HOW DO STOCK MARKET EXPERIENCES SHAPE WEALTH INEQUALITY?

XIAOWEN LEI

Abstract. This paper develops a continuous-time OLG model, incorporating rare disasters and agents who learn more from their own experiences. Disasters such as the Great Depression make investors distrustful of the market. Generations that experience disasters save in the form of safer portfolios, even if similar disasters are not likely to occur again during their lifetimes. "Fearing to attempt" therefore inhibits wealth accumulation by these "depression babies" relative to other generations. This effect is amplified in the general equilibrium since the equity premium is relatively high following a disaster. When calibrated to US data, the model can explain around half of the old-to-young wealth ratio decrease between the Great Depression and the 1980s, and around 7% of the subsequent increase. The model can also explain about a quarter of the increase of top 1% wealth shares and is consistent with observations of life cycle portfolio choices, and changes in asset returns following disasters.

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)

Date: April 16, 2024.

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 Klaus Adam, Rhys Bidder, Corina Boar, Benjamin Born, Sarah Clifford, 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, Laura Veldkamp, Rick Van der Ploeg, Mirko Wiederholt, and Francesco Zanetti. I would also like to thank Hanna Thompson Schumann and John P. Dowdall for their excellent research assistance work. Xiaowen Lei is from the Department of Economics and Finance, University of Guelph. The author acknowledges that part of the work of this paper was completed while she was at Nuffield College, University of Oxford, UK. Email: leixiaowen04@gmail.com.

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 heard the popular meme "ok boomer", and are well aware of the resentment that inspired it. The source of this resentment is clear. For the first time in history, most of the younger generations are in danger of being poorer than their parents (Chetty, Grusky, Hell, Hendren, Manduca, and Narang (2017)).

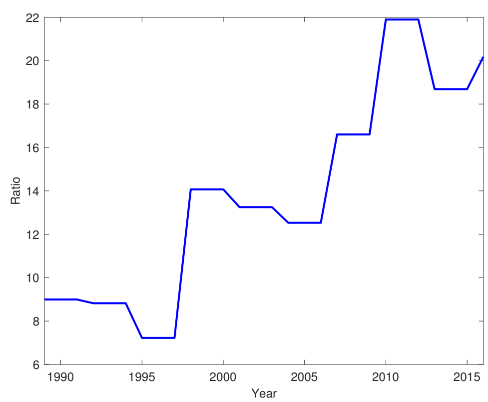

Figure 1. Median Net Worth Ratio of 65 and over vs. 35 and under (SCF)

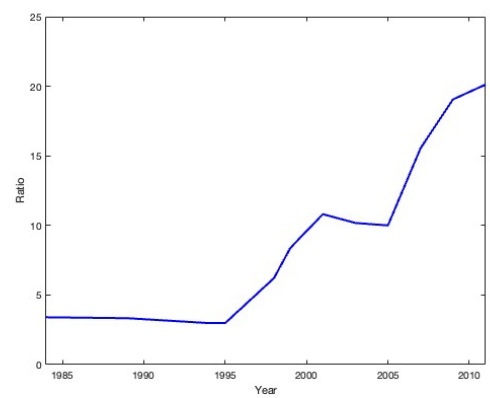

Figure 2. Median Net Worth Ratio of 65 and over vs. 35 and under (PSID)

Figures 1 and 2 plot the ratio of median net worth (per household) for those over age 65 and those under age 35, from the Survey of Consumer Finances and the Panel Study of Income Dynamics, respectively. The SCF defines net worth by total financial and non-financial assets, less the value of debt. Unsurprisingly, the old have always been wealthier than the young, in 1989 their net worth was, on average, 9 times greater. However, over the course of the next 27 years, this ratio more than doubled, to over 20.[1] One might

[1]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 the 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 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 mildly when compared to other countries. According to OECD data, it rose from 75.1 in 1989 to 78.6 in 2016. One should also note that this measure of old to young-wealth ratio is at the per-household level, which controls the population size. In reality, generational cohort size has also changed significantly over this period. This paper is silent about this.

argue that baby boomers had the potential to amass substantial wealth through the appreciation of real estate. However, a similar increase is verified using the PSID measure, where net worth is defined as the sum of six types of assets and net debt, excluding housing.

Most inequality literature focuses on the recent increase in *overall* inequality. This increase reflects a combination of within- and between-cohort inequality. Evidence suggests that more than half of the increase in overall inequality is driven by between-cohort inequality. For example, using PSID data, it is clear that the between-cohort wealth Gini was 57.2% of the overall wealth Gini in 1984, and 61.9% in 2017. One might argue that within-cohort inequality is more significant than between-cohort inequality, since between-cohort redistributions can be offset by inter-generational transfers. Evidence suggests, however, that intergenerational redistributions are not completely offset by transfers (e.g., Altonji, Hayashi, and Kotlikoff (1997)). Moreover, while parental wealth undoubtedly plays a valuable insurance role for young adults (Kaplan (2012)), prolonged financial dependence on parents can also produce adverse psychological and sociological consequences (Mortimer, Kim, Staff, and Vuolo (2016), Caputo (2020), Hill, van der Geest, and Blokland (2017)).

Standard inequality models cannot explain Figures 1 and 2 because they generate *stationary* age/wealth distributions. Of course, one could always inject an exogenous shock, and then attribute the trend in Figures 1 and 2 to transition dynamics. However, this is a rather unappealing strategy, since the trends in the two figures are a mirror image of a declining trend that took place 40 years following the Great Depression. Although direct evidence on previous generational inequality is lacking, we do know that generational inequality is highly correlated with top wealth shares, simply because the wealthy have always been relatively old. Figure 3 uses the correlation between top 1% wealth shares and generational inequality after 1984 to impute the old-to-young wealth ratio post WWII, using PSID data. According to Saez and Zucman (2016) data, the top 1% wealth share in 1945 was 32.7%. It then steadily decreased to 23.7% by 1983, which implies an old-to-young wealth ratio that declines from 16 to 3.4 during this period.[2] This suggests that baby boomers are better off than both their parents and their kids. It also suggests that you would need to introduce *two* exogenous shocks to explain the observed trends in generational inequality.

[2] I take the 5-year average of top 1% wealth share to impute the old-to-young wealth ratio so as to match the PSID survey frequency after 1984.

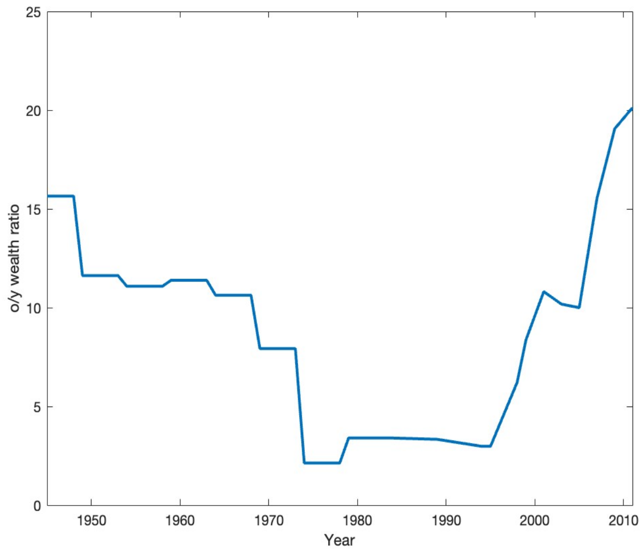

Figure 3. Imputed old to young wealth ratio (1945-2011)

What then explains this reversal? Undoubtedly, many factors are responsible, but this paper focuses on just one of them, namely, generational belief differences. It studies an economy that combines two key ingredients. First, individuals weight their personal experiences heavily when forming their beliefs, as in Malmendier and Nagel (2011). Second, an economy that is subject to rare disasters, as in Rietz (1988) and Barro (2006). When the model is calibrated to US data, it accounts for a significant portion of the recent increase in the relative wealth of older generations and also explains why this ratio decreased following the Great Depression. The model also illustrates how equilibrium feedbacks operating in financial markets amplify these changes.

Although introducing rare disasters may seem similar to introducing exogenous shocks, there is a crucial difference. Although rare, disaster shocks in this study's model are *recurrent*, and the anticipation of this recurrence influences behaviour, both before and after the shock. This anticipation explains why rare disasters literature has been successful at resolving the Equity Premium Puzzle in the first place. However, literature on asset-pricing rare disasters rely on a representative agent. My primary contribution is to show that when rare disasters are combined with overlapping generations and experiential learning, a powerful force for heterogeneity and inequality is ignited.

Specifically, this paper argues that different generations have different beliefs about market returns due to their limited experiences. This influences their risk-taking behaviour which, in turn, influences the growth rate of their wealth. For instance, a 65-year-old in 1989 would have been born in 1924, and at an early age, experienced the Great Depression. By contrast, a 65-year-old in 2016 would have been a lucky baby boomer, who skipped the Great Depression and faced more positive experiences in the stock market. Due to the rare nature of disasters, it was not likely that Depression Babies would experience another Great Depression. But its salience with their own experience cast a long shadow throughout the remainder of their lives. In other words, they were "scarred". Therefore, it is natural that investors 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 choice. Malmendier and Nagel (2011) provides strong empirical support that macroeconomic experiences in the stock market have a *lasting* impact on how much households invest in risky assets later in their lives, even long after experiencing a rare disaster. They find that "depression babies" were much less likely to participate in the stock market, and, if they did, they tended to invest less into risky assets compared to other generations. Using SCF data, they find that an increase from the 10th to the 90th percentile in experienced return 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 fraction of wealth allocated to stocks.[3] There has also been independent empirical evidence which shows that older people nowadays are more optimistic relative to younger 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) use survey data on the recent Covid-19 crisis, and show that the current older generation worries less about the health risks induced by the pandemic, despite evidence suggesting 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 wealth consistent with the data (See Gabaix, Lasry, Lions, and Moll (2016)). From the perspective of experiential learning, this is not so surprising. As households age, they witness more

[3]The potentially important distinction between liftetime experiences and financial market experiences is not present in my model, since I assume everyone participates in the financial markets.

data, and become more confident of their own estimates, which encourages them to invest a higher fraction of their wealth in risky assets. This is true for a normal economic climate, 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 millennials relative to older generations. From the experiential-learning angle, millennials have had less experience with a normal economic climate. As such, they "over-react" to the crisis and become relatively pessimistic about future stock market returns compared to their experienced elders.

Second, while most inequality 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 decreased because the old were more invested in risky assets. However, as 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 less risks in the financial market, while 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 between 1965 and 1980) and the millennials (born between 1981 and 1994) have just entered the financial market and started to accumulate experiences in the financial market, while the baby boomers have already grown a lot more confident about the market due to their memory of the earlier tranquil episodes when they were younger. Therefore, the old-to-young wealth ratio has increased, and a U-shaped inequality pattern naturally emerges over the last century.

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 implications for asset prices look similar to a representative agent economy with constant gain learning. This provides 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). [4]

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 following 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. Right after the Great Depression, increased pessimism produced a rise in equity premium. However, over time, as the "depression babies" died out, the market became dominated by the baby boomers. Since the boomers did not experience the Great Depression, they invested aggressively in risky assets and bid up asset prices, which then led to a declining equity premium. These trends in the (ex ante) equity premium are consistent with the empirical evidence provided by Blanchard, Shiller, and Siegel (1993) and Jagannathan, McGrattan, and Scherbina (2001). While both partial and general equilibrium effects might appear intuitive and simple, it is not straightforward 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 difficult to characterize in discrete time. My model attempts to disentangle the partial and general equilibrium effects of experiential learning by solving a continuous time overlapping generation model with heterogeneous agents, and providing closed form solutions for policy functions, prices, and wealth dynamics.

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

[4]This belief channel does not rule out the possibility that households' risk attitude could change in response to disasters. In fact, Cohn, Engelmann, Fehr, and Maréchal (2015) provides experimental evidence of counter-cyclical risk aversion.Dillenberger and Rozen (2015) develop a model of history dependent risk attitudes. However, given the direct evidence from survey expectations on experienced and expected returns, we know that the belief channel also exists.

the model to US data, and shows that the model can explain the observed U-shaped pattern in postwar generational inequality. Section 6 provides further 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 also embeds rare disaster risk in the tradition of Rietz (1988) and Barro (2006). The goal is to solve for portfolio allocations, asset prices, and the distribution of wealth when the arrival rate of disasters is unknown, and agents must learn about it from their own experiences.

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 birth, agents receive a constant fraction τ^* of the aggregate wealth, so that the share of wealth of the newborn is constant over time. Agents have no bequest motive. There is a representative firm that pays out dividend Y_t at each instant. In order to focus on between-cohort inequality, I assume agents only differ in the timing of birth, but are otherwise identical within the same cohort. 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 [5]

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

where Y_{t^-} denotes the endowment 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

[5] In principle, we can allow agents to learn about the disaster size κ_l, κ_h instead of the disaster intensity. However, the qualitative results will be the same. Upon realization of a large-size disaster, agents become more scarred and consider large-size disaster more likely to happen. I choose the assumption where agents learn about disaster intensity instead of disaster size here due to its analytical convenience.

history as $\{\mathcal{F}_t\} = \{\mathcal{F}^B \times \mathcal{F}^N\}$. The jump process N_t follows

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

That is, at each instant, there is λ_t probability that a disaster happens. When it happens, the jump size κ_t can take on two values. With p^* probability, the realization of a disaster size is κ_h (a severe disaster), and with $(1-p^*)$ probability, the disaster size is κ_l (a mild disaster). I assume that $\kappa_t \in (-1, 0)$, which captures the fact that there is a decline in endowment value 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 also take on 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, [6]

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

Since the focus here is on financial income instead of labor income, I follow the assumption in Benhabib, Bisin, and Zhu (2016) and assume that capital gains are subject to a linear tax rate τ. The government collects them to finance the initial endowment of the newborn. Therefore, the following accounting equation must hold at all times,

$$\tau W_t = \delta w_{t,t} \tag{2.4}$$

where W_t denotes the aggregate wealth at time t. I further assume that the market is dynamically complete and that investors can trade continuously in the capital market to hedge against both regular and disastrous economic risks. To complete the market, agents need three securities (in addition to their life insurance policies): a bond, an equity, and a disaster-contingent asset. The bond value follows

$$dB_t = r_t B_t dt \tag{2.5}$$

The risky asset value follows

$$\frac{dS_t + D_t dt}{S_{t^-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{2.6}$$

where r_t, μ_t^S, σ^S as well as κ_t^S are endogenous objects, and are determined in equilibrium. I also assume that aggregate income flow is entirely through dividend payment so that $D_t = Y_t$. Finally, the disaster-contingent security value P_t, follows the stochastic process

[6]The assumption of a two point markov process of λ_t is purely technical, and does not affect the learning process described below. However, it will soon be clear that such setup allows for a clean closed form characterization of the learning process.

$$\frac{dP_t}{P_{t^-}} = \mu_t^P dt + \kappa_t^P dN_t(\lambda_t) \tag{2.7}$$

This asset is in zero net supply. By convention, I assume the disaster-contingent security pays off during normal times, but suffers a loss during disasters. That is, by holding the disaster-contingent security, the investor 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. The real world counterpart of this security would be a catastrophe bond or a hybrid security whose value depend on the adverse state of the economy [7].

Investors observe the aggregate endowment process and know the values of μ, σ, λ_h, λ_l and κ_t. However, they do not observe π^*, and must learn about it from their own limited lifetime experience. The specific choice of which parameters to learn about is supported 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 learning works, we need to consider optimal filtering of a jump-diffusion process.

2.2. **Filtering and Information Processing.** When an investor 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 $\mathbb{E}_{s,t}[\lambda_t] = \pi_{s,t}\lambda_h + (1-\pi_{s,t})\lambda_l$ [8] 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 us 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*

$$d\pi_{s,t}|_{dN_t=0} = -(\lambda_h - \lambda_l)\pi_{s,t}(1-\pi_{s,t})dt \tag{2.8}$$

$$d\pi_{s,t}|_{dN_t=1} = \frac{\lambda_h \pi_{s,t}}{\bar{\lambda}_{s,t}} - \pi_{s,t} \tag{2.9}$$

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

[8] Learning about the weight parameter $\pi_{s,t}$ is the same as learning about the disaster likelihood $\lambda_{s,t}$ due to the linearity. They will be used interchangeably below.

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). □

Notice that 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 disaster 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 high disaster intensity is suddenly amplified by a magnitude of $\frac{\lambda_h}{\bar{\lambda}_{s,t}}$. [9]

2.3. **Optimization.** Agents continuously choose a non-negative consumption process $c_{s,t}$, the fraction of wealth allocated to the risky asset market $\alpha^S_{s,t}$, and the fraction of wealth devoted to the disaster-contingent security $\alpha^P_{s,t}$. They continuously update their beliefs about disaster risk, and dynamically trade assets given the return process and their beliefs, in order to maximize a logarithmic flow utility over consumption goods. [10] 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

$$\max_{c_{s,t}, \alpha^S_{s,t}, \alpha^P_{s,t}} \mathbb{E}_{s,t} \left[\int_s^\infty e^{-(\rho+\delta)(t-s)} \log(c_{s,t}) dt \right] \tag{2.10}$$

s.t:

$$\begin{aligned} \frac{dw_{s,t}}{w_{s,t^-}} = &\left(r_t - \tau + \delta + \alpha^S_{s,t}(\mu^S_t - r_t) + \alpha^P_{s,t}(\mu^P_t - r_t) - \frac{c_{s,t}}{w_{s,t^-}} \right) dt + \alpha^S_{s,t} \sigma^S dZ_{s,t} \\ &+ (\alpha^S_{s,t}\kappa^S_t + \alpha^P_{s,t}\kappa^P_t) dN_{s,t}(\bar{\lambda}_{s,t}) \end{aligned} \tag{2.11}$$

where $\mathbb{E}_{s,t}$ denotes the expectation of generation s evaluated at time t. 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. This allows me to employ the martingale approach (Cox and Huang (1989)). This allows me to convert the dynamic programming problem into a static problem as

[9] One might argue that Bayesian learning is contradicted by evidence of a 'recency bias'. That is, it is debatable whether agents weigh 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.

[10] 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 investor'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.

follows

$$\max_{c_{s,s}} \mathbb{E}_{s,s}\left[\int_s^\infty e^{-(\rho+\delta)(t-s)}\log\left(c_{s,t}\right)dt\right] \tag{2.12}$$

s.t:

$$\mathbb{E}_{s,s}\left[\int_s^\infty e^{-\delta(t-s)}\xi_{s,t}c_{s,t}dt\right] = w_{s,s} \tag{2.13}$$

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

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

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

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

$$c_{s,t} = c_{s,s}e^{-\rho(t-s)}\frac{\xi_{s,s}}{\xi_{s,t}} \tag{2.15}$$

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*

$$\frac{d\xi_{s,t}}{\xi_{s,t^-}} = (\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t)dt - \theta_{s,t}dZ_{s,t} + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right)dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.16}$$

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

$$\frac{d\xi_t}{\xi_{t^-}} = (\bar{\lambda}_t - \lambda^N_t - r_t)dt - \theta_t dZ_t + \left(\frac{\lambda^N_t}{\bar{\lambda}_t} - 1\right)dN_t(\bar{\lambda}_t) \tag{2.17}$$

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

$$\frac{d\eta_{s,t}}{\hat{\eta}_{s,t^-}} = \left(\frac{1}{1+\bar{\kappa}}\left(\lambda_{s,t} - \mathbb{E}(\lambda_{s,t})\right)\right)dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2}{1+\bar{\kappa}} - 1) - 1\right) - 1\right]dN(\bar{\lambda}_t) \tag{2.18}$$

where $\bar{\kappa} = p^*\kappa_h + (1-p^*)\kappa_l$.

Proof. See Appendix A.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 as likely to happen, as well as on the market price of disaster risk. Since we know that $c_{s,t} = e^{-\rho(t-s)}(y_s\xi_{s,t})^{-1}$, knowing the process of the state price density is equivalent to knowing the process of consumption. Ito's lemma then delivers

$$\frac{dc_{s,t}}{c_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_{s,t}^N + r_t - \rho)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.19}$$

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 optimally invested wealth follows

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_{s,t}^N + r_t - \rho)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.20}$$

Given the above individual optimal 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 ($c_{s,t}$, $\alpha_{s,t}^S$, $\alpha_{s,t}^P$) and a price system (r_t, μ_t^S, μ_t^P, κ_t^S, κ_t^P) such that the choice processes ($c_{s,t}$, $\alpha_{s,t}^S$, $\alpha_{s,t}^P$) maximize agents' utility subject to their budget constraints, and the market for consumption goods, bonds, risky asset and the disaster-contingent security all clear, i.e.,*

$$Y_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} c_{s,t} ds \tag{2.21}$$

$$S_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha_{s,t}^S w_{s,t} ds \tag{2.22}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha_{s,t}^P w_{s,t} ds \tag{2.23}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} (1 - \alpha_{s,t}^S - \alpha_{s,t}^P) w_{s,t} ds \tag{2.24}$$

By using the market-clearing condition for consumption goods, we can derive the stochastic processes for ξ_t. Let us conjecture that the fraction of aggregate endowment consumed by a newborn agent at time t is a fixed fraction $\beta_t = \frac{c_{t,t}}{Y_t} = \beta$. [11] We can then rewrite the goods market clearing condition as

$$\xi_t Y_t = \int_{-\infty}^{t} \beta\delta e^{-(\rho+\delta)(t-s)} \xi_s Y_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{2.25}$$

[11] Appendix B.1 verifies this conjecture, and derives an explicit expression for β.

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

$$\eta_t = \int_{-\infty}^{t} \beta\delta e^{-\beta\delta(t-s)}\eta_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{2.26}$$

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

$$\frac{d\eta_t}{\eta_t} = \bar{\mu}^{\eta}_t dt + \bar{\kappa}^{\eta}_t dN_t(\bar{\lambda}_t) = \bar{\kappa}^{\eta}_t dN_t(\bar{\lambda}_t) \tag{2.27}$$

since

$$\bar{\mu}^{\eta}_t = \mathbb{E}_{s,t}(\mu^{\eta}_{s,t}) = \int_{-\infty}^{t} f_{s,t}\mu^{\eta}_{s,t} ds = 0; \quad \bar{\kappa}^{\eta}_t = \mathbb{E}_{s,t}(\kappa^{\eta}_{s,t}) = \int_{-\infty}^{t} f_{s,t}\kappa^{\eta}_{s,t} ds \tag{2.28}$$

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

$$f_{s,t} = \beta\delta e^{-\beta\delta(t-s)} \left(\frac{\eta_s}{\eta_t}\right)\left(\frac{\eta_{s,t}}{\eta_{s,s}}\right) = \delta e^{-\delta(t-s)} \frac{c_{s,t}}{Y_t} \tag{2.29}$$

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

$$\frac{d\xi_t}{\xi_t} = \left(\sigma^2 - \mu - \rho - \delta(1-\beta)\right) dt - \sigma dZ_t + \left(\frac{1+\bar{\kappa}_\eta}{1+\bar{\kappa}} - 1\right) dN_t(\lambda_t) \tag{2.30}$$

Since we know that the state price density also has to follow eqn.(2.17), it directly gives the solution of equilibrium prices.

Proposition 1. *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*

$$r_t = \underbrace{\rho + \delta(1-\beta)}_{\textit{effective patience with OLG}} + \underbrace{\mu - \sigma^2}_{\textit{risk adjusted growth}} + \underbrace{\frac{\bar{\kappa}}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t})}_{\textit{market view of disaster risk}} ; \tag{2.31}$$

$$\theta_t = \theta = \sigma; \tag{2.32}$$

$$\lambda^N_t = \frac{\mathbb{E}_{s,t}(\lambda_{s,t})}{1+\bar{\kappa}} \tag{2.33}$$

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, effective patience lessens due to death risk. Moreover, the equilibrium interest rate increases when the endowment process has a higher rate of growth and lower volatility, which is captured in the second term. The third term reflects a flight to safety 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 $\bar{\kappa} < 0$, this implies that the equilibrium interest rate decreases with market average pessimism. The desire to save in the form of safe assets during disasters drives down the return on the safe asset, 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 terms are both constants, so variations in the interest rate are totally driven by variations in market pessimism 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, and agents have common beliefs about the regular Brownian risk, the market price of risk is therefore the same as the standard solution with log preferences, which simply equates to the volatility of the risk. Finally, the market price of disaster risk increases with the market view of the disaster likelihood. 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

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_t^N + r_t - \rho)dt + \theta_{s,t} dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{2.34}$$

Recall also that the budget constraint follows

$$\begin{aligned}\frac{dw_{s,t}}{w_{s,t^-}} = &\left(r_t - \tau + \alpha_{s,t}^S(\mu_t^S - r_t) + \delta + \alpha_{s,t}^P(\mu_t^P - r_t) - \frac{c_{s,t}}{w_{s,t^-}}\right) dt + \alpha_{s,t}^S \sigma^S dZ_{s,t} \\ &+ (\alpha_{s,t}^S \kappa_t^S + \alpha_{s,t}^P \kappa_t^P) dN_{s,t}(\bar{\lambda}_{s,t})\end{aligned} \tag{2.35}$$

Since the market is complete, we can match coefficients with the wealth process in these two stochastic differential equations. The share of wealth invested in the risky risky asset market and the disaster-contingent security at time t for an agent born at time s are given by the following expressions respectively

$$\alpha_{s,t}^S = \frac{\theta_{s,t}}{\sigma^S} = \frac{\theta_t}{\sigma^S} \tag{2.36}$$

$$\alpha_{s,t}^P = \frac{1}{\kappa_t^P}\left(\frac{\bar{\lambda}_{s,t}}{\lambda_t^N} - 1\right) - \frac{\kappa_t^S \theta_t}{\kappa_t^P \sigma^S} \tag{2.37}$$

Notice that all generations invest the same fraction of wealth in risky assets. However, pessimistic generations hold less disaster-contingent security, 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 risky asset price and the disaster-contingent security are given by*

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

$$\kappa_t^S = \kappa_t \tag{2.39}$$

$$\mu_t^S - r_t = \sigma^2 \tag{2.40}$$

$$\mu_t^P - r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t}) \tag{2.41}$$

Proof. See Appendix A.4. □

There are two interesting observations here. First, notice that the equity premium stays constant, and is unresponsive to shocks. This is because, with log utility and complete markets, portfolio holdings for the stocks stay the same. Second, the model produces an endogenous time-varying disaster asset premium. When market pessimism rises, risky assets and disaster-contingent security must pay higher average returns to clear the market. This has interesting implications for inequality. Following a disaster shock, scarred investors find safe asset investment more attractive. The increased aggregate demand for safe assets then generates a decline in the equilibrium interest rate, which then increases disaster asset premiums. This general equilibrium effect of prices amplifies the initial partial equilibrium effect. Not only does the scarred generation accumulate wealth at a slower pace due to less risk-taking, but they also sacrifice higher asset returns when it is the best time to buy the risky assets.

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

$$\alpha_{s,t}^S = 1 \tag{2.42}$$

$$\alpha_{s,t}^P = \frac{1}{\bar{\kappa}}\left(\frac{\bar{\lambda}_{s,t}}{\mathbb{E}(\bar{\lambda}_{s,t})}(1+\bar{\kappa}) - 1\right) - 1 \tag{2.43}$$

If $\lambda_{s,t} > \mathbb{E}(\lambda_{s,t})$, generation s is more pessimistic relative to the average generation and invests a lower share of their wealth in risky portfolios, vice versa.

The resulting portfolio choice solutions are rather intuitive. Due to the log utility of homogeneous beliefs on the Brownian motion risk, all investors invest all shares in risky assets. However, pessimistic generations invest a lower share of their wealth in disaster contingency assets.

2.7. **Comments on the complete market.** It is worth noting that even though the setup here assumes a complete market for analytical convenience, the belief scarring channel proposed above does not require a complete market. Without a disaster-contingent asset, investors insure themselves indirectly through the stock and bond markets (Dieckmann (2011)). The belief-scarred investor will just invest less in stocks and save more in bonds. The stocks would then price in both the regular Brownian motion risks and the disaster risk. Upon a disaster shock, the risk premium of the stock rises, similar to the rise of the risk premium of the disaster-contingent asset in the complete market setup.

2.8. **Comments on log preferences.** It is also worth discussing the logarithmic preferences assumption. Log utility is advantageous for two reasons. First, it results in a constant consumption-to-savings ratio, which simplifies our focus on the decision of portfolio selection. This assumption, however, can be relaxed by using recursive preferences with an IES of one. Second, the optimal portfolio selection is independent of estimation uncertainty for the log agent. This then allows us to concentrate on the first moment of belief differences, such as optimism and pessimism, without concern for higher-order belief differences. It could be argued that log utility implies a risk aversion coefficient of one, which is lower than typical estimates in the asset pricing literature. To address this issue, I will calibrate the prior to match the equity premium in the numerical section instead.

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 continuously. However, one can still study the long-run stationary distribution by averaging out those shocks across time and comparing its properties relative to the rational expectation economy.

Proposition 3. *The dynamics of the joint distribution of wealth and belief* $n(w, \lambda)$ *follows*

$$dn = -\frac{\partial}{\partial w}(n\hat{\mu}wdt + n\hat{\sigma}wdZ) + \frac{1}{2}\frac{\partial^2}{\partial w^2}(n\hat{\sigma}^2 w^2)dt + [n(w(1+\hat{\kappa}), t)) - n(w,t)]dN \quad (3.44)$$

Let $p(w) = \mathbb{E}_{s,t} n(w, \lambda)$ *denote the long run stationary distribution of wealth, and define* $\tilde{w}_{s,t} = \frac{w_{s,t}}{W_t}$*. To a first-order perturbation approximation, the long-run stationary distribution of* $x = \log(\tilde{w})$ *(eliminating all subscripts) is given by*

$$p(x) \approx \underbrace{Ge^{\zeta_0 x}}_{dogmatic} \underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}]}_{Learning} \quad (3.45)$$

where g_1, ζ_0 and ζ^1 are constants. Moreover,

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^0(x) \tag{3.46}$$

where $p^0(x) = Ge^{\zeta_0 x}$.

Proof. See Appendix B. □

That is, we can decompose the long-run stationary distribution into two pieces. The first piece features the standard stationary distribution of the log of wealth as an economy with dogmatic beliefs without learning. For a fair comparison, I assume that such dogmatic belief equals to the long-run average market beliefs. The second piece reflects experiential learning, which produces a fatter tail compared with the dogmatic belief economy. In the dogmatic belief economy, 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. With experiential learning, the stationary wealth distribution features a fatter right tail. This is intuitive: it is due to the "scale dependence" of wealth accumulation (See Gabaix, Lasry, Lions, and Moll (2016)). In this economy, the older are on average richer, who are also accumulating their wealth faster compared with the poorer and younger households. 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 over their lifetime, 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 generation's beliefs that become scarred the most, because they have less lifetime 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 prices derived in the previous section, and simulate sample paths, using the benchmark parameters in Table 1. The specific choice of parameters will be discussed in detail in the quantitative section. For now, let us focus on what happens to cohort behaviours after a disaster shock. To start, I shut down general equilibrium effects by fixing prices at their long-run average values. I assume that all agents start trading at age 20. One decade after the agent enters the market, I introduce a one-time disaster shock. Figure 4 plots the responses to the shock.

With log utility and complete markets, the agent invests all their wealth into risky assets and then borrows to purchase the disaster-contingent security. Inspecting the disaster-contingent security premium, its drift exceeds the risk-free rate. Therefore, shorting to

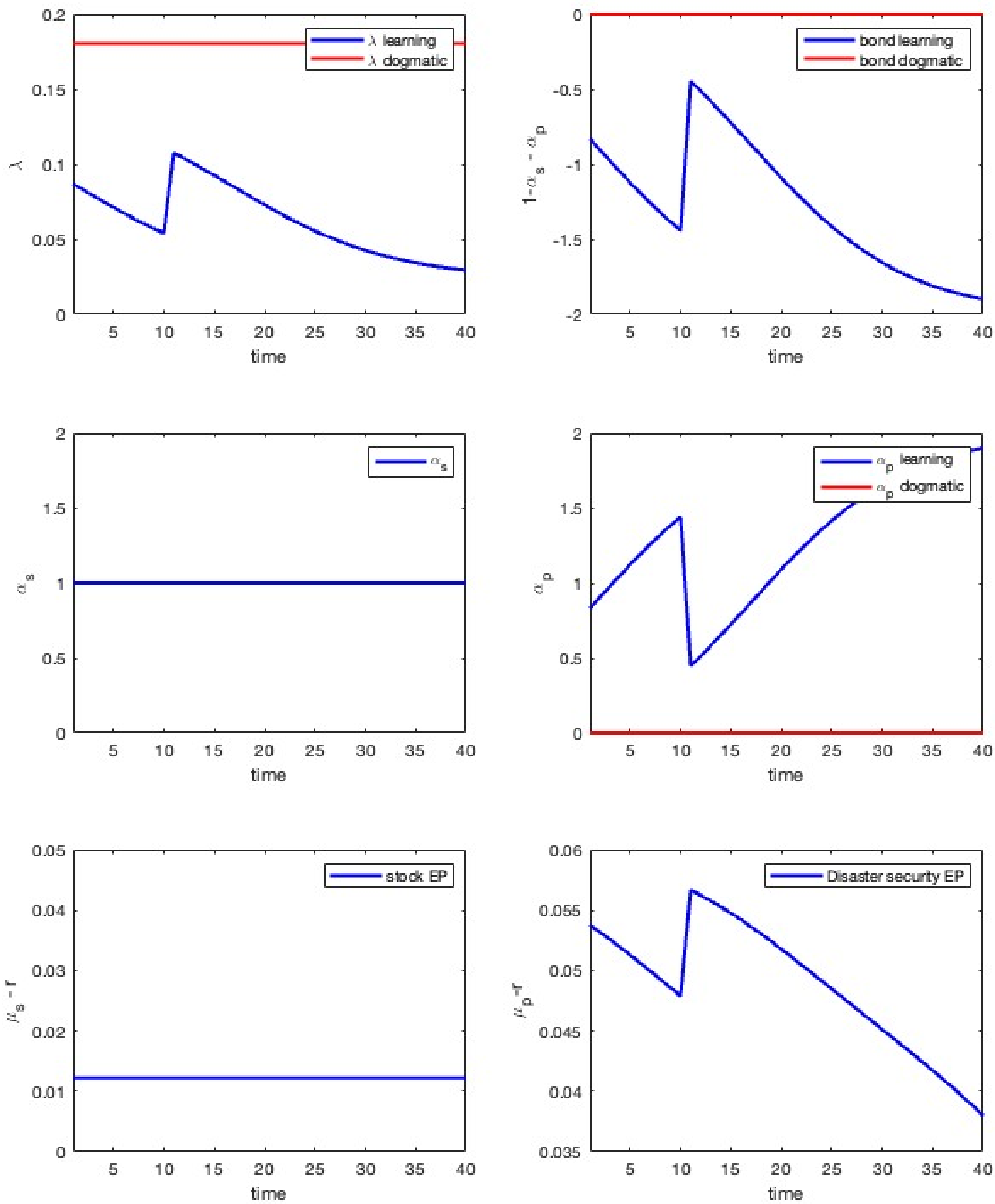

FIGURE 4. Simulated Time Paths of Policy Functions and Prices

purchase the disaster-contingent security yields positive net returns during normal times. The agent's wealth grows steadily over time. Suddenly, at $t = 10$, a disaster strikes, which brings down the endowment value. This does not affect their risky asset shares, because the risky asset only prices in regular Brownian risk, which is not affected by the disaster. However, learning from experience, the agent's pessimism rises, which triggers them to reduce their exposure to disaster-contingent security. Notice also that it takes almost 15 years for them to get back to the same level of optimism as before the disaster. For comparison, a useful benchmark economy is the case of dogmatic beliefs, in which the perceived likelihood of disasters is the same for all agents. In a complete market, this implies that nobody would be trading the disaster-contingent security, since they all have the same beliefs. The last two subplots show the response of prices after the disaster. The interest rate plummeted suddenly after the disaster due to a flight to safety. Since beliefs about Brownian motion risks are always universal, this results in no change in the stock equity premium. However, the reduction of the equilibrium interest rate drives disaster-contingent security premiums up. Quantitatively, the excess mean return spiked from 4.8% to 5.7% and again takes more than 15 years to return to its pre-crisis level.

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.

Table 1. Benchmark Parameter Values

Parameters	Value	Source
ρ	1%	Lower bound of empirical Estimate $1\% - 2\%$
δ	2.5%	average trading life expectancy of 40 years
μ	2%	Shiller's S&P 500 dividend growth
σ	11.07%	Shiler's S&P 500 dividend volatility
κ_h	-0.29	Match real GDP reduction between 1929-1933
κ_l	-0.035	Match real GDP reduction between 2008-2009
p^*	0.3261	Match 3-months US treasury bill interest rate (1989-2020)
π^*	0.89%	Match annual disaster intensity from (Barro (2006))
λ^H	24%	Upper bound of disaster intensity in (Barro (2006))
λ^L	1.5%	Lower bound of disaster intensity in (Barro (2006))
$\pi_{s,s}$	0.5	Match post-depression equity premium of 6.3% (Cogley and Sargent (2008))
τ	0.004	Match US transfer to GDP ratio (Benhabib and Zhu (2008))

The birth and death rate $\delta = 2.5\%$ is calibrated such that the average trading life is from 20 to 60 years old, implying an average trading life expectancy of 40 years. 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, following 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 of disaster frequencies from Barro (2006) of an international sample of 35 countries over 100 years. I use the real GDP growth rate reduction during the 1929-1933 Great Depression and the 2008-2009 financial crisis to calibrate $\kappa_h = -29\%$ and $\kappa_l = -3.5\%$ respectively. Next, an empirical estimate of the discount rate is around 1% to 2%. However, a 2% discount rate generates a model-implied interest rate that is too high compared with the data. Therefore, I set $\rho = 1\%$. Moreover, I calibrate the weight parameter p^* to match the annual interest rate, measured by the average 3-month US treasury bill constant maturity rate in the US between 1989 January to 2020 March, which is around 2.4% annually. Further, I assume that all agents start with a fixed prior $\pi_{s,s} = 0.5$ which roughly matches the post-1929 long-run average annual equity premium according to Cogley and Sargent (2008). Finally, the tax rate τ is taken from Benhabib and Zhu (2008), and is set to match the annual fraction of transfer to GDP ratio in the US.

Using the above parameters, I first compute the long-run average distribution of wealth and beliefs by simulation. The continuous time economy is discretized into discrete time with annual frequencies. I simulate the economy with 3000 agents for 2000 years. Each year, the newborns are endowed with a fixed fraction of aggregate wealth. The wealth share weighted average of prices is computed and fed back into the growth of wealth for each living agent. Then, δ fraction of the random sample of agents are dropped out at the end of each year, which is then replaced by the newborns again, 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 share weighted average, and the process carries on for 2000 years. At the end of the simulation, 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. Between 1929-1933, the Great Depression reduces the output by a percentage of κ_h, and between 2008-2009, the financial crisis reduces the output by a percentage of κ_l. I then re-run the simulation for 100 years to examine the response of the wealth distribution between 1920 to 2020.

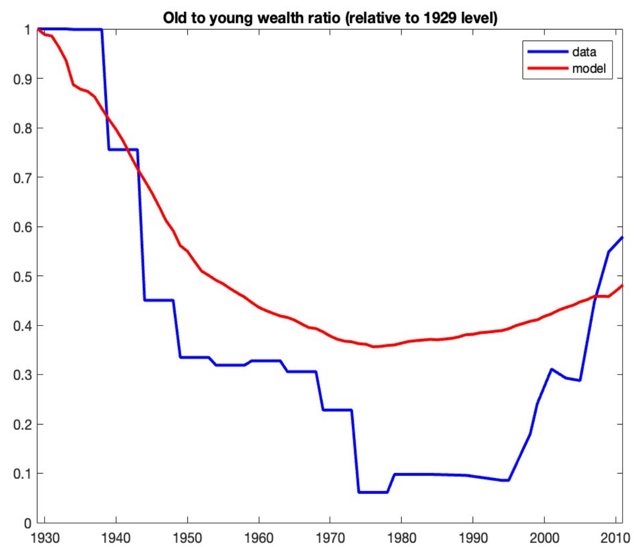

Figure 5. Calibrated Path of Old to Young Wealth Ratio

Figure 5 is the main result of this paper. It plots the calibrated path of changes in the old-to-young wealth ratio (65 and over vs. 35 and under) by normalizing its 1929 level to one. There are several interesting patterns that emerge. As one can see, right after the 1933 Depression, the old-to-young wealth ratio first went down sharply. This reflects a pure price effect, where the old generations, who were also more invested in the stock market, lost a fortune during the Great Depression. More interestingly, this initial sudden reduction is then followed by a more gradual tranquil decrease of old to young wealth ratio all the way until around 1970-1980s. This reflects the lingering "belief scarring" effect. As time goes by, the young people that experienced the Great Depression (the "Depression babies") become older. Over the life cycle, their conservative portfolio strategies cause them to lose wealth relative to the newer generations that have not experienced the Great Depression. 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. After the 1970-1980s, the optimistic boomers gradually start to take off and invest more heavily than the GenX and the Millennials. This gives rise to a mild increase in generational inequality again.

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. After all, pessimism begets pessimism. For example, even though boomers were relatively lucky during their own lifetime, they could 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. 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 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, such belief inheritance is hard to measure with data. The closest attempt has been Charles and Hurst (2003), which 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 average beliefs at the time they are born, and see how that changes the result.

Figure 6 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 decrease and then 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.

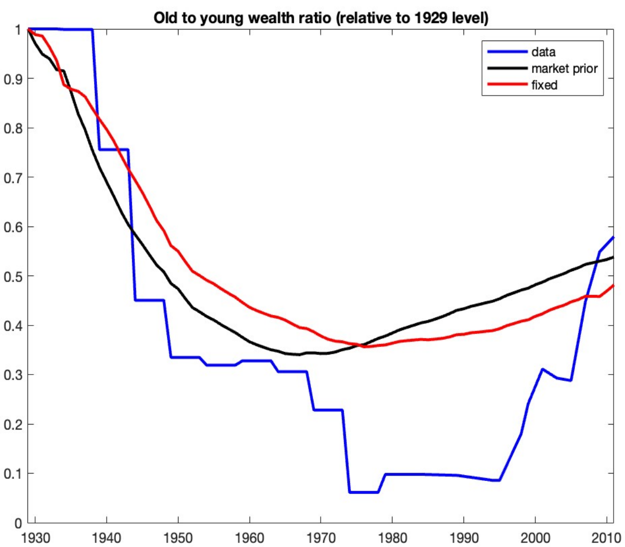

FIGURE 6. Fixed Prior vs. Market Based Prior

TABLE 2. Model vs. Data

%Δ generational inequality	1929-1984	1984-2011	1984-2011 (PE)
Data	-90.22%	492.1%	492.1%
Model (Fixed Prior)	-43.44%	29.70%	29.17%
Model (Market-Based Prior)	-59.54%	32.96%	32.42%

However, our attention is on the model's ability to explain the *changes* in generational inequality in the last century. Table 2 compares the model performance relative to the data by measuring the changes in the old-to-young wealth ratio from the beginning to the end of the sample periods measured in percentage changes in the old-to-young wealth ratio. As one can see, both the model and the data show a decreasing old-to-young wealth ratio before 1984 and then an increase after 1984, albeit the model generates relatively smaller changes. 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 change. Therefore, a better statistic to evaluate the fit of the model is to ask how much of the change in the old-to-young wealth ratio can be explained by the model. From 1929 to 1984, the old-to-young wealth ratio declined by 90.22% in the data, while the model generates a decline of 43.44%, which is around half of the decrease. In addition, the model counts for $(29.7/492.1)\% = 6\%$ of the rise of the old-to-young wealth ratio between 1984 and 2011. Using the market-based prior, the model generates an even higher increase of generational inequality of around 6.7% of the changes in this period. As a decomposition exercise, I re-run a partial equilibrium calibration by assuming that asset prices do not change after a disaster and examine its inequality prediction again. As one can see, the general equilibrium interaction produces an amplification effect on the post-1984 increase in generational inequality. However, the GE effect accounts for only 1.8% of the increase in recent generational inequality increase, and most of the changes are driven by the partial equilibrium channel. This can be explained by the fact that, even during the relatively calm period after the 1980s, the heightened optimism of the baby boomers led to an increase in demand for risky assets, thereby lowering their returns and ultimately weakening the general equilibrium amplification mechanism.

5.2. **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 belief differences influence savings rates as well, which in turn influences generational wealth inequality. Interestingly, data from Moody's Analytics shows that the savings rate has been declining for all age groups since 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. This shows that the portfolio choice channel would have been more important in recent years if savings rates are declining. To be more specific about how disasters might alter the savings rate, The Appendix further examines how the savings rate responds to experienced stock market returns, controlling other factors. In all regression specifications, there is no significant correlation between previous stock returns and the savings rate. This provides further empirical support that it is reasonable to fix the savings rate as a constant in this model.

5.3. **Other statistics.** One way to test the validity of the model is to examine its direct performance in matching changes in generational inequality and to study its implication on some other statistics in the data. Table 3 shows how the model excels in other dimensions of the data other than the old-to-young wealth ratio.[12] As stated in the benchmark calibration results, the benchmark model is able to explain half of the decline of the old-to-young wealth ratio from 1929 to 1984 and then again around 7% of the increase afterwards. The model also predicts a 14% times increase of top 14.88% wealth share increase, while in the data it's 61.95%, which implies that the model is able to explain around 24% increase of top 1% share after the 1980s. This is a fairly encouraging result, given that the model

[12]I used the Saez and Zucman top income database to get the top 1% share in the data, which ends in 2016. The risky share and age correlation is estimated from the PSID, where the 2017 data is used to approximate its value in 2016.

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, labour 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 invest a higher share of their wealth in the risky asset. A positive correlation between risky share and age are seen both in the model and in the data from PSID, albeit with different levels. In the model, such correlation amounts to 0.96, while in the data, it is only 0.36. This is not surprising, since the data also consists of many retired households who cash out from the market to finance retirement consumption, while the model focuses on before-retirement investment patterns.

Table 3. Alternative statistics: model vs. data (1984-2011)

	Data	**Benchmark**
%Δ Top 1% wealth share	61.95%	14.88%
Corr(Risky Share, Age)	0.36	0.96

6. Empirical Evidence

In this section, I provide further empirical evidence on generational belief differences, portfolio choice and wealth inequality.

6.1. **Evidence on life cycle portfolio choices.** One implication of this model is that it links portfolio choice decisions directly to experienced stock market crashes. This produces testable restrictions on observed life cycle portfolios. To examine this, I use stock market participation rate and portfolio choice data from the SCF and compare the mean risky portfolio share for all ages in the 1983 and the 2019 waves respectively. Since a direct estimate on disaster-contingent asset portfolio is not attainable (i.e.: this requires a micro-level dataset on different financial derivative asset holdings like the put options), I provide here indirect estimates by classifying asset into risk-free and risky where the counterpart of the risky asset in the model can be considered as the sum of stocks and the disaster-contingent asset following Vissing-Jorgensen (2003). The 1983 Survey of Consumer Finance wave has less information on asset positions, but still provides relatively detailed information on stock and bond holdings. I, therefore, define risk-free asset holdings as the total amount in checking accounts, money market and call accounts, savings accounts, certificates of deposits, bonds, and life insurance. Risky assets are then defined as total amounts in stock and mutual funds. The 2019 wave has richer information on detailed financial allocation data in various accounts. I define risky assets as the total amount in stock holding in the Roth IRA, roll-over IRA, regular or other IRA, Keogh accounts, stock holding in the savings accounts, direct holding in publicly traded stocks, stock holding in annuity accounts, and stock mutual funds. Risk-free assets are defined as the sum of a checking account, Certificate of deposit, non-stock savings in the savings account, bond mutual fund, government bond mutual fund, other bond mutual funds, savings bonds, other bonds, state and municipal bonds, foreign bond, corporate bonds, cash, non-stock holding in annuity accounts, life insurance. I then define wealth as the sum of risky and risk-free assets, and net debt values.

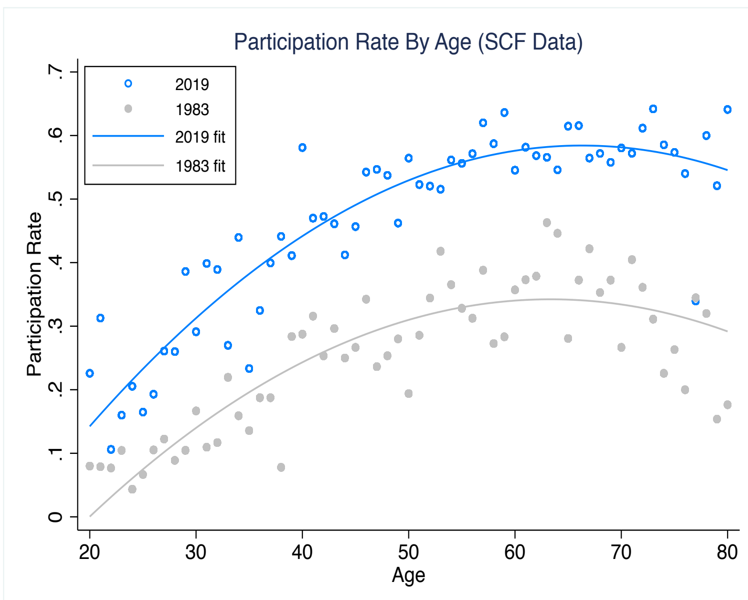

FIGURE 7. Stock Market Participation Rate By Age (SCF Data)

Figure 7 plots the stock market participation rate as a function of age. Since the life cycle plays an important role in risky portfolio share, I fit the participation rate with a quadratic function of age. However, for our purpose, we will focus our attention on the slope comparison between 1983 and 2019 instead of the life-cycle feature for each given year. The figure shows several interesting patterns: First, the stock market participation rate increases with age before retirement and decreases after retirement. Second, the difference in old vs. young participation rate is starker in 2019 than in 1983. Interestingly, my model provides a rationale for these patterns. Remember that in 1983, old people are Depression Babies. Even though they built optimism gradually after the Great Depression, they are still not as optimistic as the younger people in 1983. However, in 2019, when boomers are getting older, they are much more optimistic than the young millennials. Even though both generations experienced the recent financial crisis, the boomers were less scarred compared with the less experienced millennials. My model traces these belief changes to portfolio choice changes directly.

To further examine the differences in the participation rate and risky portfolio share across these two years, I conduct the following regression analysis to elicit the cohort effect on financial risk-taking.

$$P(participation) = \alpha_0 + \alpha_1 netwealth + \alpha_2 education + \alpha_3 age + \alpha_4 age^2 + \alpha_5 netincome + \alpha_6 race \tag{6.47}$$

$$Riskyshare = \beta_0 + \beta_1 netwealth + \beta_2 education + \beta_3 age + \beta_4 age^2 + \beta_5 netincome + \beta_6 race \tag{6.48}$$

The first equation features a logit regression, and the second regression is a standard robust ordinary least squares regression. Various factors are essential in explaining risky portfolio share. For the purpose of this paper, let us focus our attention on the coefficient in front of age. The important thing to notice here is the coefficient in front of age. As one can see, in both years, both the participation rate and the risky portfolio share increased with age significantly, albeit with a much steeper slope in 2019 than in 1983, which is consistent with the predictions of the model.

6.2. **Generational belief differences vs. Inequality.** In the model section, I consider the Great Depression and the Great Recession as the only two disasters during the last 100 years in the US. This makes the model analytically tractable, but it neglects the potential impacts of smaller disasters on wealth distribution. In this section, I provide additional empirical evidence on generational belief differences and their correlation with top wealth shares. Figure 8 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. As one can see, such events have been rather rare, and the Great Depression has so far the largest size of a stock market crash, which features an 84.76% loss of stock value in total. [13] However, even before the Great Depression, the US economy has not been tranquil. There was a 1907 banking crisis, and an 1873 stock market crash before that. However, the generations that were born between the end of the Great Depression and the 1980s have enjoyed a Golden age of the US economy, with no major crisis. In contrast, young people in recent years have

[13] In his famous book "The Greatest Generation" (Brokaw (2000)), Tom Brokaw 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.

Dependent Variables	1983	2019
(a) Prob. of participation		
net wealth	1.08e-5***	1.47e-7***
	(9.81e-07)	(7.57e-09)
education	0.530***	0.350***
	(0.045)	(6.90e-3)
age	0.099***	0.110***
	(0.022)	(0.007)
age^2	-0.001***	-0.001***
	(-3.53)	(6.68e-5)
net income	-2.99e-7	7.09e-10*
	(5.10e-7)	(5.05e-10)
(b) Risky portfolio share		
net wealth	2.19e-08***	1.35e-09***
	(6.89e-09)	(1.82e-10)
education	0.048***	0.023***
	(0.003)	(0.001)
age	0.005***	0.007***
	(0.001)	(0.001)
age^2	-2.36e-5**	5.34e-5***
	(-1.74)	(6.44e-6)
net income	1.40e-7***	3.64e-11**
	(4.19e-08)	(2.05e-11)
R-squared	0.245	0.133
Observations	3599	25330

Table 4. Cross sectional data regression on age

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Standard errors in parentheses.
Data source: Survey of Consumer Finance

witnessed more crises, from the 1987 stock market crash, to the 2000 tech bubble burst, to the financial crisis, and even more recently, the Covid crisis. Those traumatic events could have left profound mental impacts, and created strong belief scarring effects. To illustrate this, Figure 9 plots the pessimism index from 1941 to 2020 using the same data,

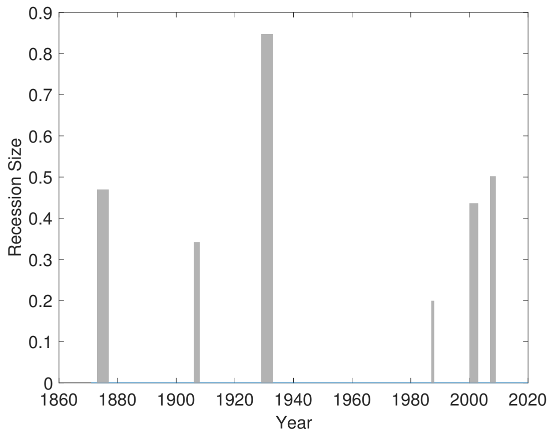

FIGURE 8. Stock Market Crash

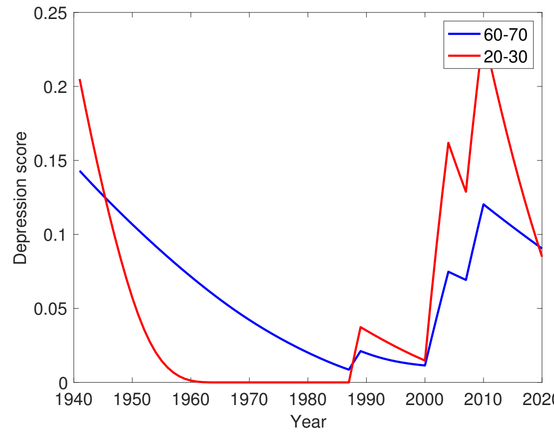

FIGURE 9. Pessimism Index

contrasting differences in pessimism between the old (60-70 years old) and the young (20-30 years old).[14] The depression score $P_{i,t}$ for generation i at time t is defined as a lifetime weighted average of depression loss, or more precisely,

$$P_{i,t}(\lambda) = \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k,\lambda)\mathbb{1}(Depression_{t-k} = 1)L_{t-k} \tag{6.49}$$

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 A.1. Here, I use the same experience weighting function to construct the pessimism index, and define disasters where the peak to trough stock market value drop of more than 20%.

[14]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.

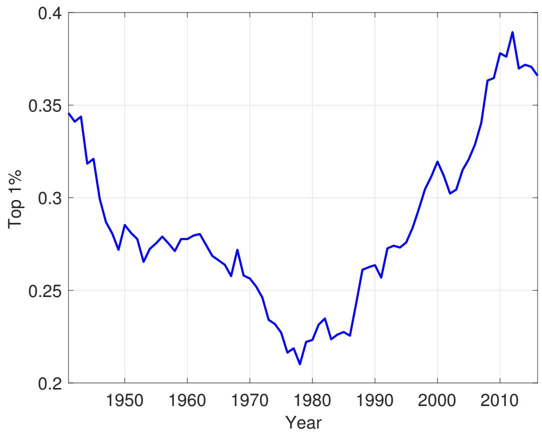

FIGURE 10. Top 1% wealth share

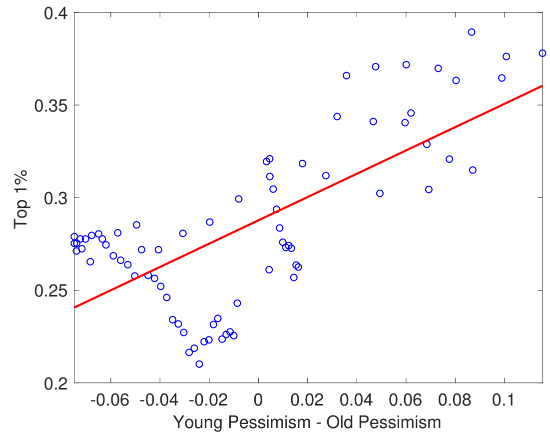

FIGURE 11. Top 1% wealth share s. Relative Optimism

Interesting patterns emerge in Figure 9. Before the mid-1980s, both the young and the old become more optimistic, but the younger generations become optimistic at a much faster rate. While the old are still digesting trauma from the Great Depression, the 1873 stock market crash and 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 the mid-1980s. Then the table turned. With smaller crashes in the 1987s, the dotcom bust, and the 2007-2009 financial crisis, doubts were raised by young people. Although both the recent young and the old generations have experienced these disasters, the young generations have less experience, and therefore would over-extrapolate from the disaster. In summary, the old were more pessimistic than the young before the 1980s but became more optimistic after the 1980s. So why is this depression score interesting? Remember, the famous U-shaped pattern of inequality also features a turning point around the 1980s!

To see the connection, Figure 10 plots the evolution of the top 1% wealth share in the United States using the Saez and Zucman (2016) data [15] Figure 11 plots the same statistics against relative optimism, defined as the difference between the young depression score and the old depression score. 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 disastrous events but also revise gradually during normal times. After all, if generations experience both boom and

[15] 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.

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 this idea, I ask the following question: for 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

$$er_{i,t} = prob(Depression = 1)_{i,t} * \kappa_t + (1 - prob(Depression = 1)_{i,t}) \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k, \lambda) r_{t-k} \tag{6.50}$$

where

$$prob(Depression = 1) = \sum_{k=1}^{age_{i,t}-1} \omega_{i,t}(k, \lambda) \mathbb{1}(Depression = 1) \tag{6.51}$$

and

$$\omega_{i,t}(k, \lambda) = \frac{(age_{i,t} - k)^{\lambda}}{\sum_{k=1}^{age_{i,t}-1} (age_{i,t} - k)^{\lambda}} \tag{6.52}$$

This captures the idea that the expected returns are the weighted average of the return during disaster times as well as normal times, 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.[16] Since there is no stock market return data before 1871.01, I compute the implied beliefs for all cohorts in 1871 assuming that no disasters happens before that, so that disaster likelihood decreases gradually with age. Figure 12 compares the expected return for old vs. young.

Up until the 1980s, the young expected higher returns than the old. This is understandable because while the old struggled with the aftermath of the Great Depression and possibly earlier crashes, the young cohort did not have those experiences. Notice that their expected return dropped in the later part of this period due to a slight downturn in the stock market in the 1960-1970s, there were no major disasters during this period, and they are still much more optimistic than the old. However, the tables turned during the 1980s. With the 1987 crash, the 2000 dot-com bubble bust, and the recent financial crisis, the new younger generation was traumatized. Taking into account possible future

[16] Malmendier and Nagel (2011) uses the arithmetic mean return to measure experienced returns. For a behavioural investor who cares about gains and losses from a reference point, this could well capture the experience of his/her investment returns. However, a more rational investor who cares about the final wealth position would take a slightly different view. Such an investor would instead take the geometric mean instead of the arithmetic mean to measure his/her return experience. In Appendix A.2, I show that although there is a slight difference in these two measures, the qualitative pattern of the expected returns of old vs. young still holds.

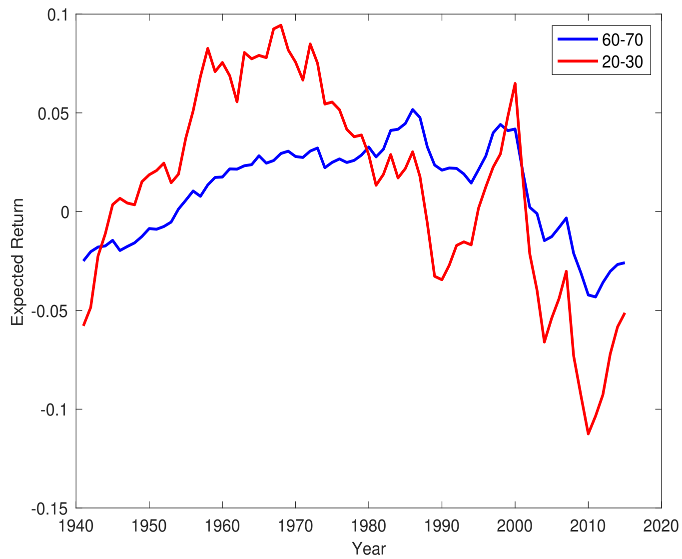

FIGURE 12. Expected Return: Old vs. Young

crashes, they even start to expect negative returns. Notice that there is a short period where young people's optimism is boosted (i.e., the stock market boom in the 1990s), but it is not enough to undo the negative effect of the two recent crises they experience. Although the old, especially the boomers, have had similar experiences, they still have the memory of the good old times and are more optimistic about the returns.

7. Alternative Mechanisms

7.1. **What about housing?** A natural question to ask might be: what about housing? After all, the last few decades have witnessed large swings in housing prices. Given that older people are more likely to be homeowners than young people, changes in housing prices and home ownership seem likely to account for the majority of changes in generational inequality (Kuhn, Schularick, and Steins (2017)), Rognlie (2016)).

To disentangle overall wealth from housing wealth, I now use the PSID data on wealth measured by the sum of six asset types' net debt value. Figure 13 plots the generational wealth gap with and without housing, measured by the median wealth ratio of the ages 65 and up and 35 and under.

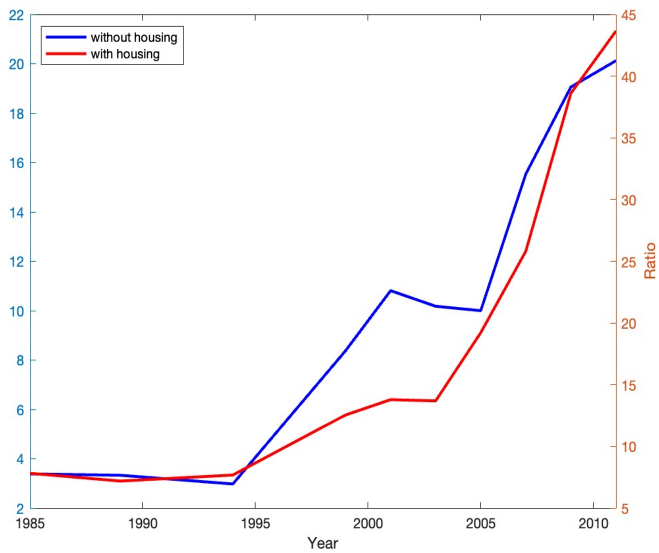

FIGURE 13. Median Net Worth Ratio with/without Housing (PSID)

As one can see, the *level* of the old-to-young wealth ratio has almost doubled by including housing. However, the *percentage* increase in this ratio is very similar. Having said that, the proposed model here aims at explaining the data without housing wealth.

7.2. **Financial Market Development.** One obvious concern could be that the financial market became much more developed after the 1980s, which produced an increase in stock market participation. This increases the growth rate of wealth for everyone but is also disproportionately benefiting the older more since they have more wealth to be invested than the young. While I acknowledge that the extensive margin of financial inclusion could be an essential aspect of generational inequality, it does not capture the intensive margin of portfolio allocation. To examine this, I now focus on stock market participants and study the life cycle behaviour of portfolio allocation in 1984 and 2017 using PSID data. If the "belief scarring" channel exists, the slope of the life cycle of risky stock shares would be very different in these two years. As expected, in both years, stock share as a fraction of wealth increases with age, and the slope has also become steeper. In 1984, the correlation between stock shares and age was only 0.2708, but in 2017, the correlation rises to 0.4579. This suggests that the extensive margin of stock market participation cannot be the only mechanism that drives the recent increase in generational inequality.

7.3. **Relaxed Borrowing Constraints.** The development in financial markets also relaxed borrowing constraints in the US since the 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 lose 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. Thus, it pays off to examine the difference between gross and net wealth. Suppose we see that gross wealth inequality has not 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 use PSID data to compute the gross wealth ratio. Again, in 1984, the wealth ratio of the two groups was 3.346 times, but in 2009 [17], the ratio has increased to 8.856 times. This suggests that there are forces other than loosening borrowing constraints that contribute to the divergence of wealth between the young and the old.

7.4. **Direct and Indirect Inter-generational Transfers.** Inheritance and other inter-generational transfers play a potentially crucial role in generational inequality (See Boar (2020)). Perhaps the millennials have nothing to worry about since they will inherit their parents' houses and bank accounts. On the other hand, the increased cost of life-extending medical treatments might 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 be 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 difference to this ratio, which has a value of 6.05 [18], while in 2013, there is only a slight difference. The old-to-young wealth ratio is 17.23 after inheritance and becomes 17.41 before inheritance. Therefore, the ratio does not differ much by varying direct transfers in the form of inheritance.

[17] PSID has a different definition of debt after that year.

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

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 parents are paying tuition for their kids, 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, with the assumption that these are the tuition paid to finance the education of their kids. Since wealth is a stock variable, but education expense is a flow variable, I adjust the cumulative education expense by four times of the yearly reported education expense assuming that these expenses occur due to the four-year college education. Interestingly, without taking into account tuition expenses, the old-to-young wealth ratio grew from 8.26 times to 13.5 times, which is about a 63% increase. If one subtracts education expenses from wealth, the ratio goes from 9.269 times to 15.756 times, which is around 70% of increase. So in fact, the rise in college tuition makes the younger generation even poorer. One possible interpretation of this is that tuition-paying parents are mostly middle-aged, and when they reach age 65, their college-educated kids have already graduated. So, even though tuition expenses might affect family budgets while parents are in their middle ages, it does not affect many after age 65. At the same time, the rising education expenses push young people to take out greater student loans, which further drags down their bank accounts. Of course, the young might recoup this expense in the form of higher future labour income, but that is uncertain.

Finally, since we are discussing generational inequality in the U.S., we must briefly consider social security. In the U.S., the social security program has been expanded hugely over the last several decades (See Bourne, Edwards, et al. (2019)). Since it primarily operates on 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 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 a significant rise in generational inequality from the 1980s.

7.5. **Increased Supply of Data.** One might ask, why include experiential learning? Wouldn't standard Bayesian learning, which incorporates all historical data, also generate wealth dispersion if everyone becomes more optimistic when more data becomes available? Perhaps pessimism induced by the Great Depression makes people invest less in the beginning, reducing inequality, but over time as optimism builds people invest 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. However, this explanation is in contrast with the data on survey expectations. If we think that investors learn from their own experience and also pay attention to historical data, then over time, as more data reaches them, their beliefs should become increasingly homogeneous - even if they start out with very different ones. The monthly Shiller's data, a stock market crash optimism index starting from 1989, shows that this is not the case. It measures the percent of the population who have low expectations (strictly less than 10%) of a stock market crash in the next six months. This 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 since 1989 to a consistent sample of respondents. Figure 14 plots the standard error of the measure for the institutional and individual data. Using standard error as a measure of belief heterogeneity 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

In this paper, inequality is generated within a complete market economy. In contrast, most other models studying inequality consider incomplete market economies (i.e., Hugget or Bewley models). Does this imply that inequality here is efficient? Perhaps not. In fact, with heterogeneous beliefs, there has been a debate about the Pareto criterion, which dates back to the 1970s Starr (1973), Harris (1978) and Hammond (1981). This early work highlighted that when beliefs are different, ex-ante efficiency might not correspond to ex-post efficiency. This issue is present in my model as well. With heterogeneous priors and experiential learning, each investor considers their own beliefs to be correct. Each thinks they would be better off with speculation ex-ante. However, ex-post consumption is excessively 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. They know that their future perceived welfare gains are likely to be spurious. Another limitation of

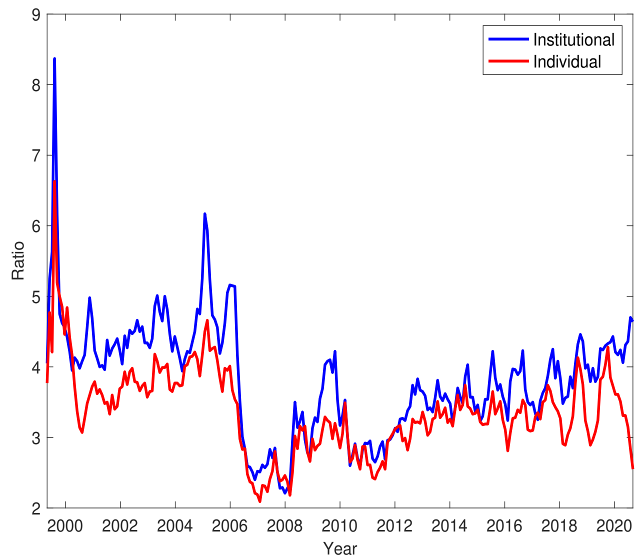

FIGURE 14. Measure of Belief Convergence: Standard Error of Cash Confidence Index

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. Recent work has proposed new Pareto criteria in evaluating efficiency with heterogeneous priors. For example, to address the problem of whose beliefs to evaluate under, Brunnermeier, Simsek, and Xiong (2014) propose 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), which proposes 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 a belief formation mechanism exhibits strong over-extrapolation behaviour (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 interacts with a rational agent in the financial market. Beaudry and Portier (2006) shows how extrapolation of information from stock prices can affect future GDP growth rate. Evidence of over-extrapolation is pervasive. In financial markets, it is supported by a seminal paper Malmendier and Nagel (2011), which uses data from the 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, labour market, the housing market, and overall business cycle conditions. (Malmendier and Nagel (2015), Wee (2016), Malmendier and Shen (2018) Kozlowski, Veldkamp, and Venkateswaran (2020) and Kuchler and Zafar (2019)). However, those papers are most suited for studying macroeconomic aggregate and asset prices, but not so much on wealth distribution. Acedański (2017) attempts to solve a heterogeneous expectations model *a la* Krusell and Smith (1998) to study wealth distribution. It focuses on exogenous forecasting rules and stationary wealth distribution, while my paper uses embedded endogenous heterogeneous beliefs and focuses on the dynamics of wealth distribution.

Second, this paper attempts to generate heterogeneous beliefs when individuals learn from their own experiences. 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 essential *endogenous*, which 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 for 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 problem 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.

This paper is related to recent literature on disaster risk in the tradition of Barro (2006). The incorporation of rare disaster risks naturally generates a disaster premium, which significantly reduces the level of risk aversion needed in matching empirically plausible equity premiums. Various extensions of disaster risk models also help 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 the 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 about 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 a response from investors that experienced it, thus generating heterogeneous generational bias in the disaster risk distribution. Although there are several interesting papers that combine 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 build 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.

Finally, this paper contributes to the 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 this framework is 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 rate after the 1980s. However, they focus on inequality within cohorts with private equities. Here, these models are generalized to solve distribution across cohorts and solve a model with aggregate shock and public equity. Finally, tracing rare disasters back to the Great Depression jointly explains the dip in wealth inequality after the Great Depression and the rise of inequality after the 80s.

10. Conclusion

We live in a world with finite lives and limited data. This paper bridges the gap between experiential learning literature, which is traditionally behavioural finance literature, and the macroeconomic literature on wealth inequality. It highlights how stock market disasters like the Great Depression could have a prolonged impact on generational inequality through the channel of experiential learning. I build and solve a general equilibrium model, with agents who learn from experience, to examine the qualitative and quantitative implications of long-run wealth differences between cohorts. To the best of my knowledge, this is the first paper that combines experiential learning with wealth inequality, which should spark interest in many possible extensions. For example, future research could extend this framework with nominal rigidity to explore the role of monetary policy when agents are learning from inflation experience (which 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. When generational belief differences matter, it opens doors for policymakers to combat inequality. An example would be a mandatory pension fund designed to improve wealth accumulation of scarred generations by helping them invest in stocks when they fear to do so by themselves.

Appendix A.

A.1. **The experience weighting function.** Figure 15 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 decrease 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 hold the hangover of the Great Depression 47-51 years ago, while a boomer would put zero weight on that.

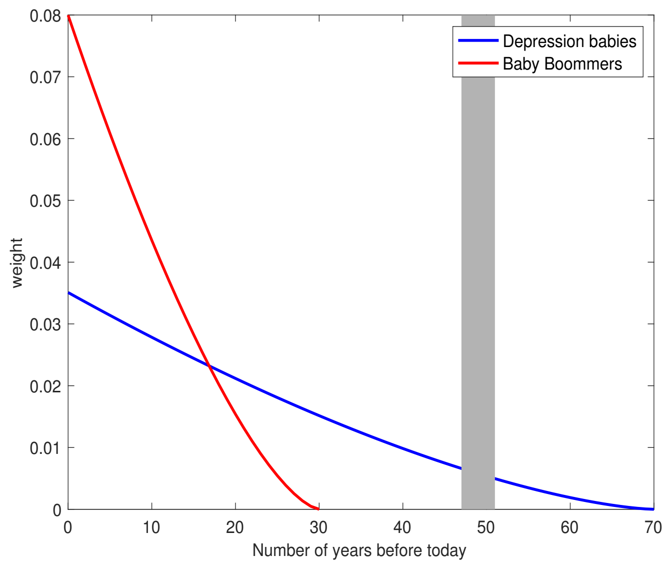

Figure 15. Historical weights: Depression babies vs. Boomers

A.2. **Robustness check on experienced return.** The following two figures plot the generational belief differences using two different measures of experienced returns.

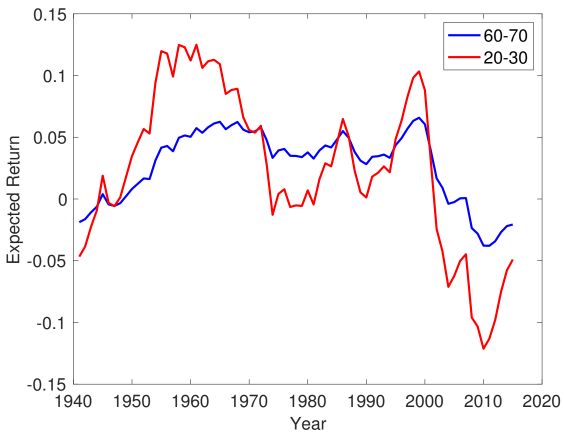

FIGURE 16. Using experienced annual return

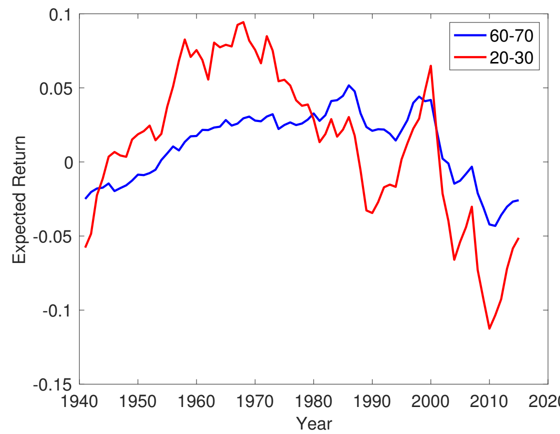

FIGURE 17. Using average cumulative annual return

A.3. **Proof of Lemma 2.2.** See Dieckmann (2011) for the proof of eqn.(2.16) and eqn.(2.17). The derivation of $\xi_{s,t}$ process follows first by applying the Girsanov theorem for the jump process, s.t:

$$dN_{s,t} - \bar{\lambda}_{s,t}dt = dN_t(\bar{\lambda}_t) - \bar{\lambda}_t dt \tag{A.53}$$

With the change of measure, we can rewrite eqn.(2.16) into

$$\frac{d\xi_{s,t}}{\xi_{s,t-}} = \left(\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t + (\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1)(\lambda_{s,t} - \bar{\lambda}_t)\right) dt - \theta_{s,t} dZ_t + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right) dN_t(\bar{\lambda}_t) \tag{A.54}$$

Then the SDE for $\eta_{s,t}$ follows directly from the application of the multidimensional jump-diffusion version of 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

$$\frac{d\eta_{s,t}}{\eta_{s,t}} = \left(\frac{1}{1+\bar{\kappa}}(\lambda_{s,t} - \mathbb{E}(\lambda_{s,t})\right) dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2}{1+\bar{\kappa}} - 1\right) - 1\right] dN(\bar{\lambda}_t) \tag{A.55}$$

A.4. **Proof of proposition 2.** To get the coefficient of the stock price, we can write down the formula for stock prices, i.e.,

$$\begin{aligned} S_t &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty \xi_u Y_u du\right] \\ &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty e^{-(\rho+\delta(1-\beta))u}\eta_u du\right] \\ &= \frac{1}{\xi_t}\eta_t\int_t^\infty e^{-(\rho+\delta(1-\beta))u}du \\ &= \frac{1}{\rho+\delta(1-\beta)}Y_t \end{aligned} \tag{A.56}$$

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

$$\frac{dS_t}{S_{t^-}} = \frac{dY_t}{Y_{t^-}} \tag{A.57}$$

Recall that the compounded stock market value follows the following process

$$\frac{dS_t + D_t dt}{S_{t^-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{A.58}$$

Matching coefficients, one get

$$\mu^S = \mu + \rho + \delta(1-\beta); \quad \sigma^S = \sigma; \quad \kappa_t^S = \kappa_t \tag{A.59}$$

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

$$\mu_t^P = -\kappa_t^P\lambda_t^N + r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t}) + r_t \tag{A.60}$$

Appendix B. Proof of Proposition 3

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

$$\frac{dw_{s,t}}{w_{s,t^-}} = \hat{\mu}_{s,t}dt + \hat{\sigma}_{s,t}dZ_t + \hat{\kappa}_{s,t}dN_t \tag{B.61}$$

where dZ_t and dN_t represent aggregate Brownian motion and jump shocks. To simplify notation, I will now eliminate all subscripts in the following texts. 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

$$\begin{aligned} A(t+dt) &= \int_{-\infty}^\infty f(w)n_{t+dt}dw \\ &= \int_{-\infty}^\infty (f(w)+df(w))n(w) - \delta f(w)n(w)dw \\ &= \int_{-\infty}^\infty f(w)(1-\delta)n(w)dw + \int_{-\infty}^\infty df(w)n(w)dw \end{aligned} \tag{B.62}$$

We then have

$$d(A(t)) = -\int_{-\infty}^{\infty} \delta n(w) f(w) dw + \int_{-\infty}^{\infty} df(w) n(w) dw. \tag{B.63}$$

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

$$df(w) = f'(w)[\hat{\mu} w dt + \hat{\sigma} w dZ] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt + [f(w(1+\hat{\kappa})) - f(w)]dN \tag{B.64}$$

Using integration by parts, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} df(w) n(w) dw &= \int_{-\infty}^{\infty} \left[f'(w) \left[\hat{\mu} w dt + \hat{\sigma} w dZ \right] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt \right] n(w) dw \\
&+ \int_{-\infty}^{\infty} \left[f(w(1+\hat{\kappa})) - f(w) \right] n(w) dN dw \\
&= \int_{-\infty}^{\infty} f(w) \left[-\frac{\partial}{\partial w} \left(n(w) \hat{\mu} w dt + n(w) \hat{\sigma} w dZ_t \right) + \frac{1}{2} f(w) \frac{\partial^2}{\partial w^2} \left(n(w) \hat{\sigma}^2 w^2 \right) dt \right] \\
&+ \int_{-\infty}^{\infty} \left[n(w(1-\hat{\kappa})) - n(w) \right] f(w) dN dw
\end{aligned} \tag{B.65}$$

Notice the way I write down changes in $A(t)$ in (B.63) fixes the density of w in the state space and calculates 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,

$$d(A(t)) = \int_{-\infty}^{\infty} f(w) \frac{\partial n}{\partial t} dt dw = \int_{-\infty}^{\infty} df(w) n(w) dw \tag{B.66}$$

Plugging in the expression in eqn. (B.65), and equating the integrands, we get

$$dn = -\frac{\partial}{\partial w}(n\hat{\mu} w dt + n\hat{\sigma} w dZ) + \frac{1}{2}\frac{\partial^2}{\partial w^2}(n\hat{\sigma}^2 w^2)dt - \delta n + [n(w(1-\hat{\kappa}), t)) - n(w,t)]dN \tag{B.67}$$

As one can see, the distribution of this variable is stochastic, and 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 $dp(w) = \mathbb{E}_t\left(dn(w)\right) = 0$? [19] By averaging out the KFP equation, we then have

$$-\frac{\partial}{\partial w}\left(\mathbb{E}(\hat{\mu}) w p(w)\right) + \frac{\partial^2}{\partial w^2}\left(\frac{\mathbb{E}(\hat{\sigma}^2)}{2} w^2 p(w)\right) - \delta p(w) + \lambda(p(w(1-\bar{\kappa})) - p) = 0 \tag{B.68}$$

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

[19]The expectation is taken as the time-series average.

wealth share defined as:

$$\tilde{w}_{s,t} = \frac{w_{s,t}}{W_t} \tag{B.69}$$

where W_t denotes the aggregate wealth. That is, the absolute wealth normalized by aggregate wealth. Since agents are born with zero financial wealth, we have $\tilde{w}_{s,s} = \frac{\tau/\delta W_s}{W_s} = \frac{\tau}{\delta}$. This variable has a stationary distribution absent aggregate shocks. Recall that, after imposing the market clearing condition, the individual wealth dynamics follow the following

$$\frac{dw_{s,t}}{w_{s,t^-}} = \left(\sigma^2 + r_t - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right)\right)dt + \sigma dZ_t + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_t^N} - 1\right)dN_t \tag{B.70}$$

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

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \left(\sigma^2 + r_t - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right) - \mu\right)dt + \left(\frac{\lambda_{s,t}}{\mathbb{E}(\lambda_{s,t})}(1+\kappa_t) - 1\right)dN_t \tag{B.71}$$

which in short-hand can be written as

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \hat{\mu}(\lambda_{s,t})dt + \hat{\kappa}(\lambda_{s,t})dN_t \tag{B.72}$$

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

$$dx = \hat{\mu}dt + \log(1+\hat{\kappa})dN_t \tag{B.73}$$

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

$$p(x) = \int_0^\infty n(x,\lambda)d\lambda \tag{B.74}$$

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.,

$$n(x,\lambda) = n_1(x|\lambda)n_2(\lambda) \tag{B.75}$$

From eqn. (B.73), we can write down the dynamics of $n_1(x|\lambda)$, i.e.,

$$0 = -\frac{\partial n_1}{\partial x}\hat{\mu} + \lambda^0\left(n_1(x - \log(1-\hat{\kappa})) - n_1\right) - \delta n_1 \tag{B.76}$$

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

$$\zeta \approx \zeta_0 + (\lambda - \lambda^0)\zeta_1 \tag{B.77}$$

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

$$d\lambda_{s,t} = (\lambda_{s,t^-} - \lambda_l)(\lambda_{s,t} - \lambda_h)dt - (\lambda_{s,t^-} - \lambda_h)(\lambda_{s,t^-} - \lambda_l)\frac{(1+\lambda_{s,t^-})}{\lambda_{s,t^-}}dN_t \tag{B.78}$$

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

$$0 = -\frac{\partial n_2}{\partial \lambda}(\lambda - \lambda_h)(\lambda - \lambda_l) - n_2(2\lambda - \lambda_l - \lambda_h + \delta) + \lambda^0(n_2^J - n_2) \tag{B.79}$$

We can guess and verify the following approximate exponential solution

$$n_2(\lambda) \approx e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} \tag{B.80}$$

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.,

$$\begin{aligned} p(x) &= G_0 e^{(\zeta_0 - \lambda^0\zeta_1)x}\int_{\lambda_l}^{\lambda_h} e^{\lambda\zeta_1 x} e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} d\lambda \\ &= \underbrace{Ge^{\zeta_0 x}}_{dogmatic}\ \underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}]}_{Learning} \end{aligned} \tag{B.81}$$

Let $p^{dogmatic}(x)$ denote the long-run stationary distribution of log normalized wealth in an economy with fixed dogmatic beliefs that equal the long-run average beliefs, we then have

$$\begin{aligned} \lim_{x\to\infty} \frac{p(x)}{p^{dogmatic}(x)} &= \lim_{x\to\infty} [\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}] \\ &= \lim_{x\to\infty} \zeta_1^{-1}\left[-(\lambda_l - \lambda^0)\zeta_1 e^{(\lambda_l - \lambda^0)\zeta_1 x}\right] \end{aligned} \tag{B.82}$$

where the second equality uses 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 when $x \to \infty$. We then have

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^{dogmatic}(x) \tag{B.83}$$

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

B.1. **Verification of Newborn Consumption Share.** We start by defining β_t, i.e.,

$$\beta_t = \frac{c_{t,t}}{Y_t} = \frac{(\rho+\delta)w_{t,t}}{Y_t} \tag{B.84}$$

where the second equality comes from the consumption smoothing of a log agent. Next, we relate initial wealth with aggregate wealth

$$w_{t,t} = \frac{\tau}{\delta} S_t = \frac{\tau}{\delta[\rho+\delta(1-\beta)]} Y_t \tag{B.85}$$

by noticing that the stock market value S_t also represents the amount of aggregate wealth.

This renders the two solutions

$$\beta_{1,2} = \frac{\rho+\delta}{2\delta} \pm \frac{\sqrt{(\rho+\delta)^2 - 4\tau(\rho+\delta)}}{2\delta} \tag{B.86}$$

Recall that $S_t = \frac{1}{\rho+\delta(1-\beta)} Y_t$, which then requires that $\beta < \frac{\delta+\rho}{\delta}$. This eliminates the larger positive root, and the smaller root is preserved, i.e.:

$$\beta = \frac{\rho+\delta}{2\delta} - \frac{\sqrt{(\rho+\delta)^2 - 4\tau(\rho+\delta)}}{2\delta} \tag{B.87}$$

B.2. **Savings rate Response to Stock Market Scarring.** The table shows the OLS regression results of contemporaneous savings rate on the 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 values 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 averages.

TABLE 5. 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

Acedański, J. (2017): "Heterogeneous expectations and the distribution of wealth," *Journal of Macroeconomics*, 53, 162–175.

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.

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1997): "Parental altruism and inter vivos transfers: Theory and evidence," *Journal of political economy*, 105(6), 1121–1166.

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.

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.

Beaudry, P., and F. Portier (2006): "Stock prices, news, and economic fluctuations," *American economic review*, 96(4), 1293–1307.

Benhabib, J., A. Bisin, and S. Zhu (2016): "The distribution of wealth in the Blanchard–Yaari model," *Macroeconomic Dynamics*, 20(2), 466–481.

Benhabib, J., and S. Zhu (2008): "Age, luck, and inheritance," Discussion paper, National Bureau of Economic Research.

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.

Blanchard, O. J., R. Shiller, and J. J. Siegel (1993): "Movements in the equity premium," *Brookings Papers on Economic Activity*, 1993(2), 75–138.

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.

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

Caputo, J. (2020): "Parental coresidence, young adult role, economic, and health changes, and psychological well-being," *Society and Mental Health*, 10(3), 199–217.

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.

Cohn, A., J. Engelmann, E. Fehr, and M. A. Maréchal (2015): "Evidence for countercyclical risk aversion: An experiment with financial professionals," *American Economic Review*, 105(2), 860–85.

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.

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

Dillenberger, D., and K. Rozen (2015): "History-dependent risk attitude," *Journal of Economic Theory*, 157, 445–477.

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.

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.

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.

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.

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.

Hill, J. M., V. R. van der Geest, and A. A. Blokland (2017): "Leaving the bank of mum and dad: Financial independence and delinquency desistance in emerging adulthood," *Journal of Developmental and Life-Course Criminology*, 3(4), 419–439.

Jagannathan, R., E. R. McGrattan, and A. Scherbina (2001): "The declining US equity premium," Discussion paper, National Bureau of Economic Research.

Kaplan, G. (2012): "Moving back home: Insurance against labor market risk," *Journal of Political Economy*, 120(3), 446–512.

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).

Krusell, P., and A. A. Smith, Jr (1998): "Income and wealth heterogeneity in the macroeconomy," *Journal of political Economy*, 106(5), 867–896.

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.

Mortimer, J. T., M. Kim, J. Staff, and M. Vuolo (2016): "Unemployment, parental help, and self-efficacy during the transition to adulthood," *Work and occupations*, 43(4), 434–465.

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.

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.

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

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.