HOW DO STOCK MARKET EXPERIENCES SHAPE WEALTH INEQUALITY

XIAOWEN LEI

Abstract. This paper develops a continuous-time overlapping generations model with rare disasters and learning-from-experience agents. Disasters such as the Great Depression permanently scar investors' trust in the market. As a consequence, generations that have experienced disasters save in the form of safer portfolios, even if similar disasters are not likely to happen again in their lifetimes. 'Fearing to attempt' therefore inhibits wealth accumulation by these "depression babies" relative to other generations. This effect is amplified in general equilibrium, because the equity premium is relatively high following a disaster. When calibrated to US data, the model can explain 18.7% of the recent increase in generational wealth inequality.

Keywords: rare disasters, heterogeneous beliefs, portfolio choice, inequality, learning

JEL Classification Numbers: D63, D81, G11, G51

"Our doubts are traitors and make us lose the good we oft might win, by fearing to attempt."

—Measure for Measure (1623, Shakespeare)

1. Introduction

Tensions between generations have existed since the last Ice Age. Perhaps Orwell (1945) said it best - "*Each generation imagines itself to be more intelligent than the one that went before it, and wiser than the one that comes after it.*" Recently, however, this tension has risen above its normal level. We've all seen the meme "ok boomer", and are well aware of the resentment that inspires it. The source of this resentment is clear. For the first

---

Date: October 13, 2020.

I thank Kenneth Kasa for continued guidance and support, and I am grateful to Benjamin Moll for providing many useful suggestions. I have also benefited from conversations with Rhys Bidder, Corina Boar, Benjamin Born, Geoffrey Dunbar, Edouard Djeutem, Martin Ellison, Gregor Jarosch, Nobuhiro Kiyotaki, John Knowles, Laurence Kotlikoff, Yang Li, Gustavo Mellior, John Muellbauer, Stefan Nagel, Laura Pilossoph, Petr Sedlacek, Vincent Sterk, Rick Van der Ploeg, Mirko Wiederholt, and Francesco Zanetti. Xiaowen Lei is from Department of Economics and Nuffield College, University of Oxford. Email: leixiaowen04@gmail.com.

time since records have been kept, most of the younger generation are in danger of being poorer than their parents (Chetty, Grusky, Hell, Hendren, Manduca, and Narang (2017)).

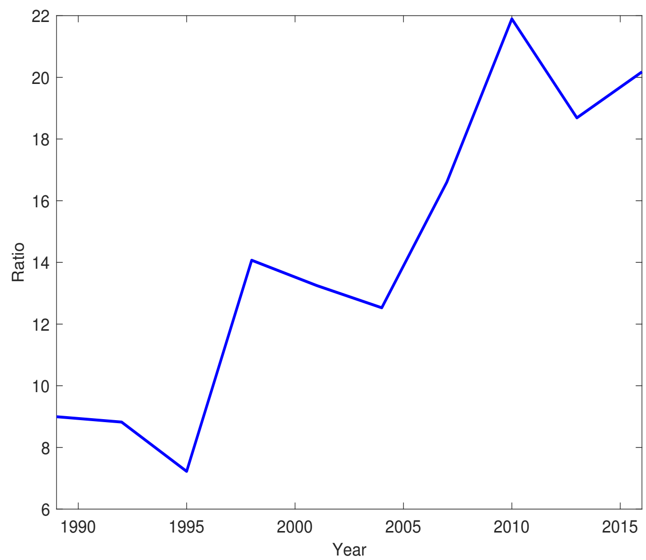

FIGURE 1. Median Net Worth Ratio of 65 and over vs. 35 and under (Survey of Consumer Finances)

Figure 1 plots Survey of Consumer Finances data on the ratio of median net worth for those over 65 years of age to those under 35.[1] Not surprisingly, the old have always been wealthier than the young. In 1989 their net worth was 9.0 times greater on average. However, over the course of the next 27 years this ratio more than doubled, to over 20.[2] Standard inequality models cannot explain the data in Figure 1 because they generate *stationary* age/wealth distributions. Of course, one could always inject an exogenous shock, and then attribute the trend in Figure 1 to transition dynamics. However, this is a rather unappealing strategy, since the trend in Figure 1 is the mirror image of a declining trend that took place during the 40 years following the Great Depression. Although direct

[1]The SCF definition of net worth includes total financial and non-financial assets, less the value of debt.
[2]SCF data are at the household level. There have been changes over time in demographics and household composition that potentially cloud the interpretation of Figure 1. First, household size has been decreasing. Data from the Current Population Survey shows that average family size decreased from 3.16 in 1989 to 3.14 in 2016. This suggests that the increase at the individual level might be even greater. Second, CPS data show that the marriage rate has also decreased, from 58% in 1995 to 53% in 2018. However, this has been offset by an equal increase in cohabitation during the same period, from 3% to 7%. Third, life expectancy has increased, which could potentially explain part of the increase in Figure 1. However, life expectancy in the US has increased relatively mildly compared to other countries. According to OECD data, it rose from 75.1 in 1989 to 78.6 in 2016.

evidence on historical generational inequality is lacking, we do know that generational inequality is highly correlated with the Top 1% wealth share. According to the Saez and Zucman (2016) data, the Top 1% wealth share in 1930 was 43.6%, and then went down to only 22.3% in 1980. Since most of rich people are old people in those older times, this suggests that the boomers are better off than both their parents and their kids. Hence, you would need to resort to *two* 'MIT shocks', not just one.[3]

What then explains this reversal? Undoubtedly, many factors lie behind this reversal. This paper focuses on just one of them, namely, generational belief differences. I study an economy that combines two key ingredients. First, individuals weight their own personal experiences more heavily when forming their beliefs, as in Malmendier and Nagel (2011) and Malmendier and Nagel (2016). Second, the economy is subject to rare disasters, as in Rietz (1988) and Barro (2006). When the model is calibrated to US data, it can not only account for a significant share of the recent increase in the relative wealth of the old generation, it can also explain why this ratio decreased following the Great Depression. The model also illustrates how general equilibrium feedback operating in financial markets contribute to these changes.

Although there has been a recent explosion of work on inequality, this work has either focused on aggregate measures, like top wealth shares, or on within-cohort inequality. For example, Deaton and Paxson (1994) show that within-cohort consumption inequality grows over time. Guvenen (2007) shows that learning about the growth rate of idiosyncratic labor income can reconcile this observation with life cycle models of inter-temporal consumption smoothing. However, standard models of top wealth shares or within-cohort inequality cannot explain the data depicted in Figure 1.

Although introducing rare disasters may seem similar to multiple MIT shocks, there is a crucial difference. Although rare, the disaster shocks in my model are recurrent, and the anticipation of this recurrence influences behavior, both before and after the shock. In fact, these anticipation effects are why the rare disasters literature has been successful at resolving the Equity Premium Puzzle. However, the asset-pricing rare disasters literature relies on a representative agent. The contribution of my paper is to show that when rare disasters are combined with overlapping generations and experiential learning, a powerful force for heterogeneity and inequality is ignited.

[3]An MIT shock is a colloquialism for a shock that was previously unanticipated, and is not expected to occur again. In other words, a measure zero event.

Specifically, I argue that different generations have different degrees of optimism/pessimism about market returns due to their own limited experiences. This influences their risk-taking behavior, which then influences the growth rate of their wealth. For instance, a 65 year old in 1989 would have been born in 1924, whom at an early age experienced the Great Depression. In contrast, a 65 year old in 2016 would have been a lucky baby boomer, who skipped the Great Depression and had more positive experiences in the stock market. Due to the rare nature of disasters, it was not likely that the depression babies would experience another Great Depression. But its salience within their own experience caused it to cast a long shadow throughout the remainder of their lives. In other words, they were "scarred". Therefore, it is natural that agents in different cohorts 'agree to disagree' about the likelihood of disasters.

Of course, this paper is not the first to propose an "experiential learning" channel in return expectations and portfolio reallocation. Malmendier and Nagel (2011) provides strong empirical support that macroeconomic experience in the stock market has a prolonged impact on how much households invest in risky assets later in their lives. They find that the "depression babies" were much less likely to participate in the stock market later in their lives. And if they did, they tended to invest a lower fraction of wealth into risky assets compared with other generations. Using SCF data, they find that an increase in experienced return from the 10th to the 90th percentile implies a 10.2% increase in the likelihood of participation in the stock market. Conditional on participation, there exists a 7.9% increase in the the fraction of wealth allocated to stocks. There has also been independent empirical evidence which shows that older people nowadays are more optimistic relative to young people. For example, Heimer, Myrseth, and Schoenle (2019) find that as households age, they grow more optimistic about longevity. Bordalo, Coffman, Gennaioli, and Shleifer (2020) uses survey data on the more recent Covid-19 crisis, and shows that the current older generation worries less about the health risk induced by the pandemic, despite the fact that evidence suggests they are the most vulnerable. This could be due to their own experience with previous pandemics.

While I do not aim to dismiss other potential mechanisms that drive between cohort inequality, the experiential-learning approach does offer several advantages. First, it microfounds "scale dependence", i.e: a positive correlation between growth and age, which is consistent with the data (See Gabaix, Lasry, Lions, and Moll (2016)). Modern life cycle portfolio choice theory *a la* Campbell, Viceira, Viceira, et al. (2002) suggests that the optimal share of risky investment should decrease with age. This is because younger households hold future labor income as a non-tradable asset, so they adjust tradeable asset holdings to compensate for the implicit holding of human wealth. However, micro evidence shows the opposite(e.g Ameriks and Zeldes (2004), Gomes and Michaelides (2005) and Fagereng, Gottlieb, and Guiso (2017)). At least before retirement, the old are more likely to participate in the stock market compared with the young, and conditional on participation, they invest a higher share of their wealth in risky assets. From the perspective of experiential learning, this is not so surprising. As households age, they witness more data, and become confident of their own estimates, which encourages them to invest a higher fraction of their wealth in risky assets. This is true during normal times, but especially so during disasters. For example, Gale, Gelfond, Fichtner, and Harris (2020) shows that the recent financial crisis has disproportionally depleted the wealth of the young millennials relative to the old. From the experiential-learning angle, millennials have had less experience with normal times, so they 'over-react' to the crisis, becoming relatively pessimistic about future stock market returns compared to their more experienced elders.

Second, while most of the literature focuses on why inequality has increased since the 1980s, the experiential learning approach provides a unified explanation of the long-run evolution of wealth inequality, tracing all the way back to 1930s. In particular, it can explain the U-shaped pattern that we see in the data. At the beginning of the Great Depression, the old to young wealth ratio at first decreased because the old were more invested in risky assets. However, as just noted, young people over-extrapolate from the disaster more than the old, since they have less experience. As these young households age, they tend to take few risks in the financial market, while the future generations are not subject to such scarring. This implies a gradual decrease of the old to young wealth ratio as time goes by. This tranquil decrease was interrupted in the 1980s, as the GenXers (born in 1965-1980) and millennial's (born in 1981-1996) experienced more recent disasters (e.g: the 1987 crash, the dotcom bubble burst, the financial crisis, and especially the more recent global pandemic!). Since the baby boomers are much less affected by these events, the old to young wealth ratio has increased. A U-shaped pattern of inequality of the last century naturally emerges.

Third, experiential learning in an overlapping generation environment can generate realistic features of asset prices. Gomez et al. (2016) studies the interaction between asset prices and wealth distribution with recursive preferences. Nakov and Nuño (2015) shows that when individuals learn from their own experience (i.e: decreasing gain learning), the aggregate implication for asset prices looks similar to a representative agent with constant gain learning, which has been shown to provide a good rationale for stock market volatility, and can explain the observed negative correlation between experienced payout growth and future excess returns (Adam, Marcet, and Nicolini (2016), Adam, Marcet, and Beutel (2017),Nagel and Xu (2019)).

Last but not least, the experiential-learning mechanism is consistent with survey data on stock return expectations. Using UBS/Gallup survey, Malmendier and Nagel (2011) find that a 1% decrease in experienced return is associated with $0.6-0.7\%$ decrease in expected returns to their own portfolio. Recent evidence that combines return expectations and portfolio choice data also shows that belief changes are indeed reflected in household portfolio choices. (See Giglio, Maggiori, Stroebel, and Utkus (2019)).

An important advantage of developing an explicit model is that it allows us to examine how these partial equilibrium effects become amplified in a general equilibrium where prices are endogenously determined. With heterogeneous beliefs and finite lives, prices reflect the wealth-weighted average beliefs of market participants. As a consequence, market pessimism induces a high equity premium right after a disaster shock. Cogley and Sargent (2008) attributes the existence of the postwar equity premium to pessimism induced by the Great Depression. This effect is endogenously generated here with overlapping generations. It causes the young cohort that experienced it to lose more wealth during a depression, not only because they fear holding stocks, but also because they fear the most when the gain from holding risky assets is the highest. While both the partial and general equilibrium effects might sound intuitive and simple, it is not an easy task to quantify them within a structural model. This is because prices depend on the wealth distribution, which is an infinite-dimensional object, whose evolution is hard to characterize in discrete time. My model attempts to disentangle the partial equilibrium and the general equilibrium channels of experiential learning by solving a continuous time overlapping generation model with heterogeneous learning from experience agents, and provides closed form solutions for policy functions, endogenous prices and the evolution of wealth.

The remainder of the paper is organized as follows. Section 2 outlines the model and solves for equilibrium prices. Section 3 uses a perturbation approximation of the Kolmogorov-Fokker-Planck (KFP) equation to characterize the dynamics of the generational wealth distribution. Section 4 provides simulation evidence. Section 5 calibrates the model to US data, and shows that the model can explain the observed U-shaped pattern in postwar generational inequality. Section 6 provides futher evidence on the connection between beliefs and stock market crashes. Section 7 discusses several alternative explanations of the rise in old/young wealth inequality, e.g., housing (Mankiw and Weil (1992)), education, inter-generational transfers, and financial market development

(Favilukis (2013)). Section 8 discusses efficiency and policy implications, while Section 9 contains a brief literature review. Finally, Section 10 concludes by discussing some possible extensions. A technical Appendix contains proofs and derivations.

2. The model

The model combines a Lucas (1978) pure exchange tree economy with a continuous-time OLG Blanchard/Yaari demographic structure. It then embeds rare disaster risk in the tradition of Rietz (1988) and Barro (2006), and tracks the distribution of portfolio allocations, asset prices, and the distribution of wealth when the arrival rate of the disaster is unknown, and agents must learn about it from their own experience.

2.1. Environment.

The economy consists of a measure 1 continuum of agents, each indexed by the time of birth s, with exponentially distributed lifetimes. Death occurs at Poisson rate δ. When an agent dies, he is instantly replaced by a new agent with zero initial financial wealth. At each instant of time $t > s$, all living agents receive an endowment flow $y_{s,t}$ where $y_{s,t} = \omega Y_t$, which can be interpreted as the agents' labor income, and that $\omega \in (0,1)$. That is, each existing agent receives a constant fraction of the aggregate endowment.[4] Agents have no bequest motive. There is a representative firm that pays out dividend $D_t = \omega Y_t$. In order to focus on between-cohort inequality, I assume agents only differ in the timing of birth, but are otherwise identical. That is, agents face only one source of idiosyncratic uncertainty, i.e: their birth and death dates. The exogenous aggregate endowment process is driven by two aggregate shocks. It is governed by the following jump-diffusion process

$$\frac{dY_t}{Y_{t^-}} = \mu dt + \sigma dZ_t + \kappa dN_t(\lambda_t) \tag{2.1}$$

where $Y(t^-)$ denotes the instantaneous dividend right before a jump occurs, if there is one, μ is the drift absent disasters, and σ denotes the volatility of the 1-dimensional Brownian motion Z_t, which satisfies the usual conditions. It is defined on a probability space $(\Omega^Z, \mathcal{F}^Z, \mathcal{P}^Z)$. N_t is a Poisson process with hazard rate λ_t, defined on a probability space $(\Omega^N, \mathcal{F}^N, \mathcal{P}^N)$. I then define $(\Omega, \mathcal{F}, \mathcal{P})$ as the product probability space, and the filtration of the combined history as $\{\mathcal{F}_t\} = \{\mathcal{F}^B \times \mathcal{F}^N\}$. The jump process N_t follows the process

$$dN_t = \begin{cases} 1, & \text{with probability } \lambda_t dt. \\ 0, & \text{with probability } 1 - \lambda_t dt. \end{cases} \tag{2.2}$$

[4]This assumption follows the tradition of Gârleanu and Panageas (2015). It is a reduced form way to capture the co-movement of the real economy and the financial market. Since the model focuses on the financial market, I abstract away from life cycle labor income profiles.

When a jump occurs, dividends jump by a magnitude of $\kappa_t Y_t$, where κ_t can take two values, κ_h (severe disasters) and κ_l (normal disasters), each could happen with 50% probability. $\kappa_t \in (-1, 0)$. This captures the drop in dividend when a disaster happens, but ensures that dividends remain strictly positive. The hazard rate λ_t itself follows a random process, and is assumed to take two values, a high hazard rate λ_h and a low hazard rate λ_l. It is characterized by an i.i.d Bernoulli distribution,

$$\lambda_t = \begin{cases} \lambda_h, & \text{with probability } \pi^*. \\ \lambda_l, & \text{with probability } 1 - \pi^*. \end{cases} \tag{2.3}$$

I assume that the market is complete, and that agents can trade continuously in the capital market to hedge against both the regular economic risk, as well as the disaster risk. To complete the market, we need one bond, one risky asset, and a disaster-contingent security. The bond value follows

$$dB_t = r_t B_t dt \tag{2.4}$$

The security 1 value (would also labeled as stock market value for the rest of the paper) follows

$$\frac{dS_t + D_t dt}{S_{t-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{2.5}$$

where r_t, μ_t^S, σ^S as well as κ^S are endogenous objects, and are determined in equilibrium. Finally, the security 2 value (would also be labeled as security 2) is P_t, and follows the stochastic process

$$\frac{dP_t}{P_{t-}} = \mu_t^P dt + \kappa_t^P dN_t(\lambda_t) \tag{2.6}$$

By convention, I assume the security 2 pays off during normal times, but suffers a loss during disasters. That is, by holding the security 2, the agent gets rewarded μ_t^P fraction of of the asset value at each instant, but the asset value drops by a magnitude of $\kappa_t^P P_t$ upon a disaster shock. The initial price $P(0)$ and the jump size κ_t^P can be chosen freely, but the drift μ_t^P is determined endogenously. This is a security that pays no dividend and is in zero net supply, with the real world counterpart of it being a catastrophe bond or a hybrid security whose value depend on the adverse state of the economy. [5].

Agents observe the dividend process and know the values of μ, σ and κ. However, they do not observe the current hazard rate (λ_h or λ_l), and must learn about it by optimal filtering. The specific choice of which parameters to learn about is supported

[5] In an incomplete market without security 2, equilibrium bond and equity returns change drastically (See Dieckmann (2011) for a comparison of asset pricing implications in compelte vs. incomplete market with rare disasters). Since the focus here is on portfolio reallocation rather than asset pricing, I focus on the simpler complete market setting.

by continuous-time filtering theory. As noted by Merton (1980), uncertainty about σ decreases as sampling frequency increases. It disappears in the continuous time limit. Although uncertainty about drift parameter μ does not dissipate, agents can still learn about it relatively quickly, and achieve asymptotic convergence. In contrast, uncertainty about disaster risk does not even disappear in an infinite horizon. To see how that works, we need to consider optimal filtering of a jump-diffusion process.

2.2. **Filtering and Information Processing.** Agents have common knowledge about the size of the disaster once it happens. However, they remain uncertain about the likelihood of disasters. They must revise their beliefs sequentially, in real-time. When an agent is born at time s, he is endowed with prior probability $\pi_{s,s}$ of the hazard rate. For $t > s$, his evolving beliefs are fully summarized by the conditional mean $\bar{\lambda}_{s,t} = \mathbb{E}_{s,t}[\lambda_t]$, where the expectation $\mathbb{E}_{s,t}[\lambda_t] = \pi_{s,t}\lambda_h + (1-\pi_{s,t})\lambda_l$ denotes the expectation with respect to the time s born agent's own filtration $\mathcal{P}_{s,t}$ at time t. I will specify how the prior is chosen in the quantitative section. For now, let's focus on belief updating.

Lemma 2.1. *The evolution of the beliefs about π^* by a Bayesian learning agent (denoted by $\pi_{s,t}$) is given by*

$$d\pi_{s,t}|_{dN_t=0} = -(\lambda_h - \lambda_l)\pi_{s,t}(1-\pi_{s,t})dt \tag{2.7}$$

$$d\pi_{s,t}|_{dN_t=1} = \frac{\lambda_h \pi_{s,t}}{\bar{\lambda}_{s,t}} - \pi_{s,t} \tag{2.8}$$

Proof. This is a direct application of the optimal filtering of a jump-diffusion process from Liptser, Shiriaev, and Shiryaev (2001) Theorem 19.6, and is later applied in Benzoni, Collin-Dufresne, and Goldstein (2011) and Koulovatianos and Wieland (2011). □

As one can see, when there is no jump, an agent's beliefs about the probability of a disaster follow a deterministic trend, with a negative drift of $-(\lambda_h - \lambda_l)(1-\pi_{s,t})$. Calm economic times gradually improve agents' optimism, albeit at a slow pace. However, when a jump occurs, beliefs shift discontinuously, and jump from $\pi_{s,t}$ to $\frac{\lambda_h \pi_{s,t}}{\bar{\lambda}_{s,t}}$. That is, the perceived likelihood of a disaster occurring is suddenly amplified by a magnitude of $\frac{\lambda_h}{\bar{\lambda}_{s,t}}$. [6]

2.3. **Optimization.** Agents continuously choose a non-negative consumption process $c_{s,t}$, the fraction of wealth allocated to the security 1 market $\alpha^S_{s,t}$, and the fraction of wealth devoted to the security 2 $\alpha^P_{s,t}$. They continuously update their beliefs about disaster risk, and dynamically trade assets given their current beliefs and their returns, in order

[6] One might argue that Bayesian learning is contradicted by evidence of a 'recency bias'. That is, it is debatable whether agents weight past observations of disasters in a statistically optimal manner. However, since I am primarily interested in generational belief differences, what matters is not the specific learning algorithm at an individual level, but the cross-sectional differences in weights on the same event.

to maximize a logarithmic flow utility over consumption goods. [7] They start with zero financial wealth, and accumulate wealth over the life cycle. An annuity contract *a la* Yaari (1965) entitles $\delta w_{s,t}$ of earnings to living agents, while a competitive insurance company collects any remaining wealth upon the unexpected death of the agent. Formally, the problem of an agent at time s can be stated as

$$\max_{c_{s,t},\alpha^S_{s,t},\alpha^P_{s,t}} \mathbb{E}_{s,t}\left[\int_s^\infty e^{-(\rho+\delta)(t-s)}\log(c_{s,t})dt\right] \tag{2.9}$$

s.t:

$$\begin{aligned}\frac{dw_{s,t}}{w_{s,t^-}} = &\left(r_t + \alpha^S_{s,t}(\mu^S_t - r_t) + \delta + \alpha^P_{s,t}(\mu^P_t - r_t) + y_{s,t} - \frac{c_{s,t}}{w_{s,t^-}}\right)dt + \alpha^S_{s,t}\sigma^S dZ_{s,t} \\ &+ (\alpha^S_{s,t}\kappa^S_t + \alpha^P_{s,t}\kappa^P_t)dN_{s,t}(\bar{\lambda}_{s,t})\end{aligned} \tag{2.10}$$

The resulting HJB equation associated with this problem is a nonlinear partial differential equation. With the presence of aggregate shocks, it is not likely to have a closed-form solution. To bypass this problem, I exploit the fact that the market is dynamically complete for all cohorts. Therefore, instead of directly tackling the decentralized problem, I instead exploit the complete market structure, and employ the martingale approach (Cox and Huang (1989)). This allows me to convert the dynamic programming problem into a static problem as follows

$$\max_{c_{s,s}} \mathbb{E}_{s,s}\left[\int_s^\infty e^{-(\rho+\delta)(t-s)}\log(c_{s,t})dt\right] \tag{2.11}$$

s.t:

$$\mathbb{E}_{s,s}\left[\int_s^\infty e^{-\delta(t-s)}\xi_{s,t}c_{s,t}\right] = \mathbb{E}_{s,s}\left[\int_s^\infty e^{-\delta(t-s)}\xi_{s,t}\omega Y_t dt\right] \tag{2.12}$$

where $\xi_{s,t}$ denotes the individual state price density.

From the first order condition (FOC) of consumption, we obtain

$$\frac{e^{-(\rho+\delta)(t-s)}}{c_{s,t}} = y_s e^{-\delta(t-s)}\xi_{s,t} \tag{2.13}$$

where y_s denotes the Lagrange multiplier associated with the agent's lifetime budget constraint. We can then relate $c_{s,t}$ to the initial consumption allocation $c_{s,s}$ using the following equation

$$c_{s,t} = c_{s,s}e^{-\rho(t-s)}\frac{\xi_{s,s}}{\xi_{s,t}} \tag{2.14}$$

[7] As we shall see later, log preferences deliver two key advantages. First, they imply a constant savings rate, which allows me to focus on the portfolio choice channel. Second, a log agent's portfolio does not need to include a hedging term (Gennotte (1986)). That is, his optimal portfolio is 'myopic'. Both these simplifications are driven by the exact offsetting of income and substitution effects

To see how the consumption process evolves, we can first solve for the stochastic process of the state price density.

Lemma 2.2. *By exploiting the fact that the regular Brownian motion and the compensated Poisson process are martingales under the agent's own filtration, one can derive the individual state price density process as follows*

$$\frac{d\xi_{s,t}}{\xi_{s,t^-}} = (\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t)dt - \theta_{s,t}dZ_{s,t} + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.15}$$

where $\theta_{s,t}$ denotes the perceived market price of risk of the regular Brownian shock, and $\lambda^N_{s,t}$ is the perceived market price of disaster risk. It then follows that the true state price density follows

$$\frac{d\xi_t}{\xi_{t^-}} = (\bar{\lambda}_t - \lambda^N_t - r_t)dt - \theta_t dZ_t + \left(\frac{\lambda^N_t}{\bar{\lambda}_t} - 1\right) dN_t(\bar{\lambda}_t) \tag{2.16}$$

Define the disagreement process $\eta_{s,t} = \frac{\xi_t}{\xi_{s,t}}$. We then have

$$\frac{d\eta_{s,t}}{\eta_{s,t}} = \left(\frac{1}{1+\bar{\kappa}}\lambda_{s,t} - \lambda^N_t\right) dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2\lambda^N_t}{\lambda_t} - 1\right) - 1\right] dN(\bar{\lambda}_t) \tag{2.17}$$

where $\bar{\kappa} = \frac{1}{2}\kappa_h + \frac{1}{2}\kappa_l$.

Proof. See Appendix 11.3. □

As expected, the disagreement process $\eta_{s,t}$ does not depend on the regular Brownian shock, but only the disaster shock. When no disaster hits, the disagreement process has a deterministic drift, which depends on how likely the agent perceives the disaster relative to the truth, as well as on the market price of disaster risk he or she is willing to bear relative to the market. Since we know that $c_{s,t} = (y_s\xi_{s,t})^{-1}$, knowing the process of the state price density is equivalent to knowing the process of consumption. Ito's lemma then delivers

$$\frac{dc_{s,t}}{c_{s,t^-}} = (\theta^2_{s,t} - \bar{\lambda}_{s,t} + \lambda^N_{s,t} + r_t)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda^N_{s,t}} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.18}$$

This is useful, because due to log utility, consumption is linear in financial wealth, i.e: $c_{s,t} = (\rho+\delta)w_{s,t}$. This implies that the stochastic process of the optimally invested wealth also follows

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta^2_{s,t} - \bar{\lambda}_{s,t} + \lambda^N_{s,t} + r_t)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda^N_{s,t}} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.19}$$

Given the above individual optimal consumption decisions, we are now ready for aggregation.

2.4. **Aggregation.** I start by defining a Walrasian equilibrium in this economy.

Definition 2.3. *Given preferences, initial endowments, and beliefs, an equilibrium is a collection of allocations* ($c_{s,t}$, $\alpha^S_{s,t}$, $\alpha^P_{s,t}$) *and a price system* (r_t, μ^S_t, μ^P_t ,κ^P_t) *such that the choice processes* ($c_{s,t}$, $\alpha^S_{s,t}$, $\alpha^P_{s,t}$) *maximize agents' utility subject to their budget constraints, and the market for consumption goods, bonds, security 1 and the security 2 all clear, i.e:*

$$Y_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} c_{s,t} ds \tag{2.20}$$

$$S_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha^S_{s,t} w_{s,t} ds \tag{2.21}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha^P_{s,t} w_{s,t} ds \tag{2.22}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} (1 - \alpha^S_{s,t} - \alpha^P_{s,t}) w_{s,t} ds \tag{2.23}$$

By using the market-clearing condition for consumption goods, we can derive the stochastic processes for ξ_t. Let's conjecture that the fraction of aggregate dividends consumed by a newborn agent at time t is a fixed fraction $\beta_t = \frac{c_{t,t}}{Y_t} = \beta$. [8] We can then rewrite the goods market clearing condition as

$$\xi_t Y_t = \int_{-\infty}^{t} \beta \delta e^{-(\rho+\delta)(t-s)} \xi_s Y_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{2.24}$$

Define $\eta_t = e^{(\rho+\delta(1-\beta))t} \xi_t D_t$, we can then rewrite the above into

$$\eta_t = \int_{-\infty}^{t} \beta \delta e^{-\beta\delta(t-s)} \eta_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{2.25}$$

Defining $\mu^\eta_{s,t}$ and $\kappa^\eta_{s,t}$ as the drift and jump coefficients of $\eta_{s,t}$ we are now ready to derive the dynamics of η_t. Applying Ito's lemma and Leibniz's rule, we obtain

$$\frac{d\eta_t}{\eta_t} = \bar{\mu}^\eta_t dt + \bar{\kappa}^\eta_t dN_t(\bar{\lambda}_t) \tag{2.26}$$

where the weighted average coefficients are defined as

$$\bar{\mu}^\eta_t = \mathbb{E}_{s,t}(\mu^\eta_{s,t}) = \int_{-\infty}^{t} f_{s,t} \mu^\eta_{s,t} ds; \quad \bar{\kappa}^\eta_t = \mathbb{E}_{s,t}(\kappa^\eta_{s,t}) = \int_{-\infty}^{t} f_{s,t} \kappa^\eta_{s,t} ds \tag{2.27}$$

and the wealth share $f_{s,t}$ is defined as

$$f_{s,t} = \beta \delta e^{-\beta\delta(t-s)} \left(\frac{\eta_s}{\eta_t}\right) \left(\frac{\eta_{s,t}}{\eta_{s,s}}\right) = \delta e^{-\delta(t-s)} \frac{c_{s,t}}{Y_t} \tag{2.28}$$

[8] Appendix 12.1 verifies this conjecture, and derives an explicit expression for β.

Since we know the dynamics of D_t, we can then back out the dynamics of the state price density.

$$\frac{d\xi_t}{\xi_t} = \left(\bar{\mu}_t^{\eta} - \mu + \sigma^2 - \rho - \delta(1-\beta)\right) dt - \sigma dZ_t + \left(\frac{1+\bar{\kappa}_\eta}{1+\bar{\kappa}} - 1\right) dN_t(\lambda_t) \tag{2.29}$$

Since we know that the state price density also has to follow eqn.(2.16), it directly gives the solution of equilibrium prices.

Proposition 1. *We are now ready to pin down all endogenous prices by matching the coefficients of the state price density in eqn.(2.16) and eqn.(2.29). In equilibrium, the short term interest rate, the market price of risk for the regular Brownian shock, and the market price of disaster risk are given by*

$$r_t = \underbrace{\rho + \delta(1-\beta)}_{\text{effective patience with OLG}} + \underbrace{\mu - \sigma^2}_{\text{risk adjusted growth}} + \underbrace{\frac{\bar{\kappa}}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\lambda_{s,t})}_{\text{market view of disaster risk}} ; \tag{2.30}$$

$$\theta_t = \theta = \sigma; \tag{2.31}$$

$$\lambda_t^N = \frac{\mathbb{E}_{s,t}(\lambda_{s,t})}{1+\bar{\kappa}} \tag{2.32}$$

The closed form solutions for prices have intuitive interpretations. Let's start with the equilibrium interest rate. As always, the risk free rate increases when agents are less patient. In a world of finite lives, the effective patience lessens due to death risk. Moreover, the equilibrium interest rate increases when the dividend process has a higher rate of growth and a lower volatility, which is captured in the second term. The third term reflects a precautionary savings motive coming from the "market view" of disaster risk, which is itself an endogenous object. It depends on the wealth-weighted distribution of beliefs. Since $\kappa < 0$, this implies that the equilibrium interest rate decreases with market average pessimism. Desire to save during disasters drives down the return on short-term bonds, leading to low equilibrium interest rates during disaster episodes, as observed in the data (See Nakamura, Steinsson, Barro, and Ursúa (2013)). Notice that the first and second term are both constants, so variations in the interest rate are totally driven by variations in market optimism about disasters. The market price of the regular Brownian risk is less interesting in this log-utility model. Since the disagreement is only about disaster risk, agents have common beliefs about the regular Brownian risk, therefore the market price of risk is the same as the standard solution with log preferences, which simply equates to the volatility of the risk. Disaster risk does not affect this part of the solution. Finally, the market price of disaster risk increases with the market view of the likelihood of the disaster. Lastly, λ_t^N also increases with the magnitude of the negative jump.

2.5. **Portfolio Allocations and Wealth Dynamics.** This subsection discusses the key predictions of the model. Namely, how does the experience of a rare disaster influence lifetime savings and portfolio allocations, and how do these decisions influence an agent's wealth accumulation? Recall that the optimally invested wealth follows

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_t^N + r_t)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.33}$$

Recall also that the budget constraint follows

$$\begin{aligned}\frac{dw_{s,t}}{w_{s,t^-}} &= \left(r_t + \alpha_{s,t}^S(\mu_t^S - r_t) + \delta + \alpha_{s,t}^P(\mu_t^P - r_t) + y_{s,t} - \frac{c_{s,t}}{w_{s,t^-}}\right) dt + \alpha_{s,t}^S \sigma^S dZ_{s,t} \\ &\quad + (\alpha_{s,t}^S \kappa_t^S + \alpha_{s,t}^P \kappa_t^P) dN_{s,t}(\bar{\lambda}_{s,t})\end{aligned} \tag{2.34}$$

Since the market is complete, we can match coefficients with the wealth process in these two SDEs, and get the following. The share of wealth invested in the risky security 1 market and the security 2 at time t for an agent born at time s are given by the following expressions respectively

$$\alpha_{s,t}^S = \frac{\theta_{s,t}}{\sigma^S} = \frac{\theta_t}{\sigma^S} \tag{2.35}$$

$$\alpha_{s,t}^P = \frac{1}{\kappa_t^P}\left(\frac{\bar{\lambda}_{s,t}}{\lambda_t^N} - 1\right) - \frac{\kappa_t^S \theta_t}{\kappa_t^P \sigma^S} \tag{2.36}$$

Notice that all generations invest the same fraction of wealth in security 1. However, pessimistic generations buy more disaster insurance (i.e: they are the sellers of security 2), as reflected in a higher $\bar{\lambda}_{s,t}$. To complete the calculation, we still need to characterize μ_t^S, σ^S, κ_t^S and κ_t^P.

2.6. **Equity Premium Dynamics.**

Proposition 2. *The equilibrium coefficients in the security 1 price and the security 2 are given by*

$$\sigma^S = \sigma \tag{2.37}$$

$$\kappa_t^S = \kappa_t \tag{2.38}$$

$$\mu_t^S - r_t = \sigma^2 + \bar{\mu}_t^\eta \tag{2.39}$$

$$\mu_t^P - r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t}) \tag{2.40}$$

Proof. See Appendix 11.4. □

That is, the model produces an endogenous time-varying equity premium, both for the security 1 as well as for the security 2. When market pessimism rises, security 1 and security 2 must pay higher average returns to clear the market. This has interesting implications for inequality. Following a disaster shock, scarred investors find bond investing to be more attractive, and shy away from the risky asset. Their collective efforts to do so lowers the return from investing in bonds. The increased equity premium can be taken advantage of by future generations. Boomers got high returns because their parents suffered through the Depression. This general equilibrium effect of prices amplifies the initial partial equilibrium effect. Not only does the scarred generation grow slower due to less risk-taking, but they also sacrifice returns when it is the best time to buy risky assets.

Corollary 2.4. *The share of wealth invested in the risky security 1 market and the security 2 at time t for an agent born at time s are given by the following expressions respectively*

$$\alpha_{s,t}^{S} = 1 \tag{2.41}$$

$$\alpha_{s,t}^{P} = \frac{1}{\bar{\kappa}}\left(\frac{\bar{\lambda}_{s,t}}{\mathbb{E}(\bar{\lambda}_{s,t})}(1+\bar{\kappa}) - 1\right) - 1 \tag{2.42}$$

If $\lambda_{s,t} > \mathbb{E}(\lambda_{s,t})$, generation s is more optimistic relative to the average generation, and become the seller of the disaster insurance, vice versa.

The resulting portfolio choice solutions are rather intuitive. Due to log utility of homogeneous beliefs on the Brownian motion risk, all investors invest all shares in security 1. However, the more pessimistic the investors are, the less investment they make on security 2.

3. Evolution of the Joint Age-Wealth Distribution

This section studies the main object of interest, i.e, the evolution of the *joint* age-wealth distribution. Note that with aggregate shocks, the Kolmogorov Forward equation, which characterizes the evolution of the wealth distribution follows a *stochastic* partial differential equation, and the distribution changes constantly. However, one can still study the long-run stationary distribution by averaging out those shocks across time, and compare its properties relative to the rational expectation economy.

Proposition 3. *Define $\tilde{w}_{s,t} = \frac{w_{s,t}}{\omega Y_t}$. The long-run stationary distribution of $x = \log(\tilde{w})$ is given by*

$$p(x) = \underbrace{Ge^{\zeta_0 x}}_{RE} \underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}]}_{Learning} \tag{3.43}$$

Moreover,

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^{RE}(x) \tag{3.44}$$

Proof. See Appendix 12. □

As one can see, we can decompose the long-run stationary distribution into two pieces. The first piece features the standard resulting distribution of log of wealth as in the rational expectation economy. The second piece reflects experiential learning, which produces a fatter tail compared with the RE economy. As wealth becomes larger, the experiential learning economy has more inequality compared with the Rational Expectation economy.

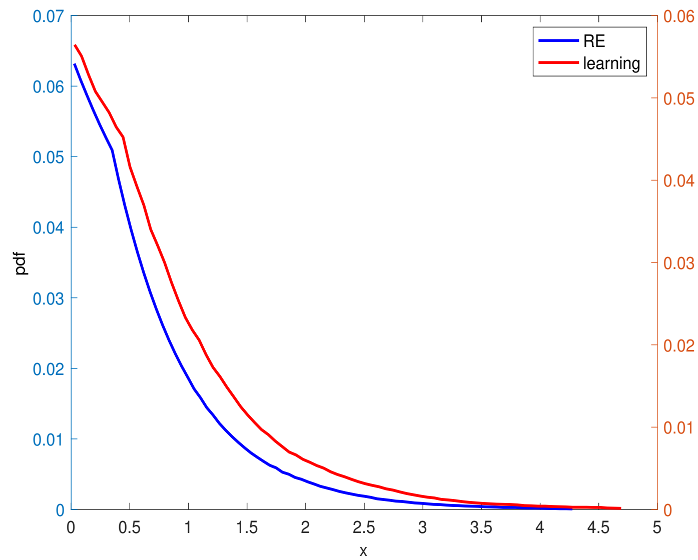

FIGURE 2. Long-Run Age Distribution of Log Normalized Wealth

We can also compare the difference by plotting the numerical solution of the long-run stationary distribution of log wealth by examining Figure 2. The blue line denotes the distribution under (full sample) Rational Expectations. In this case, the growth of wealth is homogeneous across all generations, and the stationary distribution is exponential. In this economy, the old are richer simply because they have lived longer and have had more time to accumulate wealth. The red line plots the stationary distribution under experiential learning. The reason why the experiential learning economy features a fatter tail compared with the RE economy is pretty intuitive: it is due to the scale dependence of wealth accumulation (See Gabaix, Lasry, Lions, and Moll (2016)). In this economy, older are on average richer, who are also accumulating their wealth faster compared with the poorer and younger household. This is true both in normal times as well as in disaster times. Recall that during normal times, the older households have observed more data, and therefore take on more risk compared with the younger household. During disaster times, even though all generations become more pessimistic, it is the young generations that are hit the most, because they have less life time experience, and would therefore over extrapolate information from the disaster. Therefore, "scale dependence" is even stronger during disaster times.

4. Simulations

In this section, I take the policy functions, and simulate sample paths of savings and portfolio choice for a typical agent, using the benchmark parameters in Table 1. The specific choice of parameters will be discussed in detail in the quantitative section. For now, let's focus attention on what happens to cohort behaviors after a rare disaster shock. To start, I shut down general equilibrium effects by fixing prices at their Rational Expectations equilibrium values. Death risk is eliminated so that one individual represents one generation. I assume that all agents start trading at age 20. When the trading age of the agent is 10 years old (30 years biological age) , I introduce a one time disaster shock. Figure 3 plots the responses to the shock.

As one can see, with log utility and complete markets, the agent invests all their wealth into the security 1, and then borrows to purchase the security 2. If one inspects the security 2 premium, one can see that its drift exceeds the risk free rate. Therefore, shorting to purchase the security 2 yields positive net returns during normal times. The agent's wealth grows steadily overtime. Suddenly, at $t = 10$, a disaster strikes, which drastically brings down the dividend. This does not affect his/her security 1 portfolio, because the risky security 1 only prices in regular Brownian risk, which is not affected by the disaster. However, due to learning from experience, the agent's pessimism rises, which reduces his exposure to the security 2. Notice also that it takes more than several years for him/her to reach the same level of optimism as before the disaster. A useful benchmark economy is the case of Rational Expectations, plotted in the blue line. In that world, the perceived likelihood of disasters is the same for all agents. In a complete market, this implies that nobody would be trading the security 2, since they all have the same beliefs. Therefore, this rational agent invest all his/her wealth in the risky security 1 market. The last two subplots show the response of prices after the disaster. As one can see, the interest rate plummets suddenly after the disaster due to increased precautionary savings. The reduction of the equilibrium interest rate also drives up both the security 1 risk premium and the disaster premium. However, the quantitative effects are rather small. For example, the equilibrium interest rate drops only 0.01298% after the shock. Therefore, the general equilibrium effect in this model is rather small compared to the partial equilibrium effect.

5. Calibration

In this section, I calibrate the above model to US data, and examine its quantitative implications for the dynamics of generational wealth inequality. Before presenting the results, it is important to discuss the benchmark parameters being used.

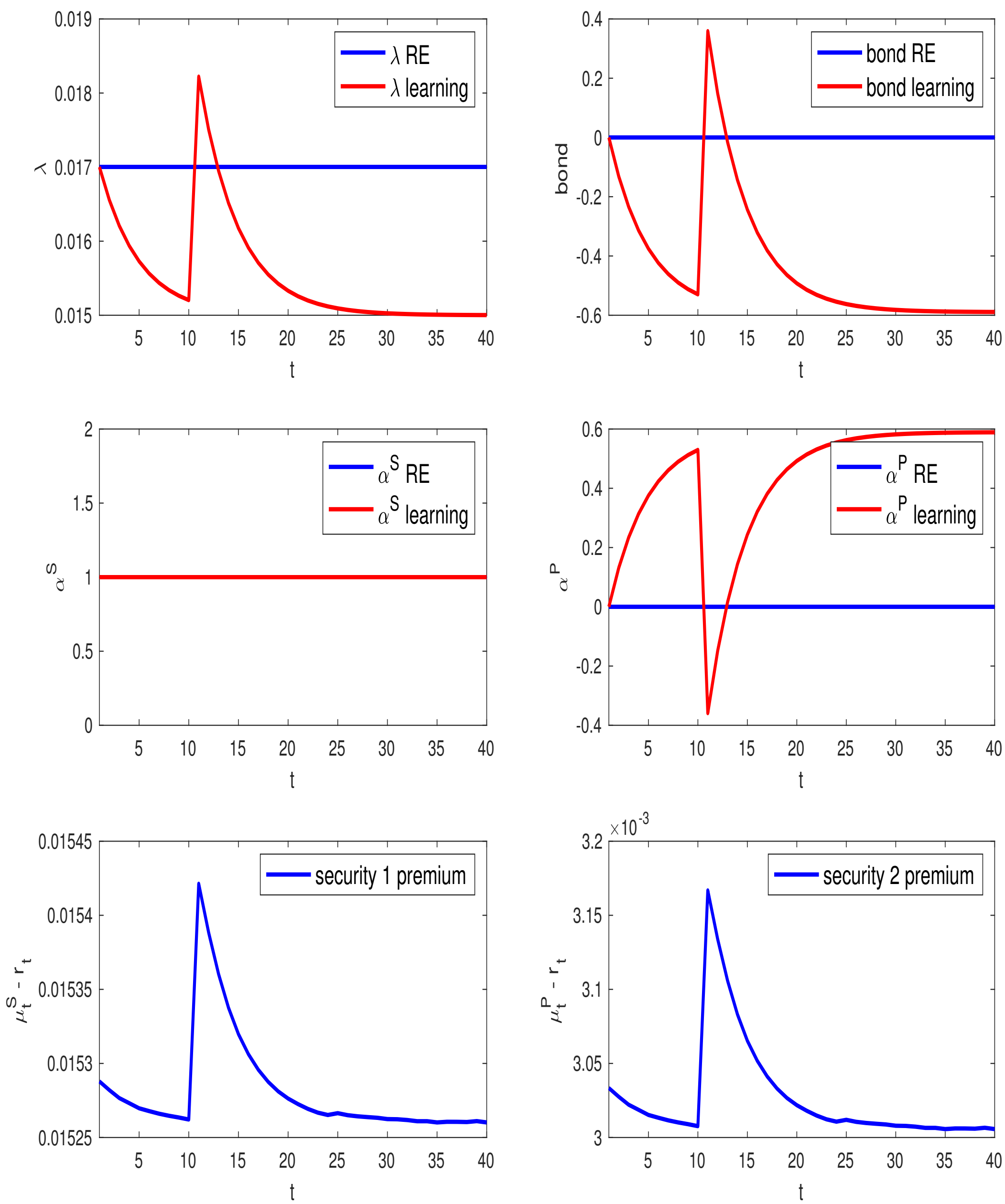

FIGURE 3. Simulated Time Paths of Policy Functions and Prices

Table 1. Benchmark Parameter Values

Parameters	ρ	δ	ω	μ	σ	κ_h	κ_l	π^*	λ^H	λ^L	$\pi_{s,s}$
Value	1.5%	1.67%	0.92	2%	11.07%	-0.35	-0.043	0.89%	24%	1.5%	0.0089

Empirical estimates of (annual) time preference are around 1% to 2%. I take the average estimate here so that $\rho = 1.5\%$. The death rate 1.67% is set such that the average trading life is from 20 to 80 years old, implying an average trading life expectancy of 60 years. The parameter ω follows from Gârleanu and Panageas (2015), which is chosen to match the fraction of capital income as a share of total income in the US. The drift coefficient μ and volatility coefficient σ is estimated using real dividend data from Shiller's data set absent disaster periods. The calibration of the two hazard rates $\lambda^H = 24\%$ and $\lambda^L = 1.5\%$ represent the upper and lower bounds of disaster rate, respectively. Those are also the hazard rate upper and lower bounds presented in Barro (2006). The weight $\pi^* = 0.89\%$ is chosen such that the average rare disaster likelihood is 1.7%, which corresponds to the empirical estimate from Barro (2006) of 35 countries over 100 years. Barro (2006) also finds that the mean contraction rate upon a disaster is about 35% after counting trend growth in GDP, so is the value of κ_h in my model. I assume that the Great Depression in 1930 features a percentage output reduction of κ_h. κ_l is then calibrated to match the percentage output reduction in the 2007 financial crisis, which features a smaller but still significant output drop. Finally, I assume that all agents start with a fixed prior that is equal to the Rational Expectations value.

Using the above parameters, I first compute the long-run average distribution of wealth and beliefs by simulation. I simulate the economy with 30000 initial agents for 2000 years. Each year, the wealth share weighted average of prices are computed, and fed back into the growth of wealth for each living agent. Then, a fraction δ of the random sample is dropped out at the end of each year, which is then replaced by newborns, who are endowed with zero financial wealth but a fixed fraction of aggregate dividend, and their beliefs are reset to the prior in the next period. For surviving agents, their beliefs and wealth are updated. Prices are again computed by the wealth weighted average, and the process carries on for 2000 years. The first 1000 years are discarded as a burn-in period, while the last 1000 years of data are used to get the average joint age-wealth distribution. This is then used as the initial distribution in 1920, where I start the calibration from. Next, I assume that two disasters happened after 1920. In 1929, the Great Depression reduces the output by a percentage of κ_h, and in 2007, the financial crisis reduces the output by a percentage of κ_l. I then re-run the simulation for 100 years to examine the model prediction between

1920-2020.

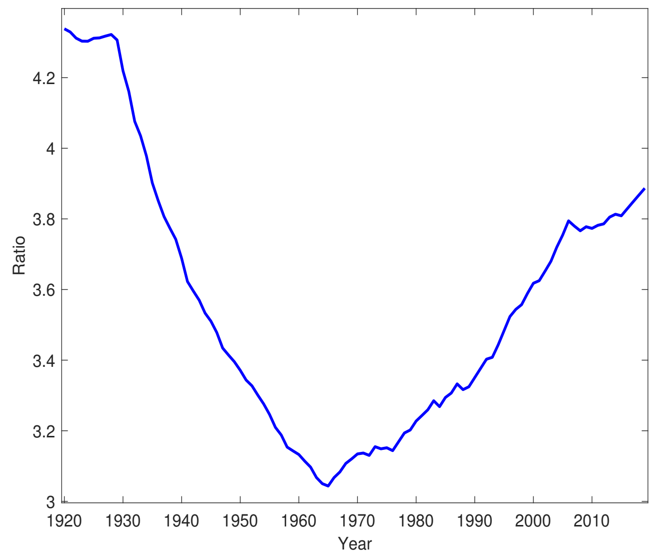

FIGURE 4. Calibrated Path of Old to Young Wealth Ratio

Figure 4 plots the calibrated path of the old to young wealth ratio (65 and over vs. 35 and under). There are several interesting patterns that emerge. As one can see, right after the 1929 Depression, the old to young wealth ratio goes down sharply and the trend continues until the late 1960s. This reflects a pure wealth reduction effect, as well as the lingering belief scarring effect, since older people have more wealth in the financial market, and suffer disproportionally more upon the stock market crash. As time goes by, the young people that experienced the Great Depression (the "Depression babies") become older, but they are more pessimistic compared with their future generations who did not experience it. Over the life cycle, their conservative portfolio strategies cause them to lose wealth relative to the newborn. This effect last quite a long while, until the "depression babies" almost disappear from the stock market scene, and finally the wealth ratio starts going back up. Interestingly, the 2007 financial crisis produces some smaller changes in the old to wealth ratio again. Although the magnitude of the output drop is much less in the 2007 financial crisis compared with the Great Depression, the old to young wealth ratio still dropped, which reflects the fact that the older generations are still more invested in the stock market. When asset value plunge, they lose a lot of wealth. However, in the last decade, this ratio mildly trended up again. The younger household that experienced the financial crisis invest even less now in the risky asset. Along with the recovery of asset prices, this dis-proportionally benefited the older household again.

Table 2. Model vs. Data

% Δ of O/Y Wealth Ratio	1989-2016	1989-2006	2007-2016
Data	124.31%	39.28%	14.13%
Model	15.16%	21.54%	1.3%

To see how the model implied old to young wealth ratio compares with the data, we can examine Table 2. In both cases, generational inequality trends up after the mid 1980s, albeit with different magnitudes. This is understandable, since the model singles out experiential learning as the only mechanism driving generational inequality, while in reality, many other channels have contributed to this increase. Therefore, a better statistic to evaluate the fit of the model is to ask how much of the rise can be explained by the model. From 1989 to 2011, the old to young wealth ratio rose by 124.31%, while the model generates an increase of 15.16%, which is around 12.2% of the increase. However, the model matches the increase of old to young wealth ratio much better before the financial crisis, generating a rise of 21.54%, which is around 55% of rise. It does less well in matching the data after the financial crisis.

5.1. **Belief inheritance, or experiential learning?** One might argue that different generations could have different priors, depending on the influence of the environment, especially their parents. Pessimism could beget pessimism. For example, even though boomers lived in a lucky environment, they might have been influenced by the pessimism of their depression era parents. Similarly, a millennial might have an optimistic boomer parent, which allows him to confront his dismal prospects with a degree of optimism. In other words, inter-generational belief transfers might dampen this paper's key mechanism. However, such belief inheritance is hard to measure with data. The closest attempt has been Charles and Hurst (2003), who uses PSID data along with survey measures to get estimates of risk tolerance across generations. However, since the PSID only asks participants to choose three levels of risk tolerance, this measure is rather rough, and it is also unclear to what extent the measure reflects risk aversion (which is intrinsic in preferences) vs. beliefs (which reflect agents' subjective estimates of the market return). Since this paper focuses on the belief channel, I continue to fix all agents' risk aversion at the same level. To see how the result might be altered by having different priors, I now set all the newborn's priors to be equal to the market view at the time they are born, and see how that changes the result.

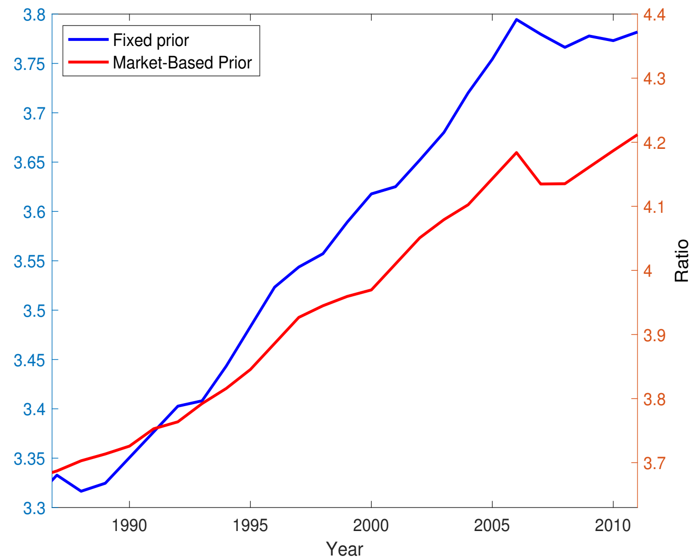

Figure 5. Old to Young Wealth Ratio : Fixed vs. Market-Based Prior

Figure 5 plots the comparison of the old to young wealth ratio by comparing the benchmark economy (with a fixed prior) to an economy where prior beliefs are equal to the market average beliefs at that time. As one can see, the qualitative increase of the old to young wealth ratio still holds, although its level is slightly different. The change in the level of inequality with a market-based prior is complicated, and in general depends on parameters. I briefly discuss forces that could increase as well decrease it. There are two main forces that generate increased inequality. First, since disasters are rare, the average market-based beliefs are more optimistic than the fixed rational expectation prior, therefore it produces more optimism for everyone, which naturally contributes to more risk taking and higher inequality. Second, a market-based prior implies that we add one more dimension of agent heterogeneity, which amplifies the heterogeneity of wealth growth differences for all agents, which also contributes to higher inequality (See Gabaix, Lasry, Lions, and Moll (2016)). On the other hand, as discussed in the previous paragraph, if the lucky generations (those that do not experience disasters in their own lifetime) happen to be born at a time when the market is pessimistic, they would have to balance between the pessimistic prior and the more optimistic lifetime experience, which could dampen generational inequality compared with the benchmark model. Therefore, the general prediction of how changing priors change generational inequality is ambiguous. However, what we are more interested in is the model's power to explain the rise in generational inequality. Compared with the data, both calibrations produce around 114.16% times of the increase of the old to young wealth ratio, despite the level difference, which is still 11.6% compaerd with the data.

5.2. **Comments on the Baby Boomers.** One might argue that the increase in overall inequality in recent decades could well be a result of an increasing cohort size of senior citizens, i.e: when the baby boomers get old, they also become on average richer. In partial equilibrium, this does not matter because the model is calibrated to the old to young wealth ratio for the *median* household, i.e: the cohort size effect is eliminated. However, in general equilibrium, the increased cohort size of the boomers matters. A large cohort could imply an increased price impact, which in turn influences everybody in the economy. After all, popular press and the media have long discussed whether the retirement of the boomers is likely to trigger a fall in stock prices, which could harm the millennials. Similar asset market meltdown hypothesis has been debated in the academic community as well [9] In the model, an increase in the cohort size of the optimistic boomers is likely to push up the equilibrium interest rate and decrease the equity premium, thus reducing the financial gains for everyone. If this is the case, generational inequality would be dampened. However, as mentioned before, such general equilibrium effects are rather small, amounting to only 0.01298% on interest rate changes from peak to trough. Therefore we are safe to take the result from the benchmark calibration as a reasonable approximation to the *median* old to young wealth ratio.

5.3. **Comments on Savings rate.** In general, wealth accumulation is driven by *two* choices, saving and portfolio allocation. By assuming log utility, this paper focuses on the portfolio allocation channel. However, it is possible that generational differences in saving rates also play a role. For example, if the savings rate of the old increased relative to the young after the 1980s, the observed increase in generational inequality might be driven by saving. Interestingly, data from Moody's Analytics shows that the savings rate has been declining for all age groups from early 1990, and went slightly back up after the financial crisis, particularly for the millennials. Therefore, if one were to examine the effect of savings on generational inequality, one would expect that the old to young wealth ratio would *decrease* during this period. If anything, this tells us that the portfolio reallocation channel examined in this model provides a lower bound on how important it is in generating recent increase in generational inequality. To be more specific about how disasters might alter the savings rate, Appendix 5.3 further examines how the savings rate responds to previous stock market returns, controlling for other factors. In all regression specifications, there is no significant correlation between previous stock returns and the

[9] See Poterba (2001), Abel (2001).

savings rate. Therefore, it seems safe to fix the savings rate here and focus on the portfolio allocation channel.

5.4. **Robustness: A US-specific experience.** The main calibration relies on the Barro (2006) estimates of disaster frequency and size, which are based on an international sample of 35 countries over 100 years. Such disasters (defined as a contraction of GDP of more than at least 15%) add up to only 60 cases in his long sample, which points to an average disaster probability of 1.7% per year. There are at least two reasons for doing this. First, since rare disasters are by definition rare, it is hopeless to just rely on the experience of US itself to "estimate" the frequency and size of disasters. Second, economic disasters are becoming increasingly global in the last century, with the main drivers being world wars, the Great Depression, the Asian financial crisis, and the Latin American debt crisis. The strong correlation of international disasters makes it defensible to use global data to infer disaster estimates for the US. Nevertheless, the US is still a relatively tranquil country. Therefore, it pays off to see how a reduced disaster size influences the results.

Table 3. Robustness: Alternative Disaster Parameters (1989-2016)

	Data	**Benchmark**	$\kappa = -0.33$
%Δ O/Y wealth ratio	124.31%	23.27%	21.83%
Δ Top 1% wealth share	61.95%	20.69%	19.04%
Corr(Risky Share, Age)	0.3644	0.9164	0.8467
Corr(Experienced Return, Expected Return)	0.6-0.7	0.6273	0.6201

Table 3 examines how well the model matches other dimensions of its predictions other than the old to young wealth ratio. [10] As stated in the main calibration results, the benchmark model is able to explain 18.7% of the increase of the old to young wealth ratio from 1989 to 2016. The model also predicts an increase of 1.2069 times of increase of the top 1% wealth share increase, while in the data it's 1.6195 times. This is a fairly encouraging result, given that the model focuses only on between-cohort heterogeneity, and has been silent about all other heterogeneity that are potentially important for explaining increases in top shares, i.e: changes in taxes, labor income, technology, etc. We can also examine the life-cycle property of portfolio shares from the model. We know that on average, the old witness more data and grow more optimistic about stock returns, which makes them to invest a higher share of their wealth in the risky share. A positive correlation between risky share and age are seen both in the model and in the data from PSID, albeit with

[10] I used the Saez and Zucman top income database to get the top 1% share in the data, which ends in 2016, therefore I look at the changes until only 2016. The risky share and age correlation is estimated from the PSID data, and I used the 2017 to approximate its value in 2016 due to the lack of data in 2016.

different levels. In the model, such correlation amounts to 0.9164, while in the data, it's only 0.3644. This is not surprising, since the data also consists of many old households who cash out from the market to finance retirement consumption, while the model focuses on before-retirement investment patterns. Finally, the model does a pretty good in explaining the correlation between experienced return and expected return. Using direct expectation survey data from UBS, Malmendier and Nagel (2011) finds that a 1% increase in the experienced return predicts a $0.6 - 0.7\%$ of increase in expected return. The model generates a correlation of 0.6273 between the experienced and the expected return, which is very close to the data.

Next, we need to check the robustness of these results for alternative parameter values. As mentioned above, the US has been a relatively tranquil country relative to others. In principle, one can either vary the disaster size of disaster frequency. However, since there is only 1-2 disasters per 100 years in the US, I will stick to the international estimates for the disaster frequency, and vary the disaster size. In Barro (2006), the per capita reduction of real GDP, adjusted by trend growth is 35% in the international sample. However, the Great depression features a relatively smaller reduction, which totals 33%. By using $\kappa = -0.33$ and redo the calibration, one can see that the predicted changes in the old to young wealth ratio is now slightly lower, albeit still amounts 10.11% of the increase. The predictions on other moments do not change much in response to the changes in κ.

6. Empirical Evidence

In the model, I consider the Great Depression as the only source of rare disasters in the last 100 years in the US history, and that disasters all have the same jump size. This makes the model analytically tractable, but it neglects the potential impacts of smaller shocks on the wealth distribution. In this section, I provide additional empirical evidence on generational belief differences and its correlation with top wealth shares. Figure 6 plots the magnitude of rare stock market crashes measured by the percentage reduction of S&P 500 values from peak to trough. It uses monthly data from Shiller's stock market index ranging from 1871.01 to 2016.12. As one can see, such events have been rather rare, and that the the Great Depression has so far the largest size of stock market crash, which features a 84.76% loss of GDP that lasted five years in total. In his famous book "The Greatest Generation" (Brokaw (2000)), Tom Broklaw dubbed the young people during that period of time as the greatest generation, who not only survived through the stock market crash, but also lived through extreme social turmoil, high income inequality, and eventually WWII.

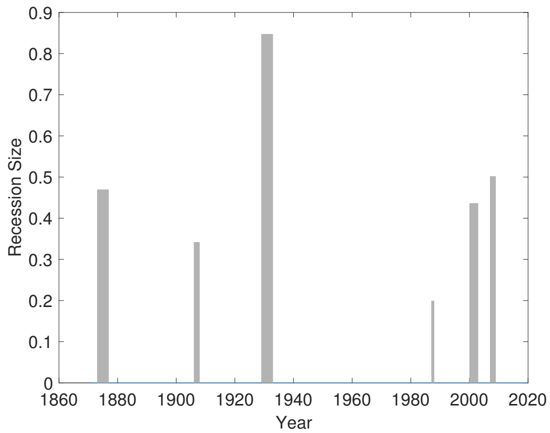

FIGURE 6. Stock Market Crash

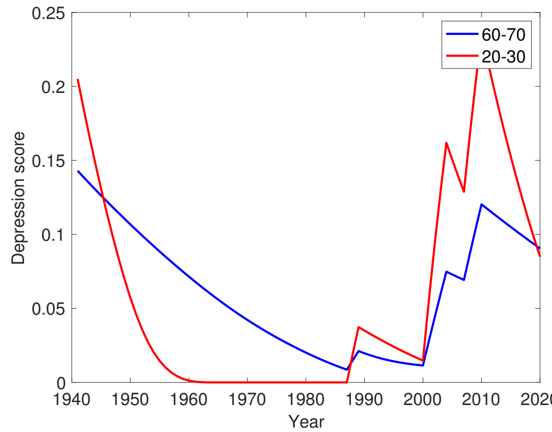

FIGURE 7. Pessimism Index

Nevertheless, those traumatic events could have left profound mental impacts, and scarred the economic optimism of those generations. To illustrate this, Figure 7 plots the pessimism index from 1941 to 2020 using the same data, contrasting differences in pessimism between the old (60-70 years old) and the young (20-30 years old).[11] The pessimism index $P_{i,t}$ for household i at time t is defined as a lifetime weighted average of depression loss, or more precisely,

$$P_{i,t}(\lambda) = \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k,\lambda)\mathbb{1}(Depression_{t-k} = 1)L_{t-k} \tag{6.45}$$

where $\omega_{i,t}(k,\lambda) = \frac{(age_{i,t}-k)^{\lambda}}{\sum_{k=1}^{age_{i,t}-1}(age_{i,t}-k)^{\lambda}}$ and L_{t-k} denotes the percentage loss in year $t-k$. The depression experience weighting function is identical to the return experience weighting function *a la* Malmendier and Nagel (2011), with the weighting parameter $\lambda = 1.5$ that they estimated using the SCF data, and is discussed in detail in Appendix 11.1. Here, I use the same experience weighting function to construct the pessimism index, and classify a downturn as a disaster if the peak to trough value drop of more than 20%.

Interesting patterns emerge in Figure 7. Before mid 1980s, pessimism for both young and the old are decreasing, but the pessimism for the young generations are decreasing at a much faster speed. While the old are still digesting trauma from the Great depression, and possibly also the 1873 stock market crash as well the 1907 panic, the young who luckily escaped those events are getting increasingly more optimistic relative to the old. This pattern continued to last until mid 1980s. Then the table turned. With smaller

[11]Note that the stock market data only goes back to 1871. Therefore, to understand the experience of a 70 year old, the index only makes sense from 1941 and onward.

crashes in 1987s, the dotcom bust, and the 2007-2009 financial crisis, doubts were raised by the young people. Although the recent old are also experiencing these disasters, they are first of all, not the depression babies that had even worse experiences, and second of all, would put less weight on these recent events relative to the young people because they have had a longer sample. To sum, the old had been much more pessimistic relative to the young before mid 1980s, but turned much more optimistic afterwards. I will show later that this pattern still holds when we consider overall return experience, rather than only experience of a depression. So why is this pessimism index interesting? Remember, the famous U-shaped pattern of inequality also features a turning point around 1980s!

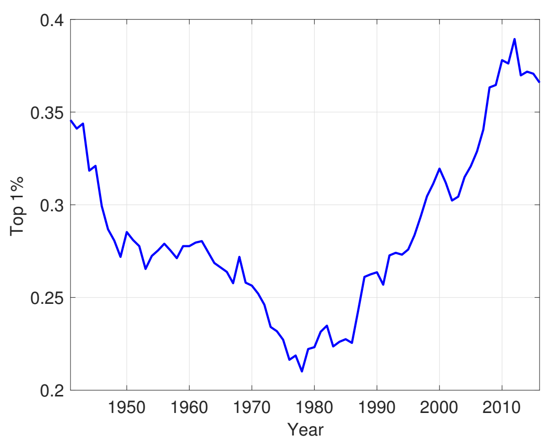

FIGURE 8. Top 1% wealth share

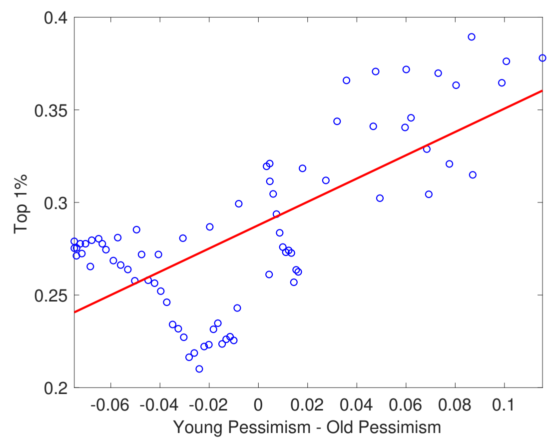

FIGURE 9. Top 1% wealth share s. Relative Optimism

To see the connection, Figure 8 plots the evolution of the top 1% wealth share in the United States using the Saez and Zucman (2016) data [12], which features a very famous U-shaped pattern with the 1980s as the turning point. Figure 9 plots the same statistics against relative optimism, defined as the difference between the young pessimism and the old pessimism. An obvious positive correlation emerges. At times when the old is more optimistic than the young, the top share is on average higher.

One might argue that households' beliefs not only react to extreme disastrous events, but could also revise gradually during normal times. After all, if generations experience both boom and bust, optimism induced by the boom might undo the depressing effect of the bust. Here, I examine in more detail if the generational belief differences are robust by considering overall experienced returns rather than only disaster experience. To capture

[12]I use the top income database top 1% net private wealth share data. Two years of missing values (1963 and 1965) are imputed with linear interpolation.

this idea, I ask the following question: In each year t, what is the subjective expected return for each cohort i implied by the model? Let r_t represent the actual realized annual return in year t, the expected annual return $er_{i,t}$, becomes

$$er_{i,t} = prob(Depression = 1)_{i,t} * \kappa + (1 - prob(Depression = 1)_{i,t}) \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k, \lambda) r_{t-k} \tag{6.46}$$

where

$$prob(Depression = 1) = \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k, \lambda) \mathbb{1}(Depression = 1) \tag{6.47}$$

and

$$\omega_{i,t}(k, \lambda) = \frac{(age_{i,t} - k)^{\lambda}}{\sum_{k=1}^{age_{i,t}-1} (age_{i,t} - k)^{\lambda}} \tag{6.48}$$

This captures the idea that the expected returns are the weighted average of the disaster return and the experienced return, with changing subjective likelihood of the disaster governed by the experience of the household. I use the monthly total real stock return of S&P 500 from Shiller's dataset, and convert returns into annual frequency. [13] Since there is no stock market return data before 1871.01, I compute the beliefs for all cohort in 1871 assuming that no disasters happens before that, so that disaster likelihood decreases gradually with age. Figure 10 compares the expected return for old vs. young.

Up until the 1980s, the young expected higher returns relative to the old. This is understandable, because while the old struggled with the aftermath of the Great depression and possibly earlier crashes, the young cohort enjoyed a good life, especially the Baby boomers. Notice that their expected return dropped in the later part of this period due to a slight downturn in the stock market in 1960-1970s, there was no major disasters during this period, and the they are still much more optimistic than the old. However, the table turned during the 1980s. With the 1987 crash, the 2000 dotcom bubble bust, and even more so the recent financial crisis, the new young generation become traumatized. Taking into account of possible future crashes, they even start to expect negative returns. Notice that there is a short period where the young people's optimism are boosted (i.e: the stock

[13] Malmendier and Nagel (2011) used the annual historical return in constructing cumulative returns. However, I find that this measure of experienced return is slightly different from actual experienced return. For example, if the market is down by $x\%$ in one year, and up by $x\%$ the year after, this definition of average return would indicate that the average return is zero. After incorporating the recency bias parameter λ, this would even indicate a positive experienced return. However, in reality, the household wealth at the end of the second year is down by $x^2\%$ already. In order to correct this bias, I instead use the average cumulative annual return to proxy experienced return. In the example, the second year experienced return would be $-x^2\%$. This reflects the lingering effect of the downturn. In Appendix 11.2, I show that although there is slight difference in these two measures, the qualitative pattern of the expected returns of old vs. young still holds.

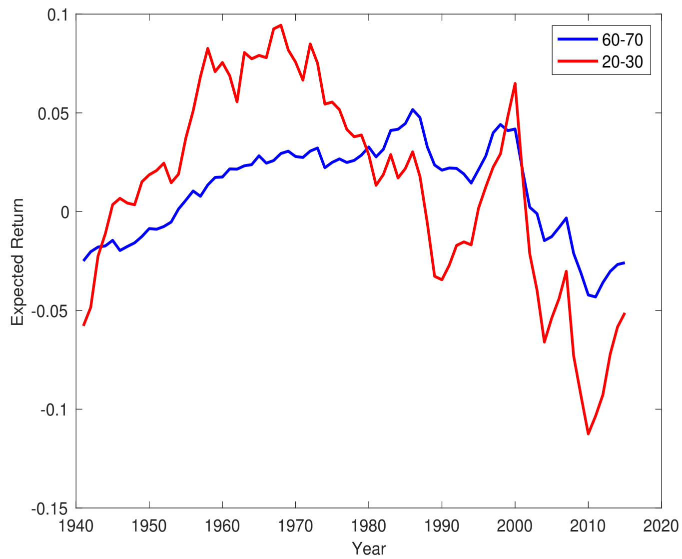

Figure 10. Expected Return: Old vs. Young

market boom in the 1990s), it is not enough to undo the negative effect of the two recent crisis they experience. Although the old, especially the boomers, have had similar experience, they still have the memory of the good old times, and are more optimistic about the returns.

Figure 11 shows that these generational belief differences translate into behaviors that generate an increase in generational wealth differences. It plots both the old to young wealth ratio in terms of overall net worth and the corporate equities and mutual fund. If one uses the blue line as a benchmark to measure of how overall generational wealth inequality trended up, the red line tells us that the old are owning disproportionally more stocks relative to the young, and that generational inequality measured by coporate equity and mutual fund is even more severe than the overall generational inequality.

7. A Comparison to Alternative Mechanisms

7.1. **What about housing?** A common question that many people ask is: is not this all about housing? After all, the last few decades have witnessed large swings in housing prices. Given that older people are more likely to be home owners than the young people, changes in housing prices and home ownership seem likely to account for the majority of

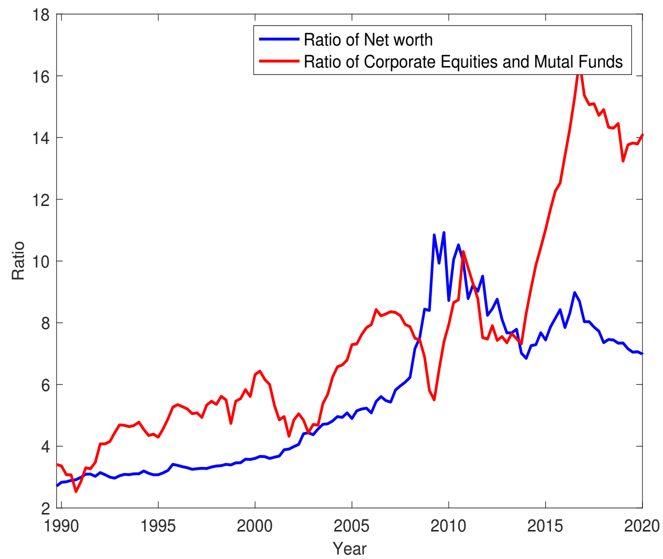

FIGURE 11. Wealth Ratio: 55-60 vs. 40 and under (SCF)

changes in generational inequality (Kuhn, Schularick, and Steins (2017)), Rognlie (2016)).

To disentangle overall wealth from housing wealth, I use PSID data from 1984 to 2011 [14], and plot again the wealth ratio of 65 years old and above vs. 35 years old and under, where the definition of wealth excludes housing wealth. Although the definition of wealth here excludes values of main residence, I construct a new measure of wealth that extracts from all other assets related to real estate (i.e: a second home, land, rental real estate, or money owed on a land contract), in order to fully eliminate the effect of housing prices changes on inequality.

Once again, we can see a striking increase of the old vs. young wealth ratio. While the wealth ratio of households that are 65 years old and above relative to the 35 years old and under was only 3.25 times in 1984, this ratio has increased to a shocking 11.49 times. This implies a 254% increase within merely 27 years. In contrast, the overall wealth ratio has increased by 126% (albeit with slightly different overlapping years). It shows that although housing could have greatly contributed to the overall increase in the *level* of generational inequality, the *rate* at which cohort inequality increased is rising at an even faster speed.

[14]Although information on overall wealth is available now until 2017, the same consistent measure of real estate wealth is not reported after 2011

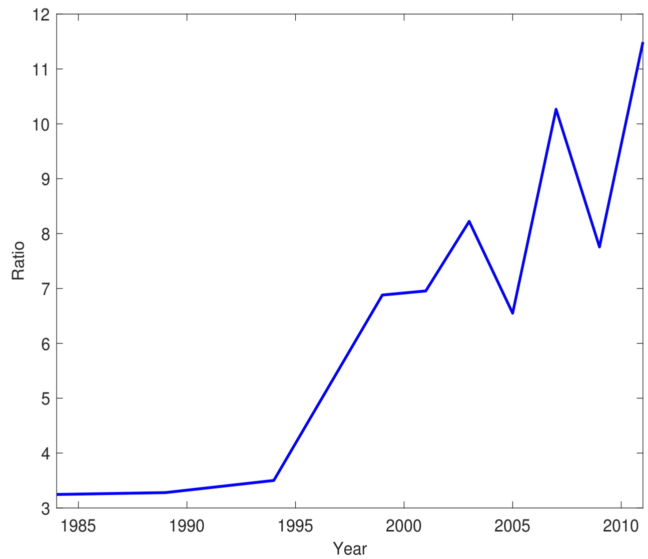

FIGURE 12. Net Worth Ratio Excluding Housing (PSID)

7.2. **Financial Market Development.** One obvious concern could be that the financial market became much more developed during the 1980s, which produced an increase in stock market participation. This increases the growth rate of wealth of everyone, while also disproportionately benefiting the older more, since they have more wealth to be invested than the young. While I acknowledge that extensive margin of financial inclusion could be an essential aspect in understanding cross-generational inequality, it does not capture the intensive margin of portfolio allocation. Here, I only focus on stock market participants, and examine the life cycle behavior of portfolio allocation in 1984 and 2017 using PSID data. If the "learning from experience" channel exists, the slope of life cycle risky stock share would be very different in these two years. As expected, in both years, stock share as a fraction of wealth increases with age. This is also widely documented in empirical papers in household finance. However, what is also striking is that these two years have very different slopes in stock share and age relationship. In 1984, the correlation of stock share and age was only 0.2708, but in 2017, the correlation rises to 0.4579. As we know from Malmendier and Nagel (2011) that the Great depression has led those who experienced to both participate less in the stock market (the extensive margin), as well as investing less conditional on participating in the stock market(the intensive margin). The PSID data suggests that, even if we attribute all the extensive margin to factors other than learning from experience, we still cannot neglect its effect on the intensive margin, which perhaps has nothing to do with financial market inclusion.

7.3. **Relaxed Borrowing Constraints.** The development in financial markets also relaxed borrowing constraint in the US since early 1980s. There are two aspects of the argument: First, since the old are usually not hand to mouth, they can leverage on existing wealth, and profit from higher returns in the stock market. Second, the loosening borrowing constraint has led the young to decumulate wealth instead of saving, whose effect on increasing wealth inequality is well documented in Favilukis (2013). Polarization occurs when the former makes the older richer, while the latter makes the younger poorer. Whichever is the channel, this implies that if we look at the gross rather than the net asset level, the generational inequality could be very different. Suppose we see that gross wealth inequality hasn't increased between cohorts, but net wealth inequality has increased, then it is more likely that loosening borrowing constraints are the main driver of cross-cohort inequality. To examine this, I go back to the PSID data, and again divide the sample into 65 years old and over, versus 35 years and under and examine their gross wealth ratio. Again, in 1984, the wealth ratio of the two groups was 3.346 times, but in 2009 (the year till which PSID has the same definition of debt), the ratio becomes 8.856 times. This suggests that there are forces other than loosening borrowing constraint that are contributing to the divergence of the young and the old.

7.4. **Direct and Indirect Inter-generational Transfers.** Inheritance and other intergenerational transfers play a potentially crucial role in generational inequality (See Boar (2020)). Maybe the millennials have nothing to worry about it, since they will inherit their parents' houses and bank accounts. On the other hand, maybe the increased cost of life extending medical treatments will cause boomers to exhaust all their wealth before they die. This section examines if the results of the paper are robust to inter-generational transfers. Evidence suggests that inheritances have doubled since the 1980s (Alvaredo, Garbinti, and Piketty (2017)). However, this rise has an equalizing effect on wealth distribution (Wolff (2002)) because even though the overall amount of inheritance has been rising, the share of wealth in inheritance has been declining dramatically during this period. One might argue that even though the overall inequality could be equalized, generational inequality might not, because older people are on average more likely to have inheritance than younger people. To examine the robustness of the old to young wealth ratio, I again use PSID data and compare the old to young wealth ratio (above 65 vs. under 35) with and without inheritance. In 1995, inheritance makes no different to this ratio, which has a value of 6.05 [15], while in 2013, there is only slight difference. The old to young wealth ratio is 17.23 after inheritance, and becomes 17.41 before inheritance. Therefore, the old

[15]The earliest information on inheritance value starts in 1995. However, there is no wealth data in that year. A linear interpolation is taken between the two surveys in 1994 and 1999 to impute the 1995 wealth level

to young wealth ratio does not differ much by varying direct transfers that in the form of inheritance. But what about indirect transfers that take the form of education expenses? After all, college tuition has become much more expensive over the last two to three decades. Capelle (2019) shows that the US higher education system has contributed greatly to increased inter-generational immobility with rising tuition fees. If the older adults are paying tuition for their children, it serves as a direct wealth transfer to the young people, which could decrease the real old to young wealth ratio. To check this, I subtract cumulative education expenses from net wealth, the assumption being that these are the tuition paid to finance the education of their kids. Since wealth is a stock variable, but education expenses are only reported as a flow variable, I adjust the cumulative education expense by four times of the yearly reported education expense by assuming that these expenses occur due to the four year college education. Interestingly, without taking into account tuition expense, the old to young wealth ratio grew from 8.26 times to 13.5 times, which is about a 63% increase. If one subtract wealth by education expense, the ratio went from 9.269 times to 15.756 times, which is around 70% of increase. So in fact, the rise in college tuition hurt the young people's wealth more than the old. One possible interpretation of this is that the tuition-paying parents are mostly middle aged instead of being over 65 years old, and when they reach 65 years and beyond, their college-educated kids have already graduated, so even though the tuition expense might affect the family budget while the parents are in the middle age, it does not affect the 65 years older group that much. At the same time, the rising education expense pushes young people to take out higher values of student loans, which hurts young people's financial wealth more than the old. Of course, the young might recoup this expense in the form of higher future labor income, but perhaps they won't. Finally, since we are discussing generational inequality in the USA, we must briefly consider social security. In the US, the social security program has been expanded a lot over the last several decades (See Bourne, Edwards, et al. (2019)). Since it primarily operates as a pay-as-you-go system, secular changes in demographics and productivity potentially induce large generational redistribution, depending on whether unfunded liabilities are financed by tax increases or benefit cuts (Kotlikoff and Burns (2005)). The type of social security that matters for generational inequality comes in the form of retirement wealth. One might argue that if we were to incorporate social security wealth into the definition of wealth, generational inequality might not be that bad, because even if young people nowadays might look poor on paper, they might still have a lot of retirement wealth to spare in the future. To examine this, I re-calculate the old to young wealth ratio in PSID in 1989 and 2013. Without retirement wealth, the old to young wealth ratio increased from 4.3 times to 17.42 times. If one adds retirement wealth into overall wealth, the increase is a little milder, which features 4.32 times in 1989 and

11.14 times in 2013. That is, even though the increase is milder, there is still significant rise in generational inequality from the 1980s.

7.5. **Increased Supply of Data.** One might ask, why learning from experience? Wouldn't standard Bayesian learning that incorporates all historical data also generate wealth dispersion, if everyone becomes more optimistic when more data become available? Perhaps pessimism induced by the great depression makes everyone more pessimistic and invest less, which reduces inequality at the beginning, and then overtime, optimism builds, everyone becomes more optimistic and invests more again, thus the economy exhibits rising inequality. This argument might sound plausible at a glance. After all, it seems consistent with the famous U-shaped pattern of inequality that we have seen in the last century. While this explanation sounds appealing at a glance (In fact, it has once inspired a whole literature on asymmetric information and return heterogeneity), it contrasts the data on expectations. If we think that investors learn not just from their own experience, but all the data available to them, and that they are only constrained by the supply of data. Then overtime, as more data reaches to them, their beliefs should become increasingly closer to each other, even if they start out having very different prior. The monthly Shiller's data starting from 1989 on stock market crash optimism index shows that this is simply not the case. It measures the percent of the population who attach little probability (strictly less than 10%) to a stock market crash in the next six months. Therefore, it is a direct measure of beliefs about stock market disaster likelihood. Each index is derived from the responses to a single question that has been asked consistently through time since 1989 to a consistent sample of respondents. Figure 13 plots the standard error of the measure for the institutional as well as the individual data. Using standard error as a measure of belief heterogeneity, Clearly, there is no evidence that beliefs are in any foreseeable future converging. If anything, it slightly diverges more after the recent financial crisis.

8. Efficiency and Policy Implications

An interesting feature of the model in this paper is that inequality is generated in a complete markets economy. This is in contrast with most other models studying inequality (e.g: Hugget or Bewley models). But does this imply that inequality is efficient? Probably not. In fact, with heterogeneous beliefs, the conventional Pareto criterion is open to debate. Several improved Pareto criteria have been proposed in recent literature. In fact, questions about the Pareto criterion date back to the 1970s Starr (1973), Harris (1978) and Hammond (1981), who first highlighted that when beliefs are different, ex-ante efficiency might not correspond to ex-post efficiency. This is true in my model too. With heterogeneous priors and experiential learning, each investor in the financial market

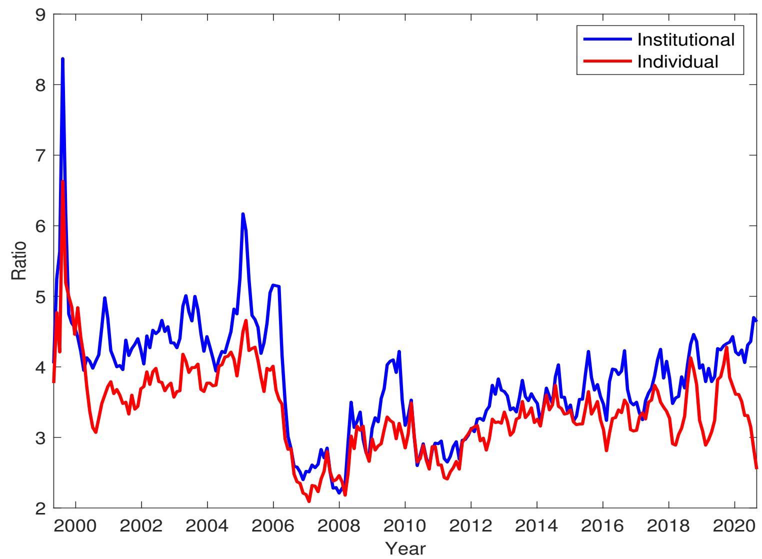

FIGURE 13. Measure of Belief Convergence: Standard Error of Cash Confidence Index

considers their own beliefs to be correct. Each thinks they would be better off with speculation ex ante. However, ex post consumption would always be more volatile from a social welfare point of view. Indeed, from behind the veil of ignorance, all investors agree that they cannot all have correct beliefs, and that their future perceived welfare gains are likely to be spurious. Another limitation of the conventional Pareto criteria lies in the assumption that the planner has the ability to know the true data generating process, which is not realistic either. An example of this is bubbles. A long period of rising asset prices might be identified as overoptimism and bubbles ex post, but there is considerable ex ante uncertainty about their presence. If no market participant knows this, it would be a too strong assumption to assume that a social planner knows this. Recent attempts to propose new Pareto criteria in evaluating efficiency are numerous. For example, to address the problem of whose beliefs to evaluate under, Brunnermeier, Simsek, and Xiong (2014) proposed an enhanced version of the Pareto criterion by suggesting a belief-neutral efficiency criterion, where an allocation is efficient if it's efficient under any convex combination of agents' beliefs. To address the problem of incomplete knowledge of the planner, Walden and Heyerdahl-Larsen (2015) proposes an incomplete knowledge efficiency criterion to evaluate efficiency and distortion from a planner's point of view. Another practical criterion related to financial regulation is Gayer, Gilboa, Samuelson, and Schmeidler (2014), who propose a no betting criterion to assess whether speculative trading should take place or not.

9. Literature review

This paper is related to four strands of literature. First, it is largely inspired by the recent macro literature that examines the implications of deviations from rational expectations. As shown in a seminal paper by Woodford (2013), although the literature hasn't reached an unequivocal verdict regarding what expectation formation rules researchers should adopt, a promising approach that relies on a statistically modest deviation from rational expectations is to assume that beliefs are refined through induction from observed history. The over-weighing of personal experiences has long been discussed in the psychology literature, named as availability bias as in Tversky and Kahneman (1974). Compared with a full Bayesian approach, such belief formation mechanism exhibits strong over extrapolation behavior (See Greenwood and Shleifer (2014) for a survey). Barberis, Greenwood, Jin, and Shleifer (2015) and Barberis, Greenwood, Jin, and Shleifer (2018) rationalize a set of asset pricing anomalies when an over-extrapolative investor interact with a rational agent in the financial market. Evidence of over extrapolation is pervasive. In financial markets, it is supported by a seminal paper Malmendier and Nagel (2011), who uses data from Survey of Consumer Finance and provides strong empirical support that personal experience in the stock market has a prolonged impact on how much they invest in risky assets later in their lives. In particular, those that experienced the 1930s great depression were less willing to participate in the stock market, and invest significantly less even if they participate. Such belief formation is not only present in the stock market, but also influences households' expectation formation of inflation, labor market, housing market as well as overall business cycle conditions. (Malmendier and Nagel (2015), Wee (2016), Malmendier and Shen (2018) Kozlowski, Veldkamp, and Venkateswaran (2020) and Kuchler and Zafar (2019)).

Second, this paper attempts to generate heterogeneous beliefs when individuals learn from their own experience. Most macro-finance models with heterogeneous beliefs focus on *exogenous* heterogeneous beliefs. Classic work includes Basak (2005), Harrison and Kreps (1979), Scheinkman and Xiong (2003) and Borovička (2020), just to name a few. Since their focus is on asset prices, belief heterogeneity could be taken as an input without having to model where it comes from. In this paper, beliefs are essentially *endogenous*, which for my purpose helps to link observable demographic structures with inequality. Nevertheless, this is not the first paper to do so. Recent advancement has studied the aggregate implication of heterogeneous generational bias stemming from learning from experience. The fact that younger people update their beliefs more frequently than the old has interesting implications on asset prices. Ehling, Graniero, and Heyerdahl-Larsen (2017) develop an elegant asset pricing model with learning from experience in a stationary diffusion environment. Malmendier, Pouzo, and Vanasco (2019) solves a similar problems in an incomplete market. Schraeder (2015) considers a noisy-rational expectation model with generational bias when agents have CARA preferences, and Collin-Dufresne, Johannes, and Lochstoer (2016) solves such model with Epstein-Zin preference, albeit with two generations.

Third, this paper is related to recent literature on disaster risk in the tradition of Barro (2006). The incorporation of risk of rare disasters naturally generates a disaster premium, which significantly reduces the level of risk aversion needed in matching empirically plausible equity premium. Various extensions of disaster risk models also helps to solve the equity premium puzzle, the volatility puzzle, return predictability,etc((See Tsai and Wachter (2015) for a survey). When disaster risk is unknown and agents must infer its distribution from historical data, Koulovatianos and Wieland (2011) shows that pessimism is triggered upon the realization of a rare disaster, and rationalizes a prolonged period of decline in P-D ratio. Moreover, they prove that although asymptotic beliefs are unbiased, one never reaches full optimism of disaster risk as one would under rational expectation. It is the slow arrival of information of disasters that keeps learning away from reaching infinite precision. In my model, the realization of a large negative shock (e.g: the Great Depression) would trigger such response from investors that experienced it, thus generating heterogeneous generational bias in the disaster risk distribution. Although there are several interesting papers that combines heterogeneous beliefs or attitudes towards disaster risk in both complete and incomplete markets (Bates (2008), Chen, Joslin, and Tran (2010), Dieckmann (2011), Chen, Joslin, and Tran (2012)), these models builds on two-agents and focus on cases with dogmatic beliefs, while my model features a continuum of heterogeneous agents with learning agents that constantly update their beliefs optimally, and focus on the evolution of wealth distribution.

Last but not least, this paper contributes to the recent advancement of HACT (heterogeneous agent continuous time) models that link distributional considerations with macroeconomics (Gabaix, Lasry, Lions, and Moll (2016), Achdou, Han, Lasry, Lions, and Moll (2017) and Ahn, Kaplan, Moll, Winberry, and Wolf (2018). However, studying belief heterogeneity in such framework is still a relatively new area. Two recent papers attempt to incorporate endogenous heterogeneous beliefs into such a framework (Kasa and Lei (2018), Lei (2019)), and rationalize"state dependence" in the growth rate of wealth, which rationalizes why inequality has been growing at such a fast speed after 1980s. However, they focus on inequality within cohort with private equities. Here, I generalize those models, and am able to solve distribution across cohort, and solve a model with aggregate shock and public equity. Finally, by tracing rare disasters all the way back to the Great depression, it allows me to jointly explain both the dip of wealth inequality after the Great depression, as well as the rise of inequality after the 80s.

10. Conclusion

The real world features finite lives and limited data for all of us to process. This paper asks whether learning from finite life experience in the stock market influences the distribution of wealth in our society, and proposes a model calibrated to US data. It bridges the gap between the experiential learning literature, which is traditionally a behavioral finance concept, and the macroeconomic literature on wealth inequality. It highlights how traumatic events like the Great depression could have a prolonged impact on cross-generational inequality through the channel of learning from experience. I build and solve a general equilibrium model with learning from experience agents, and examine the qualitative as well as quantitative implications on long-run wealth differences between cohorts. To the best of my knowledge, this is the first paper that combines learning from experience with wealth inequality, which should spark interest in many possible extensions. So far, I have built and solved a general equilibrium economy with learning and heterogeneous beliefs about stock market experience. However, one could also think of extending the framework with nominal rigidity, so that one can explore the role of monetary policy when agents are learning from inflation experience (which also exhibits strong recency bias as documented by Malmendier and Nagel (2016)). One can also generalize the current framework to incorporate features in the housing market, such as borrowing and collateral constraints, to study the distributional effect of learning from housing market experience, etc. When differences in beliefs across generation matters, it opens doors to policy makers to have a role in reducing inequality. An example would be certain semi-mandatory pension funds designed to improve the growth rate of wealth of the depressed generations by investing in stocks, when investors in those generations are themselves too scared to be engaged in risky assets.

11. Appendix

11.1. **The experience weighting function.** Figure 14 plots and compares the weights used to construct the pessimism index in 1980 by comparing a typical depression baby (age 70) and a typical boomer (age 30) as an example, with a weighting parameter $\lambda = 1.5$ estimated by Malmendier and Nagel (2011). Notice that $\lambda > 0$ implies that households exhibit recency bias, so the weights decreases with the number of days before today. Two things are noticeable. First, although both generations over-weigh recent data, the young people over-weigh even more. This is because they live through a shorter life span. Second, the depression babies still has the hangover of the Great depression happened 47-51 years ago, while a boomer would put zero weight on that.

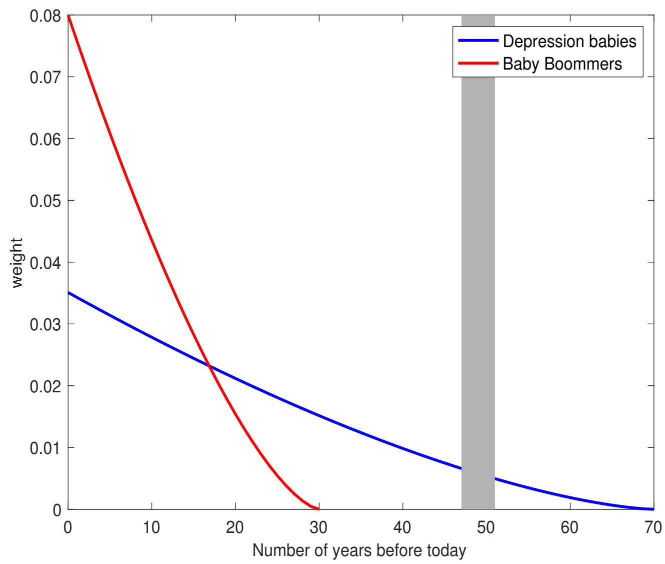

Figure 14. Historical weights: Depression babies vs. Boomers

11.2. **Robustness check on experienced return.** The following two figures plots the generational belief differences using two different measures of experienced returns.

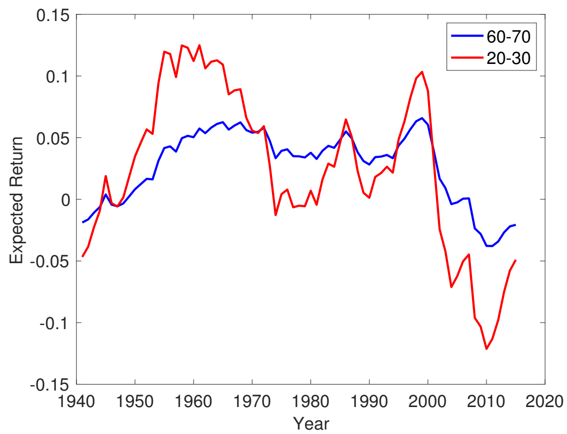

FIGURE 15. Using experienced annual return

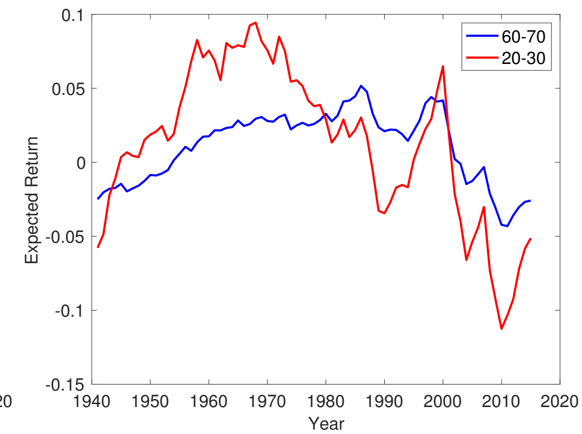

FIGURE 16. Using average cumulative annual return

11.3. **Proof of Lemma 2.2.** See Dieckmann (2011) for the proof of eqn.(2.15) and eqn.(2.16). The derivation of $\xi_{s,t}$ process follows first by applying the Girsanov theorem for the jump process, s.t:

$$dN_{s,t} - \bar{\lambda}_{s,t}dt = dN_t(\lambda_t) - \bar{\lambda}_t dt \tag{11.49}$$

With the change of measure, we can rewrite eqn.(2.15) into

$$\frac{d\xi_{s,t}}{\xi_{s,t-}} = \left(\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t + (\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1)(\lambda_{s,t} - \bar{\lambda}_t)\right) dt - \theta_{s,t} dZ_t + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right) dN_t(\bar{\lambda}_t) \tag{11.50}$$

Then the SDE for $\eta_{s,t}$ follows directly from the application of multidimensional jump-diffusion version of the Ito's lemma. Notice that all agents agree on the diffusion risk, therefore we can simplify the solution by imposing $\theta_{s,t} = \theta_t$, and that $dZ_{s,t} = dZ_t$. We can further simplify the expression by noticing that by definition, the market price of the jump risk is defined by $\lambda^N_{s,t} = \frac{\lambda_{s,t}}{1+\bar{\kappa}}$. Applying Ito's lemma again on $\eta_{s,t} = \frac{\xi_t}{\xi_{s,t}}$, we have

$$\frac{d\eta_{s,t}}{\eta_{s,t}} = \left(\frac{1}{1+\bar{\kappa}}\lambda_{s,t} - \lambda^N_t\right) dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2\lambda^N_t}{\lambda_t} - 1\right) - 1\right] dN(\bar{\lambda}_t) \tag{11.51}$$

11.4. **Proof of proposition 2.** To get the coefficient of the stock price, we can write down the formula for stock prices, i.e:

$$\begin{aligned} S_t &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty \xi_u D_u du\right] \\ &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty e^{-(\rho+\delta(1-\beta))u}\eta_u du\right] \\ &= \frac{1}{\xi_t}\eta_t\int_t^\infty e^{-(\rho+\delta(1-\beta))u}du \\ &= \frac{1}{\rho+\delta(1-\beta)}Y_t \end{aligned} \tag{11.52}$$

That is, stock price to dividend ratio is a constant, i.e:

$$\frac{dS_t}{S_{t^-}} = \frac{dY_t}{Y_{t^-}} \tag{11.53}$$

Recall that the compounded stock market value follows the following process

$$\frac{dS_t + D_t dt}{S_{t^-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{11.54}$$

Matching coefficients, one get

$$\mu^S = \mu + \rho + \delta(1-\beta); \quad \sigma^S = \sigma; \quad \kappa_t^S = \kappa_t \tag{11.55}$$

Now let's turn to the pricing of the disaster insurance product. By definition, we have

$$\mu_t^P = -\kappa_t^P\lambda_t^N + r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\lambda_{s,t}) + r_t \tag{11.56}$$

12. Proof of Proposition 3

I first derive the stationary KFP equation with a general jump diffusion process of any random variable $w_{s,t}$

$$\frac{dw_{s,t}}{w_{s,t^-}} = \hat{\mu}_{s,t}dt + \hat{\sigma}_{s,t}dZ_t + \hat{\kappa}_{s,t}dN_t \tag{12.57}$$

where dZ_t and dN_t represents aggregate Brownian motion and jump shocks. To simplify notation, I will now eliminate birth and current time combo (s,t) notations in the following text. Let $f(w)$ be any function of w, $n(w)$ be the density function of w, and let $A(t+dt)$ denotes the conditional expectation of $f(w)$ at $t+dt$. We then have

$$\begin{aligned} A(t+dt) &= \int_{-\infty}^\infty f(w)n_{t+dt}dw \\ &= \int_{-\infty}^\infty (f(w)+df(w))n(w) - \delta f(w)n(w)dw \\ &= \int_{-\infty}^\infty f(w)(1-\delta)n(w)dw + \int_{-\infty}^\infty df(w)n(w)dw \end{aligned} \tag{12.58}$$

We then have

$$d(A(t)) = -\int_{-\infty}^{\infty} \delta n(w) f(w) dw + \int_{-\infty}^{\infty} df(w) n(w) dw. \tag{12.59}$$

Applying Ito's lemma for the jump diffusion process of w, we can get

$$df(w) = f'(w)[\hat{\mu} w dt + \hat{\sigma} w dZ] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt + [f(w(1+\hat{\kappa})) - f(w)]dN \tag{12.60}$$

Using integration by parts, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} df(w) n(w) dw &= \int_{-\infty}^{\infty} \left[f'(w) \left[\hat{\mu} w dt + \hat{\sigma} w dZ \right] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt \right] n(w) dw \\
&+ \int_{-\infty}^{\infty} \left[f(w(1+\hat{\kappa})) - f(w) \right] n(w) dN dw \\
&= \int_{-\infty}^{\infty} f(w) \left[-\frac{\partial}{\partial w} \left(n(w) \hat{\mu} w dt + n(w) \hat{\sigma} w dZ_t \right) + \frac{1}{2} f(w) \frac{\partial^2}{\partial w^2} \left(n(w) \hat{\sigma}^2 w^2 \right) dt \right] \\
&+ \int_{-\infty}^{\infty} \left[n(w(1+\hat{\kappa})) - n(w) \right] f(w) dN dw
\end{aligned} \tag{12.61}$$

Notice that the way I write down changes in $A(t)$ in (12.59) fixes the density of w in the state space and calculate with Ito's Lemma how $f(w)$ will change. One can also approximate $d(A(t))$ by linearly extrapolating the density at each point, that is,

$$d(A(t)) = \int_{-\infty}^{\infty} f(w) \frac{\partial n}{\partial t} dt dw = \int_{-\infty}^{\infty} df(w) n(w) dw \tag{12.62}$$

Plugging in the expression in eqn. (12.61), and equating the integrands, we get

$$dn = -\frac{\partial}{\partial w} (n \hat{\mu} w dt + n \hat{\sigma} w dZ) + \frac{1}{2} \frac{\partial^2}{\partial w^2} (n \hat{\sigma}^2 w^2) dt + [n(w(1+\hat{\kappa}), t)) - n(w,t)]dN \tag{12.63}$$

As one can see, the distribution of this variable is stochastic, and that there is no closed form solution in general. However, we can still ask the question, what is the long-run stationary distribution of this variable in this economy? That is, what is the solution of $dp(w) = \mathbb{E}_t(dn(w)) = 0$ (where the expectation denotes the time-average)? By averaging out the KFP, we thus have

$$-\frac{\partial}{\partial w} \left(\mathbb{E}(\hat{\mu}) w p(w) \right) + \frac{\partial^2}{\partial w^2} \left(\frac{\mathbb{E}(\hat{\sigma}^2)}{2} w^2 p(w) \right) - \delta p(w) + \lambda(p^J - p) = 0 \tag{12.64}$$

I now apply this stationary KFP to the variables of interest in this model. Since the aggregate economy is growing exponentially, and the newborn gets a constant share of it, we will need to normalize wealth to get a stationary distribution. Therefore, instead of examining the stationary distribution of absolute wealth, we will instead work with the following normalized variable:

$$\tilde{w}_{s,t} = \frac{w_{s,t}}{\omega Y_t} \tag{12.65}$$

That is, the absolute wealth normalized by the newborn's endowment. Since agents are born with zero financial wealth, we have $\tilde{w}_{s,s} = \frac{\omega Y_s}{\omega Y_s} = 1$. This should have a stationary distribution absent aggregate shocks. Recall that, after imposing the market clearing condition, the individual wealth dynamics follows the following

$$\frac{dw_{s,t}}{w_{s,t^-}} = \left(\sigma^2 + r - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right)\right)dt + \sigma dZ + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_t^N} - 1\right)dN_t \tag{12.66}$$

Applying Ito's lemma for the jump-diffusion processes, we then have

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \left(\sigma^2 + r - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right) - \mu\right)dt + \left(\frac{\lambda_{s,t}}{\mathbb{E}(\lambda_{s,t})}(1+\kappa_t) - 1\right)dN_t \tag{12.67}$$

which in short-hand can be written as

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \hat{\mu}(\lambda_{s,t})dt + \hat{\kappa}(\lambda_{s,t})dN_t \tag{12.68}$$

It turns out to be easier to work with log of wealth. Define $x = \log(\tilde{w})$. With Ito's lemma, we can rewrite the above into

$$dx = \hat{\mu}dt + \log(1+\hat{\kappa})dN_t \tag{12.69}$$

Recall that our final goal is to compute the long-run average marginal density of log wealth $p(x)$, which can be seen as

$$p(x) = \int_0^\infty n(x,\lambda)d\lambda \tag{12.70}$$

Notice that we can further decompose the joint density $n(.)$ into the product of the marginal density of belief and the conditional density of wealth, i.e:

$$n(x,\lambda) = n_1(x|\lambda)n_2(\lambda) \tag{12.71}$$

From eqn. (12.69), we can write down the dynamics of $n_1(x|\lambda)$, i.e:

$$0 = -\frac{\partial n_1}{\partial x}\hat{\mu} + \lambda^0\left(n_1(\log(1+\hat{\kappa}) + x) - n_1\right) - \delta n_1 \tag{12.72}$$

We can guess and verify a solution $n_1 = Ae^{\zeta x}$, where $\zeta = \frac{\lambda^0\hat{\kappa} - \delta}{\hat{\mu}}$ and that A is the normalizing constant of the conditional distribution. We can further approximate ζ around $\lambda = \lambda^0 = 0$, and get

$$\zeta \approx \zeta_0 + (\lambda - \lambda^0)\zeta_1 \tag{12.73}$$

where $\zeta_0 = \frac{\bar{\kappa}\lambda^0 - \delta}{d}$ and $\zeta_1 = \frac{\bar{\kappa}d - \bar{\kappa}(\bar{\kappa}\lambda^0 - \delta)}{d^2}$, and where $a = \frac{1+\bar{\kappa}}{\mathbb{E}(\lambda_{s,t})}$, $c = -2 - \frac{\lambda^0}{\lambda^N}$, $d = \sigma^2 + r + \lambda^N + \delta + \lambda^0 - \mu$.

To compute $n_2(\lambda_{s,t})$, recall that

$$d\lambda_{s,t} = (\lambda_{s,t^-} - \lambda_l)(\lambda_{s,t} - \lambda_h)dt - (\lambda_{s,t^-} - \lambda_h)(\lambda_{s,t^-} - \lambda_l)\frac{(1 + \lambda_{s,t^-})}{\lambda_{s,t^-}}dN_t \tag{12.74}$$

Writing out the stationary KFP of $\lambda_{s,t}$ and again abstract away from super(sub)scripts, we can get

$$0 = -\frac{\partial n_2}{\partial \lambda}(\lambda - \lambda_h)(\lambda - \lambda_l) - n_2(2\lambda - \lambda_l - \lambda_h + \delta) + \lambda^0(n_2^J - n_2) \tag{12.75}$$

We can guess and verify the following approximate exponential solution

$$n_2(\lambda) \approx e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} \tag{12.76}$$

We can then substitute this into the above ODE, and match the constants. This ensures that the marginal density is non-negative, and that we are looking for a solution around $\lambda = 0$.

In the end, we can simply get the marginal distribution of log wealth by integrating the product of the conditional distribution of wealth and the marginal distribution of beliefs, i.e:

$$\begin{aligned} p(x) &= G_0 e^{(\zeta_0 - \lambda^0\zeta_1)x} \int_{\lambda_l}^{\lambda_h} e^{\lambda\zeta_1 x} e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} d\lambda \\ &= \underbrace{Ge^{\zeta_0 x}}_{RE} \underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}]}_{Learning} \end{aligned} \tag{12.77}$$

Let $p^{RE}(x)$ denote the long run stationary distribution of log normalized wealth in the rational expectation economy, we then have

We then have

$$\begin{aligned} \lim_{x\to\infty} \frac{p(x)}{p^{RE}(x)} &= \lim_{x\to\infty} [\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}] \\ &= \lim_{x\to\infty} \zeta_1^{-1}\left[-(\lambda_l - \lambda^0)\zeta_1 e^{(\lambda_l - \lambda^0)\zeta_1 x}\right] \end{aligned} \tag{12.78}$$

where the second equality uses the L'hopital's rule. Recall that $\zeta_1 = \frac{\bar{\kappa}d - \bar{\kappa}(\bar{\kappa}\lambda^0 - \delta)}{d^2}$. With the calibrated parameter values, we then know that $\zeta_1 < 0$. Therefore, the above expression goes to infinity wnen $x \to \infty$. We then have

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^{RE}(x) \tag{12.79}$$

That is, the experiential learning economy has a fatter right tail of wealth distribution compared with the standard RE economy.

12.1. **Verification of Newborn Consumption Share.** We start by defining β_t, i.e:

$$\beta_t = \frac{c_{t,t}}{Y_t} = \frac{(\rho+\delta)w_{t,t}}{Y_t} \tag{12.80}$$

where the second equality comes from consumption smoothing of a log agent. Since agents are born without financial wealth, $W_{t,t}$ is essentially the present value of all future earnings.

$$\begin{aligned} W_{t,t} &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty e^{-\delta(u-t)\xi_u}\omega Y_u du\right] \\ &= \omega Y_t \mathbb{E}_t\left[\int_t^\infty e^{-(\rho+\delta+\delta(1-\beta))(u-t)}\frac{\bar{\eta}_u}{\bar{\eta}_t}du\right] \\ &= \frac{\omega Y_t}{\rho+\delta+\delta(1-\beta)} \end{aligned} \tag{12.81}$$

where the second equality uses the definition of $\bar{\eta}_t$, and the third equality follows from the fact that the disagreement process $\bar{\eta}_t$ is a martingale. We then have a fixed point for β, i.e:

$$\beta = \frac{1}{\rho+\delta+\delta(1-\beta)} \tag{12.82}$$

This renders the two solutions

$$\beta_{1,2} = \frac{\rho+2\delta}{2\delta} \pm \frac{\sqrt{\rho^2+4(\rho+\delta)\delta(1-\omega)}}{2\delta} \tag{12.83}$$

However, since the stock price is $S_t = \frac{1-\omega}{\rho+\delta(1-\beta)}Y_t$, we know that $\beta < \frac{\rho+\delta}{\delta}$ has to hold. This eliminate the positive root of β, while the negative root can satisfy the constraint. So the value of β is

$$\beta = \frac{\rho+2\delta}{2\delta} - \frac{\sqrt{\rho^2+4(\rho+\delta)\delta(1-\omega)}}{2\delta} \tag{12.84}$$

12.2. **Savings rate Response to Stock Market Scarring.** The table shows the OLS regression results of contemporaneous savings rate on historical moving average of the following variables: stock return, GDP growth rate, inflation and federal funds rate. The stock return data is taken from Robert Shiller S&P 500 total real price return monthly data set, and all the rest of the variables come from St Louis Federal Reserve data set. All variables are converted to annualized value with quarterly frequency. Model 1 uses the 1 year moving average of the independent variables, while Model 2, 3 and 4 uses the 3 year, 5 year and 10 year moving average.

Table 4. Dependent Variable: Savings Rate

Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Stock return	0.277	0.078	-0.108	0.105
	(0.576)	(0.597)	(0.442)	(0.443)
GDP growth rate	0.332***	0.359***	0.273***	-0.047
	(0.081)	(0.084)	(0.070)	(0.064)
Inflation	0.518***	0.401***	0.360***	-0.300***
	(0.085)	(0.088)	(0.065)	(0.054)
Federal Fund rate	-0.094	-0.398***	-0.633***	0.232***
	(0.083)	(0.094)	(0.070)	(0.055)
Constant	6.101***	8.099***	9.604***	6.258
	(0.492)	(0.606)	(0.548)	(0.563)
N	220	196	172	112
R^2	0.225	0.156	0.373	0.272

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

References

Abel, A. B. (2001): "Will bequests attenuate the predicted meltdown in stock prices when baby boomers retire?," *Review of Economics and Statistics*, 83(4), 589–595.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2017): "Income and wealth distribution in macroeconomics: A continuous-time approach," Discussion paper, National Bureau of Economic Research.

Adam, K., A. Marcet, and J. Beutel (2017): "Stock price booms and expected capital gains," *American Economic Review*, 107(8), 2352–2408.

Adam, K., A. Marcet, and J. P. Nicolini (2016): "Stock market volatility and learning," *The Journal of Finance*, 71(1), 33–82.

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018): "When inequality matters for macro and macro matters for inequality," *NBER Macroeconomics Annual*, 32(1), 1–75.

Alvaredo, F., B. Garbinti, and T. Piketty (2017): "On the share of inheritance in aggregate wealth: Europe and the USA, 1900–2010," *Economica*, 84(334), 239–260.

Ameriks, J., and S. P. Zeldes (2004): "How do household portfolio shares vary with age," Discussion paper, working paper, Columbia University.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015): "X-CAPM: An extrapolative capital asset pricing model," *Journal of financial economics*, 115(1), 1–24.

——— (2018): "Extrapolation and bubbles," *Journal of Financial Economics*, 129(2), 203–227.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121(3), 823–866.

Basak, S. (2005): "Asset pricing with heterogeneous beliefs," *Journal of Banking & Finance*, 29(11), 2849–2881.

Bates, D. S. (2008): "The market for crash risk," *Journal of Economic Dynamics and Control*, 32(7), 2291–2321.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011): "Explaining asset pricing puzzles associated with the 1987 market crash," *Journal of Financial Economics*, 101(3), 552–573.

Boar, C. (2020): "Dynastic precautionary savings," Discussion paper, National Bureau of Economic Research.

Bordalo, P., K. B. Coffman, N. Gennaioli, and A. Shleifer (2020): "Older People are Less Pessimistic about the Health Risks of Covid-19," Discussion paper, National Bureau of Economic Research.

Borovička, J. (2020): "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences," *Journal of Political Economy*, 128(1), 206–251.

Bourne, R., C. Edwards, et al. (2019): "Exploring Wealth Inequality," .

Brokaw, T. (2000): *The greatest generation*. Random House.

Brunnermeier, M. K., A. Simsek, and W. Xiong (2014): "A welfare criterion for models with distorted beliefs," Discussion paper, National Bureau of Economic Research.

Campbell, J. Y., L. M. Viceira, L. M. Viceira, et al. (2002): *Strategic asset allocation: portfolio choice for long-term investors*. Clarendon Lectures in Economic.

Capelle, D. (2019): "The Great Gatsby Goes to College: Tuition, Inequality and Intergenerational Mobility in the US," .

Charles, K. K., and E. Hurst (2003): "The correlation of wealth across generations," *Journal of political Economy*, 111(6), 1155–1182.

Chen, H., S. Joslin, and N.-K. Tran (2010): "Affine disagreement and asset pricing," *American Economic Review*, 100(2), 522–26.

——— (2012): "Rare disasters and risk sharing with heterogeneous beliefs," *The Review of Financial Studies*, 25(7), 2189–2224.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang (2017): "The fading American dream: Trends in absolute income mobility since 1940," *Science*, 356(6336), 398–406.

Cogley, T., and T. J. Sargent (2008): "The market price of risk and the equity premium: A legacy of the Great Depression?," *Journal of Monetary Economics*, 55(3), 454–476.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016): "Asset pricing when this time is different," *The Review of Financial Studies*, 30(2), 505–535.

Cox, J. C., and C.-f. Huang (1989): "Optimal consumption and portfolio policies when asset prices follow a diffusion process," *Journal of economic theory*, 49(1), 33–83.

Deaton, A., and C. Paxson (1994): "Intertemporal choice and inequality," *Journal of political economy*, 102(3), 437–467.

Dieckmann, S. (2011): "Rare event risk and heterogeneous beliefs: The case of incomplete markets," *Journal of Financial and Quantitative Analysis*, 46(2), 459–488.

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2017): "Asset prices and portfolio choice with learning from experience," *The Review of Economic Studies*, 85(3), 1752–1780.

Fagereng, A., C. Gottlieb, and L. Guiso (2017): "Asset market participation and portfolio choice over the life-cycle," *The Journal of Finance*, 72(2), 705–750.

Favilukis, J. (2013): "Inequality, stock market participation, and the equity premium," *Journal of Financial Economics*, 107(3), 740–759.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016): "The dynamics of inequality," *Econometrica*, 84(6), 2071–2111.

Gale, W. G., H. Gelfond, J. J. Fichtner, and B. H. Harris (2020): "The Wealth of Generations, With Special Attention to the Millennials," Discussion paper, National Bureau of Economic Research.

Gârleanu, N., and S. Panageas (2015): "Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing," *Journal of Political Economy*, 123(3), 670–685.

Gayer, G., I. Gilboa, L. Samuelson, and D. Schmeidler (2014): "Pareto Efficiency with Different Beliefs," *The Journal of Legal Studies*, 43(S2), S151–S171.

Gennotte, G. (1986): "Optimal portfolio choice under incomplete information," *The Journal of Finance*, 41(3), 733–746.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2019): "Five facts about beliefs and portfolios," Discussion paper, National Bureau of Economic Research.

Gomes, F., and A. Michaelides (2005): "Optimal life-cycle asset allocation: Understanding the empirical evidence," *The Journal of Finance*, 60(2), 869–904.

Gomez, M., et al. (2016): "Asset prices and wealth inequality," *Unpublished paper: Princeton.*

Greenwood, R., and A. Shleifer (2014): "Expectations of returns and expected returns," *The Review of Financial Studies*, 27(3), 714–746.

Guvenen, F. (2007): "Learning your earning: Are labor income shocks really very persistent?," *American Economic Review*, 97(3), 687–712.

Hammond, P. J. (1981): "Ex-ante and ex-post welfare optimality under uncertainty," *Economica*, pp. 235–250.

HARRIS, R. (1978): "Ex-post efficiency and resource allocation under uncertainty," *The Review of Economic Studies*, pp. 427–436.

HARRISON, J. M., AND D. M. KREPS (1979): "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic theory*, 20(3), 381–408.

HEIMER, R. Z., K. O. R. MYRSETH, AND R. S. SCHOENLE (2019): "YOLO: Mortality beliefs and household finance puzzles," *The Journal of Finance*, 74(6), 2957–2996.

KASA, K., AND X. LEI (2018): "Risk, uncertainty, and the dynamics of inequality," *Journal of Monetary Economics*, 94, 60–78.

KOTLIKOFF, L. J., AND S. BURNS (2005): *The coming generational storm: What you need to know about America's economic future*. MIT press.

KOULOVATIANOS, C., AND V. WIELAND (2011): "Asset pricing under rational learning about rare disasters," .

KOZLOWSKI, J., L. VELDKAMP, AND V. VENKATESWARAN (2020): "The tail that wags the economy: Belief-driven business cycles and persistent stagnation," *Jounrnal of Political Economy*, (15-25).

KUCHLER, T., AND B. ZAFAR (2019): "Personal experiences and expectations about aggregate outcomes," *The Journal of Finance*, 74(5), 2491–2542.

KUHN, M., M. SCHULARICK, AND U. I. STEINS (2017): "Income and wealth inequality in america," *Journal of Political Economy*.

LEI, X. (2019): "Information and inequality," *Journal of Economic Theory*, 184, 104937.

LIPTSER, R. S., A. N. SHIRIAEV, AND A. N. SHIRYAEV (2001): *Statistics of Random Processes II: II. Applications*, vol. 2. Springer Science & Business Media.

LUCAS, R. E. (1978): "Asset prices in an exchange economy," *Econometrica: Journal of the Econometric Society*, pp. 1429–1445.

MALMENDIER, U., AND S. NAGEL (2011): "Depression babies: do macroeconomic experiences affect risk taking?," *The Quarterly Journal of Economics*, 126(1), 373–416.

——— (2015): "Learning from inflation experiences," *The Quarterly Journal of Economics*, 131(1), 53–87.

——— (2016): "Learning from inflation experiences," *The Quarterly Journal of Economics*, 131(1), 53–87.

MALMENDIER, U., D. POUZO, AND V. VANASCO (2019): "Investor experiences and financial market dynamics," *Journal of Financial Economics*.

MALMENDIER, U., AND L. S. SHEN (2018): "Scarred consumption," Discussion paper, National Bureau of Economic Research.

MANKIW, N. G., AND D. N. WEIL (1992): "The Baby Boom, the Baby Bust, and the Housing Market," *Regional Science and Urban Economics*, 21(4), 541–52.

MERTON, R. C. (1980): "On estimating the expected return on the market: An exploratory investigation," *Journal of financial economics*, 8(4), 323–361.

NAGEL, S., AND Z. XU (2019): "Asset pricing with fading memory," Discussion paper, National Bureau of Economic Research.

NAKAMURA, E., J. STEINSSON, R. BARRO, AND J. URSÚA (2013): "Crises and recoveries in an empirical model of consumption disasters," *American Economic Journal: Macroeconomics*, 5(3), 35–74.

NAKOV, A., AND G. NUÑO (2015): "Learning from experience in the stock market," *Journal of Economic Dynamics and Control*, 52, 224–239.

Orwell, G. (1945): "Review of A Coat of Many Colours: Occasional Essays by Herbert Read," *The Complete Works of George Orwell*, 17, 402–405.

Poterba, J. M. (2001): "Demographic structure and asset returns," *Review of Economics and Statistics*, 83(4), 565–584.

Rietz, T. A. (1988): "The equity risk premium a solution," *Journal of monetary Economics*, 22(1), 117–131.

Rognlie, M. (2016): "Deciphering the fall and rise in the net capital share: accumulation or scarcity?," *Brookings papers on economic activity*, 2015(1), 1–69.

Saez, E., and G. Zucman (2016): "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *The Quarterly Journal of Economics*, 131(2), 519–578.

Scheinkman, J. A., and W. Xiong (2003): "Overconfidence and speculative bubbles," *Journal of political Economy*, 111(6), 1183–1220.

Schraeder, S. (2015): "Information processing and non-bayesian learning in financial markets," *Review of Finance*, 20(2), 823–853.

Starr, R. M. (1973): "Optimal production and allocation under uncertainty," *The Quarterly Journal of Economics*, pp. 81–95.

Tsai, J., and J. A. Wachter (2015): "Disaster risk and its implications for asset pricing," *Annual Review of Financial Economics*, 7, 219–252.

Tversky, A., and D. Kahneman (1974): "Judgment under uncertainty: Heuristics and biases," *science*, 185(4157), 1124–1131.

Walden, J., and C. Heyerdahl-Larsen (2015): "Efficiency and Distortions in a Production Economy with Heterogeneous Beliefs," Discussion paper, Society for Economic Dynamics.

Wee, S. L. (2016): "Delayed learning and human capital accumulation: The cost of entering the job market during a recession," *Unpublished manuscript*, 18.

Wolff, E. N. (2002): "Inheritances and wealth inequality, 1989-1998," *American Economic Review*, 92(2), 260–264.

Woodford, M. (2013): "Macroeconomic analysis without the rational expectations hypothesis," *Annu. Rev. Econ.*, 5(1), 303–346.

Yaari, M. E. (1965): "Uncertain lifetime, life insurance, and the theory of the consumer," *The Review of Economic Studies*, 32(2), 137–150.