium prices and they just act based on their own signal $\tilde{z}_t^N$. That said, $\tilde{z}_t^N$ works as a non-fundamental uncertainty in the economy that distorts the prices and hence, prevents $z_t^I$ from being perfectly revealed to the uninformed investors. The naive investors’ posterior is denoted by $q_t^N \equiv P_t^N(\hat{g}_{t+1} = g^h)$.

**Interpretation of the naive investors:** In the literature of financial markets and market microstructure, it is commonly assumed that there is a set of “noise traders” who have an exogenous and stochastic net demand across assets traded in the market (for example, see Grossman and Stiglitz (1980) and Kyle (1985)). The main role of this assumption is to obscure the signal of the more informed investors from the less informed ones, and consequently, open the possibility of trade among a set of heterogeneously informed investors with homogeneous preferences and a common prior. In my model, the naive investors serve a similar purpose by providing an exogenous, stochastic and stationary source of liquidity across markets. The only key difference is that, in my model, the expression for the naive investors’ belief, instead of their demand schedule, is exogenously specified. This choice is made to obtain tractability. In fact, as elaborated later, the dynamics of wealth distribution can be characterized in closed-form as a function of the investors’ posterior. Therefore, exogenously specifying the beliefs, rather than the asset demands, yields a substantially more tractable framework.

Dynamic models with heterogeneous types are susceptible to extreme asymptotic outcomes and lack of a stationary distribution. To ensure the existence of a stationary solution, following Yaari (1965) and Blanchard (1985), I assume the investors are subject to random
death. Particularly, at the beginning of period $t \geq 2$, each investor dies with probability $1 - \delta \in (0, 1)$, independent of the other shocks. Each deceased investor is replaced with his “child” that has the same type.\(^\text{11}\) The inherited wealth is taxed at rate $\tau \in (0, 1)$; Hence the child endows a fraction of the parent’s wealth, as specified later once assets and wealth are introduced. The investors in such a sequence form a “family”. Each type of investors comprises an infinite number of such families. In fact, for each $k \in [0, 1] \cup \{N\}$, there is an infinite number of “families”, indexed by $f \in [0, 1]$. Therefore, each family in the economy is indexed by a pair $(f, k)$, $f \in [0, 1]$ and $k \in [0, 1] \cup \{N\}$.\(^\text{12}\)

As mentioned earlier, I assume that the investors have logarithmic and time-separable preferences over consumption, that is the terminal utility of an investor with type $k \in [0, 1] \cup \{N\}$ born in family $(f, k)$ at $s_1$ and dies at $s_2 > s_1$ from consumption path $\{c_{k}^{f,t}\}_{t=s_1}^{s_2-1}$ is as specified in (3). Note that no bequeath motive for wealth accumulation embedded in the preferences. The bequest motive for saving is examined in Section 7.2.

$$U(\{c_{k}^{f,t}\}_{t=s_1}^{s_2-1}) = \sum_{t=s_1}^{s_2-1} \beta^{t-s_1} \log c_{k}^{f,t}, \quad k \in [0, 1] \cup \{N\} \quad (3)$$

**Assets.** There are two assets in this economy. One is the shares of the tree, which is in unit supply and traded at equilibrium price $p_t$, in the unit of consumption goods at period $t$. In addition, there is a risk-free asset with a net zero supply. In particular, at period $t$, the investors trade consumption claims for period $t + 1$, at an equilibrium rate $R_t^F$. That is at time $t$, they can trade $R_t^F$ units of consumption goods at time $t + 1$ for one unit of consumption goods at time $t$.\(^\text{13}\) I denote the time-$t$ share of the tree in the hand

---

\(^\text{11}\)This assumption implies that the sophistication level is the same across all the generations of a family. I relax this assumption in Section A14. and show it modestly impacts the derivations.

\(^\text{12}\)This assumption is necessary to make sure the realization of the death shocks are independent from the wealth dynamics of each type. This property is essential to make sure the total wealth share owned by each type of investors follows a Markov Process independent of the death shocks.

\(^\text{13}\)Note that the market is dynamically incomplete. For instance, no long-term risk-free bond is available
of investor of family \((f,k)\) by \(x_{f,t}^k\). Similarly, \(s_{f,t}^k\) is defined as the consumption claims of period \(t\) (traded at \(t-1\)) that the investor owns, provided he is alive at period \(t\). Therefore, his wealth at time \(t\), in the unit of consumption goods at time \(t\), is:

\[
W_{f,t}^k = (p_t + d_t)x_{f,t}^k + s_{f,t}^k
\]

To avoid Ponzi-like consumption schemes, I assume an investor with a negative wealth cannot borrow, resulting in zero consumption and extremely low expected utility. Therefore, all investors optimally maintain a positive net worth. That said, their optimization problem at time \(t\) is:

\[
\max \left\{ c_{f,t+j}^k, x_{f,t+j+1}^k, s_{f,t+j+1}^k \right\} \mathbb{E}_t^k \left[ \sum_{j=0}^{\infty} (\beta \delta)^j \log c_{f,t+j}^k \right]
\]

\[
\text{s.t.} \quad c_{f,t+j}^k + p_t x_{f,t+j+1}^k + \frac{1}{R_{f,t+j}} s_{f,t+j+1}^k \leq (p_t + d_t)x_{f,t+j}^k + s_{f,t+j}^k \quad \forall j \geq 0 \quad (4)
\]

\[
W_{f,t+j}^k > 0 \quad \forall j \geq 0
\]

where the expectation operator \(\mathbb{E}_t^k[\cdot]\) takes expectation with respect to the information set and the belief of type-\(k\) agents at time \(t\). In (4), it is easy to see that the feasibility set scales with wealth and additionally, the preferences over different consumption paths are scale-invariant. Therefore, the optimal portfolio and consumption of all type-\(k\) investors only differ by a scalar. To simplify the expositions, it is useful to introduce the integrated variables \(c_t^k, x_t^k, s_t^k\), where

\[
c_t^k = \int_0^1 c_{f,t}^k df, \quad x_t^k = \int_0^1 x_{f,t}^k df, \quad s_t^k = \int_0^1 s_{f,t}^k df.
\]

Similarly, the total wealth in the hand of type-\(k\) investors is

\[
W_t^k = \int_0^1 W_{f,t}^k df.
\]

**Inheritance.** As mentioned earlier, each investor dies with probability \(1 - \delta\) and is replaced in this economy. However, in Section 7.3, I show this restriction is not binding. In other words, I prove even if the set of available securities fully span the space of future dividend realizations, the allocations and wealth dynamics would be exactly the same as the case I analyze here.
with his child, who has the same type. The inheritance takes place as follows: When the
investor of family \((f,k)\) with wealth \(W_{f,t}^k\) and asset holding \((x_{f,t}^k, s_{f,t}^k)\) dies at time \(t\), his
child replaces him and receives \(1 - \tau\) fraction of the assets. The remaining fraction of the
deceased’s wealth is distributed evenly among all the newborns at period \(t\). Therefore, the
initial asset holding of the replacing child in family \((f,k)\) at time \(t\) is:

\[
x_{f,t}^k = (1 - \tau)\hat{x}_{f,t}^k + \tau(x_t^N + \int_0^1 x_i^t di) = (1 - \tau)\hat{x}_{f,t}^k + \tau
\]

\[
s_{f,t}^k = (1 - \tau)\hat{s}_{f,t}^k + \tau(s_t^N + \int_0^1 s_i^t di) = (1 - \tau)\hat{s}_{f,t}^k
\]

Therefore, the initial wealth of the replacing child is \(W_{f,t}^k = (1 - \tau)\hat{W}_{f,t}^k + \tau(p_t + d_t)\).

4 Equilibrium

In this section, I solve for the equilibrium sequence of consumptions, asset allocations and
prices and the posterior beliefs, given the growth and signal realizations. The character-
ization provides the wealth dynamic for each family, used to characterize the stationary
distribution in this economy. The notion of equilibrium is Rational Expectation Equi-
librium (REE) applied to a heterogeneous beliefs framework. It means all agents make
Bayesian inference about the future growth rates based on the the available signals and
equilibrium prices, given their belief about the underlying Markov Process. I formally
define the equilibrium in Section 4.1.

To solve for the equilibrium objects, I proceed in three steps. First, in Section 4.2, I con-
struct the Bellman equation for the investors. In particular, in Lemma 1(a), I show all
investors with the same belief at time \(t\) make the same consumption and saving decisions,
regardless of their wealth level, up to a scalar. Then, I characterize the optimal intertemporal decisions given the beliefs. Finally, in section 4.3, I solve for the posterior beliefs for a given distribution of wealth and realization of signals. After these three steps, I fully specify the joint evolution of the wealth distribution and other variables in the model \{\{W^k_t\}_{k \in [0,1] \cup \{N\}, W^N_t, d_t, g_t, z^I_t, z^N_t, \lambda_t}\}_{t=1}^{\infty} in Section 5.

4.1 Definition of Equilibrium

To be more concrete, I define the equilibrium as below:

**Definition 1** (Dynamic Competitive Equilibrium).

Denote the set of all publicly observable state variables at the beginning of period \(t\) by

\[
\mathcal{F}_t^P \equiv (\{x^k_t\}_{k \in \{0,1\} \cup \{N\}}, \{s^k_t\}_{k \in \{0,1\} \cup \{N\}}, \lambda_t, g_t)
\]

The sequence of demand functions \(\{c^k_t(p_t, R^F_t, q^k_t; \mathcal{F}_t^P), x^k_{t+1}(p_t, R^F_t, q^k_t; \mathcal{F}_t^P), s^k_{t+1}(p_t, R^F_t, q^k_t; \mathcal{F}_t^P)\}_{t=1}^{\infty}\) for \(k \in [0,1] \cup \{N\}\) and the price functions \(p(z^I_t, z^N_t, \mathcal{F}_t^P), R(z^I_t, z^N_t, \mathcal{F}_t^P)\), where \(p_t = p(z^I_t, z^N_t, \mathcal{F}_t^P)\) and \(R^F_t = R(z^I_t, z^N_t, \mathcal{F}_t^P)\), constitute a dynamic competitive equilibrium if

1. **[Optimality]** The demand functions constitute a solution to the optimization problem (4) given beliefs \(\{q^k_t\}_{k \in [0,1] \cup \{N\}, t \geq 1}\).

2. **[Bayesian Inference]** All investors make Bayesian inference based on their signals and equilibrium prices \(p_t\) and \(R^F_t\) at time \(t\). In particular, if \(g_t = g^j, j \in \{h, l\}\):
\[ q^k_t = \begin{cases} 
q^i + z^I_t & k \in \lambda_t, 1 \\
q^i + z^U(z^I_t, z^N_t; F^P_t) & k \in [0, \lambda_t) \\
q^i + z^N_t & k = N 
\end{cases} \]

, where \( z^U(\cdot, \cdot; \cdot) \) is obtained from Bayes' rule.\(^{14}\)

3. **[Market Clearance]** For every \( t \), the prices \( p_t \) and \( R^F_t \) are such that the market for consumption goods, risky assets and riskless assets clear:

\[
\begin{align*}
\int_0^1 c^q_t dt + c^N_t &= d_t \\
\int_0^1 x^i_{t+1} dt + x^N_{t+1} &= 1 \\
\int_0^1 s^i_{t+1} dt + s^N_{t+1} &= 0
\end{align*}
\]

4.2 **Optimal Intertemporal Decisions**

In this section, I derive the optimal consumption and portfolio of a type-\( k \) investor with wealth \( W^k_t = (d_t + p_t)x^k_t + s^k_t \) and posterior \( q^k_t, k \in [0, 1] \cup \{N\} \). First, I prove that the value function is logarithmic in wealth. As specified in Lemma 1, it has three important implications: 1) The price-dividend ratio of the tree is constant. 2) All investors consume a fixed fraction of their wealth in every period. 3) In their portfolio choice decisions, the investors act myopically and just maximize the expected log-wealth in the next period.

In order to solve for the investors’ optimal consumption and portfolio decision, I first

\[^{14}\text{More specifically,}\]

\[ z^U(z^I_t, z^N_t; F^P_t) = \frac{\int_0^{z^N_t} \int_{z^I_t}^{\hat{z}^I_t} z^I_t \mathbf{1}\{p_t = p(z^I_t, z^N_t, F^P_t), R^F_t = R(z^I_t, z^N_t, F^P_t)\} dz^I_t dz^N_t}{\int_0^{z^N_t} \int_{z^I_t}^{\hat{z}^I_t} \mathbf{1}\{p_t = p(z^I_t, z^N_t, F^P_t), R^F_t = R(z^I_t, z^N_t, F^P_t)\} dz^I_t dz^N_t} \]
characterize their Bellman equation. Note that the state variable here is \( \mathcal{F}_t \equiv (\mathcal{F}_t^P, z^I_t, z^N_t) \), which includes the state of asset allocations at \( t \), the growth rate \( g_t \), \( \lambda_t \), and exogenous signal realizations \( z^I_t \) and \( z^N_t \). Therefore, the value function can be specified as follows:

\[
V^k(W_t; \mathcal{F}_t) = \max_{c_t, x_{t+1}, s_{t+1}} \log c_t + \beta \delta \mathbb{E}^k[V^k(W_{t+1}; \mathcal{F}_{t+1})] \tag{7}
\]

\[s.t. \quad c_t + p_t x_{t+1} + \frac{1}{R_t^F} s_{t+1} \leq W_t\]

\[W_{t+1} = (d_{t+1} + p_{t+1}) x_{t+1} + s_{t+1}\]

\[Prob(W_{t+1} > 0) = 1\]

\[p_t = p(z^I_t, z^N_t, \mathcal{F}_t^P) \quad R_t^F = R(z^I_t, z^N_t, \mathcal{F}_t^P)\]

Lemma 1.

a) For every type \( k \in [0, 1] \cup \{N\} \), there exists measurable function \( a^k(\cdot) \), with respect to the space induced by \( \mathcal{F}_t \)'s, such that

\[V^k(W_t; \mathcal{F}_t) = \frac{1}{1 - \beta \delta} \log W_t + a^k(\mathcal{F}_t)\]

b) All investors consume a fixed fraction of their wealth. In particular, \( c_t = (1 - \beta \delta) W_t \)

c) For every \( t \geq 1 \), \( p_t = \frac{\beta \delta}{1 - \beta \delta} d_t \)

Lemma 1 describes the optimal consumption and saving behavior of the investors. Part (a) shows that the value functions are logarithmic in wealth. Moreover, part (b) indicates that the investors always save a fixed fraction of their wealth, regardless of their wealth or beliefs.\(^{15}\) Hence, since the optimal consumption is independent of the future expected

\(^{15}\) The intuition lies in the Cobb-Douglas nature of the preferences. It is well-known that agents with such preferences allocate their wealth to different components of their consumption bundle in fixed fractions, determined by the parameters of the utility function. In my model, the discount rate (\( \beta \)) and survival rate (\( \delta \)) govern the households' saving ratio.
returns, the investors optimally only maximize their expected log-wealth, or equivalently their expected log-return, in the next period.\textsuperscript{16}

Part (c) of Lemma 1 shows that the price-dividend ratio is constant. The reason is given the result in Lemma 1(b), the economy’s total consumption and wealth should be at the same fixed proportion of the individual ones. Noting the total consumption is just the total dividend, $d_t$, and the total wealth is the total dividend plus the price of the tree, $p_t + d_t$, we can derive the price-dividend ratio as follows.

$$\int_0^1 W_t^i di + W_t^N = p_t + d_t = \frac{1}{1 - \beta} d_t$$

Lemma 1(c) has further implications. First, it implies that the price-dividend ratio does not depend on the prospects about the future dividends. In other words, the price of the tree contains no information about the future growth rate and the only price that partially reflects the informed investors’ information is the risk-free rate.\textsuperscript{17} It also implies that the return on the risky asset is always proportional to the growth rate. Therefore, the fundamental choice that the investors face is to what extent they want to expose their savings to the growth rate $g_{t+1}$. This optimal exposure generally depends on the risk-free rate and one’s posterior about the future growth. Lemma 2 provides the closed-form solution to the optimal portfolio problem.

\textbf{Lemma 2.}

\textsuperscript{16}Such myopic behavior is directly resulted from the logarithmic preferences over consumption. In particular, we can infer from Merton (1973) that the investors with logarithmic preferences (more specifically, when IES is one) do not hedge against the changes in the investment opportunity set. Therefore, they only maximize the (subjective) expected log-return, period by period.

\textsuperscript{17}Note that the main goal here is to study the impact of asymmetric information on wealth inequality and the model is not suited for studying the asset pricing implications of wealth inequality. This question is explored in earlier studies, such as Gollier (2001) and Gomez (2019). In fact, my model examines the interplay between the wealth distribution and information dissemination and it provides reasonable implications about the interaction, as discussed in Section 4.3.
Suppose type-k investors, \( k \in [0, 1] \cup \{N\} \), have posterior belief \( q^k_t \in (0, 1) \). Furthermore, let \( \mu^k_t = \frac{p_{x_{t+1}^k}}{p_{x^k_{t+1}+s^k_t}} \) is the fraction of their saving invested in the risky asset at time \( t \). Then\(^{18}\)

\[
\mu^k_t = 1 + \frac{q^k_t}{\beta \delta R^F_t e^{-g_t} - 1} + (1 - q^k_t) \frac{1}{\beta \delta R^F_t e^{-g_0} - 1}
\] (8)

Lemma 2 shows that the share of the risky asset in the investors’ portfolio linearly change with their beliefs about the next period’s growth. As we see next, a direct implication of this result, combined with the market clearing conditions, is that all investors can infer the wealth-weighted average belief of the other investors based on the equilibrium prices.

### 4.3 The Equilibrium Posterior of the Uninformed Investors

In this section, I discuss the dissemination of information from the informed investors to uninformed investors through the equilibrium prices. As the main result of this section, I show what determines the informativeness of the prices is the ratio between the wealth share of the informed and naive investors. Therefore, the gap between the informed and uninformed investors’ belief shrinks as either a larger fraction of the investors become informed, i.e. smaller \( \lambda_t \), or the informed investors become wealthier. A direct consequence of this result is that more inequality likely results in more informative prices and lower risk premium, consistent with the empirical findings of Bai, Philippon, and Savov (2016). Furthermore, toward the end of this section, I solve for the risk-premium and discuss its interaction with the wealth distribution.

To solve for the uninformed investors’ posterior, first, I extract the following identity from

---

\(^{18}\)The optimal portfolio is discontinuous at the extreme certainty points \( q^k_t \in \{0, 1\} \). The reason is, as long as each state has a positive probability, the investors do not take extreme positions in order to avoid reaching negative wealth values. In fact, an investor with a negative wealth gets an extremely low (\(-\infty\)) terminal payoff as he cannot borrow, due to the borrowing constraint, and hence, cannot consume.
the market clearance condition for the risky asset.

\[ 1 = \int_0^\lambda \mu_i^I w_i^I \, di + \int_1^{\lambda_c} \mu_i^I w_i^I \, di + \mu_i^N w_i^N \]  

(9)

Equation 9 states that the weighted average of the shares invested in the risky asset should add up to one, which is a direct consequence of the fact that the total savings in the economy should amount to the value of the tree.

By substituting \( \mu_k^i \)'s in (9) with the expression in (8), and after some rearrangements, we get:

\[
\left( q_i^U w_i^U + q_i^I w_i^I + q_i^N w_i^N \right) \frac{1}{\beta \delta R_t^F e^{-g_t} - 1} 
+ \left( (1 - q_i^U) w_i^U + (1 - q_i^I) w_i^I + (1 - q_i^N) w_i^N \right) \frac{1}{\beta \delta R_t^F e^{-g_t} - 1} = 0
\]  

(10)

The important implication of (10) is that all investors can infer the wealth-weighted average of the beliefs, i.e. \( \bar{q}_t \equiv w_i^U q_i^U + w_i^I q_i^I + w_i^N q_i^N \) from the risk-free rate \( R_t^F \). In fact, based on Equation 8, a hypothetical investor with posterior belief \( \bar{q}_t \) should have \( \mu \) of one, meaning he invests all of his saving in the risky asset, or equivalently, does not borrow or lend. For the rest of the paper, I refer to this hypothetical investor as the “representative investor”. An alternative interpretation of \( \bar{q}_t \) is that if the economy is replaced by a large representative investor, his belief should be \( \bar{q}_t \) in order to generate the same vector of prices.

As just mentioned, the belief of the representative investor can be inferred from the vector of prices. Therefore \( w_i^I q_i^I + w_i^N q_i^N \), and consequently, \( w_i^I z_i^I + w_i^N z_i^N \), are observable to the uninformed investors. Note that both \( w_i^I \) and \( w_i^N \) can be derived from \( F_i^P \), that is no equilibrium object used in \( w_i^I q_i^I + w_i^N q_i^N \). This point is formalized in Lemma 3.
Lemma 3.

For any \((z^I_t, z^N_t, \mathcal{F}_t^P)\), \(w^T_t z^I_t + w^N_t z^N_t\) can be inferred from the equilibrium prices. Moreover, all prices are measurable with respect to the \(\sigma\)-algebra induced by \((w^T_t z^I_t + w^N_t z^N_t, \mathcal{F}_t^P)\).

Building on this result, Lemma 4 provides a closed-form expression for the uninformed investors’ posterior, as a function of \(z^I_t\), \(z^N_t\) and the publicly observable objects.

Lemma 4.

\[ q^U_t = q(g_t) + z^U_t \]

where

\[ z^U_t = \frac{1}{2} \left[ \max\{-\bar{z}^I_t, z^I_t + \frac{w^N_t}{w^I_t} (z^N_t - \bar{z}^N_t)\}, \min\{\bar{z}^I_t, z^I_t + \frac{w^N_t}{w^I_t} (z^N_t + \bar{z}^N_t)\} \right] \]  \tag{11}

Lemma 4 provides the uninformed investors’ posterior as a function of the distribution of wealth at \(t\) and the exogenous signals \(z^I_t\) and \(z^U_t\). To get a more clear insight from (11), consider a case that \(\frac{w^N_t}{w^I_t}\) is sufficiently small, so that

\[-\bar{z}^I_t \leq z^I_t + \frac{w^N_t}{w^I_t} (z^N_t - \bar{z}^N_t) \leq z^I_t + \frac{w^N_t}{w^I_t} (z^N_t + \bar{z}^N_t) \leq \bar{z}^I_t \]  \tag{12}

In this case, the uninformed investors’ posterior boils down to \(q^U_t = q^I_t + \frac{w^N_t}{w^I_t} z^N_t\). We see that the difference between the informed and uninformed investors’ posterior is proportionate to the wealth ratio between the naive and informed investors, \(\frac{w^N_t}{w^I_t}\). That said, we can think of the ratio as a measure of noise-to-signal ratio in the vector of prices. Therefore, the prices become more informative about the future growth rate as a larger fraction of the investors become informed or the wealth share of the informed investors increases.

In the other extreme case, if \(\frac{w^N_t}{w^I_t}\) is sufficiently large, we have:
\[ z^I_t + \frac{w_t^N}{w_t^I}(z_t^N - \bar{z}^N) \leq -\bar{z}^I < z^I_t \leq z^I_t + \frac{w_t^N}{w_t^I}(z_t^N + \bar{z}^N) \]  

(13)

In this case, we have \( q_t^U = q(g_t) \), implying the prices contain no information about the future growth rates. Therefore, by reviewing the cases above, we see \( \frac{w_t^I}{w_t^N} \) determines the informativeness of the equilibrium prices. It means the wealth distribution plays an important role in the informativeness of the prices: The larger is the total wealth share of informed investors, the prices are more informative about the future growth rate. It is intuitive because the informed investors all trade in the direction that the informative signal implies. Hence, as they hold a larger fraction of the total wealth, they push the prices more, making the prices more reflective of \( z^I_t \) and hence, more informative about the future growth prospects.

After solving for the beliefs, we can find the risk-free rate by plugging the beliefs in Equation 10. In fact, one can show

\[ R^F_t = (\beta \delta)^{-1} \frac{1}{q_t e^{-gh} + (1 - q_t)e^{-gl}} \]  

(14)

Consequently, we can derive the risk premium:

\[ E_t[\frac{p_{t+1}}{p_t} + d_{t+1} - R^F_t] = \bar{q}_t(1 - \bar{q}_t) \left( \frac{e^{gh} - q_t}{\bar{q}_t e^{-gh}} - \frac{e^{-gl}}{\bar{q}_t} \right)^2 \]  

(15)

,where expectation \( E_t \) is taken with respect to the representative belief. One can see that the risk premium, computed in (15), is concave in \( \bar{q}_t \). Therefore, as the inequality increases and the wealth becomes more concentrated in the hand of the most sophisticated investors, the representative belief becomes more precise and the risk-premium decreases. This result
lends an informational justification for the negative relationship between the wealth share at top percentiles and risk-premium, documented by Gomez (2019).

Equation 15 also implies the risk premium decreases as the informative signal becomes more informative. Two forces push down the risk premium: First, the direct channel that a more informative signal reduces the investors’ uncertainty about the future growth rate, at least in average, leading to more accurate representative beliefs. The second channel is about its impact on the distribution of beliefs. In fact, it reduces the belief gap between the informed and uninformed investors. The reason being that the change widens the belief gap between the sophisticated and naive investors about future growth rates, decreasing the naive investors’ wealth share, and consequently, noise in the prices. Therefore, a more informative signal reduces the asymmetric information among the sophisticated investors and increases their wealth share, which leads to more accurate representative beliefs.

5 Evolution of the Wealth Distribution

This section includes the main results relating the information environment and wealth distribution. Lemma 5 and Corollary 1 describe the evolution of wealth shares, given the posteriors and the growth state \((g_t)\). Then, I provide some characterization of the stationary distribution in Proposition 1, where I discuss the determinants of the thickness of the right tail. Specifically, I show the thickness of the right tail increases as the informative signal \(\bar{z}_t^f\) become more informative. Finally, I discuss the distribution of expected returns and specify necessary conditions for it to be right-skewed.

**Lemma 5.**

*The wealth share of the investor in family \((f, k)\) \((f \in [0, 1] \text{ and } k \in [0, 1] \cup \{N\})\) at \(t + 1,\)
provided alive at \( t + 1 \), is:

\[
\log w_{f,t+1}^k - \log w_f^k = \begin{cases} 
\log \frac{q_k}{\bar{q}_t} & \text{if } g_{t+1} = g^h \\
\log \frac{1-q_k}{1-\bar{q}_t} & \text{if } g_{t+1} = g^l 
\end{cases} \quad \forall \ 1 \leq l \leq t \quad (16)
\]

Lemma 5 specifies the dynamics of an investor’s wealth share, as a function of his posterior belief \( q_k^t \), the representative belief \( \bar{q}_t \) and the realized growth rate \( g_t \). The right hand side in (16) specifies the log-ratio of the probabilities assigned by the investor of family \((f,k)\) and representative investors to the realized growth state \( g_{t+1} \). Therefore, Equation 16 states that an investor’s wealth share increases iff his posterior belief is more consistent with the realized growth rate, compared to the representative belief. For instance, if \( g_{t+1} = g^h \), the wealth share of the investor increases iff \( q_k^t > \bar{q}_t \), which means at \( t \), the investor was more optimistic than the representative investor about having a high growth rate at \( t + 1 \), and hence, allocated a larger fraction of his saving to the risky asset.

There are a few additional remarks on Equation 16. First, note that no equilibrium price is used in (16), which substantially simplifies our analysis, both analytically and numerically. In addition, one can show that Equation 16 is applicable for any belief process and we do not need to restrict to the set of beliefs formed by Bayes’ rule. Putting differently, the proof of Lemma 5 works for any distribution of beliefs and does not require any equilibrium condition on the cross-section of beliefs. It is effectively what I do in Section 7.1.1, where I construct the model-implied wealth dynamics with beliefs constructed based on Michigan Survey of Customers. Moreover, one can utilize this equation to study the distributional implication of non-Bayesian learning processes, such as that of Bordalo, Gennaioli, and Shleifer (2018). Nevertheless, such analysis are beyond the scope of this paper and relegated to future studies.
Now, equipped with the processes describing the wealth dynamics, we are ready to explore how the information environment impacts the wealth distribution and its dynamics. Section 5.1 examines the stationary distribution and Section A14. discusses the comparative statics, especially regarding the distributional implications of a change in the information environment.

5.1 Stationary Distribution

In this section, I study the stationary distribution of wealth shares and provide some characterizations of the stationary distribution of wealth. In particular, I mostly discuss the tail of the distribution, which captures the extent of top inequality. Before the characterizations, we first need to verify a stationary distribution indeed exists.

To verify the existence, we first need to introduce the Markov chain specifying the dynamics of the model. In particular, I show the distribution of wealth shares, combined with the growth state, constitute a Markov chain. Therefore, our state variables are:

\[
\Omega_t \equiv \{(w^i_t)_{i \in [0,1]}, w^N_t, g_t\} \in \Lambda \equiv L^1 \times (0, 1) \times \{g^h, g^l\}
\] (17)

In (17), \(L^1\) is the space of all measurable functions over \([0, 1]\) with a bounded integral. Note that the state variables here are infinite dimensional. I denote the random mapping that relates the time-\(t+1\) state \(\Omega_{t+1}\) to time-\(t\) state \(\Omega_t\) by \(Q(\Omega_t, B)\). In fact, it represents the probability of \(\Omega_{t+1}\) belonging to Borel subset \(B \in 2^\Lambda\), conditional on the time-\(t\) state \(\Omega_t\). Therefore a stationary distribution is a distribution over the set of distribution of wealth shares and the growth state.\(^{19}\) Corollary 1 can be used to specify the transition

\(^{19}\)Probability measure \(\mu : \mathcal{P}(\Lambda) \to [0, 1]\) is a stationary distribution of the Markov process if for every
Corollary 1.

Suppose $\kappa = (1 - \delta)\tau$. Then, the wealth share in the hand of type-$k$ investors evolves according to the following equation of motion:

$$w_{t+1}^k = \begin{cases} w_t^k \frac{q_t^k}{\bar{q}_t^k} (1 - \kappa) + \kappa & \text{if } g_{t+1} = g^h \quad \forall k \in [0, 1] \cup \{N\} \\ w_t^k \frac{1-q_t^k}{1-q_t^k} (1 - \kappa) + \kappa & \text{if } g_{t+1} = g^l \end{cases}$$

(18)

Note that Equation 18 introduces a mapping between the distribution of wealth shares at time $t$ and $t+1$, given the realizations $g_t, g_{t+1}, \lambda_t, z_t^I, z_t^N$. Since $\tilde{\lambda}_t$, $\tilde{z}_t^I$ and $\tilde{z}_t^N$ are i.i.d random variables, Equation 18 induces a random mapping between the state variables. After specifying the Markov chain and its transition process, we are ready to provide our existence result.

Lemma 6.

The wealth share of each family $(f, k)$ has a stationary distribution, only depending on type $k \in [0, 1] \cup \{N\}$.

Suppose the unconditional distribution of wealth shares for family $(f, i)$, with type $i \in [0, 1]$, has CDF $G^i(w) = \text{Prob}(w_{f,t}^i < w)$. Then, the “empirical” distribution of wealth shares is defined as:

$$G(w) = \int_0^1 G^i(w)di$$

Note that $G(w)$ is the unconditional fraction of the sophisticated investors with a wealth share less than $w$. The distribution is called empirical since it is the distribution that

Borel subset of $\Lambda$, like $B$, we have:

$$\mu(B) = \int_{\Omega} Q(\Omega, B)d\mu(\Omega)$$

30
an econometrician, with no knowledge of the types, would estimate for the population of sophisticated investors. Proposition 1 provides a characterization of the tail of the empirical distribution. In addition to the distribution, it also discusses the composition of types among the extremely wealthy and how it is different from the overall type composition.

**Proposition 1.**

*In any stationary distribution:*

**a)** The empirical distribution has a thick right tail and its tail parameter is the unique positive solution of the following non-linear equation:

\[
\lim_{T \to \infty} E\left[\left(\prod_{t=1}^{T} \left(\frac{q_t^I}{\bar{q}_t^I}\right)^{\gamma + 1} + \frac{(1 - q_t^I)^{\gamma + 1}}{(1 - \bar{q}_t^I)^{\gamma + 1}}\right)^\frac{1}{\gamma}\right] = (\delta + (1 - \delta)(1 - \tau)^\gamma)^{-1} \tag{19}
\]

Therefore, we have \(\lim_{w \to \infty} \frac{1 - G(w)}{w^{-\gamma}} \in (0, \infty)\), for some \(\gamma > 1\).\(^{20}\)

**b)** [*Type Composition in the Tail*] Suppose \(f(1) > 0\) and the types are private (hence random type \(\tilde{i}\) is uniformly distributed in \([0, 1]\)), then for any \(i < 1\) we have:

\[
\lim_{w \to \infty} P(\tilde{i} > i|w_{f,t}^\tilde{i} > w) = 1 \tag{20}
\]

Proposition 1(a) shows that under very general conditions, the empirical distribution has a thick right tail and the thickness is governed by the extent of asymmetric information between the most informed investors and the representative investor. To see this, note that the left hand side in (19) represents a measure of unconditional expected belief gap between the most sophisticated investors and the representative investor. In fact, loosely speaking, \(^{20}\)In fact, \(\gamma\) is the infimum all values like \(\hat{\gamma}\) for which there exists \(A(\hat{\gamma}) > 0\) such that \(G(w) > 1 - A(\hat{\gamma})w^{-\hat{\gamma}}\). Abusing the notation, I define \(A \equiv A(\gamma)\).
to compute the expected belief gap, a geometric variation of the law of large numbers is applied. Therefore, the information environment impacts the thickness parameter $\gamma$ only through the extent to which the most sophisticated investors have beliefs more accurate than the representative investor. Therefore, this simple measure can be used to gauge the impact of a change in information environment on wealth inequality, especially among the wealthiest.

Building on this result, Corollary 2 relates the tail parameter to the excess return of the most informed investors.

**Corollary 2.**

Let $r^I_t$ is the log-return of the most informed investors (sophisticated investors with type $i = 1$). Then, $\gamma$ is the unique positive solution to the following equation:

$$\lim_{T \to \infty} E[e^{\gamma \sum_{t=1}^{T}(r^I_t - g_t + \log \beta)}] \frac{1}{T} \times (\delta + (1 - \delta)(1 - \tau)^\gamma) = 1 \quad (21)$$

Corollary 2 shows that the tail parameter is related to $r^I_t - g_t$: The return of the most informed minus the growth rate. The intuition is $r^I_t - g_t$ captures how faster the wealth of the most informed grows compared to the whole economy. Note that this result resembles the well-known $r - g$ rule put forward by Piketty (2014), with only one exception. In Piketty’s rule, the return is the average return on capital, while here the return is the return of the most informed investors on their saving. That said, Corollary 2 underscores the role of asymmetric information in the top inequality and augments the Piketty’s rule to account for such heterogeneities in the population.

The proposition also indicates that the tax rate $\tau$ and the death rate $1 - \delta$ also play a counter-balancing role in the top inequality. Consistent with intuition, an increase in the tax rate or the death probability reduce the inequality, as they make the redistribution
larger and more frequent.

Part (b) in Proposition 1 specifies the composition of the wealthiest. In fact, it states that
the fraction of investors with type at most \( i \) and wealth share at least \( w \) goes to zero as \( w \) goes to infinity. In other words, as we focus on a narrower set of the top wealthiest, more
sophisticated investors are more likely to remain and less sophisticated ones asymptotically
vanish from the set. Therefore, the main message is that the tail is populated with highly
sophisticated investors. It explains the empirical observation that the individuals with a
high cognitive ability overrepresent among the wealthiest investors.

According to Equation (20), what distinguishes the investors and ranks them in the wealth
distribution is their sophistication level \( i \), not age. In other words, even though there are
plenty of less sophisticated but arbitrarily old investors that accumulate wealth for a long-
time, without getting hit by the death shock, they have a small chance to get to the top
percentiles.\(^{21}\)

Now, I discuss two implications of Proposition 1 for the distribution of expected returns in
Proposition 2.

**Proposition 2.**

\( a \) The expected return is increasing in wealth share, that is \( \mathbb{E}[r_{t+1}|w_t = w] \) is increasing
in \( w \).

\( b \) In any stationary distribution, the expected return of the most informed is bounded:

\[
\mathbb{E}[r_t^I - g_t] < -\log(1 - \delta)(1 - \tau) - \log \beta
\]

\(^{21}\) A key reason for this result could be the strong persistence in the types. When the types are highly
persistent, almost all less sophisticated families will be overtaken and fall substantially behind eventually
in the long-run. Therefore, we can expect a less extreme result to emerge when there are movements in the
types.
Proposition 2(a) implies that the distribution of expected returns is also right-skewed under certain conditions. The right-skewness is resulted from the combination of two different forces: First, the direct channel that some investors own a more accurate private signal.

The second force is the learning-channel, that is the equilibrium prices are partially informative about the informed investors’ signal. However, the informativeness depends on the fraction of investors learning the realization of the signal, which is governed by $\lambda_t$. When $\lambda_t$ is very large, there are only a few top investors that learn the realization of the informative signal. Hence, their demand would have a smaller impact on the equilibrium prices, and hence, there would be less spillover of information through the equilibrium prices. This causes a larger gap in the expected returns between the informed and uninformed investors when $\lambda_t$ is larger. Specifically, as $\lambda_t$ increases, the expected return gap between the ones with a type just above $\lambda_t$ and the ones just below $\lambda_t$ also increases. Therefore, the expected return differential is the highest in the tail of the distribution, as shown in Figure 4 and empirically observed by Fagereng, Guiso, Malacrino, and Pistaferri (2016).

Part (b) in Proposition 2 implies that the expected return of the most informed investors
in any stationary distribution is bounded by some parameters unrelated to the information environment. A direct implication is that an increase in the informativeness of signal $z_t$, while temporarily boost their return, their expected return eventually converges to a bounded level. In other words, the extent of return heterogeneity across the investors is limited, regardless of the information environment. The intuition is that as the wealth share of the informed investors increases, so does their impact on the prices, making it more informative. The more information disseminated through the prices the less expected excess return the informed investors. As a result, the wealth distribution becomes more and more skewed toward the more informed investors, up to an equilibrium point at which their expected returns balance out with their death and tax rates.

6 Endogenous Information Acquisition

In reality, economic agents are heterogeneously informed about the future returns mostly due to the costly nature of information. The cost is partly pecuniary, like that of obtaining access to datasets, or the cost of education, which is paid to earn the skills needed to process available signals (Lusardi, Michaud, and Mitchell (2017)). In this section, I endogenize the types by allowing the newborns to choose their type at a cost. Therefore, the investors choose their level of sophistication before they start trading. It can be thought of the decision to acquire permanent human capital to increase the expected returns on savings.

In short, I show that wealth inequality, captured by the tail parameter, decreases as the cost of information acquisition decreases. A direct implication of this finding is that an inequality-reducing policy is to subsidize education. In fact, wealth inequality is negatively
related with college tuitions among OECD countries, which lends support to this finding.\footnote{For details, see article “A debate is under way about the cost of higher education”, The Economist, July 18th 2019} Another important implication is that the emergence of novel, but highly expensive, ways for information production could also contribute to the rise of inequality, as only a few afford to directly or indirectly benefit from them. I formalize this point in the remainder of this section.

6.1 A Model of Wealth Dynamics With Costly Information Acquisition

In this section, I modify the setup in two aspects. First, I assume there are only two types of sophisticated investors, instead of a continuum of types: One type is the informed investors, who always learn the realization $z_t^I$, and the other type is the uninformed ones, who never learn the realization and need to imperfectly infer the signal from the equilibrium prices. Secondly, as mentioned earlier, I assume all newborns choose their type before they start trading. Particularly, they can become an informed investor for their entire lifetime by paying cost $(\beta \delta)^{-1} \chi p_t$ out of their endowed wealth, where $\chi > 0$. They remain uninformed if refrain from the payment. Then, all the payments are evenly distributed among all the newborns. Lemma 7 characterizes the newborns’ type acquisition decision.

Lemma 7.

A newborn with inherited wealth $w_{f,t} = (1 - \tau)\bar{w}_{f,t} + \tau$ becomes informed if

$$
\frac{1}{1 - \beta \delta} \log(w_{f,t} - \chi + h_t \chi) + \sum_{m=1}^{\infty} \frac{(\beta \delta)^m}{1 - \beta \delta} \mathbb{E}_t[q_{t+m-1}^I \log \frac{q_{t+m-1}^I}{q_{t+m-1}^U} + (1 - q_{t+m-1}^I) \log \frac{1 - q_{t+m-1}^I}{1 - q_{t+m-1}^U}] 
\geq \frac{1}{1 - \beta \delta} \log(w_{f,t} + h_t \chi)
$$

(22)
where $h_t$ is the fraction of newborns who decide to become informed at $t$.

Lemma 7 illustrates the trade-off a newborn faces in his information acquisition decision. On one hand, becoming informed increases the expected return on savings and increases his expected utility. On the other hand, there is a fixed cost associated with becoming informed, which is justified only for wealthy enough newborns. In fact, there is a threshold value $\bar{w}_t$ such that a newborn becomes informed iff $w_{f,t} \geq \bar{w}_t$. By inspecting (22), we see an increase in the cost of becoming informed, $\chi$, increases the threshold $\bar{w}_t$, and consequently, tends to reduce the fraction of informed investors. As a result, a higher $\chi$ widens the belief gap between the informed investors and representative investor, who takes the weighted-average belief of all investors. Due to this increase in the asymmetric information, the wealth inequality also generally increases with $\chi$. Proposition 3 formalizes this point.

**Proposition 3.**

Denote $\gamma(\chi)$ is the tail parameter of the empirical distribution corresponding to cost parameter $\chi$. Furthermore, suppose the following condition holds:

$$
\mathbb{E}[q_t^{l} \log \frac{q_t^{l}}{q(g_t)} + (1 - q_t^{l}) \log \frac{1 - q_t^{l}}{1 - q(g_t)}] + (1 - \delta) \log(1 - \tau) > 0 \tag{23}
$$

where $q(g_t) = P(g_{t+1} = g^h | g_t = q^j), j \in \{h, l\}$. Then, $\gamma(\chi) \to 1$, as $\chi$ goes to infinity.

Proposition 3 states that as the cost of information acquisition increases, the inequality also boundlessly increases, provided the informative signal is sufficiently informative. Note that $\gamma = 1$ is the lowest possible value for the tail parameter and designates the highest level of inequality. The main force leading to this result is the relationship between $\chi$ and the extent of asymmetric information between the informed investors and representative investor. As the cost of information acquisition increases, a smaller and smaller fraction of the investors opt to be informed. Therefore, the equilibrium prices continue to reflect less
of the informative signals, which levers up the expected return of the informed investors and the speed at which their wealth diverges from the uninformed investors, while the redistributive power of the taxation is fixed by parameter $\tau$.

7 Discussion

7.1 Empirical Relationships between Belief Heterogeneity and Wealth Inequality

There is a drastic heterogeneity in macroeconomic expectations, as demonstrated by several surveys of the general public and even those of professional forecasters (Mankiw, Reis, and Wolfers (2003)). Undoubtedly, the extent of this heterogeneity has a first order impact on the distribution of returns, and hence, the distribution of wealth. Equipped with the tools developed here, in this section, I quantify the extent of belief heterogeneity and evaluate their impact on the extent of wealth inequality and its dynamics.

More specifically, I construct the evolution of beliefs for different groups of population in the US, based on the rich dataset of Michigan Survey of Consumer and study their implication for the wealth distribution.\(^{23}\) In Section 7.1.1, I study the dynamic implications by utilizing Equation 18 and show the dynamics predicted by the model has a remarkable consistency with the estimated dynamics in the wealth shares over the past decades. In Section 7.1.2, I use Equation 19 to find the tail parameter implied by the beliefs and the model and show it lies in the reasonable range estimated in the empirical studies.

\(^{23}\)A more complete description of the survey is provided in Section 2.
7.1.1 Model-implied Evolution of Wealth Shares

My model delivers equation 18 that specifies the evolution of wealth shares as a function of the cross-section of beliefs and the realized growth rates. Therefore, we can find the model-implied evolution of the wealth distribution by feeding the beliefs obtained from the survey into the equation. However, since the data does not have information about the respondents’ wealth, I use their income as the proxy.

Figure 5 plots the model-implied and the actual time-series of wealth ratio between the top 1% and bottom 90%. For the model-implied series, I use the average belief among the top 1% and the bottom 90% in the distribution of the respondents’ income. Furthermore, I use the estimated wealth ratio at 1978 as the initial value. The time-series of the wealth ratio is taken from Saez and Zucman (2016). Finally, I set $\kappa = 0.005$.

Note that Figure (5) exhibits a significant co-movement between the model-implied and the actual wealth ratio, despite the fact that the model only allows for two states and the belief measures are arguably noisy. This observation is particularly important because the earlier models fall short in explaining the fast dynamics of the wealth distribution (Gabaix, Lasry, Lions, and Moll (2016)), while the figure suggests that the belief heterogeneity might, at least partly, explain the dynamics. Note that the key distinction between my model and the previous random growth models is that they ignore the role of aggregate shocks and hence, they mostly study the role idiosyncratic shocks in return or income in the wealth inequality. In contrast, there is no idiosyncratic shock in my model and I show in Proposition 6 that theoretically there is no limit on the speed of dynamics when the investors are asymmetrically informed. Therefore, the figure provides a suggestive evidence that the belief heterogeneity might have a first order effect on the dynamics of the wealth inequality.
Figure 5: Model implied vs. Actual wealth ratio between top 1% and bottom 90%, based on Michigan Survey Data for 1978-2012. The actual time-series of the wealth ratio is from Saez and Zucman (2016). To estimate the beliefs, the income proxy is used. The implied evolution is based on equation (16) for $\kappa = 0.005$.

7.1.2 Belief Heterogeneity and Implied Tail Parameter

Data suggests wealthier households tend to have more accurate beliefs about future macroeconomic conditions. Figure 2 and 3 provide supportive evidence by documenting a negative relationship between the expectation errors and two proxies of wealth levels: income level and total investment in stock market. Furthermore, Das, Kuhnen, and Nagel (2017) find that households with higher income levels have macroeconomic expectations closer to the professional forecasts, while low-income households have a larger bias. To understand the impact of such belief heterogeneity and differences in belief accuracy on wealth inequality, I use Equation 19 to estimate the tail parameter implied by the belief data.

To use the equation, we need to find estimates for the correct probabilities $q^I_t$ (conditional probabilities given the available information), the wealthiest average belief $q_{t}^{\text{Top5}}$ and the representative belief $\bar{q}_t$. I assume the most accurate belief is held by the top 5%, which
gives a lower bound on the extent of wealth inequality. Note that I am not assuming the wealthiest use all available information for their predictions, thus I allow for \( q^I_t \neq q^{Top\ 5}_t \). I construct the belief of the top 5% similar to the previous part. The representative belief is constructed by the income-weighted average of beliefs in Michigan Survey. Since the wealth inequality is larger than the income inequality in data, I weight the beliefs with incomes to the power of 1.5 \((\text{INCOME}^{1.5})\). Following Das, Kuhnen, and Nagel (2017), I construct the the true probabilities based on the Survey of Professional Forecasters. More specifically, I assume the probability of having a high growth (exceeding the historical mean) is equal to the fraction of the forecasters that predict a one-year-ahead growth rate above its historical mean.

Figure 6 exhibits the relationship between the professional forecasters’ belief with the belief of the top 5% and the representative belief, for years between 1978-2019 (overall 165 quarters). We see that the beliefs of the top 5% are more sensitive to the professional forecasts, compared to the weighted average beliefs. This is another indication of the fact that the wealthier investors, who are more likely to be among the top earners, have relatively more accurate beliefs.

Now, we are ready to estimate the tail parameter. The following equation provides an estimate for the tail parameter, given the constructed beliefs.

\[
\left\{ \prod_{t=1}^T \left( q^I_t \left( \frac{q^{Top\ 5}_t}{q_t} \right)^{\gamma} + \left( 1 - q^I_t \right) \left( \frac{1 - q^{Top\ 5}_t}{1 - q_t} \right)^{\gamma} \right) \right\}^{\frac{1}{T}} \left( \delta + (1 - \delta)(1 - \tau) \right)^{\gamma} = 1
\]

In (24), \( T = 165 \) is the number of periods in the dataset (1978q1-2019q1), \( \delta = 0.96 \) is calibrated to a 30-year expected investment period and \( \tau = 0.4 \) is used as is the federal marginal estate tax (above $1 million) in 2019. Based on these numbers, the model-implied tail parameter is \( \hat{\gamma} = 1.31 \), which is smaller than the empirical estimations of 1.48-1.55.
Figure 6: X-axis: Professional forecasters’ belief (1978q1-2019q1). It is the fraction of the forecasters that predicted the one-year-ahead growth rate would be more than the historical mean. Y-axis: The average belief of the top 5% in the income distribution (circles) and the income-weighted average of all beliefs (crosses). The lines represent the local regression lines.

(Vermeulen (2018); Klass, Biham, Levy, Malcai, and Solomon (2007)). It means the model predicts a slightly higher inequality in the tail, compared to data.

The estimation is arguably subject to different estimation errors and the choice of the parameter values. However, the key take-away from this exercise is that the extent of belief heterogeneity observed in data can generate the level of skewness observed in data under a reasonable set of parameters. In other words, this observation is a first evidence that the belief heterogeneity and asymmetric information are likely to be among the main contributors to the current level of top inequality. Therefore, it calls for policies improving the financial knowledge of the public and reducing the information inequality around the
existing investment opportunities.

### 7.2 Bequest Motives

An extensive literature highlights the role of bequest motive in wealth inequality (De Nardi (2004)). In this section, I allow the investors to have bequest motive for their intertemporal decisions. In short, I find the wealth dynamics remain unchanged if the investors are homogeneous in their bequest motive, that is all investors trade off their own and their child expected utility similarly. The intuition being that it impacts the consumption and saving decision of all investors in the same way and it does not affect the portfolio compositions.

I modify the utility function in (3) by adding a term $\Phi(\cdot)$ capturing the bequest motive. Equation (25) presents the modified utility function. In fact, $\Phi(W)$ is the utility an investor gets from leaving $W$ units of consumption goods, before tax, to his child.\(^{24}\)

$$
U\left(\{c_{f,t}^k\}_{t=s_1}^{s_2-1}, W_{s_2}\right) = \sum_{t=s_1}^{s_2-1} \beta^{t-s_1} \log c_{f,t}^k + \beta^{s_2-s_1} \Phi(W_{s_2}) \quad k \in [0,1] \cup \{N\} \quad (25)
$$

To maintain tractability, I assume $\Phi$ is logarithmic in the bequest wealth, i.e. $\Phi(W) = \phi \log W$. The following proposition shows how the bequest motives changes the consumption and saving behaviors and wealth dynamics.

**Proposition 4.**

**a)** For every $t \geq 1$, all investors consume fraction $\frac{1-\beta \delta}{1+\beta \phi(1-\delta)}$ of their wealth, i.e. $c_{f,t}^k = \frac{1-\beta \delta}{1+\beta \phi(1-\delta)} W_{f,t}^k$, where $f \in [0,1]$ and $k \in [0,1] \cup \{N\}$.

**b)** The distribution of the beliefs and evolution of wealth shares, specified in (16), are the\(^{24}\)Note that there is a one-to-one mapping between before-tax and after-tax inherited wealth.
same for all real values of $\phi$. 

Proposition 4(a) shows that the bequest motive changes the saving behavior of all investors uniformly. It means as $\phi$ increases, all investors scale their savings at the same proportion. This increase in demand for saving pushes up the asset prices, and subsequently, lowers the expected returns. Overall, this general equilibrium effect fully offsets the higher demand for saving, which renders $\phi$ irrelevant for the wealth distribution. This point is formalized in Part (b) of the proposition.

Note that Proposition 4 states the bequest motive has no role in wealth distribution when there is no heterogeneity in this motive for saving. However, a heterogeneity in the bequest motive can exacerbate or ameliorate wealth inequality, which is relegated to future studies.

7.3 The Case of Complete Markets

So far, we have assumed that the markets are dynamically incomplete, meaning that the set of available securities do not span the whole space of potential contingent transfers. For instance, a log-term risk-free bond cannot be manufactured with the one-period risk-free bonds and the shares of the tree. A natural question is how expanding the set of securities impact the allocations, and more specifically, the wealth distribution. Proposition 5 shows that the expansion has no impact on the wealth distribution.

**Proposition 5.**

*In the baseline setup described in Section 3, suppose the markets are dynamically complete, keeping the information structure unchanged. Then, Equations 16 and 18 still specify the wealth dynamics.*

Proposition 5 states expanding the set of securities does not change the wealth distribution
and its dynamics. The intuition is as follows: Standing at time $t$, all investors assign the same probability to the events after time $t+1$, once it is conditioned on the next period’s growth, $g_{t+1}$. In other words, the only state that they have disagreement about is $g_{t+1}$. In fact, in the proof, I show that all investors price all long-term assets (the ones that their value depends on a subset of $g_{t+2}, g_{t+3}, \ldots$) in the same way since all investors agree on the relative prices between any two consecutive periods. By combining this property in the baseline allocations and the abovementioned point, I conclude that all investors should assign the same price to all long-term assets, given the baseline allocation. Since the long-term assets are in zero net supply, except the tree, they are not traded in equilibrium. Therefore, the allocations remain intact.

8 Conclusion

I study a general equilibrium model with a continuum of long-lived and heterogeneously informed investors, who trade risky and risk-less claims of a Lucas tree in a financial market. I provide the equations fully specifying the joint evolution of the beliefs and wealth distribution. By exploiting those equations, I show that the expected distribution of wealth shares features a thick right-tail. As a key distinguishing feature of the informational channel, I show that the distribution of the expected returns is right-skewed as well, under some general conditions. As for the dynamic implication, the model suggests upward mobility is slower than the downward mobility. Finally, I endogenize the information environment by allowing the investors to costly acquire information about the future returns and show that the wealth share in the hand of top percentiles have an inverse U-shape relationship with the cost of information acquisition.
References


An, Li, Jiangze Bian, Dong Lou, and Donghui Shi, 2019, Wealth redistribution in bubbles and crashes, .


Campbell, John Y, Tarun Ramadorai, and Benjamin Ranish, 2018, Do the rich get richer in the stock market? evidence from india, Discussion paper, National Bureau of Economic Research.


Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, 2016, Heterogene-
ity and persistence in returns to wealth, Discussion paper, National Bureau of Economic Research.


Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen P Utkus, 2019, Five facts about beliefs and portfolios, *Available at SSRN 3336400*.

Girshina, Anastasia, 2019, Education and wealth accumulation: Evidence from sweden, *Available at SSRN 3440282*.


Hubmer, Joachim, Per Krusell, and Anthony A Smith Jr, 2018, A comprehensive quantitative theory of the us wealth distribution, .

Kacperczyk, Marcin, Jaromir Nosal, and Luminita Stevens, 2018, Investor sophistication and capital income inequality, *Journal of Monetary Economics*.


Kempf, Elisabeth, Alberto Manconi, and Oliver G Spalt, 2017, Learning by doing: The value of experience and the origins of skill for mutual fund managers, *Available at SSRN 2124896*.


Lei, Xiaowen, 2019, Information and inequality, *Journal of Economic Theory*.


Smith, Matthew, Owen Zidar, and Eric Zwick, 2019, Top wealth in the united states: New estimates and implications for taxing the rich, .


Vermeulen, Philip, 2018, How fat is the top tail of the wealth distribution?, *Review of Income and Wealth* 64, 357–387.


Appendix A: Proofs of Lemmas and Propositions

A1. Proof of Lemma 1

a,b) As mentioned earlier, the optimal strategy is scale-invariant, that is investors with the same type and different wealth levels consume and save in the same proportion. Since the investors have log-utility, the value function should be logarithmic in wealth level. Therefore, it is left to prove the coefficient in front of $\log W_t$ in the value function is $\frac{1}{1-\beta \delta}$.

I do so by guess and verification. Let $R_t^k$ be the random return on the optimal portfolio. Therefore, the optimality of the consumption implies:

$$V^k(W_t; F_t) = \max_{c_t} \log c_t + \beta \delta \mathbb{E}_t^k \left[ \frac{1}{1-\beta \delta} \log ((W_t - c_t)R_t^k) + a^k(F_{t+1}) \right]$$

$$= \max_{c_t} \log c_t + \frac{\beta \delta}{1-\beta \delta} \log (W_t - c_t) + \beta \delta \mathbb{E}_t^k \left[ \frac{1}{1-\beta \delta} \log R_t^k + a^k(F_{t+1}) \right]$$

(26)

Note that the last term in (26) is independent of $W_t$. Therefore, the first order condition for $c_t$ implies:

$$\frac{1}{c_t} = \frac{\beta \delta}{1-\beta \delta} \frac{1}{W_t - c_t} \Rightarrow c_t = (1-\beta \delta)W_t$$

By substituting $c_t = (1-\beta \delta)W_t$ in (26), we can easily verify the claim.

c) The market clearing condition for the good market implies:

$$d_t = c_t^N + \int_0^1 c_i^t di = (1-\beta \delta)(W_t^N + \int_0^1 W_i^t di)$$

$$= (1-\beta \delta)((p_t + d_t)(x_t^N + \int_0^1 x_i^t di) + s_t^N + \int_0^1 s_i^t di) = (1-\beta \delta)(p_t + d_t)$$
After a few rearrangements, we get:

\[ p_t = \frac{\beta \delta}{1 - \beta \delta} d_t \]

### A2. Proof of Lemma 2

According to Lemma 1, the investors maximize the expected log-wealth in period \( t + 1 \).

We can rewrite the optimization problem as follows:

\[
\max_{\mu} \quad \mathbb{E}_t^k [\log(\mu \frac{d_{t+1} + p_{t+1}}{p_t} + (1 - \mu) R_{t}^F)]
\]

(27)

According to Lemma 1(c), we can rewrite the optimization problem (27) as follows:

\[
\max_{\mu} \quad q_t^k \{\log e^{\gamma h} (\mu + (1 - \mu) \beta \delta e^{-\gamma h} R_{t}^F)\} + (1 - q_t^k) \{\log e^{\gamma h} (\mu + (1 - \mu) \beta \delta e^{-\gamma h} R_{t}^F)\}
\]

(28)

Due to the strict concavity of the objective function, we only need to solve for \( \mu \) that satisfies the first condition. One can verify that is the case for \( \mu_t^k \) specified in (8).

### A3. Proof of Lemma 3

Clearly \( q_t^l w_t^l + q_t^N w_t^N \) lies in the uninformed investors’ information set. Since \( q(g_t) = q_t^l - z_t^l = q_t^N - z_t^N \) and \( w_t^l + w_t^N = 1 - w_t^U \), \( z_t^U \equiv w_t^l z_t^l + w_t^N z_t^N \) should also be in the the uninformed investors’ information set, e.g. it is inferable from the equilibrium prices.

\[25\] To see this, note that the investors solve:

\[
\max_{x_{t+1}^{l}, z_{t+1}^{l}} \quad \log(1 - \beta \delta) W_t^k + \mathbb{E}_t^k \left[ \frac{\beta \delta}{1 - \beta \delta} \log W_{t+1}^k + \beta \delta a^k (\mathcal{F}_t) \right]
\]

which is equivalent to maximizing \( \mathbb{E}_t^k [\log W_{t+1}^k] \).
Similarly, both \( q_t^I w_t^I + q_t^N w_t^N \) and \((1 - q_t^I) w_t^I + (1 - q_t^n) w_t^n \) are inferable from \( \bar{z}_t^U \). Finally, \( R_t^F \) is the unique solution to (10). \( p_t \) is measurable with respect to \( \mathcal{F}_t^P \). Therefore, both equilibrium prices are measurable with respect to \( \mathcal{F}_t^P \cup \{ \bar{z}_t^U \} \).

A4. Proof of Lemma 4

Note that:

\[
q_t^U = \mathbb{E}[\mathbb{I}_{\{g_{t+1} = g^k\}} | \mathcal{F}_t^P, \bar{z}_t^U = w_t^I z_t^I + w_t^N z_t^N] = \mathbb{E}[\mathbb{E}[\mathbb{I}_{\{g_{t+1} = g^k\}} | \bar{z}_t^I] | \mathcal{F}_t^P, \bar{z}_t^U] = \mathbb{E}[\bar{z}_t^I | \mathcal{F}_t^P, \bar{z}_t^U] = \int_{-\bar{z}_t^I}^{\bar{z}_t^U} \frac{1}{2} \mathbb{I}_{\{\frac{1}{w_t^I} (\bar{z}_t^U - w_t^I z) \in [-\bar{z}_t^N, \bar{z}_t^N] \}} dz
\]

I apply the law of iterative expectations in the first line and the Bayes’ rule in the second line of the equations above.

A5. Proof of Lemma 5

The evolution of wealth for type-\( k \) investors with exposure to the risky asset \( \mu_t^k \) is given by:

\[
W_{t+1}^k = \beta \delta W_t^k (\mu_t^k \frac{d_{t+1} + p_{t+1}}{p_t} + (1 - \mu_t^k) R_t^F )
\]

Therefore, by applying Lemma 1(c), the evolution of the wealth share is:

\[
w_{t+1}^k = w_t^k (\mu_t^k + (1 - \mu_t^k) \beta \delta R_t^F e^{-g_{t+1}})
\]
Now, by some rearranging and substituting (8) for $\mu_t^k$, we have:

$$
w_{t+1}^k = w_t^k + (\mu_t^k - 1)(1 - \beta \delta R_t^F e^{-g_{t+1}})w_t^k
$$

$$
= w_t[1 + \left(\frac{q_t^k \beta \delta R_t^F e^{-g} - 1 + \kappa}{\beta \delta R_t^F e^{-g} - 1}\right)(1 - \beta \delta R_t^F e^{-g_{t+1}})]
$$

(29)

When $g_{t+1} = g^h$, we have:

$$
\frac{w_{t+1}^k}{w_t^k} = 1 - q_t^k \frac{\beta \delta R_t^F e^{-g^h} - 1 + \kappa}{\beta \delta R_t^F e^{-g^h} - 1} = q_t^k \frac{\beta \delta R_t^F (e^{-g^h} - e^{-g^l})}{\beta \delta R_t^F e^{-g^h} - 1}
$$

(30)

Note that Equation 30 holds for any agent, including the representative investor with belief $\bar{q} = w_t^N q_t^N + w_t^I q_t^I + w_t^U q_t^U$. Furthermore, the wealth share of this representative agent is one. Equations 10 and 8 imply that such representative agent always chooses $\mu_t = 1$ and Equation 29 implies that the wealth share of this representative agent is constant at one. Therefore, if we plug the belief and the wealth of the representative agent into (30), we get:

$$
\frac{\beta \delta R_t^F (e^{-g^h} - e^{-g^l})}{\beta \delta R_t^F e^{-g^h} - 1} = 1
$$

Therefore, we can rewrite (30) as follows:

$$
\frac{w_{t+1}^k}{w_t^k} = \frac{q_t^k}{q_t}
$$

For $g_{t+1} = g^l$, we can similarly show:

$$
\frac{w_{t+1}^k}{\kappa + (1 - \kappa)w_t^k} = \frac{1 - q_t^k}{1 - q_t}
$$

The specification in the Lemma directly follows from the equations above.
A6. Proof of Proposition 6

Note that:

\[
\mathbb{E}[\log \frac{w_{f_1,t+1}^{i_1}}{w_{f_2,t+1}^{i_2}}] = \mathbb{E}[\log \frac{w_{f_1,t}^{i_1}}{w_{f_2,t}^{i_2}} + q_t^I \log \frac{q_t^{i_1}}{q_t^{i_2}} + (1 - q_t^I) \log \frac{1 - q_t^{i_1}}{1 - q_t^{i_2}}] 
\]

\[
= \log \frac{w_{f_1,t}^{i_1}}{w_{f_2,t}^{i_2}} + \mathbb{E}[q_t^I \log \frac{q_t^U}{q_t^I} + (1 - q_t^I) \log \frac{1 - q_t^U}{1 - q_t^I} | \lambda_t \in (i_1, i_2)] (F(i_2) - F(i_1))
\]

(31)

From Lemma 4, we can see \( q_t^U | q_t^I, w_{i_1}^N / w_t^I \) is uniformly distributed in a subset of \( [q(g_t) - \bar{z}^I, q(g_t) + \bar{z}^I] \) \( g_t = g^j \) and, depending on the realizations, it could have a mass point at \( q(g_t) \). To simplify the expositions, suppose the distribution does not have a mass point.\(^{26}\)

Therefore, let \( \tilde{q} = q_t^U | q_t^I = q_t^I, w_{i_1}^N / w_t^I \sim U[q, \tilde{q}] \). In the next step, I find an upper bound for the expression below:

\[
\mathcal{H}(q^*; q, \tilde{q}) \equiv \mathbb{E}[q^* \log \frac{\tilde{q}}{q^*} + (1 - q^*) \log \frac{1 - \tilde{q}}{1 - q^*}] \quad \tilde{q} \sim U[q, \tilde{q}]
\]

(32)

By taking the expectation, we get:

\[
\mathcal{H}(q^*; q, \tilde{q}) = \frac{1}{q - \tilde{q}} [q^* (\log \tilde{q} - \log q - (\tilde{q} - q))
\]

\[
+ (1 - q^*) (1 - \frac{1}{2} \log(1 - q) - (1 - \tilde{q}) \log(1 - \tilde{q}) - (\tilde{q} - q))]
\]

\[
- q^* \log q^* - (1 - q^*) \log(1 - q^*)
\]

(33)

Now, let us use the Taylor series to approximate \( \mathcal{H}(q^*; q, \tilde{q}) \) around \( a^* \) up to the third order.

After some simplifications, we get:

\(^{26}\)The proof steps are exactly the same for the case with a mass point in the distribution.
\[ H(q^*; \bar{q}, \bar{q}) = -\frac{1}{6}(\frac{1}{q^*} + \frac{1}{1-q^*})[(\bar{q}-q^*)^2+(q^*-\bar{q})^2+(\bar{q}-q^*)(q^*-\bar{q})] + o((\max\{\bar{q}-q^*, q^*-\bar{q}\})^3) \]  

(34)

According to Lemma 4, the minimum possible value that the bracket above takes is \((\frac{w_N^N z^N}{w_t^N})^2\). Furthermore, note that according to Corollary 1, \(w_t^N \geq \kappa b\) for all \(t \geq 1\). Therefore, \(\frac{w_N^N}{w_t^N} \geq \frac{w_N^N}{1-w_N^N} > \frac{\kappa b}{1-\kappa b}\). The proof is completed by substituting these expressions in (31).

A7. Proof of Lemma 6

I prove the lemma in two steps.\(^{27}\) First, I prove the lemma for the case that there is only a finite number of types among the sophisticated investors. Then, using this result, I construct a sequence of approximating distributions of wealth shares for every investor by restricting \(\lambda_t\) to take value from \(\frac{j}{M}\), where \(M\) is a positive integer, and taking \(M\) to infinity. This first step yields a set of approximating distributions of wealth shares for each type of investors \(k \in [0,1) \cup \{N\}\). Then, in the second step, I show the approximating distributions have a limit, which yields the stationary distributions.

Proof for the case with finite sophisticated type: Suppose \(\lambda_t\) only takes value from finite set \(\{0, \frac{1}{M}, \ldots, \frac{M-1}{M}\}\), for some positive integer \(M\). Furthermore, suppose the probability the threshold type being \(\frac{j}{M}\), \(0 \leq j \leq M-1\), is \(F(\frac{j+1}{M}) - F(\frac{j}{M})\). Therefore, the probability that investor \(i \in [0,1]\) becomes informed is \(1 - F(\frac{|M|}{M})\). I will take \(M\) to infinity in the next step.

In this case the state variable is finite-dimensional and the transition kernel can be derived

\(^{27}\)Note that \(\Lambda\) is not totally ordered, therefore we cannot directly apply the existence and uniqueness theorem of Hopenhayn and Prescott (1992).
similar to (18). In particular, in this case, the state of state variables is as follows:

$$\Omega^M_t = \{\{\tilde{w}_j^i\}_{1 \leq j \leq M}, \tilde{w}_N^i(M), g_t\} \in \Lambda^M = [0, 1]^{M+1} \times \{g^l, g^h\}$$

$$\tilde{w}_j^i = \int_{i - 1}^{i} w_i^j di$$

where $\tilde{w}_j^i$ is the total wealth share in the hand of group $j$ and $\tilde{w}_N^i(M)$ is the wealth share of the naive investors under this information structure. Denote the corresponding transition kernel by $Q^M(\cdot, \cdot)$. Note that $Q^M(\cdot, \cdot)$ is defined on the distributions over compact set $\Lambda^M$. Due to continuity of the mapping with respect to the distribution, the process also satisfies the Feller property. Therefore, it has a stationary distribution.

For the next step, denote the implying the stationary distribution for the wealth share of the investors with type $k \in [0, 1] \cup \{N\}$ by $\zeta^M$.

**Existence of stationary distribution for every** $k \in [0, 1] \cup \{N\}$. The goal in this step is to show the sequence $\zeta^M_i$ is converging to some distribution $\zeta^*_i$. To show this, we first note that the expectation of all of these distributions is bounded by $\frac{1}{1-i}$; because the expected wealth share for type-$i$ investors, $\int w d\zeta^M_i(w)$, is weakly increasing in $i$, as the investors with a higher $i$ receive a more informative signal in Blackwell sense. The uniformly boundedness of the expectations implies the sequence is tight. Therefore, according to Prohorov’s theorem, a subsequence of this collection of distributions, $\{\zeta^M_i\}_{M=1}^\infty$ should converge to some distribution $\zeta^*_i$. Therefore, $\zeta^*_i$ is a stationary distribution for the converging sequence of type-$i$ wealth shares, which is $w^I_i$. It completes the proof.

Note that this proof can be applied to the distribution of wealth shares for any subset of investors or the joint distributions. As a result, the joint distribution of the wealth share of the informed investors $w^I_i$ and naive investors also follows a stationary distribution. It
is enough to show the beliefs also have a stationary distribution.

A8. Proof of Proposition 1

Lemma 8.

For any $i \in [0, 1]$, we have $\mathbb{E}[q_t^i \log(q_t^i) + (1 - q_t^i) \log(1 - q_t^i)] \leq -\log(1 - \kappa).$ \footnote{\( \kappa \equiv (1 - \delta)\tau \)}

Proof. To see the argument, note that Corollary 1 implies:

$$\log w_{i,t+1}^i - \log w_i^i > (\log(q_t^i))I\{g_{t+1} = g^h\} + (\log(1 - q_t^i))I\{g_{t+1} = g^l\} + \log(1 - \kappa) \quad \forall \ i \in [0, 1]$$

Therefore, if the inequality in Lemma 8 does not hold for some $i \in (0, 1)$, then the wealth share of the investors with type above $i$ would boundlessly increase and hence, it would exceed one, which is in contradiction with $\int_0^1 w_i^i di$ being bounded. \qed

The wealth share of family $(f,i)$ follows the following Markov process:

$$w_{f,t+1}^i = \begin{cases} b_t^k w_{f,t}^i & \text{w.p. } \delta \\ b_t^k (1 - \tau) w_{f,t}^i + \tau & \text{w.p. } 1 - \delta \end{cases} \quad (35)$$

where

$$b_t^i = \begin{cases} q_t^i & g_{t+1} = g^h \\ 1 - q_t^i & g_{t+1} = g^l \end{cases} \quad (36)$$

Lemma 6 implies that we can write the process in the form of $w_{f,t+1}^i = B_t^i w_{f,t}^i + A_t^i$, where $(B_t^i, A_t^i)$ follows a Markov process. In Corollary 8, we see that $\mathbb{E}[\log B_t^i] \leq 0$. Therefore
it is a Kesten process (Kesten (1973)) and its stationary distribution has a thick right tail if \( P(B_t > 1^i) > 0 \) (Roitershtein et al. (2007)), with the tail parameter \( \gamma(i) \) that satisfies:

\[
\lim_{T \to \infty} \left( \mathbb{E}[\Pi_{i=1}^T \mathbb{1}_{B_t^i}] \right)^{\frac{1}{T}} = \lim_{T \to \infty} \mathbb{E}[\Pi_{i=1}^T (\lambda_i^i)^\gamma(i)]^{\frac{1}{T}} (\delta + (1 - \delta)(1 - \tau)^\gamma(i) = 1
\]

(37)

\[
\Rightarrow \lim_{T \to \infty} \mathbb{E}[q_t^I (\frac{q_t^I}{\bar{q}_t})^\gamma(i) + (1 - q_t^I)\frac{1 - q_t^I}{1 - \bar{q}_t})^\gamma(i)]^{\frac{1}{T}} (\delta + (1 - \delta)(1 - \tau)^\gamma(i) = 1
\]

, where in (37), I used the fact that \( (\bar{\lambda}_t, z_t^t, z_t^N) \) are drawn independently across the periods and \( q_t^h = 1 - q_t^I \). For \( \gamma(i) \) that satisfies (37) we have \( P(w_{j,t}^i > w) > Q(i)\bar{w}^{-\gamma(i)} \), for some \( Q(i) > 0 \) and sufficiently large \( \bar{w} \).

Note that \( \gamma(i) \) is decreasing in \( i \), because for any \( \gamma > 1 \) and \( i_2 > i_1 \in [0,1] \), we have:

\[
\mathbb{E}[q_t^I (\frac{q_t^I}{\bar{q}_t})^\gamma + (1 - q_t^I)\frac{1 - q_t^I}{1 - \bar{q}_t})^\gamma] > \mathbb{E}[q_t^I (\frac{q_t^U}{\bar{q}_t})^\gamma + (1 - q_t^I)\frac{1 - q_t^U}{1 - \bar{q}_t})^\gamma]
\]

\[
\Rightarrow \mathbb{E}[q_t^I (\frac{q_t^I}{\bar{q}_t})^\gamma + (1 - q_t^I)\frac{1 - q_t^I}{1 - \bar{q}_t})^\gamma] > F(i_2)\mathbb{E}[q_t^I (\frac{q_t^U}{\bar{q}_t})^\gamma + (1 - q_t^I)\frac{1 - q_t^U}{1 - \bar{q}_t})^\gamma]
\]

, where the first inequality is achieved from the fact that \( q_t^I (\frac{q_t^I}{\bar{q}_t})^\gamma + (1 - q_t^I)\frac{1 - q_t^I}{1 - \bar{q}_t})^\gamma \) is a convex function of \( \gamma \) and the distribution of \( q_t^I \) is a mean-preserving spread of that of \( q_t^U \).

Therefore:

\[
1 - G^i(w) = \int_0^1 P(w_{j,t}^i \geq w) di > \int_0^1 Q(i)w^{-\gamma(i)} di \geq \bar{Q}w^{-\gamma(1)} \quad \forall \bar{Q} > 0
\]

where the last inequality is obtained from the fact that \( \gamma(1) \leq \gamma(i) \), for \( i \in [0,1] \), and Equation 5 in Gabaix (2009). Therefore, \( 1 - G(w) \sim w^{-\gamma(1)} \), which completes the proof.
**Part b.** By using the derivations in the previous part and applying Bayes’ rule on the unconditional distributions, we get:

\[
P(\tilde{i} < i | w_{f,t}^\tilde{i} > w) = \frac{\int_0^i P(w_{f,t}^{i'} > w) di'}{\int_0^1 P(w_{f,t}^{i'} > w) di'} < \frac{L_1 w^{-\gamma(i)}}{L_2 w^{-\gamma(1)}}
\]

where \( L_1 \) and \( L_2 \) are some positive real numbers. Since \( \gamma(1) < \gamma(i) \), the statement can be proven by taking \( w \) in (38) to infinity.

**A9. Proof of Proposition 2**

**Proof of Part (a)**

Note that according to Lemma 1(b),

\[
R_t^i = \frac{W_t^{i+1}}{\beta W_t^i} = \beta^{-1} e^{\eta t+1} b_t^i,
\]

where \( b_t^i \) is defined in (36). Therefore, we only need to show \( \mathbb{E}[\log b_t^i] \) is increasing and convex in \( i \). We can see that the derivative is the following:

\[
\frac{d}{di} \mathbb{E}[\log b_t^i] = \mathbb{E}[q_t^i \log \frac{q_t^i}{q_t^{i'}} + (1 - q_t^i) \log \frac{1 - q_t^i}{1 - q_t^{i'}} | \lambda_t = i]
\]

The function inside the expectation in (39) is convex in \( q_t^i \) and the minimum is achieved at \( q_t^i \). Therefore, it is strictly positive.

Now, I show it is also increasing in \( i \). Since the dissemination of information decreases as \( \lambda_t \) increases, \( \mathbb{E}[|q_t^i - q_t^U| | \Omega_t, \lambda_t] \) also increases in \( \lambda_t \). Therefore, for any \( \Omega_t \in \Lambda, \mathbb{E}[q_t^i \log \frac{q_t^i}{q_t^{i'}} + (1 - q_t^i) \log \frac{1 - q_t^i}{1 - q_t^{i'}} | \Omega_t, \lambda_t = i] \) is increasing in \( i \). By appealing to the law of iterative expectations with respect to \( \Omega_t \), we see the derivative in (39) is increasing in \( i \).

**Proof of Part (b)**

Note that the exponential function is convex and \( \gamma > 1 \). Therefore, applying the Jensen
inequality to (21), we get:

\[ \exp(\gamma \mathbb{E}[r_t^I - g_t + \log \beta])(\delta + (1 - \delta)(1 - \tau)^\gamma) < 1 \]

Therefore, by taking log from both sides and noting \( \gamma \geq 1 \):

\[ \gamma \mathbb{E}[r_t^I - g_t + \log \beta] + \gamma \log(1 - \delta) + \gamma \log(1 - \tau) < \mathbb{E}[r_t^I - g_t + \log \beta] + \log(\delta + (1 - \delta)(1 - \tau)^\gamma) < 0 \]

Since \( \gamma > 0 \), it completes the proof.

A10. Proof of Proposition 3

Similar to the baseline case, the tail is governed by the informed investors. Denote the stationary distribution for a given \( \chi \) by \( G_\chi(\cdot) \). Consider the contrary and suppose the tail parameter converges to some \( \gamma^* > 1 \). Therefore, for sufficiently small values of \( \varepsilon > 0 \), there exists positive number \( A(\varepsilon) \) such that:

\[ 1 - G_\chi(w) < A(\varepsilon)w^{-(\gamma^* - \varepsilon)} \quad (40) \]

Inequality 40 implies that the fraction of investors affording to pay the information cost \( \chi \) goes to zero as \( \chi \) goes to infinity. That said, the wealth share of the informed investors share should converge to zero. Therefore, the representative belief converges to \( q(g_t) \), almost surely. Due to the convergence, according to condition (23), there exists a sufficiently large \( \chi \) and positive number \( A \) such that

\[ \mathbb{E}_q[q_t^I \log \frac{q_t^I}{q_t} + (1 - q_t^I) \log \frac{1 - q_t^I}{1 - q_t} + (1 - \delta) \log(1 - \tau)] > A > 0 \]

A-11
Note that we have the following inequality for the wealth share of an informed investor, like $f$:

$$
\mathbb{E}[\log w_{f,t+1}] - \log w_{f,t} > \delta \mathbb{E}_t[\log \frac{q^f_t}{\bar{q}_t} (g_{t+1})] + (1 - \delta) \mathbb{E}_t[\log \frac{q^f_t}{\bar{q}_t} (g_{t+1})] + (1 - \delta) \log(1 - \tau) \\
= \mathbb{E}_t[q^f_t \log \frac{q^f_t}{\bar{q}_t} + (1 - q^f_t) \log \frac{1 - q^f_t}{1 - \bar{q}_t}] + (1 - \delta) \log(1 - \tau) > \mathcal{A}
$$

It means for such $\chi$, the wealth share of the informed investors boundlessly increases, which is a contradiction.

### A11. Proof of Proposition 4

For the proof, we only need to incorporate the bequest motive into the baseline value function, specified in (26). Similar to the proof of Proposition 1, we can guess and verify the value function is logarithmic in wealth and is in the following form:

$$
V^k_\phi(W_t; F_t) = \frac{1 + \beta(1 - \delta) \phi}{1 - \beta \delta} \log W_t + a^k_\phi(F_t) \tag{41}
$$

The optimal consumption can be similarly solved and verified that $c_t = \frac{1 - \beta \delta}{1 + \beta \delta(1 - \delta)}$. Therefore, similar to the baseline case, the investors optimally consume a fixed fraction of their wealth in every period. Moreover, due to the logarithmic form of the value function, the investors optimally maximize the expected log-return of their investment, similar to the baseline case. Therefore, a change in $\phi$ does not affect the composition of portfolios, and as a result, does not impact the information dissemination through the prices. Therefore, by following the procedure carried out in the baseline case, we can find the wealth and
belief dynamics and prove part (b).

**A12. Proof of Proposition 7**

Note that the tail parameter is the solution to the following equation:

$$\lim_{T \to \infty} \mathbb{E}[\left( \frac{q_{f,t}(g_{t+1})}{q(g_{t+1})} \right)^{1/\gamma}(\delta + (1 - \delta)(1 - \tau)\gamma)] = 1$$  \hspace{1cm} (42)

where $q_{f,t}$ is the probability that the investor from family $f$ assigned on the realized growth $g_{t+1}$ at time $t$.\(^{29}\) A higher persistence in types means that consecutive terms in sequence $\{q_{f,t}\}^\infty_{t=1}$ have a higher correlation. Therefore, one can show that the expectation of the consecutive terms in the sequence, $\mathbb{E}(\prod_{j=0}^T q_{f,t+j}(g_{t+j+1}))$, is increasing in $\rho$ by applying law of iterated expectation and noting the wealth share of the informed investors increases with $\rho$.

A higher $\rho$ implies that the informed investors can save more and, eventually, claim a larger fraction of the wealth, which implies the representative belief becomes more accurate. It leads to an increase in $\mathbb{E}(\prod_{j=0}^T q_{t+j}(g_{t+j+1}))$ for every $T$. However, if the increase is so large that $\mathbb{E}[\prod_{j=0}^T q_{f,t+j}(g_{t+j+1})] \neq \mathbb{E}[\prod_{j=0}^T q_{t+j}(g_{t+j+1})]$ decreases, it means the expected wealth share of the sophisticated investors should decrease in $\rho$, or equivalently, the wealth share of the naive investors should increase with $\rho$, which is not possible.

**A13. Proof of Lemma 7**

The expected utility of a newborn from choosing to be informed at time $t$ is:

\(^{29}\)Note that, in this case, the investors are homogeneous, ex-ante.
\[ V^I_t(w_{f,t}) = E_t \sum_{j=0}^{\infty} \beta^j \log((1 - \beta \delta) \frac{d_{t+j}}{1 - \beta \delta} (w_{f,t} - \chi + h_t \chi) \prod_{j'=0}^{j-1} \frac{q^I_{t+j'}(g_{t+j'+1})}{q^I_{t+j'}(g_{t+j'+1})} \]

\[ = \sum_{j=0}^{\infty} (\beta \delta)^j E_t[\log d_{t+j}] + \frac{1}{1 - \beta \delta} \log(w_{f,t} - \chi + h_t \chi) \]

\[ + \sum_{m=1}^{\infty} (\beta \delta)^m \frac{1}{1 - \beta \delta} E_t[q^I_{t+m-1} \log \frac{q^I_{t+m-1}}{q^I_{t+m-1}} + (1 - q^I_{t+m-1}) \log \frac{1 - q^I_{t+m-1}}{1 - q^I_{t+m-1}} | \Omega_t] \]

Inequality (43) can be verified by comparing the equations above.

\[ V^U_t(w_{f,t}) = E_t \sum_{j=0}^{\infty} \beta^j \log((1 - \beta \delta) \frac{d_{t+j}}{1 - \beta \delta} (w_{f,t} + h_t \chi) \prod_{j'=0}^{j-1} \frac{q^U_{t+j'}(g_{t+j'+1})}{q^U_{t+j'}(g_{t+j'+1})} \]

\[ = \sum_{j=0}^{\infty} (\beta \delta)^j E_t[\log d_{t+j}] + \frac{1}{1 - \beta \delta} \log(w_{f,t} + h_t \chi) \]

\[ + \sum_{m=1}^{\infty} (\beta \delta)^m \frac{1}{1 - \beta \delta} E_t[q^U_{t+m-1} \log \frac{q^U_{t+m-1}}{q^U_{t+m-1}} + (1 - q^U_{t+m-1}) \log \frac{1 - q^U_{t+m-1}}{1 - q^U_{t+m-1}} | \Omega_t] \]

Inequality (22) can be verified by comparing the equations above.

A14. Proof of Proposition 5

To prove the proposition, I show for any sequence of growth rate realizations \( G = (g^*_{t+1}, g^*_{t+2}, \ldots, g^*_{t+N}) \), all investors assign the same price to the security that pays \( \varepsilon \) at beginning of \( t + N \) following \( G \), under the allocation provided for the baseline case. In particular, we need to
\[(\beta \delta)^N \mathbb{E}^k \left[ \frac{u'(c(G))}{c_t} \mathbb{I}_{\left\{ (g_{t+1}, g_{t+2}, \ldots, g_{t+N}) = G \right\}} \right] = (\beta \delta)^N \mathbb{E}^{k'} \left[ \frac{u'(c(G))}{c_t} \mathbb{I}_{\left\{ (g_{t+1}, g_{t+2}, \ldots, g_{t+N}) = G \right\}} \right] \quad \forall k, k' \in [0, 1] \cup \{N\} \]

(45)

where \(c(g_{t+1}, g_{t+2}, \ldots, g_{t+K})\) denotes the consumption under the baseline allocation after the realization \((g_{t+1}, g_{t+2}, \ldots, g_{t+K})\) and \(u(\cdot) \equiv \log(\cdot)\). In the baseline, the allocations between any two consecutive periods are optimal, given the information available at the time of consumption and portfolio decision. Therefore, for any \(1 \leq j \leq N - 1\) and any type of investor \(k \in [0, 1] \cup \{N\}\), we have:

\[\beta \delta \frac{u'(c(\ldots, g_{t+j}', g_{t+j+1}'))}{u'(c(\ldots, g_{t+j}'))} P^k_{t+j} (\tilde{g}_{t+j+1} = g_{t+j+1}^*) = 1\]

(46)

Now, by multiplying all of these Euler equations, we get:

\[(\beta \delta)^N \frac{u'(c(G))}{u'(c(g_{t+1}^*)))} \prod_{j=1}^{N-1} P^k_{t+j} (\tilde{g}_{t+j+1} = g_{t+j+1}^*) = 1\]

(47)

After taking time \(t\) expectation, we have:

\[(\beta \delta)^N \frac{u'(c(G))}{u'(c(g_{t+1}^*)))} \prod_{j=1}^{N-1} P(\tilde{g}_{t+j+1} = g_{t+j+1}^* | g_{t+j}^*) = 1\]

(48)

By multiplying the last equation with the Euler equation at time \(t\), we have:

\[(\beta \delta)^N \frac{u'(c(G))}{u'(c_{t})} P^k_{t} (\tilde{g}_{t+1} = g_{t+1}^*) \prod_{j=1}^{N-1} P(\tilde{g}_{t+j+1} = g_{t+j+1}^* | g_{t+j}^*) = (\beta \delta)^N \frac{u'(c(G))}{u'(c_{t})} P^k_{t} (G) = 1\]

(49)
Equation 49 shows that all investors assign the same price to the security that pays contingent on $\mathcal{G}$. Since this security is in zero net supply, it would not be traded. Therefore, the dynamic incompleteness imposed in the baseline does not change the allocations.
Appendix B: Additional Results

B1. Dynamic Implications of Asymmetric Information

Overall, three elements of the information environment impact the wealth distribution. First, the precision of available signals about the future states. Second, the allocation of available signals among the economic agents, capturing how accessible, or even understandable, they are for the investors. Third, the non-fundamental noises that enter into the prices and hamper the dissemination of information through the prices. In this section, I investigate the role of these three elements in wealth distribution.

To this end, I generate some economic insights by analyzing the wealth gaps resulted from information gaps, or sophistication differences, between pairs of investors. I do this by combining Lemma 4, which characterizes the distribution of posteriors, with Lemma 5, which characterizes the wealth dynamics as a function of the posteriors. Combining these two results, Proposition 6 provides a helpful characterization of the dynamics of the wealth gaps, for a given information environment. Then, I employ the lemma to study how the tail parameter is affected by the these changes in the information environment.

Proposition 6.

Suppose $q_t^i \leq 1 - \frac{nb}{1-nb}z^N$ with probability one and $q(g_t)$ is the conditional probability given the growth state ($g_t = g^j$). Furthermore, suppose the sophisticated investors in families $(f_1, i_1)$ and $(f_2, i_2)$ have wealth shares $w_{f_1, t}^i$, $i_2 > i_1$, and $w_{f_2, t}^i$ at $t$ respectively. Then, provided neither of them receive the death shock at $t + 1$, we have:

$$-\frac{1}{3} \left( \frac{w_t^N}{\int_{i_2} w_t^j dj} \right)^2 \leq \frac{\mathbb{E}_t[\log \frac{w_{f_1, t+1}^i}{w_{f_2, t+1}^i}] - \log \frac{w_{f_1, t}^i}{w_{f_2, t}^i}}{(\log \frac{q(g_t) + z^j}{q(g_t) - z^j}) (1 - q(g_t) + z^j)(F(i_2) - F(i_1))(z^N)^2} \leq -\frac{1}{12} \left( \frac{w_t^N}{\int_{i_1} w_t^j dj} \right)^2 \quad (50)$$
Inequality 50 provides a lower and upper bound for the speed of wealth divergence among the sophisticated investors. We see the wealth ratio between a less informed investor (here \((f_1, i_1)\)) and a more informed investor (here \((f_2, i_2)\)) decreases in expectation, as long as they survive. Moreover, the expected amount of the decrease depends on all three factors mentioned earlier: Informativeness of the available signal, captured by \(\bar{z}^I\), the allocation of the informative signal, captured by CDF \(F(\cdot)\) and the non-fundamental noise, captured by \(\bar{z}^N\). Furthermore, one can show that the bounds in (50) are bounded for any \(i_1\) and \(i_2\) since \(w^N_t > \kappa b\) and \(w^I_t > \kappa\) for any \(i \in [0, 1]\).\(^{30}\) Now, I discuss these channels in turn.

Regarding the role of the informative signal, inequality 50 implies that a sufficiently large increase in the informativeness of the signal widens the wealth gap, and hence, increases the inequality. We can see this by noting that the expression \(\log\frac{q(g_t) + \bar{z}^I}{q(g_t) - \bar{z}^I}\) can take any arbitrarily large value, while the bounds in (50) are bounded. This shows as the available information resources become more informative, the information inequality, namely the investors having unequal access to such resources, becomes more pronounced in the wealth inequality. Strikingly, the wealth inequality increases even though a more informative signal also makes the prices more informative, which reduces the belief gap between the informed and uninformed investors.\(^{31}\)

To see the intuition, note that the information and beliefs impact the wealth distribution through the portfolio decisions. Facing less uncertainty, the investors take more aggressive positions and react to their beliefs more strongly when the signal is more informative.

\(^{30}\)Every type of investors have total wealth share of at least \(\kappa\) per unit mass, because the newborns constitute fraction \(1 - \delta\) of each type of investors and each newborn has the wealth of share at least \(\tau\). As a result, the upper bound in (50) is bounded by \(-\frac{1}{12}(\frac{\kappa b}{1 - \bar{a}^2})^2\) and the lower bound is bounded by \(-\frac{1}{3}(\frac{1 - \kappa(1 - i_2)(1 - b)}{1 - \kappa(1 - i_2)(1 - b)})^2\).

\(^{31}\)To see this, note that a more informative signal increases the belief gap between the sophisticated and naive investors. Therefore, loosely speaking, the wealth share of the naive investors decreases, which implies \(w^N_t\) and hence \(|q^N_t - q^I_t|\), also decrease.
This can be well illustrated in the investors’ optimal portfolio formula, specified in (8). According to the equation, the sensitivity of an investor’s portfolio ($\mu_k^t$) to his belief ($q_k^t$) is given by the expression below:

$$\frac{\partial \mu_k^t}{\partial q_k^t} = \frac{1}{\beta \delta R^F_t e^{-g_t} - 1} - \frac{1}{\beta \delta R^F_t e^{-g_h} - 1}$$

Note that the return on the risky asset is $\frac{p_{t+1} + d_{t+1} + p_t}{p_t} = (\beta \delta)^{-1} e^{g_{t+1}}$. Moreover, as shown earlier, the risk premium also decreases as the signal becomes more informative, leading to an increase in $\frac{\partial \mu_k^t}{\partial q_k^t}$. Therefore, the same belief dispersion results in a larger portfolio heterogeneity as the risk premium shrinks. As a result, having superior information matters even more when the prices are already more informative.

Many policies and regulations aim to improve transparency in financial markets and lower the barrier to access information about available investment opportunities. Informing the distributional impact of such policies, inequality 50 states that there is almost a one-to-one relationship between the wealth divergence and extent of asymmetric information, more specifically, the probability that an investor receives strictly superior information than another investor ($F(i_2) - F(i_1)$). In other words, it recommends to reduce inequality, the informative signals should become publicly available and understandable, possibly through subsidizing access to information. Especially, it can massively reduce inequality when distribution $F(\cdot)$ itself is right-skewed, meaning the informative signals are virtually only available to a small group of investors.

Inequality 50 also underscores the role of non-fundamental noises in the wealth distribution. An increase in the magnitude of non-fundamental noise, represented by $\tilde{z}^N$, reduces the price informativeness, and hence, hampers the dissemination of information through the prices. It enlarges the belief gap, resulting in a faster wealth divergence, and consequently,
a larger wealth inequality.

Finally, inequality 50 implies that the wealth divergence is the fastest among the most informed. We can see this by noting the bounds increase as we focus on the more informed groups. The intuition lies in the way the information is disseminated in the economy: If only few most informed investors learn the realization of the informative signal, which happens when $\lambda_t$ is large, their impact on the prices is less than the case of a smaller $\lambda_t$, in which a larger fraction of the investors learn the realization and all trade according to the same signal realization. Thus, the prices are less informative when $\lambda_t$ is larger. Therefore, conditional on $\lambda_t$ being large, the asymmetric information between the informed and uninformed is also larger, which leads to faster wealth dynamics among the most informed investors.

**B2. Moving Types**

Data suggests billionaires’ descendants are not as good as their fathers in wealth creation and maintenance. In fact, there is a tendency to mean in the average returns that different family generations get and this mean-reversion has amplified over the past decades. For example Kaplan and Rauh (2013) find that only one-third of the richest individuals in Forbes 400 list at 2011 have grown wealthy, compared to two-third at 1988. Relatedly, Fagereng, Guiso, Malacrino, and Pistaferri (2016) find an economically small intergenerational correlation in returns to financial wealth and net worth. These findings suggest that the investment sophistication is imperfectly transmitted to the next generations.

Motivated by these findings, in this section, I analyze a modification of the model in which the newborns imperfectly inherit their parents type. Particularly, suppose there are only two types of sophisticated investors: informed and uninformed, similar to Section 6. How-
ever, a newborn inherits the type of his parent with probability \( \rho \in (0, 1) \). Therefore, the child of an informed investor is also informed with probability \( \rho \) and becomes uninformed with probability \( 1 - \rho \). The following proposition shows the wealth inequality is larger for higher values of \( \rho \).

**Proposition 7.**

*The tail parameter \( \gamma(\rho) \) is at least weakly decreasing in \( \rho \).*

Higher persistence in the sophistication types increases the inequality since it amplifies the compounding effect, that is the ones with a higher return today are more likely to get a higher return tomorrow, which further increases the inequality. This result is another indication of the importance of education in reducing inequality. Chetty, Friedman, Saez, Turner, and Yagan (2017) find that after students attending the same college land in relatively similar income percentiles, regardless of their parents’ income status. It underscores the role of higher education in social mobility.