Three Essays in Macroeconomics and Finance

by

Yang Li

M.A., Simon Fraser University, 2014
M.A., Toulouse School of Economics, 2010
B.A., University of International Business and Economics, 2007

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Economics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY
Spring 2022

Declaration of Committee

Name: **Yang Li**

Degree: **Doctor of Philosophy**

Thesis title: **Three Essays in Macroeconomics and Finance**

Committee: **Chair:** Simon Woodcock
Associate Professor, Economics

Kenneth Kasa
Co-Supervisor
Professor, Economics

John Knowles
Co-Supervisor
Professor, Economics

Luba Petersen
Committee Member
Associate Professor, Economics

Minjie Deng
Examiner
Assistant Professor, Economics

Jun Nie
External Examiner
Senior Economist
Federal Reserve Bank of Kansas City

Abstract

Chapter 1 develops a continuous-time, heterogeneous agents version of the Barro-Rietz rare disasters model. Following Gabaix (2012), the disaster probability is assumed to be time-varying. The economy consists of two types of agents: (1) a “rational” agent, who updates his beliefs using Bayes Rule, and (2) a “robust” agent, who updates his beliefs using a pessimistically distorted prior. Following Hansen and Sargent (2008), pessimism is disciplined using detection error probabilities. Disaster risk is assumed to be nontradeable.

The model is calibrated to US data, and focuses on three disaster episodes: (1) The Great Depression of 1929-33, (2) The Financial Crisis of 2008-09, and (3) The Covid Pandemic of 2020. The key contribution of the paper is to show that the model can replicate the observed spike in trading volume that occurs during disasters. Trading produces endogenous low frequency dynamics in the distribution of wealth. The relative wealth of robust agents gradually declines during normal times, but rises sharply during disasters. These results sound a note of caution when interpreting short-run movements in the distribution of wealth.

Chapter 2 examines the market selection hypothesis in a continuous time asset pricing model with jumps. It is shown that the hypothesis is valid when agents have log preferences. The result is robust as it does not depend on whether markets are incomplete. Jumps affect long-run wealth dynamics through a redistribution channel: Disasters lead to large wealth redistribution as agents with heterogeneous beliefs about disasters have different exposures to risky assets. Using tools from ergodic theory, I prove a novel result that generalizes the rationality concept in the existing literature: an agent endowed with the optimal filter will outperform other agents in complete financial markets asymptotically.

Chapter 3, a joint paper with Xiaowen Lei, develops a continuous-time overlapping generations model with rare disasters and agents who learn from their own experiences. Using microdata about household finance in China, we establish that economic disasters such as the Great Leap Forward make investors distrustful of the market. Generations that experience disasters invest a lower fraction of their wealth in risky assets, 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.

Keywords: Trading Volume; Heterogeneous Beliefs; Rare Disasters; Market Selection; Intergenerational Inequality

Dedication

To My Family

Acknowledgements

It is a marathon, not a sprint. This may be a cliché. But for me, getting a PhD in economics is truly a journey. Thanks to Simon Fraser University for allowing me to embark on this journey that was filled with excitement. Over the years, I have received help and advice from many people in the Department of Economics.

Ken Kasa, my senior supervisor, is definitely a role model. It is no exaggeration to say that I am a different person (and a better economist) because of him. It is hard to enumerate the various advantages of learning under the guidance of such a knowledgeable and insightful advisor. His passion for economic research and the ease with which he can uncover the intricate links among different papers in the macroeconomics and finance literature still amaze me to this day. Ken also taught me to not be satisfied with a superficial understanding and to always dig deeper. That is something I will benefit from throughout my career.

I am also lucky enough to learn from other members of my thesis committee. John Knowles asked me regularly about what papers and books I had been reading and was always ready to serve as a sounding board for research ideas. Luba Petersen, in addition to being generous with her time, hired me as an RA and introduced me to the fascinating world of experimental macroeconomics.

I would like to thank all my friends and classmates. Xiaowen Lei was not only a coauthor but also a good friend. She provided constructive feedback on all my papers. I am indebted to Haiyun Kevin Chen, Liang Diao, and Huiqian Louise Song for their camaraderie and support. Garrett Petersen was a wonderful host of board game nights.

Many thanks to all the people who worked hard to ensure that I had a smooth experience in the graduate programs: the graduate chairs, Simon Woodcock, Christoph Luelfesmann, and Alexander Karaivanov; and the staff, Gwen Wild, Kathleen Vieira-Ribeiro, and Lisa Agosti.

Chris Bidner, Zachary Bethune, Lucas Herrenbrueck, Scott Schuh, Xin Thomson Tong and Ramon van Handel were all kind enough to give comments on my papers or answer my questions.

Last but certainly not least, I wish to express my gratitude to my family. Without your tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study.

Table of Contents

List of Tables

List of Figures

Chapter 1

Fear and Trading

1.1 Introduction

On September 29, 2008, the S&P 500 dropped more than 8%. By itself, this was not surprising. On that same day, Congress failed to pass a bank bailout bill, which raised doubts about the stability of the US financial system. In response, risk premia increased and future earnings prospects deteriorated. The canonical Lucas (1978) model can easily explain why such adverse news could trigger a market crash. However, more puzzling is the fact that trading volume spiked on that day as well. It more than doubled. The Lucas model predicts *zero* trading in response to news.

A simple way to account for trading in a Lucas-style model is to introduce heterogeneous prior beliefs. Agents "agree to disagree". Observed trading in response to *public* signals, such as Congressional deliberations, provides strong evidence in support of heterogeneous priors (Kandel and Pearson (1995)). Unfortunately, existing heterogeneous beliefs models suffer from two drawbacks. First, since they concern *priors*, almost by definition they lack discipline, and are therefore difficult to test. Where do these prior differences come from? Second, why doesn't learning cause these prior differences to "merge" (Blackwell and Dubins (1962), Morris (1996))? Shouldn't trading gradually dissipate in the presence of learning?

This paper responds to both these criticisms. It responds to the first by engaging in a disciplined retreat from Bayesian Decision Theory. In particular, I assume some agents adhere to the "robust" decision theory of Hansen and Sargent (2008).[1] A robust agent's prior is *endogenous*. It is the solution of a dynamic zero sum game, and will change in response to changes in the environment. I respond to the second criticism by making the learning environment challenging. Specifically, I place agents into the rare disasters setting of Rietz (1988), Barro (2006), Gabaix (2012), and Wachter (2013), in which agents must learn about the ever changing arrival rate of infrequent disasters. True, if agents had data

[1] Gilboa and Schmeidler (1989), Klibanonoff, Marinacci, and Mukerji (2005), Maccheroni, Marinacci, and Rustichini (2006), and Strzalecki (2011) provide axiomatic foundations for robust decision theory.

going back to Roman times, then perhaps their priors would merge, but using the detection error probabilities of Hansen and Sargent (2008), I show that even with 100 years of data the robust agent has little reason to doubt his prior.[2]

If all we cared about were prices and aggregate quantities, then one might argue that trading volume is irrelevant. Versions of the complete markets models of Lucas (1978) and Barro (2006) are all you need. Even with homogeneous beliefs, one could generate at least *some* portfolio rebalancing trade by introducing forms of preference heterogeneity (see, e.g., Dumas (1989)). However, with complete markets, trading volume is necessarily indeterminant, as it depends on which of an arbitrarily large number of financial market structures are used to complete the market. The more complex and state-contingent the asset, the less it needs to be traded to support the Pareto Optimal allocation (Arrow (1964)).[3]

However, in keeping with much of the recent literature in macroeconomics, this paper is not just interested in prices and aggregate quantities. It is also interested in how belief heterogeneity influences the distribution of wealth (Coury and Sciubba (2012), Kasa and Lei (2018), Lei (2020)). Asset trade is the key conduit by which belief heterogeneity influences the distribution of wealth. Intuitively, one might suspect that agents with distorted beliefs would eventually be driven from the market. Blume and Easley (2006) show that with complete markets this is generally the case. However, with incomplete markets, this need no longer be true (Beker and Chattopadhyay (2010), Cogley, Sargent, and Tsyrennikov (2014)). In fact, wealth dynamics can be reversed. This is important for the results in this paper, since I assume markets are incomplete. There are two independent shocks, a discrete jump process and a continuous Brownian motion. So even with continuous trading, one would need at least three independent assets to complete the market. However, here agents can only trade a (zero net supply) bond and an equity claim. No options or other disaster-contingent assets are allowed. Moreover, for analytic simplicity, I adopt time-additive log preferences, so that income and substitution effects exactly offset each other. I show that the resulting wealth distribution dynamics feature recurrent long-run cycles, in which the relative wealth of the robust agent gradually declines during normal times, but then spikes up following disasters.[4]

[2]In a forward-looking control context, Hansen and Sargent (2008) show that doubts can persist, even with an infinite sample, if agents discount future increments to relative entropy in the same way they discount future utility. In a backward-looking robust filtering context, presumably the same result could be obtained if agents discounted old data.

[3]Interestingly, trading volume has grown in recent decades, despite (or perhaps because of) rapid growth of derivatives and other state-contingent assets. This suggests that preference induced portfolio rebalancing is not the only factor at work behind trading volume.

[4]More generally, Borovicka (2020) shows that long-run survival depends on the ability to separate intertemporal substitution, which determines savings behavior, from risk aversion, which determines portfolio allocation.

Although heterogeneous priors are perhaps the most natural way to generate asset trade, the recent behavioral finance literature typically argues that *overconfidence* is what motivates trade (see, e.g., Statman, Thorley, and Vorkink (2006)). Of course, if everyone is overconfident in exactly the same way, there would be no trade. So this literature also argues that confidence varies across individuals, e.g., men are supposedly more overconfident (Barber and Odean (2001)). Interestingly, this paper instead argues that trade is based on *fear*, not confidence. When some agents fear model misspecification, and in response formulate robust savings and portfolio policies, they will trade with those who are less fearful. Trade only requires a difference of opinion. The sign of this difference is immaterial.

1.1.1 Related Literature

The famous "No Trade Theorem" (Milgrom and Stokey (1982), Tirole (1982)) states that rational investors with common priors will not trade if they start from a Pareto efficient allocation, even in the presence of asymmetric information. One strand of the literature that circumvents the "No Trade Theorem" is the heterogeneous beliefs (differences of opinion) literature. These models rely on the assumption that investors have heterogeneous priors and differ in their interpretations of information. Harrison and Kreps (1978), Scheinkman and Xiong (2003), Basak (2005) and Dumas et al. (2009) focus on the implications of heterogeneous expectations, and how disagreement among investors influences prices. Harris and Raviv (1993), Kandel and Pearson (1995), Cao and Ou-Yang (2008) and Banerjee and Kremer (2010) emphasize the importance of trading volume in differences of opinion models. There is ample empirical evidence in support of the heterogeneous priors assumption. For example, Hong and Stein (2007) report that trading volume spikes right after earnings are announced, and remains high up to a week. This contradicts the prediction of a common priors model: public information should reduce disagreement, not increase it. Another approach that can generate trading is asymmetric information and liquidity shocks. Liquidity shocks ensure that these models are not subject to the "No Trade Theorem" because investors trade for non-informational reasons. A shortcoming of this approach is that the volume of liquidity trades is exogenous.

Many papers study market crashes under the assumption of complete markets. Bates (2008) investigates the importance of heterogeneous attitudes toward crash risk and concludes that the less crash averse agents sell insurance in options markets. Benzoni et al. (2011) explain market crashes without corresponding jumps in fundamentals (Black Monday in 1987) using a continuous time version of the long run risk model. The focus in these papers is on option valuation and price changes. In contrast, this paper focuses on trading volume in an incomplete markets/rare event context. Market completeness supported by disaster insurance tends to pacify stock market fluctuations, as agents "place their bets" in the insurance market. This contradicts the empirical patterns that we observe during disasters.

This paper is most closely related to the rare disasters literature. This literature focuses on resolving traditional asset pricing puzzles. Based on calibrated disaster probabilities and jump distributions from the 20th century, Barro (2006) explains the equity premium puzzle, the risk-free rate puzzle and the excess volatility puzzle by incorporating rare disasters into the Lucas model. Barro and Ursua (2008) provide an exhaustive summary of macroeconomic crises since 1870. One interesting feature is that the distributions of consumption and GDP are very similar during the disaster episodes included in the paper. Dieckman (2011) discusses the interaction between heterogeneous beliefs and rare event risk with incomplete markets. Dieckman (2011) is probably the closest to the present paper, but it centers on asset price changes and the effects of market completion through a disaster insurance. The main difference is that the evolution of beliefs is not a mere input to derive asset allocations in the present paper. The central message in the present paper is the effects of belief revision on market crashes and trading volume. Gabaix (2012) recognizes that time-varying disaster probabilities can provide an explanation for a large set of asset pricing puzzles concerning stocks, bonds, and options. Gourio (2012) emphasizes the importance of time-varying risk premia, and demonstrates that a real business cycle model with disaster risk can match the relations between macroeconomic aggregates and asset prices.

This paper also draws from the literature at the intersection of robust control and asset pricing. Cagetti et al. (2002) build a stochastic growth model with Brownian motion and infrequent jumps in the drift term. Robustness dictates that the decision maker chooses to accumulate a larger stock of capital as a form of precautionary saving. Maenhout (2004) introduces concern for model uncertainty into the Merton model and obtains results that lead to a reduction in the demand for risky assets and an observational equivalence to recursive preferences. Liu et al. (2005) investigate the effects of a concern for model uncertainty in a jump diffusion model and demonstrate the implications for index option smirks. Hansen and Sargent (2010) note that a pessimist thinks that good news will be short-lived, while bad news will be persistent. They show that fear of model misspecification can explain countercyclical risk premia. A novel direction in the robust control literature is time-varying fear. Sbuelz and Trojani (2008) introduce time-varying fear as a time-varying local bound on the size of ambiguity in a recursive multiple-priors utility context and demonstrate that it is useful in reconciling the equity premium puzzle and the risk-free rate puzzle. Drechsler (2013) shows that a model with time-varying fear can match the variance premium and the volatility skew associated with index options.

The remainder of the paper is organized as follows. The next section provides motivation by presenting some stylized facts on market crashes, trading volume, and belief heterogeneity. Section 3 presents the model. It shows how the martingale method can be extended to an incomplete markets setting. Explicit analytical expressions are derived for market prices, portfolios, and the distribution of wealth. Section 4 computes detection error probabilities using Monte Carlo simulation methods. For the model's benchmark parameter values, the robust agent's detection error probability is in excess of 15%, even with 100 years of data. Section 5 calibrates the model to US data, and provides plots of prices, trading, and wealth dynamics. It shows that the model is able to replicate observed trading volume during historical disaster episodes. Section 6 contains a variety of robustness checks, while Section 7 contains a brief conclusion. A technical Appendix provides proofs and derivations.

1.2 Empirical Facts About Trading Volume and Market Crashes

Two general patterns documented in the trading volume literature are relevant to understanding market crashes: (1) The correlation between absolute price changes and trading volume is positive in equity markets, and (2) On average, weekly turnover on the NYSE is about 2%, but spikes sharply during crashes (Lo and Wang (2010)).

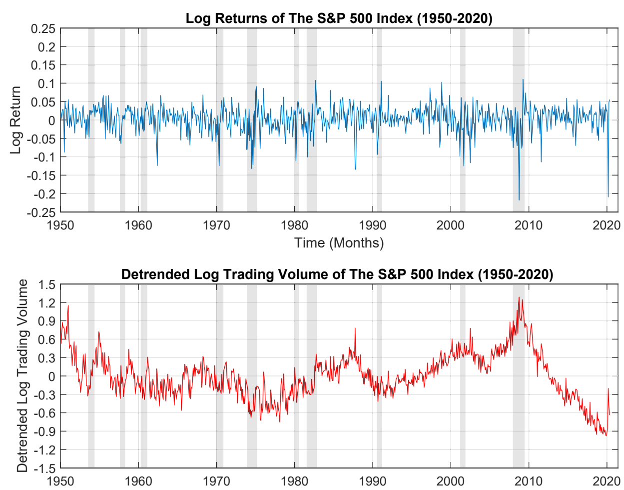

Figure 1.1: Log Return of The S&P 500 and Its Monthly Trading Volume (1950M1-2020M5 Detrended Natural Logs). Shaded areas are NBER recessions.

Figure 1 shows that crashes in the S&P 500 coincide with large trading volume.[5] During both the Financial Crisis of 2008 and the Covid Crisis of 2020, the market crashed by about 20%. At the same time, trading volume spiked by about 60-80%.[6] Cross section data on the S&P 500 index constituents allows us to have a better picture of the crash episodes in 2008. On September 29, the S&P 500 index dropped by 8.79%. The prices of 499 out of the 500 constituents decreased on that day. The single day price decrease was as high as

[5] The log of trading volume appears to have a linear trend. A linear trend is therefore removed from the series.

[6] The Covid spike is harder to see in the Figure, as it followed a period of declining trading volume, and occurs at the boundary of the plot.

80% (see Figure A.1 in the Appendix). The constituents' daily turnover rates were mostly higher than 2% (see Figure A.2 in the Appendix).

Since S&P 500 stocks are obviously large cap stocks, using raw trading volume to capture market activity is perhaps misleading because the total number of shares outstanding has increased over time. Turnover rates address this concern. I obtain data on trading volume and total number of shares outstanding for the S&P 500 index constituents from the Center for Research in Security Prices (CRSP). I calculate the turnover rate as

$$\text{Turnover} = \frac{\text{total trading volume of all the constituents}}{\text{total number of shares outstanding of all the constituents}}$$

Figure 2 presents monthly turnover for S&P 500 stocks from 1950 to 2020. Before 2000, monthly turnover is mostly below 10%. After 2000, monthly turnover starts to increase at an increasing rate until it reaches 50% during the Great Recession.[7] I also fit a quadratic trend to the monthly turnover series. The detrended monthly turnover during the Great Recession can be as high as 30%. Although turnover decreased after the Great Recession, the COVID-19 market crash of March 2020 was accompanied by heightened trading. Since the turnover of S&P 500 stocks also spikes during the Great Recession, the same pattern emerges whether I choose to measure market activity in terms of trading volume or turnover.[8]

[7]Chordia, Roll, and Subrahmanyam (2011) document that monthly turnover for S&P 500 stocks start from below 6% per month in 1993 and reach 40% per month in 2008. They calculate the value-weighted monthly turnover for S&P 500 stocks. The lower numbers they obtain seem to reflect the effects of market capitalization.

[8]In the model, the stock is an asset with unit supply. Hence, turnover and trading volume are the same.

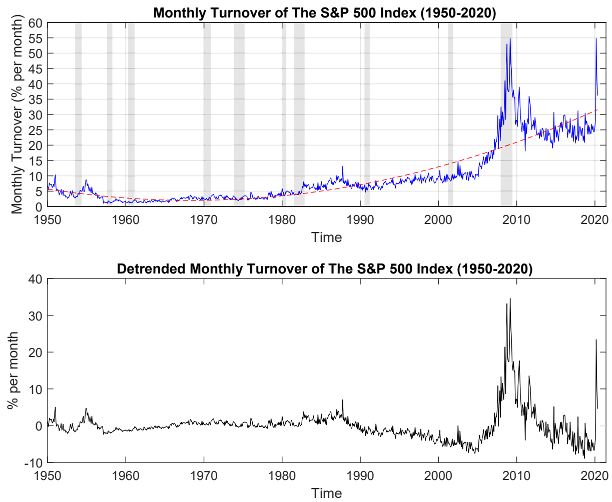

Figure 1.2: Monthly Turnover of the S&P 500 Index (1950M1-2020M5) Source: CRSP. The dashed line is a quadratic trend.

Table 1 reports correlations between monthly log price changes and trading. The full sample correlations between trading and absolute price changes are indeed positive. The correlations between trading and absolute price changes are much higher during the Great Recession. This confirms that trading is very active when the S&P 500 index decreases. As emphasized by Scheinkman and Xiong (2003), a common feature of bubble episodes is that they also tend to coincide with frenzied trading. Hence, I also report the correlations from April 1991 to February 2000, since the dot-com era is often described as a stock market bubble. The correlations between absolute log price changes and trading volume during this boom period are above 0.3.

	Absolute Log Price Changes	Log Price Changes
Detrended Log Volume (full sample)	0.1953	0.0487
Detrended Turnover (full sample)	0.3170	-0.1309
Log Volume (2007.12-2009.06)	0.6354	-0.3418
Turnover (2007.12-2009.06)	0.6583	-0.3459
Log Volume (1991.04-2000.02)	0.3813	0.1319
Turnover (1991.04-2000.02)	0.3407	0.1486

Table 1.1: Correlations Between Log Price Changes and Trading (1950M1-2020M5)

The key mechanism driving trading in the model is heterogeneous beliefs. Unfortunately, beliefs are not directly observable. However, survey data is at least suggestive. Forecast disagreement inferred from the Survey of Professional Forecasters and the Livingston Survey exhibit some interesting patterns. Figure 3 reports the interquartile range of next quarter

GDP forecasts in the Survey of Professional Forecasters.[9] Despite the fact that disagreement was higher in the 1970s and 1980s, a general pattern is that disagreement tends to increase dramatically during recessions. The disagreement spike during the second quarter of 2020 is spectacular: the interquartile range of forecasts is 19.33% in annualized terms. The fact that this disagreement spike dwarfs all the previous disagreement spikes is a testament to the elevated uncertainty caused by the COVID-19 pandemic. The Livingston Survey provides valuable information regarding economists' forecasts about the growth rate of the S&P 500 index (see Figure 4). The interquartile range of the forecasts reach 20.9% (in annualized terms) during the COVID-19 pandemic. This is evidence that both professional forecasters and economists hold wide-ranging views about the prospects of the economy and the stock market during recessions, even though they have access to the same data and models for forecasting purposes.

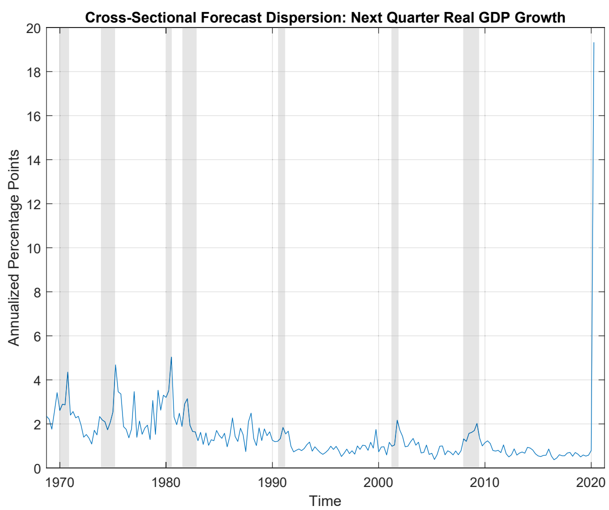

Figure 1.3: Interquartile Range of Next Quarter GDP Forecasts in the Survey of Professional Forecasters (1968Q4-2020Q2). Source: Survey of Professional Forecasters, Federal Reserve Bank of Philadelphia.

[9]Some authors use the standard deviation to measure forecast dispersion. An advantage of the interquartile range is that it limits the effects of outliers. This is important as reporting errors are inevitable in surveys.

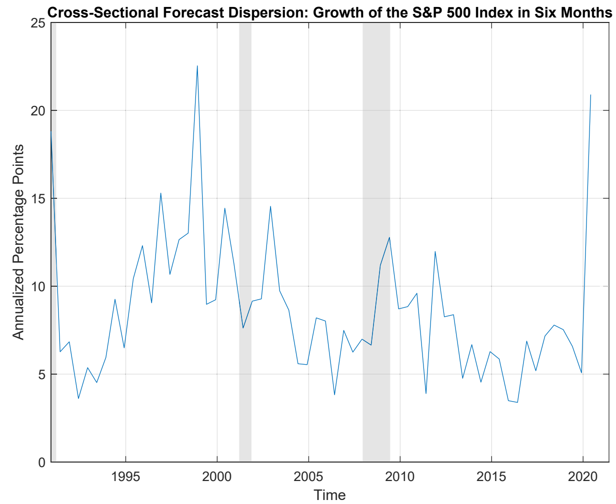

Figure 1.4: Interquartile Range of the Forecasts for Growth of the S&P 500 Index (1990-2020). Source: Livingston Survey, Federal Reserve Bank of Philadelphia. Shaded areas are NBER recessions.

The empirical patterns suggest that economic disasters, heightened trading volume, and disagreement spikes tend to occur simultaneously.

1.3 The Model

There are four main ingredients in the model: (1) Aggregate consumption and dividends are exogenous, and follow jump diffusion processes; (2) The jump intensity of the common jump component is governed by a continuous time Markov chain; (3) Robust filtering and control determine how some investors' beliefs about the jump intensity change; (4) Markets are incomplete, yet log utility produces closed-form solutions to the agents' portfolio choice problems. Equilibrium trading dynamics can be inferred from these portfolio policies.

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}(t)\}_{t\geq 0}, P)$. Let $\mathcal{F}^W(t)$ be the σ-algebra generated by a Brownian motion $W(s)$; $s \leq t$ and let $\mathcal{F}^N(t)$ be the σ-algebra generated by a Poisson jump process $N(s)$; $s \leq t$. $W(t)$ and $N(t)$ are assumed to be independent. More precisely, the σ-algebra $\mathcal{F}(t)$ is defined as $\mathcal{F}(t) = \mathcal{F}^W(t) \times \mathcal{F}^N(t)$. I consider a pure-exchange economy with a finite horizon $[0, T]$. There are two types of agents. Agent 1 (Bayesian) has no doubts. Agent 2 (robust) has an endogenous time-varying fear of model misspecification.

The dynamics of consumption and dividend follow jump diffusion processes.

$$\frac{dC(t)}{C(t-)} = \mu dt + \sigma dW(t) + j_C dN(t) \tag{1.1}$$

$$D(t) = C(t)^{\phi} \tag{1.2}$$

Using Ito's lemma, we obtain

$$\frac{dD(t)}{D(t-)} = \mu_D dt + \sigma_D dW(t) + j_D dN(t) \tag{1.3}$$

The jump sizes are denoted by $j_C = exp(Z_t) - 1$ and $j_D = exp(\phi Z_t) - 1$, where Z_t is a random variable that determines the jump sizes. $\mu_D = \phi\mu + \frac{1}{2}\phi(\phi - 1)\sigma^2$ and $\sigma_D = \phi\sigma$. The jump sizes are restricted to be in the interval $(-1, 0)$ to ensure that the consumption and dividend processes remain positive. The agents have equivalent probability measures and commonly observe the aggregate dividend.[10] They only have incomplete information about the potential jump intensities of the Poisson process. To simplify the analysis, I assume the agents know the true values of μ, σ, j_C, ϕ and j_D. The jump sizes j_C and j_D are assumed to be constant. The parameter ϕ determines the relationship between consumption and dividends in the economy. The separation of consumption and dividends is empirically motivated. Longstaff and Piazzesi (2004) report that dividends are 10 times more volatile than consumption during the post war period. More strikingly, aggregate consumption dropped by 10% while aggregate earnings fell by 103% during the early stages of the Great Depression.

In order to highlight the importance of business cycles, the jump intensity, $\lambda(t)$, is assumed to be governed by a n-state Markov chain in continuous time, with generator matrix $\mathbb{Q}$. Without loss of generality, assume the n states are $\Lambda_1 < \Lambda_2 < ... < \Lambda_n$. Let $P_{ij}(t) = Pr(\lambda(t+s) = \Lambda_j | \lambda(s) = \Lambda_i)$.[11] $\mathbb{P}(t)$ is the transition probability matrix with $P_{ij}(t)$ as its (i, j)-th entry and the time t derivative of the transition probability matrix is $\mathbb{P}'(t) = \mathbb{P}(t)\mathbb{Q}$. The jumps are rare events by definition. Agents in the economy do not have access to enough data to obtain good estimates of the statistical properties of the rare events. The robust agent has 2 types of concerns about model misspecification: (1) Misspecification about the potential states (jump intensities). (2) Misspecification about the transition rates between different states (the entries in the generator matrix $\mathbb{Q}$).

PROPOSITION 1.3.1. *The filtering equation is*

$$\hat{\lambda}(t) = E[\lambda(t)|\mathcal{F}(t)] = \sum_{k=1}^{n} \Lambda_k \hat{p}_k(t) \tag{1.4}$$

[10] The assumption of finite horizon is important if we want to stop merging of opinions. Blackwell and Dubins (1962) prove that beliefs converge almost surely when two Bayesians have equivalent probability measures.

[11] I assume that the Markov chain is time homogeneous: the probability $Pr(\lambda(t+s) = \Lambda_j|\lambda(s) = \Lambda_i)$ does not depend on s.

$$d\hat{p}_k(t) = \sum_j q_{jk}\hat{p}_j(t)dt + \hat{p}_k(t)\Big(\frac{\Lambda_k - \hat{\lambda}(t)}{\hat{\lambda}(t)}\Big)d\hat{\eta}(t) \tag{1.5}$$

$$\hat{\eta}(t) = N(t) - \int_0^t \hat{\lambda}(s)ds \tag{1.6}$$

where $\hat{p}_k(t)$ denotes the conditional probability that $\lambda(t) = \Lambda_k$.

The filtering equation in Proposition 1.3.1 is intuitive. Consider a potential state Λ_k that is higher than the current estimate of jump intensity $\hat{\lambda}(t)$. If no jumps are observed, then the second term in (5) becomes negative and the growth of the conditional probability of being in Λ_k decreases as well. If a jump is observed, then the conditional probability of Λ_k increases. The filtering equation implies that investors' beliefs will decrease smoothly as long as they do not observe jumps.

The following two definitions allow us to measure the set of possible model misspecifications in a jump diffusion context.

DEFINITION 1.3.1. *For two probability measures Q and P in a given measure space such that $Q \in \mathcal{Q}$ (Q is absolutely continuous with respect to P), the relative entropy of Q with respect to P is defined as*

$$H = \int log(dQ/dP)dQ \tag{1.7}$$

where dQ/dP is the Radon-Nikodym derivative of Q with respect to P. To simplify notation, let L=dQ/dP. The change in the relative entropy from t to $t + \Delta t$ is

$$H(t, t + \Delta t) = E_t^L[logL(t + \Delta t)] - logL(t) \tag{1.8}$$

The instantaneous growth rate of the relative entropy [12] is

$$R(L(t)) = \lim_{\Delta t \to 0} \frac{H(t, t + \Delta t)}{\Delta t} \tag{1.9}$$

DEFINITION 1.3.2. *Let $\theta(t)$ be an adapted process and $\xi(t)$ a predictable process. t_n is a sequence of jump times. In a jump-diffusion context, the change of measure is*

$$L_W(t) = exp\Big(- \int_0^t \theta(s)dW(s) - \frac{1}{2}\int_0^t \theta(s)^2 ds\Big) \tag{1.10}$$

$$L_N(t) = exp\Big(- \int_0^t \lambda(s)\xi(s)ds\Big) \prod_{n, t_n \le t} (1 + \xi(t_n)) \tag{1.11}$$

$$L(t) = L_W(t)L_N(t) \tag{1.12}$$

[12] Anderson et al. (2003) define the time derivative of relative entropy as a statistical measure of discrepancy between models.

The filtering problem for the robust agent is [13]

$$\min_{\tilde{\lambda}(t)} \max_{Q\in\mathcal{Q}} E_Q[(\tilde{\lambda}(t)-\lambda(t))^2|\mathcal{F}(t)] \tag{1.13}$$

subject to the time-varying fear constraint

$$R(L(t)) = \lim_{\Delta t\to 0} \frac{H(t,t+\Delta t)}{\Delta t} \leq \eta_t \tag{1.14}$$

where η_t is a time-varying bound on the instantaneous growth rate of the relative entropy. The robust filtering problem dictates that the robust agent finds the best estimate in mean square of the jump intensity under a worst case probability measure induced by the time-varying fear constraint.

If the alternative models are Markov chains as well, the robust agent will be concerned about misspecifications of the bad states and their transition rates.[14] He will consider different values of the states ($\tilde{\Lambda}_k \neq \Lambda_k$) or different transition rates ($\tilde{q}_{jk} \neq q_{jk}$). Time-varying fear puts a constraint on the relative entropy between the reference model and the set of alternative models. The entropy constraint will limit the agent's choices in terms of potential values of $\tilde{\Lambda}_k$ and $\tilde{q}_{jk}$ under consideration.

PROPOSITION 1.3.2. *The robust agent's beliefs evolve according to the robust filter*

$$\tilde{\lambda}(t) = E_Q[\lambda(t)|\mathcal{F}(t)] = \sum_{k=1}^{n} \tilde{\Lambda}_k \tilde{p}_k(t) \tag{1.15}$$

$$d\tilde{p}_k(t) = \sum_j \tilde{q}_{jk}\tilde{p}_j(t)dt + \tilde{p}_k(t)\left(\frac{\tilde{\Lambda}_k - \tilde{\lambda}(t)}{\tilde{\lambda}(t)}\right)d\tilde{\eta}(t) \tag{1.16}$$

$$\tilde{\eta}(t) = N(t) - \int_0^t \tilde{\lambda}(s)ds \tag{1.17}$$

where $\tilde{p}_k(t)$ denotes the conditional probability that $\lambda(t) = \tilde{\Lambda}_k$.

Robust filtering amplifies belief heterogeneity about rare events after a jump is observed. The robust agent experiences a larger revision of his expected jump intensity than the Bayesian agent due to concern for model uncertainty.

[13] Hansen and Sargent (2001) refer to this as 'constraint preferences', as opposed to 'penalty preferences'.

[14] This is an example of what Hansen and Sargent (2019) call 'structured uncertainty'.

COROLLARY 1.3.1. *When $n = 2$, the condition under which the revision is larger for the robust agent is*

$$\sum_j \tilde{q}_{j2}\tilde{p}_j(t)dt - \tilde{p}_2(t)(\tilde{\Lambda}_2 - \tilde{\lambda}(t))dt + \tilde{p}_2(t)\left(\frac{\tilde{\Lambda}_2 - \tilde{\lambda}(t)}{\tilde{\lambda}(t)}\right) >$$

$$\sum_j q_{j2}\hat{p}_j(t)dt - \hat{p}_2(t)(\Lambda_2 - \hat{\lambda}(t))dt + \hat{p}_2(t)\left(\frac{\Lambda_2 - \hat{\lambda}(t)}{\hat{\lambda}(t)}\right) \quad (1.18)$$

When a jump happens, if the worst case jump intensity is well above the current estimate of the jump intensity, the terms $\tilde{p}_2(t)\left(\frac{\tilde{\Lambda}_2 - \tilde{\lambda}(t)}{\tilde{\lambda}(t)}\right)$ and $\hat{p}_2(t)\left(\frac{\Lambda_2 - \hat{\lambda}(t)}{\hat{\lambda}(t)}\right)$ will dominate the other terms.

Since multiple equivalent martingale measures exist in incomplete markets and the agents have heterogeneous beliefs, we need to derive their agent specific stochastic discount factors. Let $r(t)$ be the equilibrium interest rate. The agent specific stochastic discount factor is $(i = 1, 2)$

$$M^i(t) = \frac{L^i(t)}{exp(\int_0^t r(s)ds)} \quad (1.19)$$

where $L^i(t)$ is agent i's change of measure from Definition 1.3.2. The stochastic discount factor has the following differential form:

$$dM^i(t) = -M^i(t-)\left[r(t)dt + \theta^i(t)dW^i(t) - (\lambda^i(t) - \lambda^i_{RN}(t))dt - \left(\frac{\lambda^i_{RN}(t)}{\lambda^i(t)} - 1\right)dN^i(t)\right] \quad (1.20)$$

where $\lambda^i(t)$ is agent i's best estimate of the jump intensity from his filter. $\theta^i(t)$ is agent i's market price of risk and $\lambda^i_{RN}(t) = \lambda^i(t)(1 + \xi^i(t))$ is the market price of rare event risk (the risk-neutral jump intensity).

In our incomplete markets economy, agents have access to 2 types of assets: a bond $B(t)$ with 0 net supply and a risky stock $S(t)$ with net supply of 1. Their price dynamics are as follows

$$dB(t) = B(t)r(t)dt \quad (1.21)$$

$$dS(t) + D(t)dt = S(t-)[\mu_S(t)dt + \sigma_S dW(t) + j_S dN(t)] = S(t-)[\mu^i_S(t)dt + \sigma_S dW^i(t) + j^i_S dN^i(t)] \quad (1.22)$$

where j_S is the jump size of the stock price process.

Each investor's problem is[15]

[15] Most papers in the martingale approach literature abstract from time preference. Time preference will influence the equilibrium interest rate. If we assume the two agents have the same time preference, then time preference will have symmetric effects on the two agents. Time preference seems unlikely to be a major factor that drives trading volume.

$$\max_{c_i} E^i \left[\int_0^T u(c_i(t))dt \right] \tag{1.23}$$

subject to

$$E^i \left[\int_0^T M^i(t) c_i(t) dt \right] \leq w_i(0) \tag{1.24}$$

where $w_i(0)$ is agent i's initial endowment.

The essence of the martingale approach is that it allows us to transform a dynamic problem into a static one. It delivers the following optimality condition:

$$u'(c_i(t)) = y_i M^i(t) \tag{1.25}$$

where y_i is the Lagrange multiplier associated with agent i's lifetime budget constraint. An equivalent condition is

$$c_i(t) = (u')^{-1}(y_i M^i(t)) \tag{1.26}$$

Given the assumption of time-additive log utility:

$$c_i(t) = (u')^{-1}(y_i M^i(t)) = \frac{1}{y_i M^i(t)} \tag{1.27}$$

$$y_i = \frac{T}{M^i(0) w_i(0)} \tag{1.28}$$

$$w_i(t) = \frac{E^i[\int_t^T M^i(s) c_i(s) ds]}{M^i(t)} \tag{1.29}$$

$$c_i(t) = \frac{w_i(t)}{T - t} \tag{1.30}$$

By Ito's lemma for jump diffusion

$$dw_i(t) = [...]dt + w_i(t-)\theta^i(t) dW^i(t) + w_i(t-)\left(\frac{\lambda^i(t)}{\lambda^i_{RN}(t)} - 1 \right) dN^i(t) \tag{1.31}$$

The wealth process $w_i(t)$ is

$$dw_i(t) = (1 - \pi_i(t)) w_i(t-) \frac{dB(t)}{B(t)} + \pi_i(t) w_i(t-) \frac{dS(t) + D(t)dt}{S(t)} - c_i(t)dt \tag{1.32}$$

where $\pi_i(t)$ is the share of wealth allocated to the stock market by agent i at time t.

$$dw_i(t) = (1 - \pi_i(t)) w_i(t-) r(t) dt + \pi_i(t) w_i(t-) [\mu^i_S(t) dt + \sigma_S dW^i(t) + j^i_S dN^i(t)] - c_i(t) dt \tag{1.33}$$

DEFINITION 1.3.3. *An equilibrium is a collection of asset prices {B, S} and a set of consumption and portfolio choices $\{c_i, \pi_i\}$ such that*

1. Each agent's consumption is optimal: the consumption plan c_i maximizes lifetime utility subject to the budget constraint and is financed by the portfolio choice π_i.

2. Goods market and financial markets clear at all times. The market clearing conditions are

$$c_1(t) + c_2(t) = D(t) \tag{1.34}$$

$$w_1(t) + w_2(t) = S(t) \tag{1.35}$$

$$\pi_1(t)w_1(t) + \pi_2(t)w_2(t) = S(t) \tag{1.36}$$

The bond and stock prices are endogenous, and Proposition 1.3.3 characterizes each investor's perceived stock price process.

PROPOSITION 1.3.3. *The stock price process is given by*

$$S(t) = D(t)(T - t) \tag{1.37}$$

$$\mu_S^i(t) = \mu_D \tag{1.38}$$

$$\sigma_S = \phi\sigma \tag{1.39}$$

$$j_S = j_D \tag{1.40}$$

The stock price process has the same drift term, volatility term, and jump size as the dividend process.

With incomplete markets there exists multiple equivalent martingale measures, so it can be difficult to pin down a suitable stochastic discount factor for each agent. Using the martingale approach, we can characterize the equilibrium with incomplete markets by searching for the "least favorable fictitious completion".[16] When demands for non-marketed assets are zero, the solutions under complete markets and incomplete markets coincide (see He and Pearson (1991), Cvitanic and Karatzas (1992), Basak and Croitoru (2000), and Dieckmann (2011)). The incomplete market setup is equivalent to a portfolio constraint that limits the demand (and the supply) for the non-marketed assets to be zero. In the absence of the portfolio constraint, the investors could have achieved higher utility through access to more assets. There are two sources of uncertainty, a Brownian motion and a Poisson jump process. An ideal setup with complete markets should allow the agents to hedge the risks in any way they want. More precisely, the introduction of a "disaster insurance" will

[16] In Karatzas et al. (1991), the fictitious completion is implemented by introducing additional stocks so that the number of random sources and the number of risky assets match. Then the incomplete market equilibrium is obtained by searching for a complete market equilibrium in which the investors choose not to invest in the fictitious stocks.

allow the agents to fully hedge the risks. Incomplete markets dictate that the market for the disaster insurance is nonexistent and take away the agents' ability to hedge risks. Therefore incomplete markets decrease the maximum attainable utility. As not buying or selling the disaster insurance is always an option, any equilibrium allocations that involve trading of the disaster insurance should lead to higher utility. Therefore the incomplete markets equilibrium indeed corresponds to the minimum of all equilibrium allocations in terms of maximum attainable utility.

I follow Dieckmann (2011) and solve a minimax problem that arises from the "least favorable fictitious completion":

$$\min_{\theta^i,\lambda^i_{RN}} \left[\max_{c_i} E^i\left[\int_0^T u(c_i(t))dt\right] \quad \text{subject to} \quad E^i\left[\int_0^T M^i(t)c_i(t)dt\right] \le w_i(0)\right] \tag{1.41}$$

The search for the "least favorable fictitious completion" requires finding a stochastic discount factor that minimizes the maximum attainable utility across all potential stochastic discount factors (the outer minimization), given that the agents maximize utility through consumption choices (the inner maximization). The solution to this minimax problem determines the agents' portfolio holdings.

PROPOSITION 1.3.4. *The portfolio holdings are* [17]

$$\pi_i(t) = \frac{\theta^i(t)}{\sigma_S} = \frac{1}{j^i_S}\left(\frac{\lambda^i(t)}{\lambda^i_{RN}(t)} - 1\right) \tag{1.42}$$

$$\theta^i(t) = -\frac{1}{2j^i_S\sigma_S}\left(j^i_S(r(t)-\mu^i_S(t)) + {\sigma_S}^2 - \sqrt{4{j^i_S}^2\lambda^i(t){\sigma_S}^2 + (-j^i_S(r(t)-\mu^i_S(t)) + {\sigma_S}^2)^2}\right) \tag{1.43}$$

$$\lambda^i_{RN}(t) = \frac{1}{2{j^i_S}^2}\left(j^i_S(r(t)-\mu^i_S(t)) - {\sigma_S}^2 + \sqrt{4{j^i_S}^2\lambda^i(t){\sigma_S}^2 + (-j^i_S(r(t)-\mu^i_S(t)) + {\sigma_S}^2)^2}\right) \tag{1.44}$$

In order to derive the market equilibrium, we use the martingale approach by forming a representative agent with stochastic weights:

$$U(c,k) = \max_{c_1+c_2=c} u(c_1) + ku(c_2) \tag{1.45}$$

[17] If we let $\lambda^i(t) = 0$, then the solution becomes $\theta^i(t) = \frac{\mu^i_S(t)-r(t)}{\sigma_S}$ and $\pi_i(t) = \frac{\mu^i_S(t)-r(t)}{\sigma^2_S}$. We obtain the Merton solution if we disable the jump component completely.

where $k(t)$ is the wealth ratio between agent 2 and agent 1. From the first order condition of the Pareto problem, we have

$$k(t) = \frac{u'(c_1(t))}{u'(c_2(t))} = \frac{y_1 M^1(t)}{y_2 M^2(t)} = \frac{w_2(t)}{w_1(t)} \tag{1.46}$$

PROPOSITION 1.3.5. *The optimal consumption is given by*

$$c_1(t) = \frac{D(t)}{1+k(t)} \tag{1.47}$$

$$c_2(t) = \frac{D(t)k(t)}{1+k(t)} \tag{1.48}$$

The wealth ratio is governed by

$$\frac{dk(t)}{k(t-)} = [\theta^2(t)^2 - \theta^1(t)\theta^2(t) + (\lambda^1(t) - \lambda^1_{RN}(t)) - (\lambda^2(t) - \lambda^2_{RN}(t))]dt + [\theta^2(t) - \theta^1(t)]dW(t) + \left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)} - 1\right)dN^i(t) \tag{1.49}$$

The equilibrium interest rate is

$$r(t) = \mu^i_S(t) - \sigma_S\theta^i(t) + j^i_S\lambda^i_{RN}(t) \tag{1.50}$$

$$\sigma_S - \frac{1}{1+k(t)}\theta^1(t) = \frac{k(t)}{1+k(t)}\theta^2(t) \tag{1.51}$$

$$\frac{\lambda^1(t)}{\lambda^1_{RN}(t)}\frac{1}{1+k(t)} + \frac{\lambda^2(t)}{\lambda^2_{RN}(t)}\frac{k(t)}{1+k(t)} = j_S + 1 \tag{1.52}$$

In principle, three nonredundant securities will complete the market as there are two random sources. For example, introducing an index option in addition to the bond and the stock will complete the market. But an index option may not be a good choice in terms of tractability. We consider a security that delivers tractability and fulfills the function of disaster insurance at the same time. This security is a disaster insurance in the sense that its payoff only depends on the jump component. Its price $I(t)$ is:

$$dI(t) = I(t-)[\mu_I(t)dt + j^i_I dN^i(t)] \tag{1.53}$$

where j^i_I is the jump size of the disaster insurance. The wealth process $w_i(t)$ is

$$dw_i(t) = (1-\pi_i(t)-\pi^I_i(t))w_i(t-)\frac{dB(t)}{B(t)} + \pi_i(t)w_i(t-)\frac{dS(t)+D(t)dt}{S(t)} + \pi^I_i(t)w_i(t-)\frac{dI(t)}{I(t)} - c_i(t)dt \tag{1.54}$$

where $\pi^I_i(t)$ is the wealth share allocated to the disaster insurance by agent i.

PROPOSITION 1.3.6. *Under complete markets, the portfolio holdings are*

$$\theta^i(t) = \sigma_S \tag{1.55}$$

$$\pi_i(t) = \frac{\theta^i(t)}{\sigma_S} = 1 \tag{1.56}$$

$$\pi_i^I(t) = \frac{1}{j_I}\Big(\frac{\lambda^i(t)}{\lambda_{RN}^i(t)} - 1\Big) - \frac{j_S\theta^i(t)}{j_I\sigma_S} \tag{1.57}$$

$$\lambda_{RN}(t) = \frac{1}{1+k(t)}\frac{\lambda^1(t)}{1+j_S} + \frac{k(t)}{1+k(t)}\frac{\lambda^2(t)}{1+j_S} \tag{1.58}$$

$$r(t) = \mu_S(t) - {\sigma_S}^2 + j_S\lambda_{RN}(t) \tag{1.59}$$

The wealth ratio is governed by

$$\frac{dk(t)}{k(t-)} = [\lambda^1(t) - \lambda^2(t)]dt + \Big(\frac{\lambda^2(t)}{\lambda^1(t)} - 1\Big)dN^i(t) \tag{1.60}$$

COROLLARY 1.3.2. *Under complete markets, the wealth ratio at time t is*

$$k(t) = k(0)exp\Bigg(\int_0^t(\lambda^1(s) - \lambda^2(s))ds + \int_0^t log\Big(\frac{\lambda^2(s)}{\lambda^1(s)}\Big)dN^i(s)\Bigg) \tag{1.61}$$

Under incomplete markets, the wealth ratio at time t is

$$k(t) = k(0)exp\Bigg(\int_0^t(\theta^2(s)^2 - \theta^1(s)\theta^2(s) + (\lambda^1(s) - \lambda_{RN}^1(s)) - (\lambda^2(s) - \lambda_{RN}^2(s)) - \frac{1}{2}(\theta^2(s) - \theta^1(s))^2)ds$$
$$+ \int_0^t(\theta^2(s) - \theta^1(s))dW(s) + \int_0^t log\Big(\frac{\lambda_{RN}^1(s)\lambda^2(s)}{\lambda^1(s)\lambda_{RN}^2(s)}\Big)dN^i(s)\Bigg) \tag{1.62}$$

and the expectation of the wealth ratio is

$$E^i[k(t)] = k(0)exp\Bigg(\int_0^t(\theta^2(s)^2 - \theta^1(s)\theta^2(s) + (\lambda^1(s) - \lambda_{RN}^1(s)) - (\lambda^2(s) - \lambda_{RN}^2(s)))ds$$
$$+ \int_0^t \lambda^i(s)\Big(\frac{\lambda_{RN}^1(s)\lambda^2(s)}{\lambda^1(s)\lambda_{RN}^2(s)} - 1\Big)ds\Bigg) \tag{1.63}$$

In complete markets, the wealth share allocated to the stock is always one. Heterogeneous beliefs are completely inactive in the determination of stock market trading volume. Bond holdings and disaster insurance holdings are mirror images of each other. When we compare the two wealth ratio equations, we notice that the main difference is that the wealth ratios are not affected by the Brownian motion under the assumption of complete markets. But the drift term dictates that as long as no jumps happen the wealth ratio is decreasing in complete markets. The robust agent's consumption share has a tendency to decrease because the disaster insurance will not pay off as frequently as expected. The Bayesian agent receives higher returns and accumulates wealth at a higher rate in the absence of jumps.

1.4 Detection Error Probabilities

Suppose we have two models, A and B, with equal prior probabilities. Model A is the reference model and model B is the worst case model. Detection error probabilities enable us to constrain the robust agent's concern for model misspecification in such a way that he only considers empirically plausible alternative models. This prevents him from being unduly pessimistic.

Specifically, we suppose the robust agent behaves as a trained econometrician. Competing models are compared using likelihood ratio tests. Defining L_A as the likelihood under the reference model and L_B as the likelihood under the worst-case model, the log-likelihood ratio is

$$\ell = logL_A - logL_B$$

We are interested in the following error probabilities:

$$P_A = Prob(\ell < 0|A)$$

$$P_B = Prob(\ell > 0|B)$$

The detection error probability (DEP) is

$$DEP = \frac{1}{2}(P_A + P_B)$$

In practice, P_A and P_B can be approximated using Monte Carlo simulations.

1.4.1 Detection Error Probabilities: Jump Intensities

The change of measure is

$$L_W^A(t) = exp\Big(-\int_0^t \theta(s)dW^A(s) - \frac{1}{2}\int_0^t \theta(s)^2 ds\Big) \tag{1.64}$$

$$L_N^A(t) = exp\Big(\int_0^t (\lambda^A(s) - \lambda^B(s))ds\Big) \prod^{N(t)} \Big(\frac{\lambda^B(s)}{\lambda^A(s)}\Big) \tag{1.65}$$

$$L^A(t) = L_W^A(t)L_N^A(t) \tag{1.66}$$

Since we are interested in the misspecifications of the worst state in the continuous time Markov chain, we can focus on the case where $\theta(s)$ and $\frac{\lambda^B(s)}{\lambda^A(s)}$ are constants. The error probabilities P_A and P_B can be expressed as

$$P_A = Pr(-\theta W^A(t) - \frac{1}{2}\theta^2 t - (\lambda^A - \lambda^B)t + N(t)log(\frac{\lambda^B}{\lambda^A}) > 1) \tag{1.67}$$

$$P_B = Pr(\theta W^A(t) + \frac{1}{2}\theta^2 t + (\lambda^A - \lambda^B)t - N(t)log(\frac{\lambda^B}{\lambda^A}) > 1) \tag{1.68}$$

If agents are not concerned about the misspecifications of the drift term in the consumption process, they agree about the drift term, then $\theta = 0$ in the above change of measure.

The error probabilities simplify to

$$P_A = Pr(N(t)log(\frac{\lambda^B}{\lambda^A}) - (\lambda^A - \lambda^B)t > 1) = \sum_{n=0}^{\infty} \frac{(\lambda^A t)^n}{n!} e^{-\lambda^A t} Pr(nlog(\frac{\lambda^B}{\lambda^A}) - (\lambda^A - \lambda^B)t > 1) \tag{1.69}$$

$$P_B = Pr((\lambda^A - \lambda^B)t - N(t)log(\frac{\lambda^B}{\lambda^A}) > 1) = \sum_{n=0}^{\infty} \frac{(\lambda^B t)^n}{n!} e^{-\lambda^B t} Pr((\lambda^A - \lambda^B)t - nlog(\frac{\lambda^B}{\lambda^A}) > 1) \tag{1.70}$$

Figure 5 depicts the results when we vary the potential jump intensity associated with the worst case model for different sample lengths. We can see that the DEP is still around 0.15 for a potential jump intensity of 0.2 with 100 years of data.

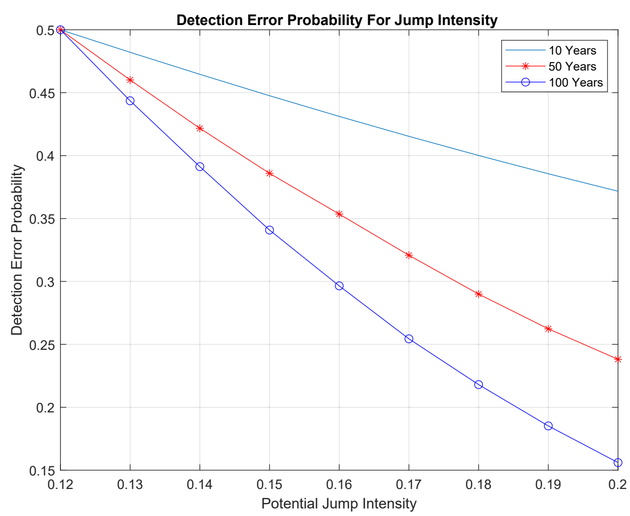

Figure 1.5: This figure plots detection error probabilities for different potential jump intensities in the bad state. The reference model has a jump intensity of 0.12 in the bad state.

1.4.2 Detection Error Probabilities: Transition Probabilities

The probability of observing a particular sample of data $g^t = \{g_s\}_{s=1}^t$ is proportional to the product of binomial densities

$$Pr(g^t|P_{11}, P_{12}) \propto P_{11}^{n_{11}}(1-P_{11})^{n_{12}}(1-P_{21})^{n_{21}}P_{22}^{n_{22}} \tag{1.71}$$

where n_{ij} denotes the number of transitions from state i to state j in the sample for $i, j = 1, 2$.

In the benchmark case, the annual transition probability matrix is

$$\mathbb{P}(1) = \begin{bmatrix} 0.9 & 0.1 \\ 0.45 & 0.55 \end{bmatrix}$$

Now consider an alternative transition probability matrix

$$\mathbb{P}^*(1) = \begin{bmatrix} 0.9 & 0.1 \\ 0.25 & 0.75 \end{bmatrix}$$

The DEP for these two transition probability matrices is 0.1667 based on 50000 samples.

1.5 Calibration

For annual US data from 1889 to 2009, the growth rate of real per capita consumption has a mean of 0.0208 and a standard deviation of 0.0357.[18] The sample skewness is -0.1979 and the sample excess kurtosis is 0.9245. These numbers are consistent with occasional disasters and fat tails.

In the model, the jump intensities are governed by a Markov chain that is consistent with a Markov-regime switching model. Cogley and Sargent (2008) note that the estimate for the transition probability that a recession will continue in Cecchetti et al. (2000) is 0.515 with a standard error of 0.264. This leads to a rather wide 90% confidence interval of $[0.079, 0.951]$. The expected duration of the recession state ranges from 1 year to 20 years. This example demonstrates that considerable model uncertainty persists even with 100 years of US data (1890-1994). Using a large data set containing consumption data from 24 countries, Nakamura et al. (2013) estimate a probability of remaining in a disaster state to be 0.835 with a standard error of 0.027. The two standard deviation confidence interval is therefore $[0.781, 0.889]$ and the expected length of disasters ranges from 4.5 to 9 years. The uncertainty surrounding the growth rates associated with consumption disasters is also large. The short-run and long-run shocks to consumption are normally distributed in Nakamura et al. (2013). The mean of the short-run shock and the mean of the long-run

[18] The data are from Robert Shiller's website. This is an annual series called long term stock, bond, interest rate and consumption data.

shock are estimated to be -0.111 and -0.025 respectively. But the corresponding standard deviations are 0.083 and 0.121.

The benchmark model calibration is based on the following parameters.

Parameter	Value (Annual Terms)
Drift term for consumption μ	0.02
Volatility for consumption σ	0.02
Jump size for consumption j_C	-3.45%
Jump size for dividends j_D	-10%
Jump size for disaster insurance j_I	-10%
Leverage ϕ	3
Jump intensity in the good state Λ_1	0.01
Jump intensity in the bad state Λ_2	0.12
Jump intensity in the bad state $\tilde{\Lambda}_2$	0.15 (without jumps) or 0.20 (with jumps)
Transition probability P_{11}	0.9
Transition probability P_{22}	0.55
Prior probabilities	$[\frac{5}{6}, \frac{1}{6}]$
Time horizon T	200 years

Table 1.2: Calibration: Model Parameters

The time unit is t=1 for a year. To transform this into a daily frequency, the time step is 1/365 in the calibration results. The drift term and the volatility term in the consumption process are consistent with historical US data, and are adopted from the rare disasters literature. The jump sizes are chosen so that the magnitude of the drop in consumption matches the data from 2008.[19] The parameter ϕ determines the relationship between consumption and dividends. In the existing literature, this parameter is usually in the range of 3 to 5. Here we choose the lower bound to ensure a conservative choice. The investors start with the same prior probabilities about the 2 potential states. Their disagreement comes from different specifications of the bad state (Λ_2 for the Bayesian and $\tilde{\Lambda}_2$ for the robust agent). The choices of different jump intensities are based on the detection error probabilities in the previous section. Time-varying fear constraint is in its most simple form: If the robust agent observes a jump, the set of potential missspecifications under consideration will expand and his worst case scenario choice of $\tilde{\Lambda}_2$ will increase from 0.15 to 0.2. The concern for model uncertainty will slowly dissipate: the worst case scenario returns to 0.15 in 3 years. At time 0, each agent is endowed with 0.5 shares of stock so that the initial wealth ratio is

[19]De Nardi et al. (2012) report that consumption dropped 3.4% from peak to trough during the Great Recession. One potential problem is that we use the peak-to-trough consumption drop during the Great Recession to determine the jump sizes. Usually, economic disasters unfold over several years. This issue has been raised in Constantinides (2008). As dividend jump size affects the wealth share allocated to the stock in the model, it will also influence the magnitude of trading volume spikes. But the results in this paper will not change significantly if we use different jump sizes.

1. The transition probabilities mean that an expansion will last 10 years on average and the average length of recessions is 2.2 years. The true jump intensity that governs the dividend process is 0.12 in the bad state. In the calibration, we choose a sample path that features three jumps in dividends on day 853, day 30033, and day 34109 in 100 years. This sample path is chosen so that the occurrences of disasters resemble the three disasters episodes in the US.

A sample path with 3 jumps in 100 years is a representative sample given the true model used in calibration. We are interested in the steady state probability distribution ($\bar{P} = [\bar{P}_1, \bar{P}_2]$) of the Markov chain. The steady state distribution must satisfy $\bar{P}\mathbb{Q} = 0$. Solving the set of equations and we get $\bar{P}_1 = \frac{9}{11}, \bar{P}_2 = \frac{2}{11}$. These probabilities characterize the long-run behavior of the Markov chain. The probability $\bar{P}_i$ is the long-run proportion of time that the chain spends in state i, $i = 1, 2$. These probabilities also allow us to calculate the expectation of the jump process $E[N(t)] = \bar{P}_1\Lambda_1 t + \bar{P}_2\Lambda_2 t = 0.03t$. The expected number of jumps in 100 years is indeed 3.

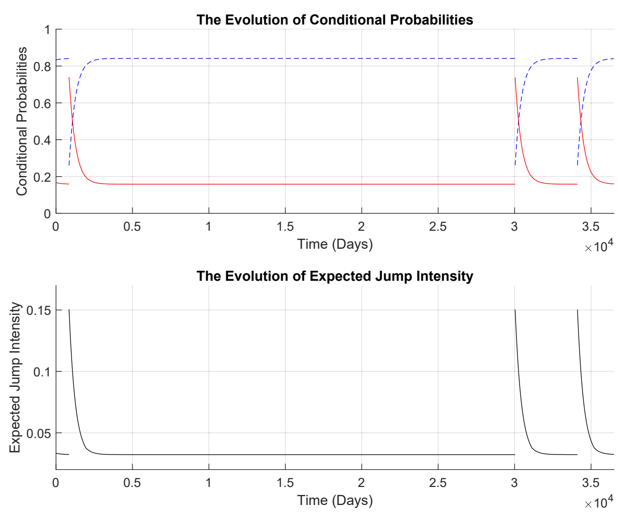

Figure 1.6: Robust Agent: Evolution of Conditional Probabilities and Expected Jump Intensity ($\Lambda_1 = 0.01, \tilde{\Lambda}_2 = 0.15$ (0.2 after a jump) , Priors = $[5/6, 1/6]$, Dashed Line: State 1, Solid Line: State 2.)

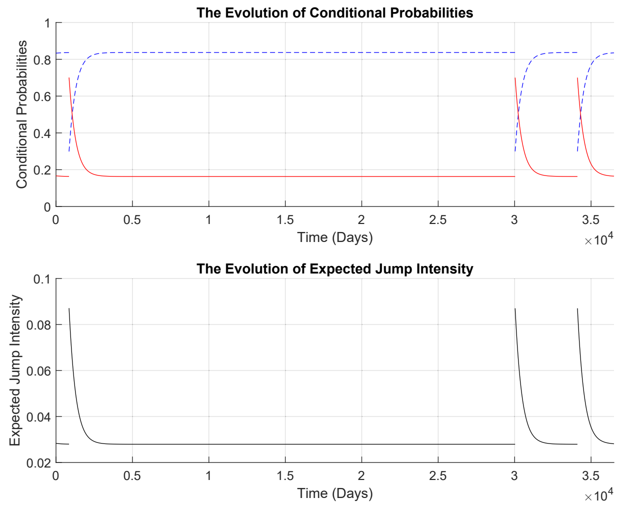

Figure 1.7: Pure Bayesian: Evolution of Conditional Probabilities and Expected Jump Intensity ($\Lambda_1 = 0.01, \Lambda_2 = 0.12, \text{Priors} = [5/6, 1/6]$, Dashed Line: State 1, Solid Line: State 2.)

Figures 6 and 7 document the changes in the conditional probabilities of the 2 states and the corresponding expected jump intensity for the two agents. When they do not observe any jumps, they both revise their beliefs such that their expected jump intensity keeps falling. But once a jump is observed, the disagreement between the two agents will dramatically increase due to model uncertainty. The difference in belief revision has important implications since it affects how the disagreement evolves. The disagreement shrinks at a lower rate before a jump. But the rate increases by a lot after a jump as the speed of belief revision is much higher for the robust agent. The elevated disagreement will lead to different portfolio holdings and changes in the equilibrium interest rate and wealth ratio. But more importantly, the higher disagreement will result in abnormal trading volume.

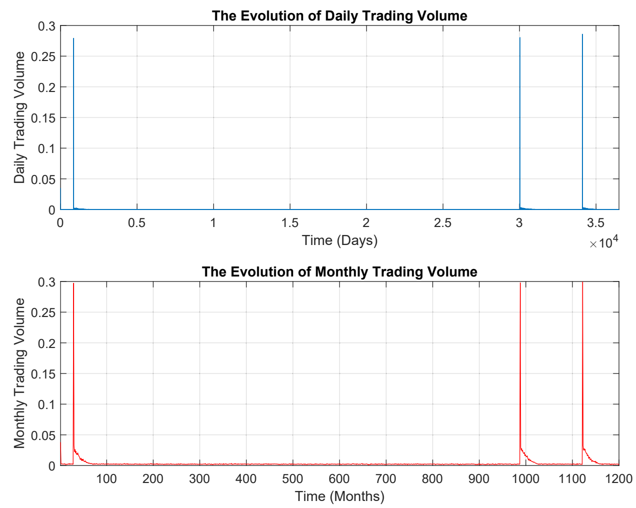

Figure 1.8: Evolution of Trading Volume: Incomplete Markets

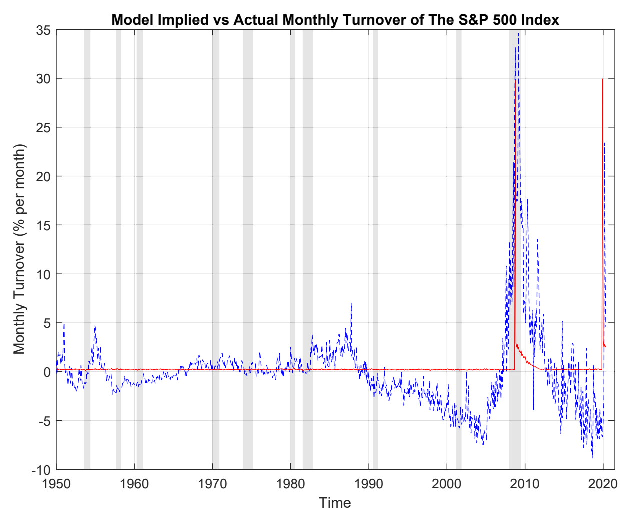

Figure 1.9: Model Implied vs Actual Monthly Turnover of The S&P 500 Index (1950M1-2020M5) (Dashed Line: Actual Detrended Turnover, Solid Line: Model Implied Turnover.)

The robust agent decreases his exposure to the stock market while the Bayesian agent increases his wealth share allocated to the stock market right after a jump. These portfolio holdings reflect the magnitude of their disagreement, but trading volume is directly related to how fast the disagreement changes. There are 3 large spikes in trading volume due to jumps in dividends and belief revisions. The monthly trading volume is 0.3 during the month when a jump happens. This spike in trading is consistent with the detrended monthly turnover data in which the monthly turnover spikes can be as high as 30%. Trading volume is generally very small without jumps in the dividends. But the monthly trading volume after a jump is higher than the pre-jump monthly trading volume. The key mechanism that determines how trading volume changes is the speed of belief revision. Figure 9 plots the model implied monthly turnover and the actual detrended monthly turnover of the S&P 500 index from 1950 to 2020.[20] The model can generate reasonable trading volume spikes when there are jumps in the dividends. For example, the detrended monthly turnover in October 2008 is 33%. The model implied monthly turnover for October 2008 is 30%. Heterogeneous beliefs based on robustness can account for 90% of the trading volume spike.

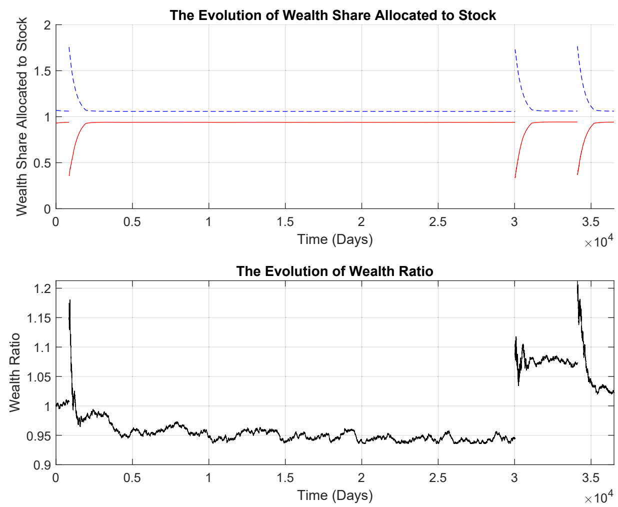

Figure 1.10: Evolution of Wealth Share Allocated to Stock and Wealth Ratio: Incomplete Markets (Dashed Line: Pure Bayesian, Solid Line: Robust Agent.)

The wealth ratio is governed by a jump diffusion under incomplete markets (see Proposition 1.3.5). Consequently, the wealth ratio fluctuates even when there are no jumps. The changes in dividends caused by the movements of the Brownian motion produces relative wealth changes because the agents have different exposures to the stock market. A jump in the dividends induces a larger loss to the Bayesian agent. The wealth ratio therefore increases in favor of the robust agent. When a jump in dividends happens, the wealth transfer is large because of the stock price drop. The negative jump in the wealth ratio is mainly caused by the stock price decrease. These results are in line with the literature that focus on the long-run survival of agents in incomplete markets.[21] The market selec-

[20]The first model implied spike is matched to the October 2008 spike in the data.

[21]Due to the finite horizon assumption, we cannot rigorously examine the asymptotic properties of the wealth ratio. However, the relevant channels that determine long-run survival are nonetheless present in this paper.

tion hypothesis dictates that financial markets select agents with correct beliefs in the long run. Blume and Easley (2006) demonstrate that the market selection hypothesis may fail if market completeness is violated. For example, an agent with pessimistic beliefs will generally allocate his financial assets in a suboptimal way. However, the pessimistic agent might choose to save enough to compensate for his poor asset allocation. More generally, incomplete markets sever the close link between correct beliefs and long-run survival. Beker and Chattopadhyay (2010) show that either some agent vanishes or the consumption of both agents is arbitrarily close to zero infinitely often in a two agent incomplete market framework. Coury and Sciubba (2012) prove that survival does not even require agents' beliefs to merge with the truth in incomplete markets. Both Beker and Chattopadhyay (2010) and Coury and Sciubba (2012) provide examples in which agents with corrrect beliefs vanish or agents with incorrect beliefs survive. We can therefore conclude that correct beliefs are neither necessary nor sufficient for long-run survival in incomplete markets. In a similar vein, Cogley, Sargent and Tsyrennikov (2014) study the effects of market incompleteness on the distribution of wealth when agents have heterogeneous beliefs. In their model, one agent knows the true endowment process while another agent learns about it through Bayesian updating. With complete markets, the learning agent loses wealth as the asset purchased for the recession state pays off less often than he expects. When markets are incomplete, precautionary savings become more important as the learning agent accumulates risk-free bonds. Eventually, both the well-informed agent and the learning agent survive but the more knowledgeable agent is pushed to his debt limit. Market incompleteness thus acts as a mechanism that stops the agent with incorrect beliefs from placing bets through disaster insurance type securities. It is important to note that the mechanism in this paper is slightly different from the precautionary saving channel in Cogley, Sargent and Tsyrennikov (2014). In their setup, incomplete markets are implemented through a single bond, agents' saving choice and portfolio choice coincide: all their savings are invested in the bond. In our incomplete markets, agents' saving decisions are not different. Their heterogeneous beliefs lead to different portfolio choices: the robust agent invest more heavily in the bond while the Bayesian focuses on the stock market.

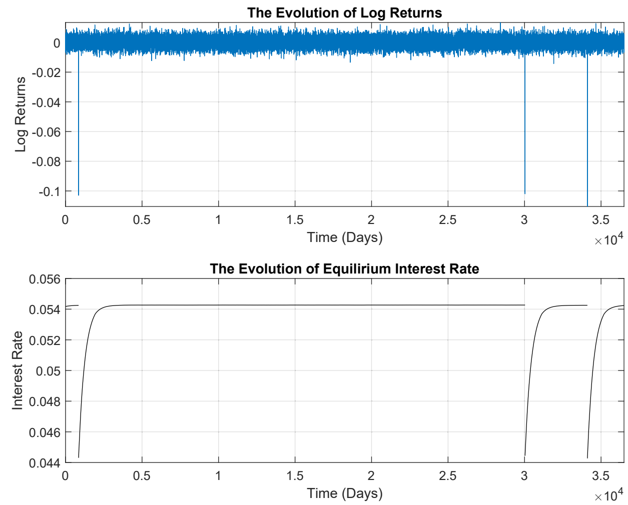

Figure 1.11: Evolution of Log Returns and Interest Rate

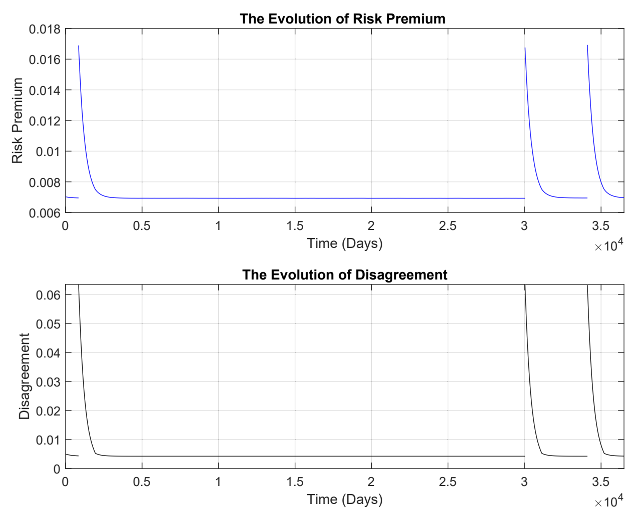

Figure 1.12: Evolution of Risk Premium and Disagreement

The stock price reflects changes in the dividend process closely, as its expression is $S(t) = D(t)(T - t)$. The equilibrium interest rate decreases in a way that is consistent with a ”flight to safety“ effect: when a jump happens, demand for safe asset will increase as the precautionary saving motive strengthens. More precautionary savings will bring down the equilibrium interest rate. Two types of effects affect the interest rates. The first one is a wealth effect: when the dividends decrease, there is a transfer of wealth from the Bayesian to the robust agent because the robust agent suffers less from the decrease in the stock price. The interest rate has a tendency to decrease as now the overall rare event risk is higher. The second one is belief revision: when both agents revise their beliefs downward, the interest rate increases as the overall rare event risk is lower. At the moment a jump happens, these two effects reinforce each other as both agents significantly increase their expected jump intensity in addition to a wealth transfer from the Bayesian to the robust agent. This is why there is a large negative jump in the equilibrium interest rate. When the dividends fluctuate due to the movements of the Brownian motion, an increase of the dividend leads to a wealth transfer in favor of the Bayesian and both agents become more optimistic, the two effects all contribute to a higher interest rate. But when the dividends decrease during normal times, the two effects operate in opposite directions and the overall effect is ambiguous. The negative jump in the equilibrium interest rate also explains the counter cyclical risk premium. The risk premium is too small compared to the data. But this is understandable as the risk aversion is fairly small under log utility.

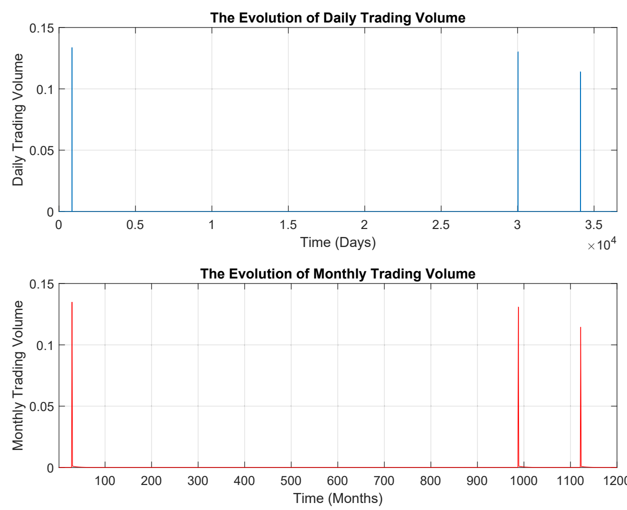

Figure 1.13: Evolution of Trading Volume: Complete Markets

From Proposition 1.3.6, we know that the wealth shares allocated to stock are both one in complete markets. Heterogeneous beliefs do not affect stock market trading volume. Stock trading is only driven by wealth fluctuations in this case. Figure 13 plots the daily trading volume and monthly trading volume in complete markets : the trading volume spikes are lower than 0.15. The lower spikes in complete markets are not surprising as large wealth fluctuations occur when jumps happen. But the effects of belief revision are absent as heterogeneous beliefs dictate that agents have different holdings of disaster insurance instead of stock in complete markets.

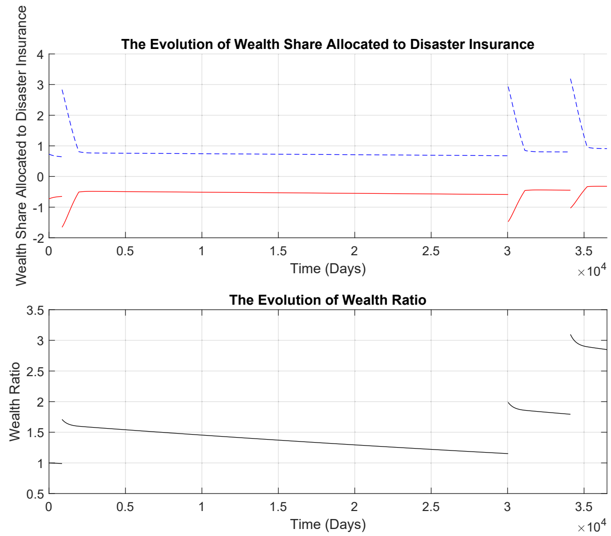

Figure 1.14: Evolution of Wealth Share Allocated to Disaster Insurance and Wealth Ratio: Complete Markets (Dashed Line: Pure Bayesian, Solid Line: Robust Agent.)

The wealth ratio is governed by the jump process given in Proposition 1.3.6. The drift term reflects the main mechanism behind the long-run survival results in complete markets: agents with more accurate beliefs invest in securities with higher returns. The jump term appears as agents have different holdings of disaster insurance. Once a jump happens, there is a relative wealth transfer from the Bayesian agent to the robust agent. These two mechanisms determine how the wealth ratio evolves in complete markets. In Figure 14, we can see that agents have very different holdings of disaster insurance. The robust agent shorts the disaster insurance while the Bayesian acts as a buyer of the insurance.[22] This enables the Bayesian to focus on investing in securities with higher returns. Without jumps, the wealth ratio is decreasing as the Bayesian is accumulating wealth at a higher rate. But when a jump happens, there is a large drop of the wealth ratio due to the relative wealth transfer. These results are also consistent with the long-run survival literature. Sandroni (2000) examines the market selection hypothesis in complete markets and establishes that agents with correct beliefs are indeed selected in the long run. The first theorem of welfare economics provides a link between competitive equilibrium and Pareto optimality: equilibrium allocations in economies with complete markets are Pareto optimal. Blume and Easley (2006) therefore show that the market selection hypothesis is valid in the sense that agents with correct beliefs will be selected in Pareto optimal economies with bounded endowment. A consequence of Pareto optimality is that heterogeneous beliefs lead to different marginal utilities. The central question in the long-run survival literature is whether differences in

[22]This is because the jump size of the disaster insurance is also negative. In this case, the robust agent shorts the disaster insurance as he expects the price to drop more frequently.

marginal utilities translate into different consumption shares. Kogan et al. (2017) develop necessary and sufficient conditions for long-run survival in complete market pure exchange economies. The only restriction is that utility functions have to be time-separable.[23] The market selection hypothesis holds as long as the curvature of the utility functions drops fast enough when consumption drops (the relative risk aversion coefficient is bounded) or the aggregate endowment process is bounded. Cogley and Sargent (2009) notice that various wealth allocations can emerge in a complete market setup, depending on the initial wealth shares and the occurrence of contractions. For example, a particular sample path that features no contractions is clearly favorable to the agent with correct beliefs. On the other hand, the better informed agent will suffer greatly on a sample path with many contractions. [24]

1.6 Sensitivity Analysis

In this section, I discuss how the choices of different prior probabilities and concerns for model misspecification affect trading volume. I consider four different scenarios that involve changing one parameter in the benchmark calibration with incomplete markets.

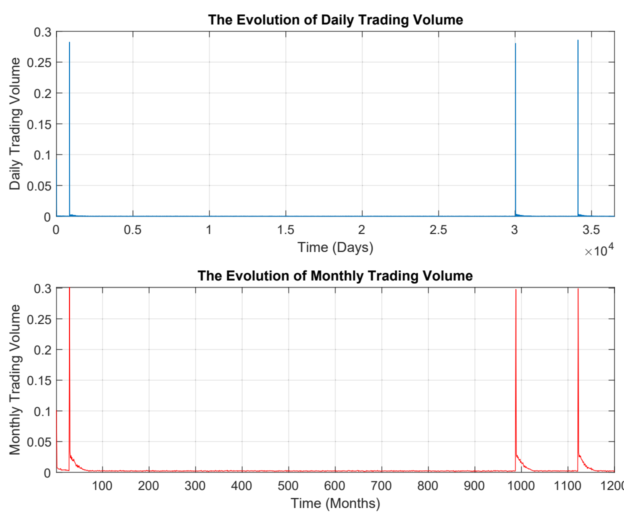

Figure 1.15: Evolution of Trading Volume: Different Prior Probabilities ($\Lambda_1 = 0.01, \Lambda_2 = 0.12, \tilde{\Lambda}_2 = 0.15$ (0.2 after a jump) Priors=[1/2, 1/2])

[23]The results in Kogan et al. (2017) do not extend to the recursive preference case. Borovička (2020) study the long-run survival problem with recursive preferences and heterogeneous beliefs. There are three channels that determine long-run wealth dynamics: the risk premium channel, the speculative volatility channel and the saving channel. The saving channel is relevant since a high enough intertemporal elasticity of substitution dictates that the agent with negligible wealth share chooses to save more and avoids extinction.

[24]It is important to distinguish between survival on a path and survival in the almost surely sense. Since we choose a sample path in the calibration exercise, our focus is on survival on a path (instead of almost sure survival).

If we change the prior probabilities from $[5/6, 1/6]$ to $[1/2, 1/2]$, then the spike in monthly trading volume is still 0.3 (see Figure 15), the major difference from the benchmark calibration is that the trading volume at the beginning is higher in this case because the agents have larger disagreement due to the different priors. Their adjustment from the initial allocation (0.5 shares of stock each) is therefore higher. A similar pattern emerges if we allow the robust agent to have constant fear instead of time-varying fear. When the robust agent sets the jump intensity for the bad state to be 0.2 all the time, the trading volume spike is still 0.3 and the trading volume during the first month is higher than the benchmark case (see Figure 16).

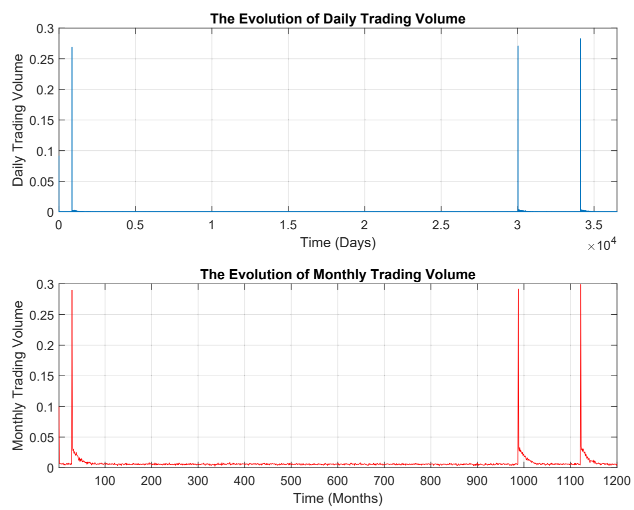

Figure 1.16: Evolution of Trading Volume: Higher Constant Fear ($\Lambda_1 = 0.01, \Lambda_2 = 0.12, \tilde{\Lambda}_2 = 0.20, \text{Priors} = [5/6, 1/6]$)

In stark contrast, if the robust agent only sets the jump intensity for the bad state to be 0.15 all the time, the trading volume spike is only around 0.12 (see Figure 17). Therefore the model's ability to match the trading volume spike in the data hinges on the magnitude of disagreement spikes after jumps in dividends.

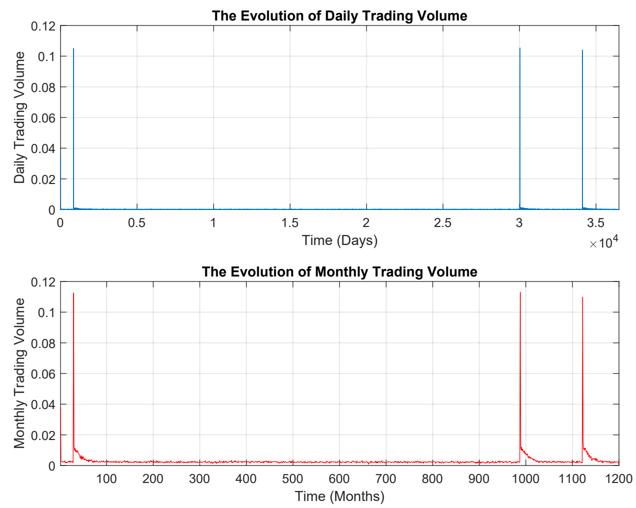

Figure 1.17: Evolution of Trading Volume: Lower Constant Fear ($\Lambda_1 = 0.01, \Lambda_2 = 0.12, \tilde{\Lambda}_2 = 0.15, \text{Priors} = [5/6, 1/6]$)

Figure 18 plots the evolution of daily and monthly tading volume when the robust agent is also concerned with the persistence of the bad state. If the robust agent sets the probability that the economy will stay in the bad state to be 0.75. The expected duration of the recession is 4 years in this case. The concern for persistence misspecification does not alter the benchmark results. We can conclude that misspecification about the potential states (jump intensities) is much more important in terms of explaining trading activities.

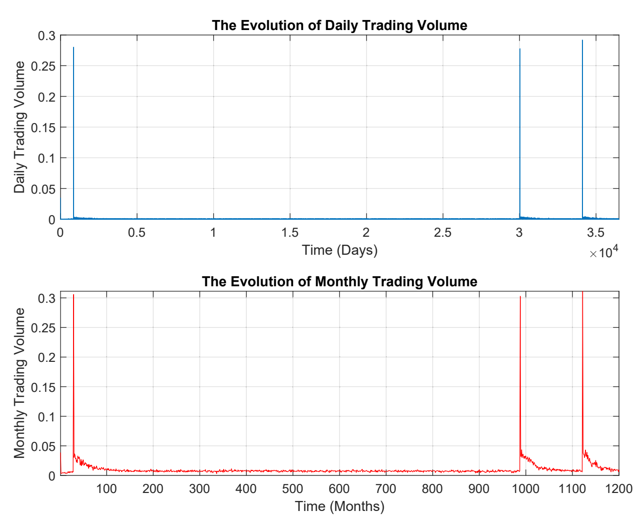

Figure 1.18: Evolution of Trading Volume: Higher Persistence ($\Lambda_1 = 0.01, \Lambda_2 = 0.12, \tilde{\Lambda}_2 = 0.15$ (0.2 after a jump) Priors $= [5/6, 1/6], P_{22} = 0.75$)

1.7 Conclusion

This paper introduces a novel approach that relates heterogeneous beliefs to rare disasters. The key mechanism is a time-varying fear of model misspecification. The model generates plausible dynamics for both prices *and* trading volume. The results also show that trading in the stock market in response to rare events has important distributional consequencs, since investors choose different exposures to financial risks based on their beliefs. An important advantage of the model in this paper relative to other heterogeneous beliefs models is that robust decision theory lends discipline to the specification of belief heterogeneity.

I used log preferences in this paper to obtain closed-form solutions, which helps to clarify the underlying dynamics. Unfortunately, the price we pay for this clarity is that we cannot fully capture the historical equity premium and risk-free rate. A larger equity premium would likely reinforce the paper's results. Therefore, a useful extension would be to numerically solve a version with CRRA or recursive preferences.

Chapter 2

The Market Selection Hypothesis and Rare Disasters

2.1 Introduction

An important question in economics is whether agents with inaccurate beliefs can survive and influence prices in financial markets. If they cannot, then only agents with accurate beliefs will survive and dictate market prices. This is the market selection hypothesis. The intuition that markets should eventually select agents whose beliefs are closest to the truth is certainly appealing to economists. This conjecture provides support for both the efficient market hypothesis and rational expectations.

One major insight in the existing literature is that the market selection hypothesis only holds in complete markets. Another insight is that maximizing expected utility is not the same as maximizing wealth. Wealth accumulation depends on expected log returns and saving behavior. Therefore, expected returns, portfolio volatility and saving are the main channels that determine long run survival in financial markets.

Why do incomplete markets have different implications for the market selection hypothesis? Which channel is responsible for the differences? In this paper, I test the market selection hypothesis in a continuous time asset pricing model with two agents that accommodates both complete and incomplete markets. An essential feature is that the endowments can jump. I show that the market selection hypothesis is valid when agents have log preferences.

One of the most notable features of the stock market is that stock market crashes happen. Figure 1 plots the monthly log returns of the S&P 500 between 1950 and 2020. The data show that although the monthly log returns are mostly between 5% and -5%; market crashes happen sporadically in the US. The market crashes from the Great Recession and the COVID-19 pandemic really stand out. They both produced a roughly 20% drop in the index. Another well-known empirical fact about stocks is that they consistently outperform

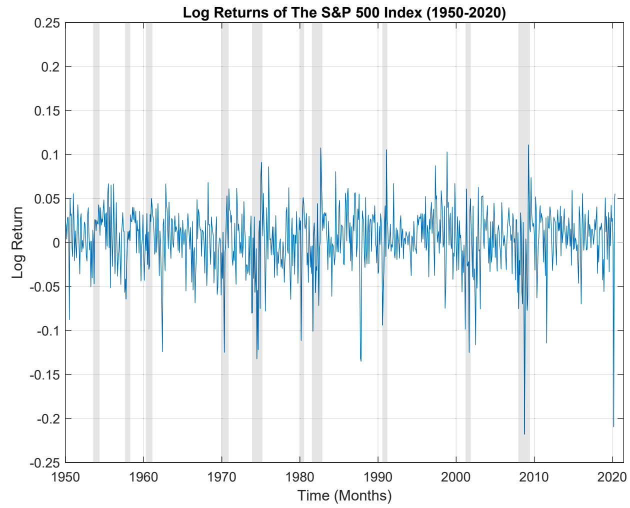

Figure 2.1: Log Returns of The S&P 500 (1950M1-2020M5). Shaded areas are NBER recessions.

bonds by a high margin. This is the equity premium puzzle.[1] The above empirical patterns suggest that an important trade-off is involved in stock investments. To reap the benefits of higher returns, investors need to allocate more wealth to the stock market. But a higher exposure to the stock market may translate into large losses when stock market crashes.

Do people have different exposures to the stock market? Wolff (2021) report that the top 1% US households in terms of wealth hold 25.8% of their total assets in the form of stocks. The households in the middle three quintiles are less exposed to the stock market as stocks only represent 8.6% of their assets.[2] In this paper, I focus on the case in which heterogeneous beliefs lead to differential exposures to the stock market. Moreover, market crashes will trigger large wealth redistribution.

Jumps are important in the sense that they influence wealth dynamics in the long run.[3] But in this paper, the wealth redistribution channel is not strong enough to invalidate the market selection hypothesis. When both agents have log preferences, wealth accumulation and utility maximization are no longer separated, long-run survival is determined by the log optimal rule. An immediate implication is that an agent with irrational beliefs will vanish as his portfolio choices are further away from those dictated by the log optimal rule. A

[1]Interestingly, some economists view rare disasters as a potential explanation for the puzzle. See Barro (2006) and Barro (2009).

[2]The data are from the 2019 Survey of Consumer Finances.

[3]In comparison, a Brownian motion will not affect wealth dynamics when asset prices are governed by diffusion processes. This is a result of the strong law of large numbers for Brownian motion. An example is provided in section IV.

rational agent, on the other hand, will survive as his portfolio choices follow the log optimal rule. This result is robust as it holds in both complete and incomplete markets.

On the surface, the above result seems to contradict the many existing results in the literature suggesting that the market selection hypothesis tends to fail in incomplete markets. But the results in this paper reveal that the main channel driving the existing results is the saving channel. When both agents have log preferences, their saving behavior is identical given that they have the same discount rate. So if the saving channel is deactivated, only the agent with rational beliefs survive. Without the possibility of oversaving, the agent with irrational beliefs cannot reverse the effects of poor asset allocation.

Using results from ergodic theory, I establish that the market selection hypothesis is valid even if both agents cannot observe the true state. If an agent is not endowed with the optimal filter, then he will vanish in the long run. This result is relevant as hidden Markov models become more and more popular in the asset pricing literature. The intuition behind this result is that the optimal filter's time average loss is asymptotically smaller than other strategies on almost all sample paths under certain conditions. The time average loss determines long-run survival in a hidden Markov model with complete markets.

2.2 Literature

The market selection hypothesis is a long standing hypothesis in the economics literature. The conjecture that agents with rational beliefs should prevail in the long run is certainly plausible. The existing literature, however, establishes that the fate of agents in financial markets is determined by a number of factors: beliefs, market structure and preferences.

In complete markets, beliefs turn out to be the most important factor. Blume and Easley (1992) establish that agents who follow the log optimal rule survive in the long run when saving is exogenous. Sandroni (2000) proves that agents with accurate beliefs are indeed selected in complete markets with bounded endowments. Blume and Easley (2006) further generalize the result: the market selection hypothesis is valid in Pareto optimal economies with bounded endowments. Yan (2008) shows that elasticity of intertemporal substitution also matters if endowment is unbounded. The main insight in the literature is that agents with accurate beliefs will allocate more consumption to paths with high probability according to the true model and this mechanism determines their survival in the long run. But this insight only applies in the complete market setup.

When markets are incomplete, it is more difficult to obtain clear-cut results about long-run survival. Blume and Easley (2006) demonstrate that the market selection hypothesis may fail if market completeness is violated. They provide examples in which agents with irrational beliefs might choose to save enough to compensate for poor asset allocation. Beker and Chattopadhyay (2010) show that either some agent vanishes or the consumption of both agents is arbitrarily close to zero infinitely often in a two agent incomplete market framework. Coury and Sciubba (2012) prove that survival does not even require agents' beliefs to merge with the truth in incomplete markets. Both Beker and Chattopadhyay (2010) and Coury and Sciubba (2012) provide examples in which agents with corrrect beliefs vanish or agents with incorrect beliefs survive. We can therefore conclude that correct beliefs are neither necessary nor sufficient for long-run survival in incomplete markets. In a similar vein, Cogley, Sargent and Tsyrennikov (2014) study the effects of market incompleteness on the distribution of wealth when agents have heterogeneous beliefs. When markets are incomplete, precautionary savings become more important as the learning agent accumulates risk-free bonds. Eventually, both the well-informed agent and the learning agent survive but the more knowledgeable agent is pushed to his debt limit.

Borovicka (2020) shows that preferences also play an important role in determining survival. In a continuous time asset pricing model with complete markets, two agents are endowed with identical recursive preferences and heterogeneous beliefs about the growth rate of the aggregate endowment. For wide ranges of preference parameter values, the agent with less accurate beliefs can survive or even dominate, especially in the case where risk aversion is sufficiently higher than the inverse of intertemporal elasticity of substitution.

2.3 The Model

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}(t)\}_{t\geq 0}, P)$. Let $\mathcal{F}^W(t)$ be the σ-algebra generated by a Brownian motion $W(s)$; $s \leq t$ and let $\mathcal{F}^N(t)$ be the σ-algebra generated by a Poisson jump process $N(s)$; $s \leq t$. $W(t)$ and $N(t)$ are assumed to be independent. More precisely, the σ-algebra $\mathcal{F}(t)$ is defined as $\mathcal{F}(t) = \mathcal{F}^W(t) \times \mathcal{F}^N(t)$. I consider a pure-exchange economy with infinite horizon .

The dynamics of dividend follow a jump diffusion process.

$$\frac{dD(t)}{D(t-)} = \mu_D dt + \sigma_D dW(t) + j_D dN(t) \tag{2.1}$$

The jump size j_D is restricted to be in the interval $(-1, 0)$ to ensure that the dividend process remains positive. There are two types of agents: 1 and 2. The agents have equivalent probability measures and commonly observe the aggregate dividend. They only have incomplete information about the potential jump intensities of the Poisson process $\lambda(t)$. To simplify the analysis, I assume the agents know the true values of μ_D, σ_D and j_D. The jump size j_D is assumed to be constant.

Denote agent i's beliefs about $\lambda(t)$ as $\lambda^i(t), i = 1, 2$. The agents have to agree about the aggregate dividend. The following consistency requirement holds.

$$\frac{dD(t)}{D(t-)} = \mu_D dt + \sigma_D dW(t) + j_D dN^1(t) = \mu_D dt + \sigma_D dW(t) + j_D dN^2(t) \tag{2.2}$$

The above consistency condition tells us that the agents have to agree about whether a jump happens[4]:

$$dN^1(t) = dN^2(t) = \begin{cases} 0 & \text{if no jumps happen at time } t, \\ 1 & \text{if a jump happens at time } t, \end{cases}$$

The following definition is useful in the derivation of stochastic discount factors.

DEFINITION 2.3.1. *Let $\theta(t)$, $\lambda(t)$ and $\xi(t)$ be predictable processes. t_n is a sequence of jump times. In a jump-diffusion context, the change of measure is*

$$L_W(t) = \exp\Big(- \int_0^t \theta(s) dW(s) - \frac{1}{2}\int_0^t \theta(s)^2 ds \Big) \tag{2.3}$$

$$L_N(t) = \exp\Big(- \int_0^t \lambda(s)\xi(s) ds \Big) \prod_{n, t_n \le t} (1 + \xi(t_n)) \tag{2.4}$$

$$L(t) = L_W(t) L_N(t) \tag{2.5}$$

Since multiple equivalent martingale measures exist in incomplete markets and the agents have heterogeneous beliefs, we need to derive their agent specific stochastic discount factors. Let $r(t)$ be the equilibrium interest rate. The agent specific stochastic discount factor is $(i = 1, 2)$

$$M^i(t) = \frac{L^i(t)}{\exp(\int_0^t r(s) ds)} \tag{2.6}$$

where $L^i(t)$ is agent i's change of measure from Definition 2.3.1. The stochastic discount factor has the following differential form:

$$dM^i(t) = -M^i(t-)\Big[r(t)dt + \theta^i(t) dW^i(t) - (\lambda^i(t) - \lambda^i_{RN}(t))dt - \Big(\frac{\lambda^i_{RN}(t)}{\lambda^i(t)} - 1 \Big) dN^i(t) \Big] \tag{2.7}$$

where $\theta^i(t)$ is agent i's market price of risk and $\lambda^i_{RN}(t) = \lambda^i(t)(1 + \xi^i(t))$ is the market price of rare event risk (the risk-neutral jump intensity).

In our incomplete markets economy, agents have access to 2 types of assets: a bond $B(t)$ with 0 net supply and a risky stock $S(t)$ with net supply of 1. Their price dynamics are as follows

$$dB(t) = B(t) r(t) dt \tag{2.8}$$

$$dS(t) = S(t-)[\mu_S(t)dt + \sigma_S dW(t) + j_S dN(t)] = S(t-)[\mu^i_S(t)dt + \sigma_S dW^i(t) + j^i_S dN^i(t)] \tag{2.9}$$

where j_S is the jump size of the stock price process.

[4] According to the Girsanov theorem, if N is a Poisson process with jump intensity λ under a probability measure P, then under a new probability measure Q, N has a different jump intensity λ^Q. This is different from the diffusion case where a "drift adjustment" is required.

Each investor's problem is

$$\max_{c_i} E^i\left[\int_0^\infty e^{-\delta t}u(c_i(t))dt\right] \tag{2.10}$$

subject to

$$E^i\left[\int_0^\infty M^i(t)c_i(t)dt\right] \leq w_i(0) \tag{2.11}$$

where $w_i(0)$ is agent i's initial endowment and δ is agent i's discount rate.

The essence of the martingale approach is that it allows us to transform a dynamic problem into a static one. It delivers the following optimality condition:

$$e^{-\delta t}u'(c_i(t)) = y_i M^i(t) \tag{2.12}$$

where y_i is the Lagrange multiplier associated with agent i's intertemporal budget constraint. An equivalent condition is

$$c_i(t) = (u')^{-1}(y_i M^i(t)e^{\delta t}) \tag{2.13}$$

Given the assumption of time-additive log utility:

$$c_i(t) = (u')^{-1}(y_i M^i(t)e^{\delta t}) = \frac{1}{y_i M^i(t)e^{\delta t}} \tag{2.14}$$

$$y_i = \delta w_i(0) \tag{2.15}$$

$$w_i(t) = \frac{E^i[\int_t^\infty M^i(s)c_i(s)ds]}{M^i(t)} = \frac{1}{\delta^2 w_i(0)M^i(t)e^{\delta t}} \tag{2.16}$$

$$c_i(t) = \delta w_i(t) \tag{2.17}$$

By Ito's lemma for jump diffusion

$$dw_i(t) = [...]dt + w_i(t-)\theta^i(t-)dW^i(t) + w_i(t-)\left(\frac{\lambda^i(t-)}{\lambda^i_{RN}(t-)} - 1\right)dN^i(t) \tag{2.18}$$

The wealth process $w_i(t)$ is

$$dw_i(t) = (1-\pi_i(t))w_i(t-)\frac{dB(t)}{B(t)} + \pi_i(t)w_i(t-)\frac{dS(t)+D(t)dt}{S(t)} - c_i(t)dt \tag{2.19}$$

where $\pi_i(t)$ is the share of wealth allocated to the stock market by agent i at time t.

$$dw_i(t) = (1-\pi_i(t))w_i(t-)r(t)dt + \pi_i(t)w_i(t-)[\mu^i_S(t)dt + \sigma_S dW^i(t) + j^i_S dN^i(t)] - c_i(t)dt \tag{2.20}$$

DEFINITION 2.3.2. *An equilibrium is a collection of asset prices {B, S} and a set of consumption and portfolio choices* $\{c_i, \pi_i\}$ *such that*

1. Each agent's consumption is optimal: the consumption plan c_i *maximizes lifetime utility subject to the budget constraint and is financed by the portfolio choice* π_i.

2. Goods market and financial markets clear at all times. The market clearing conditions are

$$c_1(t) + c_2(t) = D(t) \tag{2.21}$$

$$w_1(t) + w_2(t) = S(t) \tag{2.22}$$

$$\pi_1(t)w_1(t) + \pi_2(t)w_2(t) = S(t) \tag{2.23}$$

PROPOSITION 2.3.1. *The stock price process is given by*

$$S(t) = \frac{D(t)}{\delta} \tag{2.24}$$

$$\mu_S^i(t) = \mu_D \tag{2.25}$$

$$\sigma_S = \sigma_D \tag{2.26}$$

$$j_S = j_D \tag{2.27}$$

The stock price process has the same drift term, volatility term, and jump size as the dividend process.

PROPOSITION 2.3.2. *In incomplete markets, the portfolio holdings are*

$$\pi_i(t) = \frac{\theta^i(t)}{\sigma_S} = \frac{1}{j_S^i}\left(\frac{\lambda^i(t)}{\lambda_{RN}^i(t)} - 1\right) \tag{2.28}$$

The wealth ratio $k(t) = \frac{w_2(t)}{w_1(t)}$ *is governed by*

$$\frac{dk(t)}{k(t-)} = [\theta^2(t)^2 - \theta^1(t)\theta^2(t) + (\lambda^1(t) - \lambda_{RN}^1(t)) - (\lambda^2(t) - \lambda_{RN}^2(t))]dt + [\theta^2(t) - \theta^1(t)]dW(t) + \left(\frac{\lambda_{RN}^1(t)\lambda^2(t)}{\lambda^1(t)\lambda_{RN}^2(t)} - 1\right)dN(t) \tag{2.29}$$

The equilibrium interest rate is

$$r(t) = \delta + \mu_S^i(t) - \sigma_S\theta^i(t) + j_S^i\lambda_{RN}^i(t) \tag{2.30}$$

$$\sigma_S - \frac{1}{1+k(t)}\theta^1(t) = \frac{k(t)}{1+k(t)}\theta^2(t) \tag{2.31}$$

$$\frac{\lambda^1(t)}{\lambda_{RN}^1(t)}\frac{1}{1+k(t)} + \frac{\lambda^2(t)}{\lambda_{RN}^2(t)}\frac{k(t)}{1+k(t)} = j_S + 1 \tag{2.32}$$

In principle, three nonredundant securities will complete the market as there are two random sources. For example, introducing an index option in addition to the bond and the stock will complete the market. But an index option may not be a good choice in terms of tractability. We consider a zero net supply security that delivers tractability and fulfills the function of disaster insurance at the same time. This security is a disaster insurance in the sense that its payoff only depends on the jump component. Its price $I(t)$ is:

$$dI(t) = I(t-)[\mu_I(t)dt + j_I^i dN^i(t)] \tag{2.33}$$

where j_I^i is the jump size of the disaster insurance. The wealth process $w_i(t)$ is

$$dw_i(t) = (1-\pi_i(t)-\pi_i^I(t))w_i(t-)\frac{dB(t)}{B(t)}+\pi_i(t)w_i(t-)\frac{dS(t)+D(t)dt}{S(t)}+\pi_i^I(t)w_i(t-)\frac{dI(t)}{I(t)}-c_i(t)dt \tag{2.34}$$

where $\pi_i^I(t)$ is the wealth share allocated to the disaster insurance by agent i.

PROPOSITION 2.3.3. *Under complete markets, the portfolio holdings are*

$$\theta^i(t) = \sigma_S \tag{2.35}$$

$$\pi_i(t) = \frac{\theta^i(t)}{\sigma_S} = 1 \tag{2.36}$$

$$\pi_i^I(t) = \frac{1}{j_I}\left(\frac{\lambda^i(t)}{\lambda_{RN}^i(t)} - 1\right) - \frac{j_S\theta^i(t)}{j_I\sigma_S} \tag{2.37}$$

$$\lambda_{RN}(t) = \frac{1}{1+k(t)}\frac{\lambda^1(t)}{1+j_S} + \frac{k(t)}{1+k(t)}\frac{\lambda^2(t)}{1+j_S} \tag{2.38}$$

$$r(t) = \delta + \mu_S(t) - {\sigma_S}^2 + j_S\lambda_{RN}(t) \tag{2.39}$$

The wealth ratio is governed by

$$\frac{dk(t)}{k(t-)} = [\lambda^1(t) - \lambda^2(t)]dt + \left(\frac{\lambda^2(t)}{\lambda^1(t)} - 1\right)dN(t) \tag{2.40}$$

COROLLARY 2.3.1. *Under complete markets, the wealth ratio at time t is*

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s) - \lambda^2(s))ds + \int_0^t log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)dN(s)\right) \tag{2.41}$$

Under incomplete markets, the wealth ratio at time t is

$$k(t) = k(0)\exp\Bigg(\int_0^t (\theta^2(s)^2 - \theta^1(s)\theta^2(s) + (\lambda^1(s) - \lambda^1_{RN}(s)) - (\lambda^2(s) - \lambda^2_{RN}(s)) - \frac{1}{2}(\theta^2(s) - \theta^1(s))^2)ds$$

$$+ \int_0^t (\theta^2(s) - \theta^1(s))dW(s) + \int_0^t log\Big(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\Big)dN(s)\Bigg) \quad (2.42)$$

2.4 Long Run Survival in Complete and Incomplete Markets

The equations (41) and (42) describe how the wealth ratio fluctuates in the economy. In this section, it is assumed that agent 1 has rational beliefs, i.e., he knows the true jump intensity of the Poisson process $\lambda(t)$. The survival, extinction, and dominance of an agent can be defined in terms of the wealth ratio.[5]

DEFINITION 2.4.1. *Almost sure extinction,survival and dominance.*

1. *Agent 2 becomes extinct if* $\lim_{t\to\infty} k(t) = 0$ *a.s.*
2. *Agent 2 survives if* $\limsup_{t\to\infty} k(t) > 0$ *a.s.*
3. *Agent 2 dominates if* $\lim_{t\to\infty} k(t) = \infty$ *a.s.*

When asset prices can jump, there is a redistribution channel that leads to large wealth fluctuations when market crashes happen. Is this channel important enough in the examination of the market selection hypothesis? In order to emphasize the importance of jumps, I show in the following example that a Brownian motion does not affect wealth dynamics if the endowment process is governed by a diffusion process.

2.4.1 A Motivating Example

Here I consider a simple case without jumps. Aggregate dividend is governed by a geometric Brownian motion.

$$\frac{dD(t)}{D(t)} = \mu_D dt + \sigma_D dW(t) \quad (2.43)$$

I assume that agents have access to two assets, a bond and a stock. Since the only random source is the Brownian motion $W(t)$, markets are complete. The asset price dynamics are as follows

$$dB(t) = B(t)r(t)dt \quad (2.44)$$

$$dS(t) + D(t)dt = S(t)[\mu_S(t)dt + \sigma_S dW(t)] \quad (2.45)$$

[5]Although the definition is only related to agent 2, the survival, extinction, and dominance of agent 1 can be defined analogously.

The two agents have heterogeneous beliefs about the drift term $\mu_S(t)$. They also have to agree about the stock price dynamics. We define $\Delta^i(t) = \frac{\mu_S^i(t)-\mu_S(t)}{\sigma_S}$ and $\theta^i(t) = \frac{\mu_S^i(t)-r(t)}{\sigma_S}$ to be the agent specific disagreement process and market price of risk. Therefore the following consistency requirement holds.

$$dS(t) + D(t)dt = S(t)[\mu_S^1(t)dt + \sigma_S dW^1(t)] = S(t)[\mu_S^2(t)dt + \sigma_S dW^2(t)] \tag{2.46}$$

$$dW^2(t) = dW^1(t) - \frac{\mu_S^2(t) - \mu_S^1(t)}{\sigma_S} = dW^1(t) - (\theta^2(t) - \theta^1(t)) \tag{2.47}$$

Agent 1's stochastic discount factor is

$$dM^1(t) = -M^1(t)\Big[r(t)dt + \theta^1(t)dW^1(t)\Big] \tag{2.48}$$

Agent 2's stochastic discount factor is

$$dM^2(t) = -M^2(t)\Big[r(t)dt+\theta^2(t)dW^2(t)\Big] = -M^2(t)\Big[r(t)dt+\theta^2(t)(dW^1(t)-(\theta^2(t)-\theta^1(t))dt)\Big] \tag{2.49}$$

Using the martingale approach, we can derive the wealth ratio

$$k(t) = \frac{u'(c_1(t))}{u'(c_2(t))} = \frac{y_1 M^1(t)}{y_2 M^2(t)} = \frac{w_2(t)}{w_1(t)} \tag{2.50}$$

$$\frac{dk(t)}{k(t)} = [\theta^2(t) - \theta^1(t)]dW(t) \tag{2.51}$$

For simplicity, it is assumed that agent 1 knows the true drift term $\mu_S(t)$ and $\Delta^2(t)$ is also constant. In this case, we have $\theta^2(t) - \theta^1(t) = \Delta^2(t)$. The wealth ratio at time t is

$$k(t) = k(0)\exp[(\theta^2(t) - \theta^1(t))W(t) - \frac{1}{2}(\theta^2(t) - \theta^1(t))^2 t] \tag{2.52}$$

By the strong law of large numbers for Brownian motion

$$\lim_{t\to\infty} \frac{W(t)}{t} = 0 \quad \text{a.s.} \tag{2.53}$$

We can conclude that as long as agent 2 does not have rational beliefs
($\theta^2(t) \neq \theta^1(t)$ or $\Delta^2(t) \neq 0$),

$$\lim_{t\to\infty} k(t) = 0 \quad \text{a.s.} \tag{2.54}$$

It is not surprising that an irrational agent cannot survive in the long run when markets are complete. But it is worth noticing that the long-run effects of the Brownian motion are zero.

2.4.2 Complete Markets

When asset prices can jump, jumps lead to large wealth redistribution as agents have different portfolio choices. The redistribution channel affects the long-run wealth dynamics in the economy. The result is driven by the fact that the law of large numbers for Poisson process dictates that the long-run effects of jumps are not zero.

The mechanism of the redistribution channel can be illustrated with the following proposition that focuses on the case in which agent have constant beliefs.

PROPOSITION 2.4.1. *In complete markets with constant beliefs, an agent with inaccurate beliefs will not survive in the long run.*

Proof. To simplify the analysis, I assume that the true jump intensity of the Poisson process and both agents' beliefs are constant: $\lambda(t) = \lambda$, $\lambda^1(t) = \lambda_1$, $\lambda^2(t) = \lambda_2$. Since agent 1 has rational beliefs, we also have $\lambda^1(t) = \lambda_1 = \lambda$. When beliefs are constant, the wealth ratio in complete markets is governed by

$$\frac{dk(t)}{k(t-)} = (\lambda_1 - \lambda_2)dt + \left(\frac{\lambda_2}{\lambda_1} - 1\right)dN(t) \tag{2.55}$$

the solution to the above stochastic differential equation is

$$k(t) = k(0)\exp\left((\lambda_1 - \lambda_2)t + \log\left(\frac{\lambda_2}{\lambda_1}\right)N(t)\right) \tag{2.56}$$

By the strong law of large numbers for Poisson process (Çinlar (1975)):

$$\lim_{t\to\infty} \frac{N(t)}{t} = \lambda = \lambda_1 \quad \text{a.s.} \tag{2.57}$$

survival is determined by the sign of

$$(\lambda_1 - \lambda_2) + \log\left(\frac{\lambda_2}{\lambda_1}\right)\lambda_1 \tag{2.58}$$

If $\lambda_1 \neq \lambda_2$, the above term is always negative. We can conclude that

$$\lim_{t\to\infty} k(t) = 0 \quad \text{a.s.}$$

To discuss the learning case where agents can learn about the jump intensity of the Poisson process, we need to make use of the theory of stochastic integral for Poisson processes and the law of large numbers for martingales. In the remainder of this paper, the concept of predictability plays an important role.[6]

LEMMA 2.4.1. *If $N(t)$ is a Poisson process with predictable intensity $\lambda(t)$, then $\tilde{N}(t) = N(t) - \int_0^t \lambda(s)ds$ is an $\mathcal{F}(t)$ martingale.*

Proof. $E[\tilde{N}(t)|\mathcal{F}(t_0)] = \tilde{N}(t_0) + E[\tilde{N}(t) - \tilde{N}(t_0)|\mathcal{F}(t_0)] = \tilde{N}(t_0)$

PROPOSITION 2.4.2. *Let $\tilde{N}$ be the martingale defined in lemma 2.4.1 and h is a predictable process. If they satisfy the following condition*

$$E\left[\int_0^t |h(s)|^2 d\tilde{N}(s)\right] < \infty$$

for all $t \geq 0$. Then the process X given by

$$X(t) = \int_0^t h(s) d\tilde{N}(s)$$

is a martingale and

$$\lim_{t\to\infty} \frac{X(t)}{t} = 0 \quad a.s.$$

Proof. See Rosenkrantz and Simha (1992).

Proposition 2.4.2 can be viewed as the law of large numbers for martingales. The martingales in this context are closely linked to the Poisson process: if a stochastic integral is defined based on $\tilde{N}$, then it is a martingale and its long-run effects are negligible.[7]

When agents can learn about the jump intensity, the survival results can be summarized as follows: an agent with irrational beliefs survives if he can learn the truth quickly enough.

PROPOSITION 2.4.3. *1. If the agent with incorrect beliefs cannot learn the truth, then he will be driven out of the market.*

2. If the learning agent can learn the truth quickly enough, then both agents survive. The rate of convergence must satisfy the following condition: $|(\lambda^1(t) - \lambda^2(t)) + \log(\frac{\lambda^2(t)}{\lambda^1(t)})\lambda^1(t)| < \frac{1}{t^\alpha}$ for all $t > t_0$ and $\alpha > 1$.

Proof. See appendix.

[6] Recall that in Ito integration theory, we require the integrands to be adapted. But in point-process integration theory, the requirement is stronger: the integrands have to be predictable.

[7] This does not mean that jumps are negligible or the redistribution channel is not working. We can decompose a stochastic integral into two parts: $\int_0^t h(s)dN(s) = \int_0^t h(s)d\tilde{N}(s) + \int_0^t h(s)\lambda(s)ds$. The first part is negligible; the second part is not.

In complete markets with learning, the wealth ratio at time t is

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s)-\lambda^2(s))ds + \int_0^t log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)dN(s)\right) \tag{2.59}$$

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s)-\lambda^2(s))ds + \int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)ds + \int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)\right) \tag{2.60}$$

Proposition 2.4.2 allows us to focus on the following term which determines long-run survival in complete markets with learning.

$$\int_0^t \lambda^1(s)-\lambda^2(s)+\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)ds \tag{2.61}$$

In the existing literature, a link between long-run survival and relative entropy has already been established. For example, Blume and Easley (2006) show that agents whose beliefs are closest to the truth in terms of relative entropy survive in complete markets when both the states and beliefs are independent and identically distributed. Here a similar interpretation is possible with the help of the following definition.

DEFINITION 2.4.2. *For two probability measures Q and P in a given measure space such that $Q \in \mathcal{Q}$ (Q is absolutely continuous with respect to P), the relative entropy of Q with respect to P is defined as*

$$H = \int log(dQ/dP)dQ \tag{2.62}$$

where dQ/dP is the Radon-Nikodym derivative of Q with respect to P. To simplify notation, let L=dQ/dP. The change in the relative entropy from t to $t+\Delta t$ is

$$H(t,t+\Delta t) = E_t^L[logL(t+\Delta t)] - logL(t) \tag{2.63}$$

The instantaneous growth rate of the relative entropy is

$$R(L(t)) = \lim_{\Delta t\to 0}\frac{H(t,t+\Delta t)}{\Delta t} \tag{2.64}$$

Notice that the integrand in (61) can be decomposed as[8]

$$\left((\lambda(t)-\lambda^2(t))+\log\left(\frac{\lambda^2(t)}{\lambda(t)}\right)\lambda(t)\right) - \left((\lambda(t)-\lambda^1(t))+\log\left(\frac{\lambda^1(t)}{\lambda(t)}\right)\lambda(t)\right)$$

[8]Recall that agent 1 has rational beliefs, $\lambda(t)=\lambda^1(t)$.

where $\left((\lambda(t)-\lambda^2(t))+\log\left(\frac{\lambda^2(t)}{\lambda(t)}\right)\lambda(t)\right)$ is the negative of the instantaneous growth rate of the relative entropy. If we adopt the instantaneous growth rate of the relative entropy as a measure of distance between beliefs, then the natural interpretation of (61) is that long-run survival is determined by the "cumulative" distance of beliefs. More precisely, an agent with irrational beliefs is further away from the truth than a rational agent in the "cumulative" distance sense. An irrational agent can only survive if the integrand in (61) converges to zero quickly enough. If the irrational agent can learn the truth quickly enough, the integral in (61) is finite and the wealth ratio converges to a finite number when t approaches infinity.

The term in (61) can also be connected to portfolio choices. Why does an agent with irrational beliefs always vanish in complete markets? A general answer is that in complete markets irrational agents will allocate more consumption to paths with low probability according to the true model. But in the current setup, a specific answer is that irrational agents will make portfolio choices that deviate from the log optimal rule. And there are some subtle differences between optimists and pessimists even though they both vanish in the long run. For an optimist who believes that the jump intensity is lower than the truth, the first term $\lambda^1(s)-\lambda^2(s)$ in (61) is positive. This is an indication that the optimist holds more risky assets and earn higher expected returns. When there are no jumps, the wealth ratio has a tendency to increase. But the second term $\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)$ in (61) is negative. This indicates that the optimist's wealth decreases when market crashes happen. The optimist can earn higher returns in normal times. But once a disaster strikes, the return becomes negative. The overall effect is that the volatility of the optimist's portfolio is higher and this undermines his long-run survival. The optimist can also bid up the price of the risky asset and effectively depress his portfolio returns. Lower returns will further diminish his chances of survival in the long run. There are two competing forces that shape the prospects of long-run survival for agents in this economy: the first one is belief difference (represented by $\lambda^1(s)-\lambda^2(s)$); the second one is the redistribution channel (represented by $\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)$).

A pessimist, on the other hand, will acquire less risky asset compared to the rational agent. He cannot reap the full benefits of the disaster insurance and earn lower returns from his portfolio. But an advantage of this strategy is that his portfolio returns have lower volatility: he earns lower positive returns during normal times and large positive returns during market crashes. Unfortunately the lower volatility is not enough to compensate for lower returns.

2.4.3 Incomplete Markets

In incomplete markets with constant beliefs, the laws of large numbers for both Poisson process and Brownian motion suggest that the two random sources have different long-run effects: the jumps are not negligible while the Brownian has no impact on wealth dynamics in the long run.

PROPOSITION 2.4.4. *In incomplete markets with constant beliefs, the agent with inaccurate beliefs will not survive in the long run.*

Proof. In incomplete markets, the wealth ratio is governed by

$$\frac{dk(t)}{k(t-)} = [\theta^2(t)^2 - \theta^1(t)\theta^2(t) + (\lambda^1(t) - \lambda^1_{RN}(t)) - (\lambda^2(t) - \lambda^2_{RN}(t))]dt + [\theta^2(t) - \theta^1(t)]dW(t) + \Big(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)} - 1\Big)dN(t) \quad (2.65)$$

The portfolio holdings are

$$\pi_i(t) = \frac{\theta^i(t)}{\sigma_S} = \frac{1}{j^i_S}\Big(\frac{\lambda^i(t)}{\lambda^i_{RN}(t)} - 1\Big) \quad (2.66)$$

The equilibrium interest rate satisfies

$$r(t) = \delta + \mu^i_S(t) - \sigma_S\theta^i(t) + j^i_S\lambda^i_{RN}(t) \quad (2.67)$$

We have

$$\theta^2(t) - \theta^1(t) = \frac{j_S}{\sigma_S}(\lambda^2_{RN}(t) - \lambda^1_{RN}(t))$$

and

$$\theta^2(t) = \frac{\sigma_S}{j_S}\Big(\frac{\lambda^2(t)}{\lambda^2_{RN}(t)} - 1\Big)$$

Therefore

$$\frac{dk(t)}{k(t-)} = \Big[\frac{\lambda^2_{RN}(t))\lambda^1(t) - \lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)}\Big]dt + [\theta^2(t) - \theta^1(t)]dW(t) + \Big(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)} - 1\Big)dN(t) \quad (2.68)$$

then

$$k(t) = k(0)\exp\Bigg(\Big[\frac{\lambda^2_{RN}(t))\lambda^1(t) - \lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} - \frac{1}{2}(\theta^2(t) - \theta^1(t))^2\Big]t + [\theta^2(t) - \theta^1(t)]W(t) + \log\Big(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\Big)N(t)\Bigg) \quad (2.69)$$

By the strong laws of large numbers for Poisson process and Brownian motion, survival is determined by the sign of

$$\left[\frac{\lambda^2_{RN}(t))\lambda^1(t) - \lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} - \frac{1}{2}(\theta^2(t) - \theta^1(t))^2\right] + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t) \qquad (2.70)$$

Notice that

$$\frac{\lambda^2_{RN}(t))\lambda^1(t) - \lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t)$$
$$= \lambda^1(t)\left(\left(1 - \frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right) + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\right)$$

is always negative. We can conclude that

$$\lim_{t\to\infty} k(t) = 0 \quad a.s.$$

Using the the theory of stochastic integral for Poisson processes and the law of large numbers for martingales, we obtain the following results that are similar to those in the complete markets case.

PROPOSITION 2.4.5. *1. If the agent with incorrect beliefs cannot learn the truth, then he will be driven out of the market.*

2. If the learning agent can learn the truth quickly enough, then both agents survive. The rate of convergence must satisfy the following condition: $\left|\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t)\right| < \frac{1}{t^\alpha}$ *for all* $t > t_0$ *and* $\alpha > 1$.

Proof. See appendix.

Why does an agent with irrational beliefs vanish in incomplete markets? The key mechanism is that wealth accumulation and utility maximization coincide when agents have log preferences. An agent survives if his choices are closest to the log optimal rule. In the presence of a rational agent, the irrational agent's portfolio choices are further away from the log optimal rule. Recall that in incomplete markets, agents only have access to a stock and a bond. An optimist will allocate more wealth share to the stock market and earn higher expected returns because of the equity premium. But at the same time, higher portfolio volatility is detrimental to his survival. A pessimist will have lower exposure to the stock market and lower portfolio volatility. But he is missing out on the equity premium. An interesting pattern in the incomplete market wealth dynamics is that the wealth ratio also fluctuates with the Brownian motion . But since the Ito integral is a martingale, the long-run effects of the fluctuations induced by the Brownian motion are still zero. Here the same two competing forces, belief difference and the redistribution channel, determine agents' fate in the long run. In the existing literature, there are plenty of examples in which irrational agents survive. The major insight behind these examples is that irrational agents can save enough to avoid extinction.[9] But in the current setup, the saving channel is no longer active when agents have log preferences and the same discount rate.

2.5 Filtering in Complete Markets

In this section, the primary focus is on a hidden Markov model in which the jump intensity of the Poisson process is governed by a Markov chain. Since the true state is not observable, both agents will rely on filtering techniques. This is an ideal setup to generalize the notion of "rationality" in the market selection hypothesis: an agent endowed with the optimal filter will outperform other agents in financial markets asymptotically. Using the tools from the robust control literature pioneered by Hansen and Sargent (2008), I discuss the wealth dynamics when the economy is populated by two agents: a "Bayesian" agent (agent 1) who is endowed with the optimal filter and a robust agent (agent 2) who relies on robust filtering.[10]

The jump intensity, $\lambda(t)$, is assumed to be governed by a n-state Markov chain in continuous time, with generator matrix $\mathbb{Q}$. Without loss of generality, assume the n states are $\Lambda_1 < \Lambda_2 < ... < \Lambda_n$. Let $P_{ij}(t) = Pr(\lambda(t+s) = \Lambda_j|\lambda(s) = \Lambda_i)$.[11] $\mathbb{P}(t)$ is the transition probability matrix with $P_{ij}(t)$ as its (i,j)-th entry and the time t derivative of the transition probability matrix is $\mathbb{P}'(t) = \mathbb{P}(t)\mathbb{Q}$. The jumps are rare events by definition. Agents in the economy do not have access to enough data to obtain good estimates of the statistical properties of the rare events. The robust agent has 2 types of concerns about model misspecification: (1) Misspecification about the potential states (jump intensities). (2) Misspecification about the transition rates between different states q_{jk} (the entries in the generator matrix $\mathbb{Q}$).

PROPOSITION 2.5.1. *The optimal filter is*

$$\lambda^1(t) = E[\lambda(t)|\mathcal{F}(t)] = \sum_{k=1}^{n} \Lambda_k p_k^1(t) \tag{2.71}$$

[9] An important exception is Blume and Easley(1992). They show that survival is determined by an agent's distance to the log optimal rule if saving is exogenous. An irrational agent can happen to be closer to the log optimal rule.

[10] The robust control literature focuses on model uncertainty which is relevant in a rare disaster context. When agents have heterogeneous beliefs about rare disaster probabilities, it is reasonable to acknowledge that they face substantial model uncertainty in forming beliefs.

[11] I assume that the Markov chain is time homogeneous: the probability $Pr(\lambda(t+s) = \Lambda_j|\lambda(s) = \Lambda_i)$ does not depend on s.

$$dp_k^1(t) = \sum_j q_{jk} p_j^1(t)dt + p_k^1(t)\left(\frac{\Lambda_k - \lambda^1(t)}{\lambda^1(t)}\right)d\eta^1(t) \tag{2.72}$$

$$\eta^1(t) = N(t) - \int_0^t \lambda^1(s)ds \tag{2.73}$$

where $p_k^1(t)$ *denotes the conditional probability that* $\lambda(t) = \Lambda_k$*.*

Proof. See appendix.

The filtering equation in Proposition 2.5.1 is intuitive. Consider a potential state Λ_k that is higher than the current estimate of jump intensity $\lambda^1(t)$. If no jumps are observed, then the second term in (72) becomes negative and the growth of the conditional probability of being in Λ_k decreases as well. If a jump is observed, then the conditional probability of Λ_k increases. The filtering equation implies that investors' beliefs will decrease smoothly as long as they do not observe jumps.

Following Anderson et al. (2003), I define the time derivative of relative entropy as a statistical measure of discrepancy between models.[12] The filtering problem for the robust agent is [13]

$$\min_{\lambda^2(t)} \max_{Q\in\mathcal{Q}} E_Q[(\lambda^2(t) - \lambda(t))^2|\mathcal{F}(t)] \tag{2.74}$$

subject to the time-varying fear constraint

$$R(L(t)) = \lim_{\Delta t\to 0} \frac{H(t, t+\Delta t)}{\Delta t} \leq \eta_t \tag{2.75}$$

where η_t is a time-varying bound on the instantaneous growth rate of the relative entropy. The robust filtering problem dictates that the robust agent finds the best estimate in mean square of the jump intensity under a worst case probability measure induced by the time-varying fear constraint.

If the alternative models are Markov chains as well, the robust agent will be concerned about misspecifications of the bad states and their transition rates.[14] He will consider different values of the states ($\Lambda_k^2 \neq \Lambda_k$) or different transition rates ($q_{jk}^2 \neq q_{jk}$). Time-varying fear puts a constraint on the relative entropy between the reference model and the set of alternative models. The entropy constraint will limit the agent's choices in terms of potential values of Λ_k^2 and q_{jk}^2 under consideration.

[12] See definition 2.4.2.

[13] Hansen and Sargent (2001) refer to this as 'constraint preferences', as opposed to 'penalty preferences'.

[14] This is an example of what Hansen and Sargent (2019) call 'structured uncertainty'.

PROPOSITION 2.5.2. *The robust agent's beliefs evolve according to the robust filter*

$$\lambda^2(t) = E_Q[\lambda(t)|\mathcal{F}(t)] = \sum_{k=1}^{n} \Lambda_k^2 p_k^2(t) \tag{2.76}$$

$$dp_k^2(t) = \sum_j q_{jk}^2 p_j^2(t)dt + p_k^2(t)\bigg(\frac{\Lambda_k^2 - \lambda^2(t)}{\lambda^2(t)}\bigg)d\eta^2(t) \tag{2.77}$$

$$\eta^2(t) = N(t) - \int_0^t \lambda^2(s)ds \tag{2.78}$$

where $p_k^2(t)$ denotes the conditional probability that $\lambda^2(t) = \Lambda_k^2$.

Robust filtering amplifies belief heterogeneity about rare events after a jump is observed. The robust agent experiences a larger revision of his expected jump intensity than the Bayesian agent due to concern for model uncertainty.

Recall that in complete markets , the wealth ratio at time t is

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s) - \lambda^2(s))ds + \int_0^t log\bigg(\frac{\lambda^2(s)}{\lambda^1(s)}\bigg)dN(s)\right) \tag{2.79}$$

Since the agents do not know the true state, the following integrand determines long-run survival:

$$(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda(t) \tag{2.80}$$

One difficulty is that the loss function is based on squared errors in the derivation of the filtering equation. The connection between the squared error loss function and the above integrand is not obvious. In order to discuss wealth dynamics, we need to prove that the same filtering equation can be derived under a different loss function.

PROPOSITION 2.5.3. *Let x be a square-integrable random variable in the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}(t)\}_{t\geq 0}, P)$. Let $\mathcal{G}$ be a sub sigma algebra of $\mathcal{F}$. If we consider the following Poisson loss function*

$$l(x, \hat{x}) = x\log\left(\frac{x}{\hat{x}}\right) - x + \hat{x}$$

and the squared error loss function

$$l_{SE}(x, \hat{x}) = (x - \hat{x})^2$$

then the conditional expectation $E(x|\mathcal{G})$ is the unique estimator minimizing the mean loss under the Poisson loss function l and square errors.

Proof. See appendix.

According to Proposition 2.5.3, the optimal filter can also be derived using the Poisson loss function in the current hidden Markov setup.

Notice that

$$
\begin{aligned}
&(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda(t) \\
&\quad = \left((\lambda^1(t) - \lambda(t)) + \log\left(\frac{\lambda(t)}{\lambda^1(t)}\right)\lambda(t)\right) - \left((\lambda^2(t) - \lambda(t)) + \log\left(\frac{\lambda(t)}{\lambda^2(t)}\right)\lambda(t)\right)
\end{aligned} \tag{2.81}
$$

The integrand is indeed connected to the loss function l.

Another difficulty is that we cannot guarantee that the integrand is always negative. To discuss wealth dynamics, we need to check the pathwise optimality of the optimal filter. The following result from van Handel (2014) establishes that the optimal filter is pathwise optimal in discrete time, given that both the signal and the filter are uniquely ergodic.

PROPOSITION 2.5.4. *In discrete time hidden Markov models, if the signal and filtering processes are both uniquely ergodic, then the filter x^* is pathwise optimal:*

$$
\liminf_{t\to\infty} \frac{1}{t}\sum_{k=1}^{t} l(x_k, \tilde{x}_k) - \frac{1}{t}\sum_{k=1}^{t} l(x_k, x_k^*) > 0 \quad a.s. \tag{2.82}
$$

for every strategy $\tilde{x} \neq x^*$.

Proof. See theorem 3.16 and corollary 3.17 in van Handel (2014).

We can prove the continuous time counterpart using the discrete time result. The key idea is to write the continuous time average as a discrete time average over intervals of length one and a vanishing remainder.

PROPOSITION 2.5.5. *In continuous time hidden Markov models, if the signal and filtering processes are both uniquely ergodic, then the filter x^* is pathwise optimal:*

$$
\liminf_{t\to\infty} \frac{1}{t}\int_0^t l(x_s, \tilde{x}_s)ds - \frac{1}{t}\int_0^t l(x_s, x_s^*)ds > 0 \quad a.s. \tag{2.83}
$$

for every strategy $x \neq x^*$.

Proof. Let $\zeta_k = \int_k^{k+1}(l(x_s, \tilde{x}_s) - l(x_s, x_s^*))ds$ and $\psi_k = \int_k^{k+1} |l(x_s, \tilde{x}_s) - l(x_s, x_s^*)|ds$.

$$
\frac{1}{K}\psi_k \leq \left|\frac{1}{K}\sum_{k=0}^{K-1}\psi_k - \frac{1}{K-1}\sum_{k=0}^{K-2}\psi_k\right| + \frac{1}{K(K-1)}\sum_{k=0}^{K-2}\psi_k \tag{2.84}
$$

$$
\frac{1}{t}\int_0^t (l(x_s, \tilde{x}_s) - l(x_s, x_s^*))ds = \frac{1}{[t]}\int_0^{[t]}(l(x_s, \tilde{x}_s) - l(x_s, x_s^*))ds\frac{[t]}{t} + \frac{1}{t}\int_{[t]}^t (l(x_s, \tilde{x}_s) - l(x_s, x_s^*))ds \tag{2.85}
$$

where $[t]$ is the largest integer less than or equal to t.

We have

$$\frac{1}{[t]}\int_0^{[t]}(l(x_s,\tilde{x}_s)-l(x_s,x_s^*))ds=\frac{1}{[t]}\sum_{k=0}^{[t]-1}\zeta_k \tag{2.86}$$

and

$$\left|\frac{1}{t}\int_{[t]}^{t}(l(x_s,\tilde{x}_s)-l(x_s,x_s^*))ds\right|\leq\frac{1}{[t]}\psi_{[t]} \tag{2.87}$$

we can conclude that

$$\liminf_{t\to\infty}\frac{1}{t}\int_0^t(l(x_s,\tilde{x}_s)-l(x_s,x_s^*))ds>0 \quad \text{a.s.} \tag{2.88}$$

Before proving the main result in this section, we need to make sure that the signal and filter are both uniquely ergodic. van Handel (2009) and van Handel (2014) prove that three conditions must be satisfied.

1. The Markov chain is absolutely regular: its transition kernel must converge to a unique invariant measure in total variation. In the current setup, we need the Markov chain to be irreducible.

2. The observation process is nondegenerate. In the current setup, this is a consequence of Girsanov theorem for Poisson process, as long as the jump intensity does not equal zero.

3. The loss function should be equimeasurable. In the current setup, this is also satisfied as the loss function is continuous.

All three conditions are met if we assume that the Markov chain is irreducible and all the potential jump intensities are positive.

PROPOSITION 2.5.6. *In complete markets, the robust agent vanishes. More generally, if an agent is not endowed with the optimal filter, he will not survive in the long run.*

Proof. In complete markets with filtering, the wealth ratio is governed by

$$k(t)=k(0)\exp\left(\int_0^t(\lambda^1(s)-\lambda^2(s))ds+\int_0^t\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)dN(s)\right) \tag{2.89}$$

$$k(t)=k(0)\exp\left(\int_0^t(\lambda^1(s)-\lambda^2(s))ds+\int_0^t\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda(s)ds+\int_0^t\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)\right) \tag{2.90}$$

According to proposition 2.4.2, $\int_0^t\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)$ is a martingale and $\lim_{t\to\infty}\frac{\int_0^t\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)}{t}$ is zero almost surely.

We have

$$(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda(t)$$
$$= \left((\lambda^1(t) - \lambda(t)) + \log\left(\frac{\lambda(t)}{\lambda^1(t)}\right)\lambda(t)\right) - \left((\lambda^2(t) - \lambda(t)) + \log\left(\frac{\lambda(t)}{\lambda^2(t)}\right)\lambda(t)\right) \quad (2.91)$$

then by pathwise optimality in Proposition 2.5.5

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t (\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda(s)ds < 0 \quad \text{a.s.} \quad (2.92)$$

$$\lim_{t\to\infty} k(t) = 0 \quad \text{a.s.}$$

There are not many results in the literature that discuss the relationship between long-run survival and robust control. A result with a similar flavor to Proposition 2.5.6 is found in Anderson (2005). Anderson (2005) proves that all risk-sensitive agents have zero Pareto weights in the long run when the economy is made up by both agents with time-additive preferences and risk-sensitive agents. The intuition is that a social planning problem with risk-sensitive agents can be viewed as a social planning problem with heterogeneous beliefs in which risk-sensitive agents have irrational beliefs. In Proposition 2.5.6, if an agent is not endowed with the optimal filter, then his time average loss is asymptotically larger. In this sense, the robust agent also has irrational beliefs.

2.6 Conclusion

In this paper, the redistribution channel plays a prominent role. In the aftermath of major market crashes (like the COVID-19 crash or the market crashes in 2008), investors who were less exposed to the stock market might believe that they had made the right decision to avoid large losses. But a natural question is whether they can thrive in the long run. In this paper, it has been established that agents who are less exposed to the stock market are not guaranteed to do well in the long run. This result tells us that we should exercise caution when interpreting short-run trends in inequality as they may not generalize to the long run case. Another insight in this paper is that the market selection hypothesis is surprisingly resilient despite the fact that many results suggest that it does not generally hold in incomplete markets. When agents have log preferences, it holds in both complete and incomplete markets. In complete markets, it is valid even if agents cannot observe the true state and rely on filtering techniques.

Chapter 3

Portfolio Choice and Intergenerational Inequality in China: Theory and Evidence

"Fate has not been kind to China's generation of baby-boomers. It will be even less kind as they grow old."

—The Economist

3.1 Introduction

China's baby boomers are truly the unlucky generation.[1] They are less healthy than later generations as their early years coincided with the Great Leap Forward (the Chinese economy contracted by 27.3% in 1961). Due to the Cultural Revolution, their school years can only be described as unproductive. One cannot help but wonder if those were the only disadvantages of being a baby boomer in China. Using data from the China Household Finance Survey (CHFS), we establish that the baby boomers in China are less wealthy than later generations. They also have a lower willingness to take financial risks: they are less likely to participate in the stock market and invest a lower fraction of their wealth in the stock market.

The millennials in China have very different macroeconomic experiences. They grew up during the reform and opening-up era. The Chinese economy has not experienced a single recession since 1976. College degrees or even advanced degrees are no longer inaccessible. It is therefore not surprising that they outperform the baby boomers in China in terms of health and education. The CHFS data confirm that they exhibit more interest in the stock market in addition to being wealthier than the baby boomers in China.

[1]The generations that interest us are baby boomers (born in 1946-1964), GenXers (born in 1965-1980) and millennials (born in 1981-1996).

What are the connections between macroeconomic experiences, risk-taking behavior and intergenerational wealth dynamics? This paper focuses on generational belief differences and proposes an "experiential learning" channel in return expectations and portfolio choice. More specifically, we argue that different generations have different beliefs about market returns due to their own limited experiences. This influences their risk-taking behavior which, in turn, influences the growth rate of their wealth.

The CHFS is a comprehensive survey about household finance in China. The respondents answer questions about assets holdings, debt, saving, consumption and financial literacy. We use the data from the 2015, 2017 and 2019 waves of the CHFS.[2] The 2015 survey includes 37289 households in 29 provinces and 350 counties. The sample size of the 2017 survey increases to 40011 households from 29 provinces and 350 counties. 34643 households (representing 29 provinces and 345 counties) participate in the 2019 survey. Since the CHFS collects detailed information about household financial decisions, we are able to investigate the differences between the baby boomers and the millenials in China, especially in household wealth and risk-taking behavior.

Why should people be interested in the Chinese data when we already have many papers documenting the portfolio choices of different generations in the US? By comparing China and the US, we notice that the macroeconomic experiences of the two generations are very different. The baby boomers in the US skipped the Great Depression and enjoyed the prosperity of the post-war years. The millenials in the US, on the other hand, experienced the Great Recession and the COVID-19 pandemic within two decades.

[2]The CHFS has been conducted since 2011. We focus on recent data for representativeness and completeness.

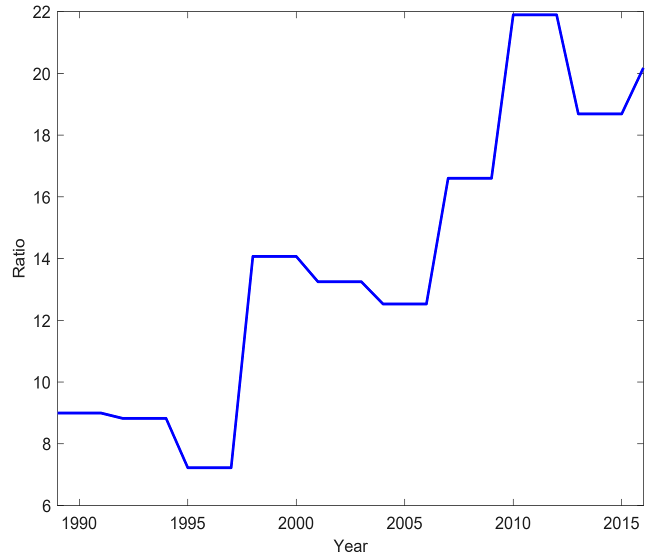

Figure 3.1: Median Net Worth Ratio of 65 and over vs. 35 and under in the US (Survey of Consumer Finances)

Figure 1 plots Survey of Consumer Finances data on the ratio of median net worth for those over 65 years of age to those under 35. For simplicity, we will refer to the households in which the household head is older than 65 years as the old generation. Accordingly, the households with household head under 35 years old are the young generation. In the US, the old have always been wealthier than the young. In 1989 their net worth was 9 times greater on average. However, over the course of the next 27 years this ratio more than doubled, to over 20.

Figure 2 shows the ratio of median net worth for those over 65 years of age to those under 35 in the CHFS. In China, the young households are wealthier than the old households and the gap seems to be widening as well. Figure 1 and Figure 2 paint an interesting picture: the old generation has been doing really well in the US while their counterparts in China have been lagging behind. What factors are responsible for the staggering differences in the two countries? The data reveal that generational belief differences induced by “experiential learning” seem to be at work in both countries.

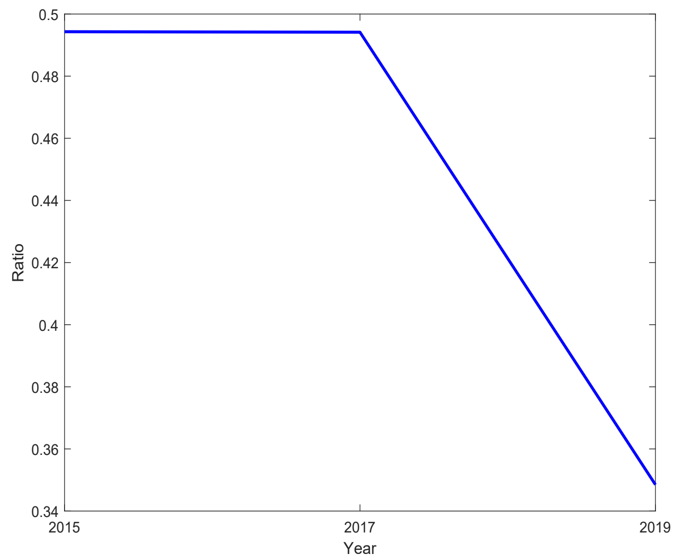

Figure 3.2: Median Net Worth Ratio of 65 and over vs. 35 and under in China (China Household Finance Survey)

The "experiential learning" channel is first introduced in Malmendier and Nagel (2011). A crucial piece of evidence in their paper is that differences in stock market participation rates between different generations in the US are positively correlated with differences in experienced stock market returns. For comparison purposes, we calculate experienced returns for households in China and plot them against the differences in their stock market participation rates in Figure 3.[3] A superficial conclusion would be that the correlation is not strong in the Chinese data. But we cannot overlook the fact that the Chinese stock market did not come into existence until 1990. Weighted returns will not accurately capture the old generation's macroeconomic experiences due to the short history of the stock market in China.

To overcome the shortcomings of weighted returns in China, we use average real GDP growth rates as a proxy for different generations' macroeconomic experiences. More precisely, we calculate the average real GDP growth rate in the past 55 years for the old generation and the average real GDP growth rate in the past 20 years for the young generation.

[3]Following Malmendier and Nagel (2011), we calculate weighted returns with a weighting function that put more weights on recent returns.

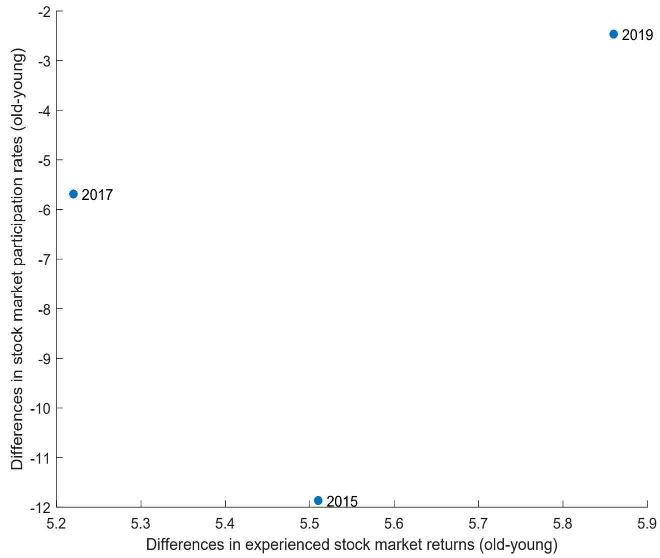

Figure 3.3: Differences in Stock Market Participation Rates of Old and Young Individuals Plotted against Differences in Experienced Stock Market Returns in China (in percentage points)

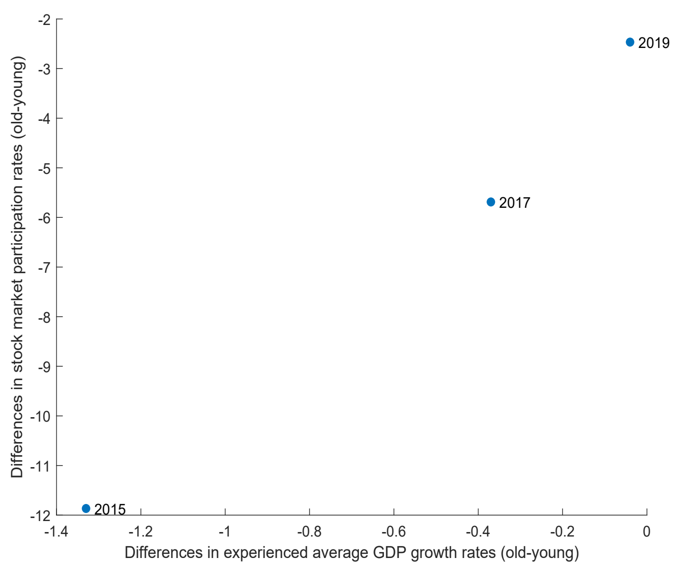

Figure 3.4: Differences in Stock Market Participation Rates of Old and Young Individuals Plotted against Differences in Experienced Average GDP Growth Rates in China (in percentage points)

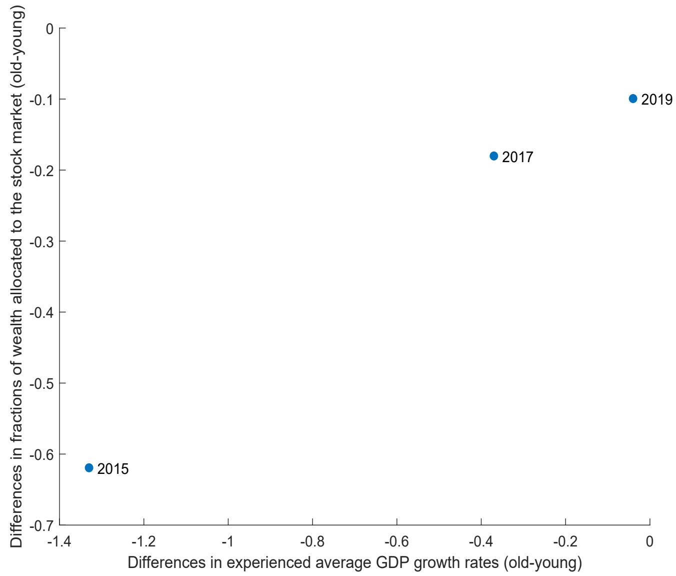

Figure 3.5: Differences in Fractions of Wealth Allocated to Stocks of Old and Young Individuals Plotted against Differences in Experienced Average GDP Growth Rates in China (in percentage points)

In Figure 4 and Figure 5, we plot differences in stock market participation rates and exposures to the stock market against differences in experienced average GDP growth rates in China. Stock market participation and exposures seem to be closely connected to macroeconomic experiences in the Chinese data. The patterns are consistent with the intuition that economic disasters determine people's attitude towards stock investments. The generations that skipped severe economic disasters are more interested in the stock market. The unlucky generations that experienced severe recessions become more pessimistic. Due to their lower exposures to the stock market, they cannot fully benefit from the equity premium. If they had been more optimistic about the stock market, their wealth would have been increasing more rapidly.

3.2 Literature Review

This paper is related to four strands of literature. First, it is largely inspired by the recent macro literature that examines the implications of deviations from rational expectations. As shown in a seminal paper by Woodford (2013), although the literature hasn't reached an unequivocal verdict regarding what expectation formation rules researchers should adopt, a promising approach that relies on a statistically modest deviation from rational expectations is to assume that beliefs are refined through induction from observed history. The over-weighing of personal experiences has long been discussed in the psychology literature, named as availability bias as in Tversky and Kahneman (1974). Compared with a full Bayesian approach, such belief formation mechanism exhibits strong over extrapolation behavior (see Greenwood and Shleifer (2014) for a survey). Barberis, Greenwood, Jin, and Shleifer (2015) and Barberis, Greenwood, Jin, and Shleifer (2018) rationalize a set of asset pricing anomalies when an over-extrapolative investor interact with a rational agent in the financial market. Evidence of over extrapolation is pervasive. In financial markets, it is supported by a seminal paper Malmendier and Nagel (2011), who uses data from Survey of Consumer Finances and provides strong empirical support that personal experience in the stock market has a prolonged impact on how much people invest in risky assets later in their lives. In particular, those that experienced the 1930s great depression were less willing to participate in the stock market, and invest significantly less even if they participate. Such belief formation is not only present in the stock market, but also influences households' expectation formation of inflation, labor market, housing market as well as overall business cycle conditions (see 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. Acedanski (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 this paper embeds 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 experience. Most macro-finance models with heterogeneous beliefs focus on exogenous heterogeneous beliefs. Classic work includes Basak (2005), Harrison and Kreps (1979), Scheinkman and Xiong (2003) and Borovicka (2020), just to name a few. Since their focus is on asset prices, belief heterogeneity could be taken as an input without having to model where it comes from. In this paper, beliefs are essentially endogenous, which helps to link observable demographic structures with inequality. Nevertheless, this is not the first paper to do so. Some recent papers have studied the aggregate implications of heterogeneous generational bias induced by learning from experience. The fact that younger people update their beliefs more frequently than the old has interesting implications on asset prices. Ehling, Graniero, and Heyerdahl-Larsen (2017) develops 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.

Third, this paper is related to recent literature on disaster risk in the tradition of Barro (2006). The incorporation of risk of rare disasters naturally generates a disaster premium, which significantly reduces the level of risk aversion needed in matching empirically plausible equity premium. Various extensions of disaster risk models also helps to solve the equity premium puzzle, the volatility puzzle, and return predictability (see Tsai and Wachter (2015) for a survey). When disaster risk is unknown and agents must infer its distribution from historical data, Koulovatianos and Wieland (2011) shows that pessimism is triggered upon the realization of a rare disaster, and rationalizes a prolonged period of decline in P-D ratio. Moreover, they prove that although asymptotic beliefs are unbiased, one never reaches full optimism of disaster risk as one would under rational expectation. It is the slow arrival of information of disasters that keeps learning away from reaching infinite precision. In this paper, the realization of a large negative shock (e.g., the Great Leap Forward) would trigger such response from investors that experienced it, thus generating heterogeneous generational bias in the disaster risk distribution. Although there are several interesting papers that combines heterogeneous beliefs or attitudes towards disaster risk in both complete and incomplete markets (see Bates (2008), Chen, Joslin, and Tran (2010), Dieckmann (2011), Chen, Joslin, and Tran (2012)), these models only include two agents and focus on cases with dogmatic beliefs, while the model in this paper features a continuum of heterogeneous agents with learning agents that constantly update their beliefs optimally, and focus on the evolution of wealth distribution. Last but not least, this paper contributes to the recent advancement of HACT (heterogeneous agent continuous time) models that link distributional considerations with macroeconomics (Gabaix, Lasry, Lions, and Moll (2016), Achdou, Han, Lasry, Lions, and Moll (2017) and Ahn, Kaplan, Moll, Winberry, and Wolf (2018). However, studying belief heterogeneity in such framework is still a relatively new area. Two recent papers attempt to incorporate endogenous heterogeneous beliefs into such a framework (Kasa and Lei (2018), Lei (2019)), and rationalize "state dependence" in the growth rate of wealth, which rationalizes why inequality has been growing at such a fast speed after 1980s. However, they focus on inequality within cohort with private equities. Here, we generalize those models, and are able to solve distribution across cohort, and solve a model with aggregate shock and public equity.

3.3 The Model

The model combines a Lucas (1978) pure exchange 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.

3.3.1 Environment

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

$$\frac{dY_t}{Y_{t^-}} = \mu dt + \sigma dZ_t + \kappa_t dN_t(\lambda_t) \tag{3.1}$$

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 $(\Omega^N, \mathcal{F}^N, \mathcal{P}^N)$. We then define $(\Omega, \mathcal{F}, \mathcal{P})$ as the product probability space, and the filtration of the combined history as $\{\mathcal{F}_t\} = \{\mathcal{F}^B \times \mathcal{F}^N\}$. The jump process N_t follows

$$dN_t = \begin{cases} 1, & \text{with probability } \lambda_t dt. \\ 0, & \text{with probability } 1 - \lambda_t dt. \end{cases} \tag{3.2}$$

That is, at each instant, there is λ_t probability that a disaster happens. We assume that the jump size $\kappa_t \in (-1, 0)$, 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,

$$\lambda_t = \begin{cases} \lambda_h, & \text{with probability } \pi^*. \\ \lambda_l, & \text{with probability } 1 - \pi^*. \end{cases} \tag{3.3}$$

We assume that the market is dynamically complete, and that investors can trade continuously in the capital market to hedge against both regular economic risk, as well as disaster risk. 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

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

follows

$$dB_t = r_t B_t dt \tag{3.4}$$

The risky asset value follows

$$\frac{dS_t + D_t dt}{S_{t^-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{3.5}$$

where r_t, μ_t^S, σ^S as well as κ_t^S are endogenous objects, and are determined in equilibrium. Finally, the disaster-contingent security value is P_t, and follows the stochastic process

$$\frac{dP_t}{P_{t^-}} = \mu_t^P dt + \kappa_t^P dN_t(\lambda_t) \tag{3.6}$$

This asset is in zero net supply. By convention, we 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.[5]

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.

3.3.2 Filtering and Information Processing

Investors have common knowledge about the size of the disaster. However, they remain uncertain about the likelihood of disasters. They must revise their beliefs sequentially, in real-time. 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

[5]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, we focus on the benchmark complete market setting.

$\bar{\lambda}_{s,t} = \mathbb{E}_{s,t}[\lambda_t]$, where the expectation $\mathbb{E}_{s,t}[\lambda_t] = \pi_{s,t}\lambda_h + (1-\pi_{s,t})\lambda_l$ denotes the expectation with respect to the time s born agent's own filtration $\mathcal{P}_{s,t}$ at time t. We will specify how the prior is chosen in the quantitative section. For now, let us focus on belief updating.

LEMMA 3.3.1. *The evolution of the beliefs about π^* by a Bayesian learning agent (denoted by $\pi_{s,t}$) is given by*

$$d\pi_{s,t}|_{dN_t=0} = -(\lambda_h - \lambda_l)\pi_{s,t}(1-\pi_{s,t})dt \tag{3.7}$$

$$d\pi_{s,t}|_{dN_t=1} = \frac{\lambda_h \pi_{s,t}}{\bar{\lambda}_{s,t}} - \pi_{s,t} \tag{3.8}$$

Proof. This is a direct application of the optimal filtering of a jump-diffusion process from Liptser 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 disaster occurring is suddenly amplified by a magnitude of $\frac{\lambda_h}{\bar{\lambda}_{s,t}}$.[6]

3.3.3 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. [7] They start with zero financial wealth, and accumulate wealth over the life cycle. An annuity contract *à 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

$$\max_{c_{s,t},\alpha^S_{s,t},\alpha^P_{s,t}} \mathbb{E}_{s,t}\left[\int_s^\infty e^{-(\rho+\delta)(t-s)} \log{(c_{s,t})} dt\right] \tag{3.9}$$

[6] One might argue that Bayesian learning is contradicted by evidence of a 'recency bias'. That is, it is debatable whether agents weight past observations of disasters in a statistically optimal manner. However, since we are 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.

[7] As we shall see later, log preferences deliver two key advantages. First, they imply a constant savings rate, which allows us 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.

s.t:

$$\frac{dw_{s,t}}{w_{s,t^-}} = \left(r_t + \delta + \alpha^S_{s,t}(\mu^S_t - r_t) + \alpha^P_{s,t}(\mu^P_t - r_t) + y_{s,t} - \frac{c_{s,t}}{w_{s,t^-}}\right)dt + \alpha^S_{s,t}\sigma^S dZ_{s,t} \\ + (\alpha^S_{s,t}\kappa^S_t + \alpha^P_{s,t}\kappa^P_t)dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.10}$$

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, we exploit the fact that the market is dynamically complete for all cohorts. This allows us to employ the martingale approach (Cox and Huang (1989)). This allows us to convert the dynamic programming problem into a static problem as follows

$$\max_{c_{s,s}} \mathbb{E}_{s,s}\left[\int_s^\infty e^{-(\rho+\delta)(t-s)}\log(c_{s,t})dt\right] \tag{3.11}$$

s.t:

$$\mathbb{E}_{s,s}\left[\int_s^\infty e^{-\delta(t-s)}\xi_{s,t}c_{s,t}\right] = \mathbb{E}_{s,s}\left[\int_s^\infty e^{-\delta(t-s)}\xi_{s,t}\omega Y_t dt\right] \tag{3.12}$$

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

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

$$\frac{e^{-(\rho+\delta)(t-s)}}{c_{s,t}} = y_s e^{-\delta(t-s)}\xi_{s,t} \tag{3.13}$$

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

$$c_{s,t} = c_{s,s}e^{-\rho(t-s)}\frac{\xi_{s,s}}{\xi_{s,t}} \tag{3.14}$$

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

LEMMA 3.3.2. *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*

$$\frac{d\xi_{s,t}}{\xi_{s,t^-}} = (\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t)dt - \theta_{s,t}dZ_{s,t} + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right)dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.15}$$

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

$$\frac{d\xi_t}{\xi_{t^-}} = (\bar{\lambda}_t - \lambda_t^N - r_t)dt - \theta_t dZ_t + \left(\frac{\lambda_t^N}{\bar{\lambda}_t} - 1\right) dN_t(\bar{\lambda}_t) \tag{3.16}$$

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

$$\frac{d\eta_{s,t}}{\eta_{s,t^-}} = \left(\frac{1}{1+\bar{\kappa}}\lambda_{s,t} - \lambda_t^N\right) dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2\lambda_t^N}{\lambda_t} - 1\right) - 1\right] dN(\bar{\lambda}_t) \tag{3.17}$$

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

Proof. See appendix.

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

$$\frac{dc_{s,t}}{c_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_{s,t}^N + r_t)dt + \theta_{s,t} dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.18}$$

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

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_{s,t}^N + r_t)dt + \theta_{s,t} dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.19}$$

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

3.3.4 Aggregation

We start by defining a Walrasian equilibrium in this economy.

DEFINITION 3.3.1. *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.,*

$$Y_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} c_{s,t} ds \tag{3.20}$$

$$S_t = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha_{s,t}^S w_{s,t} ds \tag{3.21}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} \alpha_{s,t}^P w_{s,t} ds \tag{3.22}$$

$$0 = \int_{-\infty}^{t} \delta e^{-\delta(t-s)} (1 - \alpha_{s,t}^S - \alpha_{s,t}^P) w_{s,t} ds \tag{3.23}$$

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 $\beta_t = \frac{c_{t,t}}{Y_t} = \beta$. [8] We can then rewrite the goods market clearing condition as

$$\xi_t Y_t = \int_{-\infty}^{t} \beta \delta e^{-(\rho+\delta)(t-s)} \xi_s Y_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{3.24}$$

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

$$\eta_t = \int_{-\infty}^{t} \beta \delta e^{-\beta\delta(t-s)} \eta_s \frac{\eta_{s,t}}{\eta_{s,s}} ds \tag{3.25}$$

Defining $\mu_{s,t}^\eta$ and $\kappa_{s,t}^\eta$ 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

$$\frac{d\eta_t}{\eta_t} = \bar{\mu}_t^\eta dt + \bar{\kappa}_t^\eta dN_t(\bar{\lambda}_t) \tag{3.26}$$

where the weighted average coefficients are defined as

$$\bar{\mu}_t^\eta = \mathbb{E}_{s,t}(\mu_{s,t}^\eta) = \int_{-\infty}^{t} f_{s,t} \mu_{s,t}^\eta ds; \qquad \bar{\kappa}_t^\eta = \mathbb{E}_{s,t}(\kappa_{s,t}^\eta) = \int_{-\infty}^{t} f_{s,t} \kappa_{s,t}^\eta ds \tag{3.27}$$

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

$$f_{s,t} = \beta \delta e^{-\beta\delta(t-s)} \left(\frac{\eta_s}{\eta_t}\right) \left(\frac{\eta_{s,t}}{\eta_{s,s}}\right) = \delta e^{-\delta(t-s)} \frac{c_{s,t}}{Y_t} \tag{3.28}$$

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

$$\frac{d\xi_t}{\xi_t} = \Big(\bar{\mu}_t^\eta - \mu + \sigma^2 - \rho - \delta(1-\beta)\Big) dt - \sigma dZ_t + \left(\frac{1+\bar{\kappa}_\eta}{1+\bar{\kappa}} - 1\right) dN_t(\lambda_t) \tag{3.29}$$

Since we know that the state price density also has to follow eqn.(3.16), it directly gives the solution of equilibrium prices.

[8] Appendix C.2 verifies this conjecture, and derives an explicit expression for β.

PROPOSITION 3.3.1. *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*

$$r_t = \underbrace{\rho + \delta(1-\beta)}_{\text{effective patience with OLG}} + \underbrace{\mu - \sigma^2}_{\text{risk adjusted growth}} + \underbrace{\frac{\bar{\kappa}}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t})}_{\text{market view of disaster risk}} \quad ; \tag{3.30}$$

$$\theta_t = \theta = \sigma; \tag{3.31}$$

$$\lambda_t^N = \frac{\mathbb{E}_{s,t}(\lambda_{s,t})}{1+\bar{\kappa}} \tag{3.32}$$

The closed form solutions for prices have intuitive interpretations. Let's start with the equilibrium interest rate. As always, the risk free rate increases when agents are less patient. In a world of finite lives, the effective patience lessens due to death risk. Moreover, the equilibrium interest rate increases when the endowment process has a higher rate of growth and a lower volatility, which is captured in the second term. The third term reflects a flight to safty 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 asset during disasters drives down the return on the safe asset, leading to a low equilibrium interest rates during disaster episodes, as observed in the data (see Nakamura, Steinsson, Barro, and Ursùa (2013)). Notice that the first and second term 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.

3.3.5 Portfolio Allocations and Wealth Dynamics

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

$$\frac{dw_{s,t}}{w_{s,t^-}} = (\theta_{s,t}^2 - \bar{\lambda}_{s,t} + \lambda_t^N + r_t)dt + \theta_{s,t}dZ_{s,t} + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_{s,t}^N} - 1\right)dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.33}$$

Recall also that the budget constraint follows

$$\frac{dw_{s,t}}{w_{s,t^-}} = \left(r_t + \alpha^S_{s,t}(\mu^S_t - r_t) + \delta + \alpha^P_{s,t}(\mu^P_t - r_t) + y_{s,t} - \frac{c_{s,t}}{w_{s,t^-}}\right) dt + \alpha^S_{s,t}\sigma^S dZ_{s,t} \\ + (\alpha^S_{s,t}\kappa^S_t + \alpha^P_{s,t}\kappa^P_t) dN_{s,t}(\bar{\lambda}_{s,t}) \tag{3.34}$$

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

$$\alpha^S_{s,t} = \frac{\theta_{s,t}}{\sigma^S} = \frac{\theta_t}{\sigma^S} \tag{3.35}$$

$$\alpha^P_{s,t} = \frac{1}{\kappa^P_t}\left(\frac{\bar{\lambda}_{s,t}}{\lambda^N_t} - 1\right) - \frac{\kappa^S_t \theta_t}{\kappa^P_t \sigma^S} \tag{3.36}$$

Notice that all generations invest the same fraction of wealth in risky asset. 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^S_t, σ^S, κ^S_t and κ^P_t.

3.3.6 Equity Premium Dynamics

PROPOSITION 3.3.2. *The equilibrium coefficients in the risky asset price and the disaster-contingent security are given by*

$$\sigma^S = \sigma \tag{3.37}$$

$$\kappa^S_t = \kappa_t \tag{3.38}$$

$$\mu^S_t - r_t = \sigma^2 + \bar{\mu}^\eta_t \tag{3.39}$$

$$\mu^P_t - r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t}) \tag{3.40}$$

Proof. See appendix.

The model produces an endogenous time-varying equity premium, both for the risky asset as well as for the disaster-contingent security. When market pessimism rises, risky asset 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 of safe asset then generates a decline in equilibrium interest rate, which then increases equity premium. 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 asset and the disaster-contingent security.

COROLLARY 3.3.1. *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*

$$\alpha^S_{s,t} = 1 \tag{3.41}$$

$$\alpha^P_{s,t} = \frac{1}{\bar{\kappa}}\left(\frac{\bar{\lambda}_{s,t}}{\mathbb{E}(\bar{\lambda}_{s,t})}(1+\bar{\kappa}) - 1\right) - 1 \tag{3.42}$$

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

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

3.3.7 Evolution of the Joint Age-Wealth Distribution

This subsection 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 compares its properties relative to the rational expectation economy.

PROPOSITION 3.3.3. *The dynamics of the joint distribution of wealth and belief $n(w,\lambda)$ follows*

$$dn = -\frac{\partial}{\partial w}(n\hat{\mu}wdt + n\hat{\sigma}wdZ) + \frac{1}{2}\frac{\partial^2}{\partial w^2}(n\hat{\sigma}^2w^2)dt + [n(w(1+\hat{\kappa}),t)) - n(w,t)]dN \tag{3.43}$$

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

$$p(x) \approx \underbrace{Ge^{\zeta_0 x}}_{RE}\underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h-\lambda^0)\zeta_1 x} - e^{(\lambda_l-\lambda^0)\zeta_1 x}]}_{Learning} \tag{3.44}$$

where ζ_0 and ζ^1 are constants. Moreover,

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^{RE}(x) \tag{3.45}$$

Proof. See appendix.

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

We need to emphasize one feature of the experiential learning economy called the "scale dependence" of wealth accumulation (see Gabaix, Lasry, Lions, and Moll (2016)). In this economy, the older households 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 are hit the most, because they have less life time experience, and would therefore over-extrapolate information from the disaster. Therefore, "scale dependence" is even stronger during disaster times.

3.4 Calibration

In this section, we calibrate the above model to the Chinese 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 3.1: Benchmark Parameter Values

Parameters	Value	Source
ρ	1%	Empirical estimate 1%-2%, chosen to match interest rate
δ	1.67%	Average trading life expectancy of 60 years
ω	0.92	Dividend income share from NIPA
μ	8.76%	Average real GDP growth rate in China
σ	4.94%	Real GDP growth rate volatility in China
κ	-0.273	Drop in Chinese real GDP in 1961
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)

The birth and death rate $\delta = 1.67\%$ is calibrated such that the average trading life is from 20 to 80 years old, implying an average trading life expectancy of 60 years. The parameter ω follows from Garleanu and Panageas (2015), which is chosen to match the fraction of capital income from the total income in the US. The drift coefficient μ and volatility coefficient σ are estimated using real GDP data in China. The two hazard rates $\lambda^H = 24\%$ and $\lambda^L = 1.5\%$ represent the upper and lower bounds of disaster probabilities, 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 an international sample of 35 countries over 100 years in Barro (2006). The jump size κ is chosen to match the real GDP drop in 1961. Next, empirical estimates of discount rate are around 1% to 2%. However, a 2% discount rate generates a model implied interest rate that is too high compared with the data. Therefore, we set $\rho = 1\%$. Finally, we assume that all agents start with a fixed prior that is equal to the Rational Expectations value.

Using the above parameters, we 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. We simulate the economy with 8000 initial agents for 2000 years. Each year, each living agent is endowed with ω fraction of aggregate endowment, and the wealth share weighted average of prices are 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, who are endowed with zero financial wealth but a fixed fraction of aggregate dividend, and their beliefs are reset to the prior in the next period. For surviving agents, their beliefs and wealth are updated. Prices are again computed by the wealth weighted average, and the process carries on for 2000 years. 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 1951, the starting point of our calibration. Next, we assume that only one disaster happened after 1951. In 1961, the Great Leap Forward reduces the output by κ. We then re-run the simulation for 68 years to examine the response of the wealth distribution between 1951 to 2019.

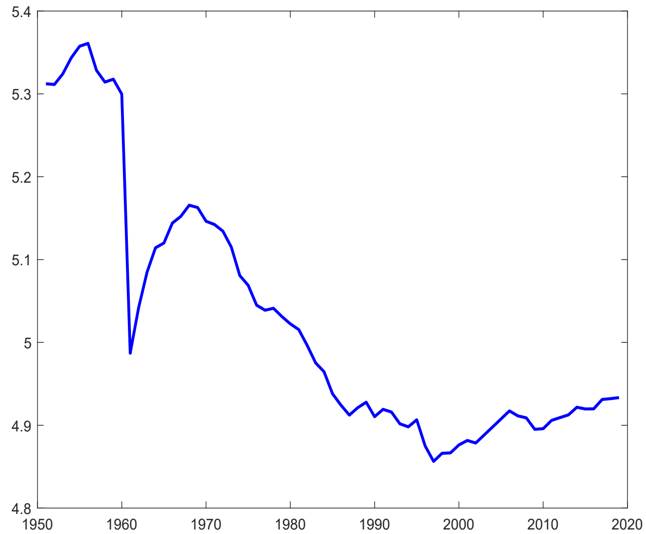

Figure 3.6: Calibrated Path of Old to Young Wealth Ratio

Figure 3.6 plots the calibrated path of the old to young wealth ratio (65 and over vs. 35 and under). After the initial disaster, there is a gradual decrease of old to young wealth ratio that lasts about 30 years. This reflects the lingering “belief scarring” effect. As time goes by, the young people that experienced the Great Leap Forward 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 disaster.

3.5 Conclusion

The CHFS data reveal that the baby boomers are in general poorer than the millennials in China. In addition to wealth inequality, the two generations have very different risk-taking behavior. Their stock market participation and exposures are highly correlated with their experienced average GDP growth rates. The empirical evidence suggests that experiential learning is an important channel in determining intergenerational wealth dynamics in the economy. With the help of a model that combines rare disasters and overlapping generations, we demonstrate that learning from macroeconomic experiences can have long-lasting effects on portfolio choices and wealth accumulation.

Bibliography

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

Acedański, J., 2017. Heterogeneous expectations and the distribution of wealth. Journal of Macroeconomics, 53, pp.162-175.

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

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

Adam, K., Marcet, A. and Nicolini, J.P., 2016. Stock market volatility and learning. Journal of Finance, 71(1), pp.33-82.

Ahn, D., Choi, S., Gale, D. and Kariv, S., 2014. Estimating ambiguity aversion in a portfolio choice experiment. Quantitative Economics, 5(2), pp.195-223.

Ahn, S., Kaplan, G., Moll, B., Winberry, T. and Wolf, C., 2018. When inequality matters for macro and macro matters for inequality. NBER macroeconomics annual, 32(1), pp.1-75.

Aït-Sahalia, Y., Cacho-Diaz, J. and Hurd, T.R., 2009. Portfolio choice with jumps: A closed-form solution. Annals of Applied Probability, 19(2), pp.556-584.

Aït-Sahalia, Y. and Matthys, F., 2019. Robust consumption and portfolio policies when asset prices can jump. Journal of Economic Theory, 179, pp.1-56.

Altonji, J.G., Hayashi, F. and Kotlikoff, L.J., 1997. Parental altruism and inter vivos transfers: Theory and evidence. Journal of Political Economy, 105(6), pp.1121-1166.

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

Ameriks, J. and Zeldes, S.P., 2004. How do household portfolio shares vary with age. Working paper, Columbia University.

Anderson, E.W., 2005. The dynamics of risk-sensitive allocations. Journal of Economic Theory, 125(2), pp.93-150.

Anderson, E.W., Hansen, L.P. and Sargent, T.J., 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. Journal of the European Economic Association, 1(1), pp.68-123.

Arrow, K.J., 1964. The role of securities in the optimal allocation of risk-bearing. Review of Economic Studies, 31(2), pp.91-96.

Atar, R. and Weissman, T., 2012. Mutual information, relative entropy, and estimation in the Poisson channel. IEEE Transactions on Information theory, 58(3), pp.1302-1318.

Banerjee, S. and Kremer, I., 2010. Disagreement and learning: Dynamic patterns of trade. Journal of Finance, 65(4), pp.1269-1302.

Barber, B.M. and Odean, T., 2001. Boys will be boys: Gender, overconfidence, and common stock investment. Quarterly Journal of Economics, 116(1), pp.261-292.

Barberis, N., Greenwood, R., Jin, L. and Shleifer, A., 2015. X-CAPM: An extrapolative capital asset pricing model. Journal of Financial Economics, 115(1), pp.1-24.

Barberis, N., Greenwood, R., Jin, L. and Shleifer, A., 2018. Extrapolation and bubbles. Journal of Financial Economics, 129(2), pp.203-227.

Bardhan, I. and Chao, X., 1996. Stochastic multi-agent equilibria in economies with jump-diffusion uncertainty. Journal of Economic Dynamics and Control, 20(1-3), pp.361-384.

Barro, R.J., 2006. Rare disasters and asset markets in the twentieth century. Quarterly Journal of Economics, 121(3), pp.823-866.

Barro, R.J., 2009. Rare disasters, asset prices, and welfare costs. American Economic Review, 99(1), pp.243-64.

Barro, R.J. and Ursua, J.F., 2008. Macroeconomic crises since 1870. Brookings Papers on Economic Activity.

Barro, R.J. and Ursua, J.F., 2017. Stock-market crashes and depressions. Research in Economics, 71(3), pp.384-398.

Barro, R.J. and Ursua, J.F. and Weng, J., 2020. The coronavirus and the Great Influenza epidemic: Lessons from the "Spanish Flu" for the coronavirus' potential effects on mortality and economic activity. NBER Working Paper.

Basak, S., 2005. Asset pricing with heterogeneous beliefs. Journal of Banking and Finance, 29(11), pp.2849-2881.

Basak, S. and Croitoru, B., 2000. Equilibrium mispricing in a capital market with portfolio constraints. Review of Financial Studies, 13(3), pp.715-748.

Basak, S. and Cuoco, D., 1998. An equilibrium model with restricted stock market participation. Review of Financial Studies, 11(2), pp.309-341.

Bates, D.S., 2008. The market for crash risk. Journal of Economic Dynamics and Control, 32(7), pp.2291-2321.

Bäuerle, N. and Rieder, U., 2007. Portfolio optimization with jumps and unobservable intensity process. Mathematical Finance, 17(2), pp.205-224.

Beker, P. and Chattopadhyay, S., 2010. Consumption dynamics in general equilibrium: A characterisation when markets are incomplete. Journal of Economic Theory, 145(6), pp.2133-2185.

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

Bhandari, A., Borovička, J. and Ho, P., 2017. Survey data and subjective beliefs in business cycle models. manuscript, University of Minnesota.

Blackwell, D. and Dubins, L., 1962. Merging of opinions with increasing information. The Annals of Mathematical Statistics, 33(3), pp.882-886.

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

Blume, L.E., Cogley, T., Easley, D.A., Sargent, T.J. and Tsyrennikov, V., 2018. A case for incomplete markets. Journal of Economic Theory, 178, pp.191-221.

Blume, L. and Easley, D., 1992. Evolution and market behavior. Journal of Economic Theory, 58(1), pp.9-40.

Blume, L. and Easley, D., 2006. If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets. Econometrica, 74(4), pp.929-966.

Boar, C., 2020. Dynastic precautionary savings. NBER Working Paper.

Bordalo, P., Coffman, K.B., Gennaioli, N. and Shleifer, A., 2020. Older people are less pessimistic about the health risks of COVID-19. NBER Working Paper.

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

Bossaerts, P., Ghirardato, P., Guarnaschelli, S. and Zame, W.R., 2010. Ambiguity in asset markets: Theory and experiment. Review of Financial Studies, 23(4), pp.1325-1359.

Branger, N., Konermann, P. and Schlag, C., 2020. Optimists and pessimists in (in)complete markets. Journal of Financial and Quantitative Analysis, 55(8), pp.2466-2499.

Brémaud, P., 1981. Point Processes and Queues: Martingale Dynamics. New York: Springer-Verlag.

Brokaw, T., 2000. The greatest generation. Random House.

Brunnermeier, M.K., Simsek, A. and Xiong, W., 2014. A welfare criterion for models with distorted beliefs. Quarterly Journal of Economics, 129(4), pp.1753-1797.

Buraschi, A. and Jiltsov, A., 2006. Model uncertainty and option markets with heterogeneous beliefs. Journal of Finance, 61(6), pp.2841-2897.

Cagetti, M., Hansen, L.P., Sargent, T. and Williams, N., 2002. Robustness and pricing with uncertain growth. Review of Financial Studies, 15(2), pp.363-404.

Campbell, J.Y., 2003. Consumption-based asset pricing. Handbook of the Economics of Finance, 1, pp.803-887.

Campbell, J.Y. and Viceira, L.M., 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. OUP Oxford.

Cao, H.H. and Ou-Yang, H., 2008. Differences of opinion of public information and speculative trading in stocks and options. Review of Financial Studies, 22(1), pp.299-335.

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), pp.199-217.

Cecchetti, S.G., Lam, P.S. and Mark, N.C., 2000. Asset pricing with distorted beliefs: are equity returns too good to be true?. American Economic Review, 90(4), pp.787-805.

Charles, K.K. and Hurst, E., 2003. The correlation of wealth across generations. Journal of political Economy, 111(6), pp.1155-1182.

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

Chen, H., Joslin, S. and Tran, N.K., 2012. Rare disasters and risk sharing with heterogeneous beliefs. Review of Financial Studies, 25(7), pp.2189-2224.

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

Cho, I.K. and Kasa, K., 2018. Doubts, inequality, and bubbles. Working paper, Simon Fraser University.

Chordia, T., Roll, R. and Subrahmanyam, A., 2011. Recent trends in trading activity and market quality. Journal of Financial Economics, 101(2), pp.243-263.

Çinlar, E., 1975. Introduction to Stochastic Processes. Prentice-Hall.

Cochrane, J.H., 2002. Stocks as money: Convenience yield and the tech-stock bubble . NBER Working Paper.

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

Cogley, T. and Sargent, T.J., 2009. Diverse beliefs, survival and the market price of risk. The Economic Journal, 119(536), pp.354-376.

Cogley, T., Sargent, T.J. and Tsyrennikov, V., 2014. Wealth dynamics in a bond economy with heterogeneous beliefs. The Economic Journal, 124(575), pp.1-30.

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

Colacito, R., Croce, M.M. and Liu, Z., 2019. Recursive allocations and wealth distribution with multiple goods: Existence, survivorship, and dynamics. Quantitative Economics, 10(1), pp.311-351.

Collin-Dufresne, P., Johannes, M. and Lochstoer, L.A., 2017. Asset pricing when 'this time is different'. Review of Financial Studies, 30(2), pp.505-535.

Constantinides, G.M., 2008. Comment on 'Macroeconomic Crises since 1870'. Brookings Papers on Economic Activity, 1, pp.341-50.

Coury, T. and Sciubba, E., 2012. Belief heterogeneity and survival in incomplete markets. Economic Theory, 49(1), pp.37-58.

Cox, J.C. and Huang, C.F., 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. Journal of Economic Theory, 49(1), pp.33-83.

Cvitanic, J. and Karatzas, I., 1992. Convex duality in constrained portfolio optimization. Annals of Applied Probability, 2(4), pp.767-818.

De Nardi, M., French, E. and Benson, D., 2012. Consumption and the Great Recession. Economic Perspectives, Federal Reserve Bank of Chicago, .36(1), pp.1-16.

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

Dillenberger, D. and Rozen, K., 2015. History-dependent risk attitude. Journal of Economic Theory, 157, pp.445-477.

Dindo, P., 2019. Survival in speculative markets. Journal of Economic Theory, 181, pp.1-43.

Drechsler, I., 2013. Uncertainty, time-varying fear, and asset prices. Journal of Finance, 68(5), pp.1843-1889.

Dumas, B., 1989. Two-person dynamic equilibrium in the capital market. Review of Financial Studies, 2(2), pp.157-188.

Dumas, B., Kurshev, A. and Uppal, R., 2009. Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. Journal of Finance, 64(2), pp.579-629.

Easley, D. and Yang, L., 2015. Loss aversion, survival and asset prices. Journal of Economic Theory, 160, pp.494-516.

Edwards, C. and Bourne, R., 2019. Exploring wealth inequality. Policy Analysis No. 881, Cato Institute.

Ehling, P., Graniero, A. and Heyerdahl-Larsen, C., 2018. Asset prices and portfolio choice with learning from experience. Review of Economic Studies, 85(3), pp.1752-1780.

Epstein, L.G. and Schneider, M., 2007. Learning under ambiguity. Review of Economic Studies, 74(4), pp.1275-1303.

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

Favilukis, J., 2013. Inequality, stock market participation, and the equity premium. Journal of Financial Economics, 107(3), pp.740-759.

Gabaix, X., 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. Quarterly Journal of Economics, 127(2), pp.645-700.

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

Gale, W.G., Gelfond, H., Fichtner, J.J. and Harris, B.H., 2020. The wealth of generations, with special attention to the millennials. NBER Working Paper.

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

Gayer, G., Gilboa, I., Samuelson, L. and Schmeidler, D., 2014. Pareto efficiency with different beliefs. The Journal of Legal Studies, 43(S2), pp.S151-S171.

Gennotte, G., 1986. Optimal portfolio choice under incomplete information. The Journal of Finance, 41(3), pp.733-746.

Giglio, S., Maggiori, M., Stroebel, J. and Utkus, S., 2019. Five facts about beliefs and portfolios. NBER Working Paper.

Gilboa, I. and Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics, 18(2), pp.141-153.

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

Gomez, M., 2016. Asset prices and wealth inequality. Unpublished paper: Princeton University.

Gourio, F., 2012. Disaster risk and business cycles. American Economic Review, 102(6), pp.2734-66.

Greenwood, R. and Shleifer, A., 2014. Expectations of returns and expected returns. Review of Financial Studies, 27(3), pp.714-746.

Guerdjikova, A. and Sciubba, E., 2015. Survival with ambiguity. Journal of Economic Theory, 155, pp.50-94.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, pp.357-384.

Hammond, P.J., 1981. Ex-ante and ex-post welfare optimality under uncertainty. Economica, 48(191), pp.235-250.

Hansen, L.P. and Sargent, T.J., 2001. Robust control and model uncertainty. American Economic Review, 91(2), pp.60-66.

Hansen, L.P. and Sargent, T.J., 2008. Robustness. Princeton University Press.

Hansen, L.P. and Sargent, T.J., 2010. Fragile beliefs and the price of uncertainty. Quantitative Economics, 1(1), pp.129-162.

Hansen, L.P. and Sargent, T.J., 2011. Wanting Robustness in Macroeconomics. Handbook of Monetary Economics, 3, pp.1097-1157.

Hansen, L.P. and Sargent, T.J., 2011. Robustness and ambiguity in continuous time. Journal of Economic Theory, 146(3), pp.1195-1223.

Hansen, L.P. and Sargent, T.J., 2019. Structured uncertainty and model misspecification. Working Paper.

Hansen, L.P., Sargent, T.J., Turmuhambetova, G. and Williams, N., 2006. Robust control and model misspecification. Journal of Economic Theory, 128(1), pp.45-90.

Hanson, F.B., 2007. Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation. SIAM.

Harris, M. and Raviv, A., 1993. Differences of opinion make a horse race. Review of Financial Studies, 6(3), pp.473-506.

Harris, R., 1978. Ex-post efficiency and resource allocation under uncertainty. The Review of Economic Studies, 45(3), pp.427-436.

Harrison, J.M. and Kreps, D.M., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. Quarterly Journal of Economics, 92(2), pp.323-336.

Harrison, J.M. and Kreps, D.M., 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory, 20(3), pp.381-408

Harrison, J.M. and Pliska, S.R., 1981. Martingales and stochastic integrals in the theory of continuous trading. Stochastic Processes and Their Applications, 11(3), pp.215-260.

He, H. and Pearson, N.D., 1991. Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. Journal of Economic Theory, 54(2), pp.259-304.

Heimer, R.Z., Myrseth, K.O.R. and Schoenle, R.S., 2019. YOLO: Mortality beliefs and household finance puzzles. Journal of Finance, 74(6), pp.2957-2996.

Hill, J.M., van der Geest, V.R. and Blokland, A.A., 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), pp.419-439.

Hong, H. and Stein, J.C., 2003. Differences of opinion, short-sales constraints, and market crashes. Review of Financial Studies, 16(2), pp.487-525.

Hong, H. and Stein, J.C., 2007. Disagreement and the stock market. Journal of Economic Perspectives, 21(2), pp.109-128.

Huang, J. and Wang, J., 2008. Liquidity and market crashes. Review of Financial Studies, 22(7), pp.2607-2643.

Jagannathan, R., McGrattan, E. and Scherbina, A., 2001. The declining US equity premium. NBER Working Paper.

Jin, X., Luo, D. and Zeng, X., 2021. Tail risk and robust portfolio decisions. Management Science, 67(5), pp.3254-3275.

Jouini, E. and Napp, C., 2007. Consensus consumer and intertemporal asset pricing with heterogeneous beliefs. Review of Economic Studies, 74(4), pp.1149-1174.

Meyn, S.P. and Tweedie, R.L., 1993. Stability of Markovian processes II: Continuous-time processes and sampled chains. Advances in Applied Probability, 25(3), pp.487-517.

Kandel, E. and Pearson, N.D., 1995. Differential interpretation of public signals and trade in speculative markets. Journal of Political Economy, 103(4), pp.831-872

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

Karatzas, I., Lehoczky, J.P., Shreve, S.E. and Xu, G.L., 1991. Martingale and duality methods for utility maximization in an incomplete market. SIAM Journal on Control and Optimization, 29(3), pp.702-730.

Kasa, K., 2006. Robustness and information processing. Review of Economic Dynamics, 9(1), pp.1-33.

Kasa, K. and Lei, X., 2018. Risk, uncertainty, and the dynamics of inequality. Journal of Monetary Economics, 94, pp.60-78.

Klibanoff, P., Marinacci, M. and Mukerji, S., 2005. A smooth model of decision making under ambiguity. Econometrica, 73(6), pp.1849-1892.

Kogan, L., Ross, S.A., Wang, J. and Westerfield, M.M., 2017. Market selection. Journal of Economic Theory, 168, pp.209-236.

Kotlikoff, L.J. and Burns, S., 2005. The coming generational storm: What you need to know about America's economic future. MIT press.

Koulovatianos, C. and Wieland, V., 2011. Asset pricing under rational learning about rare disasters.

Kozlowski, J., Veldkamp, L. and Venkateswaran, V., 2020. The tail that wags the economy: Beliefs and persistent stagnation. Journal of Political Economy, 128(8), pp.2839-2879.

Krusell, P. and Smith, Jr, A.A., 1998. Income and wealth heterogeneity in the macroeconomy. Journal of Political Economy, 106(5), pp.867-896.

Kuchler, T. and Zafar, B., 2019. Personal experiences and expectations about aggregate outcomes. Journal of Finance, 74(5), pp.2491-2542.

Kuhn, M., Schularick, M. and Steins, U.I., 2020. Income and wealth inequality in america, 1949–2016. Journal of Political Economy, 128(9), pp.3469-3519.

Liptser, R.S., Shiryaev, A.N., 2001. Statistics of Random Processes II: Applications. Springer.

Lei, X., 2019. Information and inequality. Journal of Economic Theory, 184, p.104937.

Lei, X., 2020. Experiential Learning and Intergenerational Inequality. Working paper, University of Oxford.

Lo, A.W. and Wang, J., 2010. Stock market trading volume. In Handbook of Financial Econometrics: Applications (pp. 241-342). Elsevier.

Longstaff, F.A. and Piazzesi, M., 2004. Corporate earnings and the equity premium. Journal of Financial Economics, 74(3), pp.401-421.

Lucas Jr, R.E., 1978. Asset prices in an exchange economy. Econometrica, 46, pp.1429-1445.

Maccheroni, F., Marinacci, M. and Rustichini, A., 2006. Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica, 74(6), pp.1447-1498.

Malmendier, U. and Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk taking?. Quarterly Journal of Economics, 126(1), pp.373-416.

Malmendier, U. and Nagel, S., 2016. Learning from inflation experiences. Quarterly Journal of Economics, 131(1), pp.53-87.

Malmendier, U., Pouzo, D. and Vanasco, V., 2020. Investor experiences and financial market dynamics. Journal of Financial Economics, 136(3), pp.597-622.

Malmendier, U. and Shen, L.S., 2018. Scarred consumption. NBER Working Paper.

Mankiw, N.G. and Weil, D.N., 1989. The baby boom, the baby bust, and the housing market. Regional Science and Urban Economics, 19(2), pp.235-258.

Martin, I.W., 2013. Consumption-based asset pricing with higher cumulants. Review of Economic Studies, 80(2), pp.745-773.

Milgrom, P. and Stokey, N., 1982. Information, trade and common knowledge. Journal of Economic Theory, 26(1), pp.17-27.

Merton, R.C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics, 8(4), pp.323-361.

Morris, S., 1996. Speculative investor behavior and learning. Quarterly Journal of Economics, 111(4), pp.1111-1133.

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

Nagel, S. and Xu, Z., 2019. Asset pricing with fading memory. NBER Working Paper.

Nakamura, E., Steinsson, J., Barro, R. and Ursua, J., 2013. Crises and recoveries in an empirical model of consumption disasters. American Economic Journal: Macroeconomics, 5(3), pp.35-74.

Nakov, A. and Nuño, G., 2015. Learning from experience in the stock market. Journal of Economic Dynamics and Control, 52, pp.224-239.

Øksendal, B. and Sulem, A., 2019. Applied Stochastic Control of Jump Diffusions. Springer.

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

Poterba, J.M., 2001. Demographic structure and asset returns. Review of Economics and Statistics, 83(4), pp.565-584.

Rietz, T.A., 1988. The equity risk premium a solution. Journal of Monetary Economics, 22(1), pp.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), pp.1-69.

Rosenkrantz, W.A. and Simha, R., 1992. Some theorems on conditional Pasta: A stochastic integral approach. Operations Research Letters, 11(3), pp.173-177.

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

Sandroni, A., 2000. Do markets favor agents able to make accurate predictions?. Econometrica, 68(6), pp.1303-1341.

Sbuelz, A. and Trojani, F., 2008. Asset prices with locally constrained-entropy recursive multiple-priors utility. Journal of Economic Dynamics and Control, 32(11), pp.3695-3717.

Scheinkman, J.A. and Xiong, W., 2003. Overconfidence and speculative bubbles. Journal of Political Economy, 111(6), pp.1183-1220.

Schraeder, S., 2016. Information processing and non-bayesian learning in financial markets. Review of Finance, 20(2), pp.823-853.

Shleifer, A. and Vishny, R., 2011. Fire sales in finance and macroeconomics. Journal of Economic Perspectives, 25(1), pp.29-48.

Skiadas, C., 2003. Robust control and recursive utility. Finance and Stochastics, 7(4), pp.475-489.

Skiadas, C., 2007. Dynamic portfolio choice and risk aversion. Handbooks in Operations Research and Management Science, 15, pp.789-843.

Skiadas, C., 2013. Smooth ambiguity aversion toward small risks and continuous-time recursive utility. Journal of Political Economy, 121(4), pp.775-792.

Starr, R.M., 1973. Optimal production and allocation under uncertainty. Quarterly Journal of Economics, 87(1), pp.81-95.

Statman, M., Thorley, S. and Vorkink, K., 2006. Investor overconfidence and trading volume. Review of Financial Studies, 19(4), pp.1531-1565.

Strzalecki, T., 2011. Axiomatic foundations of multiplier preferences. Econometrica, 79(1), pp.47-73.

Tirole, J., 1982. On the possibility of speculation under rational expectations. Econometrica, pp.1163-1181.

Tong, X.T. and van Handel, R., 2012. Ergodicity and stability of the conditional distributions of nondegenerate Markov chains. Annals of Applied Probability, 22(4), pp.1495-1540.

Tsai, J. and Wachter, J.A., 2015. Disaster risk and its implications for asset pricing. Annual Review of Financial Economics, 7, pp.219-252.

Tversky, A. and Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science, 185(4157), pp.1124-1131.

van Handel, R., 2009. The stability of conditional Markov processes and Markov chains in random environments. Annals of Probability, 37(5), pp.1876-1925.

van Handel, R., 2012. On the exchange of intersection and supremum of σ-fields in filtering theory. Israel Journal of Mathematics, 192(2), pp.763-784.

van Handel, R., 2014. Ergodicity, decisions, and partial information. In Séminaire de Probabilités XLVI (pp. 411-459). Springer.

Wachter, J.A., 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility?. Journal of Finance, 68(3), pp.987-1035.

Walden, J. and Heyerdahl-Larsen, C., 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.

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

Wolff, E.N., 2021. Household wealth trends in the United States, 1962 to 2019: Median wealth rebounds... But not enough. NBER Working Paper.

Woodford, M., 2013. Macroeconomic analysis without the rational expectations hypothesis. Annu. Rev. Econ., 5(1), pp.303-346.

Yaari, M.E., 1965. Uncertain lifetime, life insurance, and the theory of the consumer. Review of Economic Studies, 32(2), pp.137-150.

Yan, H., 2008. Natural selection in financial markets: Does it work?. Management Science, 54(11), pp.1935-1950.

Appendix A

Fear and Trading

A.1 Proofs

Proof of Proposition 1.3.1.

Proof.

$$Pr(\lambda(t+h) = \Lambda_k | N(t+h) - N(t) = 0) = \frac{Pr(\lambda(t+h) = \Lambda_k, N(t+h) - N(t) = 0)}{Pr(N(t+h) - N(t) = 0)} \quad \text{(A.1)}$$

$$Pr(\lambda(t+h) = \Lambda_k, N(t+h) - N(t) = 0) = \sum_{j \neq k} q_{jk} \hat{p}_j(t) h + \hat{p}_k(t)(1 - q_{kk} h - \Lambda_k h) + o(h) \quad \text{(A.2)}$$

Sum the above equation over all possible states, we get

$$Pr(N(t+h) - N(t) = 0) = 1 - \sum_k \hat{p}_k(t) \Lambda_k h + o(h) \quad \text{(A.3)}$$

If no jumps happen in this interval, then

$$\hat{p}_k(t+h) = Pr(\lambda(t+h) = \Lambda_k | N(t+h) - N(t) = 0) \quad \text{(A.4)}$$

$$\lim_{h \to 0} \frac{\hat{p}_k(t+h) - \hat{p}_k(t)}{h} = \sum_j q_{jk} \hat{p}_j(t) - \hat{p}_k(t)(\Lambda_k - \hat{\lambda}(t)) \quad \text{(A.5)}$$

$$d\hat{p}_k(t) = \sum_j q_{jk} \hat{p}_j(t) dt - \hat{p}_k(t)(\Lambda_k - \hat{\lambda}(t)) dt \quad \text{(A.6)}$$

If one jump happens in this interval, then

$$\hat{p}_k(t+h) = Pr(\lambda(t+h) = \Lambda_k | N(t+h) - N(t) = 1) = \frac{\hat{p}_k(t) \Lambda_k h}{\hat{\lambda}(t) h} = \frac{\hat{p}_k(t) \Lambda_k}{\hat{\lambda}(t)} \quad \text{(A.7)}$$

$$\hat{p}_k(t+h) - \hat{p}_k(t) = \hat{p}_k(t) \frac{\Lambda_k - \hat{\lambda}(t)}{\hat{\lambda}(t)} \quad \text{(A.8)}$$

If we let $h \to 0$, then right after the moment when jump happens, we add the increment in (A.8) to (A.6) to obtain (5).

Proof of Proposition 1.3.3.

Proof.

$$c_i(t) = I(y_i M^i(t)) = \frac{1}{y_i M^i(t)} \tag{A.9}$$

$$y_i = \frac{T}{M^i(0) w_i(0)} \tag{A.10}$$

$$w_i(t) = \frac{E^i[\int_t^T M^i(s) c_i(s) ds]}{M^i(t)} \tag{A.11}$$

$$c_i(t) = \frac{w_i(t)}{T-t} \tag{A.12}$$

$$c_1(t) + c_2(t) = D(t) \tag{A.13}$$

$$w_1(t) + w_2(t) = S(t) \tag{A.14}$$

$$S(t) = D(t)(T-t) \tag{A.15}$$

Using Ito's lemma, we get

$$dS(t) + D(t)dt = S(t-)[\mu_D dt + \phi\sigma dW(t) + j_D dN^i(t)] \tag{A.16}$$

Proof of Proposition 1.3.4.

Proof. (42) is obtained by matching the coefficients associated with the Brownian motion and the Poisson jump terms in the two budget constraints (31) and (33).

We can plug the condition

$$c_i(t) = (u')^{-1}(y_i M^i(t)) = \frac{1}{y_i M^i(t)} \tag{A.17}$$

into the minimax problem, and we obtain

$$\min_{\theta^i, \lambda^i_{RN}} E^i\Big[\int_0^T -log(M^i(t))dt\Big] \tag{A.18}$$

Use the expression for $M^i(t)$ in (20), this simplifies to a pointwise minimization problem:

$$\min_{\theta^i, \lambda^i_{RN}} \left[r(t) + 0.5\theta^i(t)^2 - (\lambda^i(t) - \lambda^i_{RN}(t)) - log\Big(\frac{\lambda^i_{RN}(t)}{\lambda^i(t)}\Big)\lambda^i(t) \right] \tag{A.19}$$

Solve the above minimization problem, we obtain expressions (43) and (44).

Proof of Proposition 1.3.5.

Proof.
$$k(t) = \frac{u'(c_1(t))}{u'(c_2(t))} = \frac{y_1 M^1(t)}{y_2 M^2(t)} = \frac{w_2(t)}{w_1(t)} \tag{A.20}$$

$$k(t) = \frac{1/c_1(t)}{1/c_2(t)} \tag{A.21}$$

and

$$c_1(t) + c_2(t) = D(t) \tag{A.22}$$

We obtain

$$c_1(t) = \frac{D(t)}{1+k(t)} \tag{A.23}$$

$$c_2(t) = \frac{D(t)k(t)}{1+k(t)} \tag{A.24}$$

The representative agent's utility is

$$log(c_1(t)) + k(t)log(c_2(t)) \tag{A.25}$$

The differential form of the stochastic discount factor is

$$dM^i(t) = -M^i(t-)[r(t)dt + \theta^i(t)dW^i(t) - (\lambda^i(t) - \lambda^i_{RN}(t))dt - \left(\frac{\lambda^i_{RN}(t)}{\lambda^i(t)} - 1\right)dN^i(t)] \tag{A.26}$$

Using Ito's lemma, we get

$$\begin{aligned} \frac{dk(t)}{k(t-)} = {} & [\theta^2(t)^2 - \theta^1(t)\theta^2(t) + (\lambda^1(t) - \lambda^1_{RN}(t)) - (\lambda^2(t) - \lambda^2_{RN}(t))]dt + [\theta^2(t) - \theta^1(t)]dW(t) \\ & + \left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)} - 1\right)dN^i(t) \end{aligned} \tag{A.27}$$

We know that the representative agent's marginal utility $\frac{1+k(t)}{D(t)}$ should equal agent 1's stochastic discount factor from the envelope theorem. This gives us another expression for the evolution of agent 1's stochastic discount factor as $dM^1(t) = d\left(\frac{1+k(t)}{D(t)}\right)$.

Matching the corresponding terms associated with the drift term, the volatility term and the jump term leads to the expressions in Proposition 1.3.5.

Proof of Proposition 1.3.6.

Proof. (56) and (57) are obtained by matching the coefficients associated with the Brownian motion and the Poisson jump terms in the two budget constraints (31) and (54).

Agreement on the prices dictate that $\theta^1(t) = \theta^2(t)$ and $\lambda^1_{RN}(t) = \lambda^2_{RN}(t)$. Therefore we obtain

$$dM^i(t) = -M^i(t-)[r(t)dt + \theta(t)dW^i(t) - (\lambda^i(t) - \lambda_{RN}(t))dt - \left(\frac{\lambda_{RN}(t)}{\lambda^i(t)} - 1\right)dN^i(t)] \quad (A.28)$$

Using Ito's lemma, we get

$$\frac{dk(t)}{k(t-)} = [\lambda^1(t) - \lambda^2(t)]dt + \left(\frac{\lambda^2(t)}{\lambda^1(t)} - 1\right)dN^i(t) \quad (A.29)$$

The representative agent's marginal utility should equal agent 1's stochastic discount factor.

The expressions in Proposition 6 can be obtained by matching the relevant terms.

A.2 Additional Tables and Graphs

Largest Daily Percentage Decreases in the S&P 500 Index				
Rank	Date	Value Weighted Return	Equal Weighted Return	Recession
1	1987-10-19	-19.46	-18.42	N
2	2020-03-16*	-11.98	-11.98	Y
3	1929-10-28	-11.68	-10.86	Y
4	1929-10-29	-10.03	-11.53	Y
5	1929-11-06	-9.70	-9.22	Y
6	2020-03-12*	-9.51	-9.51	Y
7	1937-10-18	-9.26	-9.85	Y
8	2008-10-15	-9.00	-9.87	Y
9	1933-07-21	-8.99	-11.17	N
10	1933-07-20	-8.90	-9.85	N
11	2008-12-01	-8.86	-10.08	Y
12	2008-09-29	-8.71	-9.20	Y
13	1932-10-05	-8.39	-9.80	Y
14	1987-10-26	-8.29	-9.05	N
15	1932-08-12	-8.29	-8.21	Y
16	1934-07-26	-7.97	-10.16	N
17	1931-09-24	-7.83	-7.91	Y
18	1932-05-31	-7.74	-8.24	Y
19	2020-03-09*	-7.60	-7.60	Y
20	2008-10-09	-7.57	-8.21	Y
21	1932-09-12	-7.47	-9.66	Y
22	1940-05-14	-7.19	-8.87	N

Table A.1: Market Crashes in the US. Source: CRSP. The ranking is based on value weighted return. The returns on the dates marked by * are returns on the S&P 500 index.

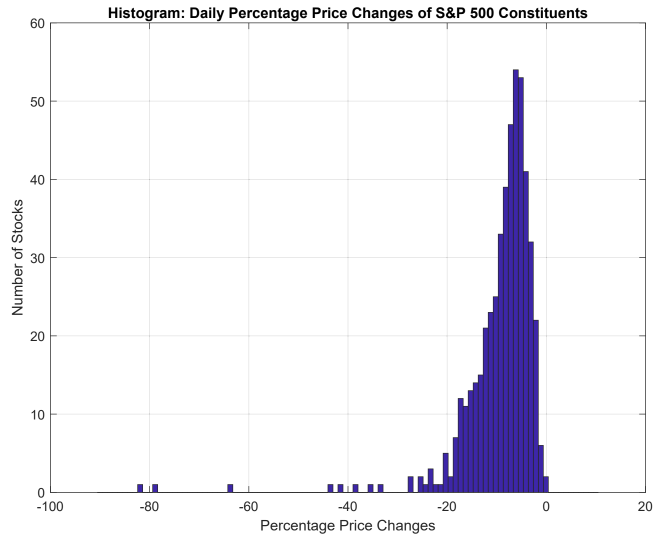

Figure A.1: The Market Crash on September 29, 2008: Price Changes

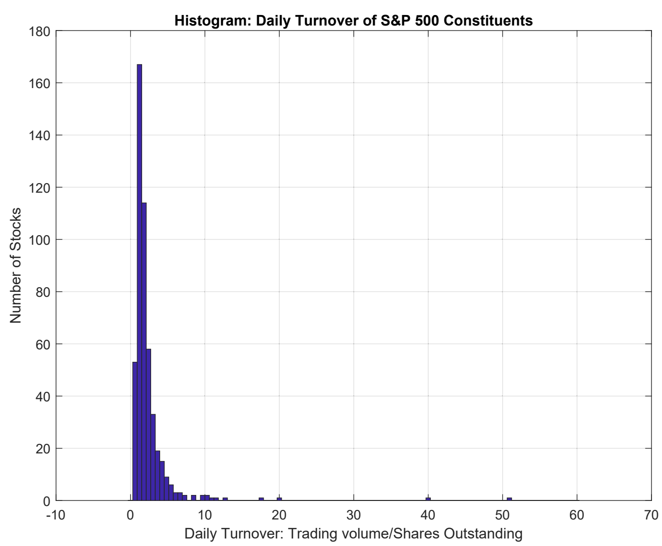

Figure A.2: The Market Crash on September 29, 2008: Daily Turnover

Appendix B

The Market Selection Hypothesis and Rare Disasters

Proof of Proposition 2.3.1.

Proof.

$$c_i(t) = I(y_i M^i(t) e^{\delta t}) = \frac{1}{y_i M^i(t) e^{\delta t}} \tag{B.1}$$

$$y_i = \delta w_i(0) \tag{B.2}$$

$$w_i(t) = \frac{E^i[\int_t^\infty M^i(s) c_i(s) ds]}{M^i(t)} = \frac{1}{\delta^2 w_i(0) M^i(t) e^{\delta t}} \tag{B.3}$$

$$c_i(t) = w_i(t)\delta \tag{B.4}$$

$$c_1(t) + c_2(t) = D(t) \tag{B.5}$$

$$w_1(t) + w_2(t) = S(t) \tag{B.6}$$

$$S(t) = \frac{D(t)}{\delta} \tag{B.7}$$

Using Ito's lemma, we get

$$dS(t)dt = S(t-)[\mu_D dt + \sigma_D dW(t) + j_D dN^i(t)] \tag{B.8}$$

Proof of Proposition 2.3.2.

Proof.

$$k(t) = \frac{u'(c_1(t))}{u'(c_2(t))} = \frac{y_1 M^1(t)}{y_2 M^2(t)} = \frac{w_2(t)}{w_1(t)} \tag{B.9}$$

$$k(t) = \frac{1/c_1(t)}{1/c_2(t)} \tag{B.10}$$

and

$$c_1(t) + c_2(t) = D(t) \tag{B.11}$$

We obtain

$$c_1(t) = \frac{D(t)}{1 + k(t)} \tag{B.12}$$

$$c_2(t) = \frac{D(t)k(t)}{1 + k(t)} \tag{B.13}$$

Following Basak and Cuoco (1998), we construct a representative agent that consumes the aggregate dividend. The representative agent's utility is

$$log(c_1(t)) + k(t)log(c_2(t)) \tag{B.14}$$

The differential form of the stochastic discount factor is

$$dM^i(t) = -M^i(t-)[r(t)dt + \theta^i(t)dW^i(t) - (\lambda^i(t) - \lambda^i_{RN}(t))dt - \left(\frac{\lambda^i_{RN}(t)}{\lambda^i(t)} - 1\right)dN^i(t)] \tag{B.15}$$

Using Ito's lemma, we get

$$\frac{dk(t)}{k(t-)} = [\theta^2(t)^2 - \theta^1(t)\theta^2(t) + (\lambda^1(t) - \lambda^1_{RN}(t)) - (\lambda^2(t) - \lambda^2_{RN}(t))]dt + [\theta^2(t) - \theta^1(t)]dW(t)$$
$$+ \left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)} - 1\right)dN^i(t) \tag{B.16}$$

We know that the representative agent's marginal utility $e^{-\delta t}\frac{1+k(t)}{D(t)}$ should equal agent 1's stochastic discount factor from the envelope theorem. This gives us another expression for the evolution of agent 1's stochastic discount factor as $dM^1(t) = d\left(e^{-\delta t}\frac{1+k(t)}{D(t)}\right)$.

Matching the corresponding terms associated with the drift term, the volatility term and the jump term leads to the expressions in Proposition 2.3.2.

Proof of Proposition 2.3.3.

Proof. (36) and (37) are obtained by matching the coefficients associated with the Brownian motion and the Poisson jump terms in the two budget constraints (18) and (34).

Agreement on the prices dictate that $\theta^1(t) = \theta^2(t)$ and $\lambda^1_{RN}(t) = \lambda^2_{RN}(t)$. Therefore we obtain

$$dM^i(t) = -M^i(t-)[r(t)dt + \theta(t)dW^i(t) - (\lambda^i(t) - \lambda_{RN}(t))dt - \left(\frac{\lambda_{RN}(t)}{\lambda^i(t)} - 1\right)dN^i(t)] \tag{B.17}$$

Using Ito's lemma, we get

$$\frac{dk(t)}{k(t-)} = [\lambda^1(t) - \lambda^2(t)]dt + \left(\frac{\lambda^2(t)}{\lambda^1(t)} - 1\right)dN^i(t) \tag{B.18}$$

The representative agent's marginal utility should equal agent 1's stochastic discount factor. The expressions in Proposition 2.3.3 can be obtained by matching the relevant terms.

Proof of Proposition 2.4.3

Proof. Part 1. In complete markets with learning, the wealth ratio is governed by

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s) - \lambda^2(s))ds + \int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)dN(s)\right) \tag{B.19}$$

$$k(t) = k(0)\exp\left(\int_0^t (\lambda^1(s) - \lambda^2(s))ds + \int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)ds + \int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)\right) \tag{B.20}$$

According to proposition 2.4.2, $\int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)$ is a martingale and $\lim_{t\to\infty} \frac{\int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)}{t}$ is zero almost surely.

We assume that $|\lambda^1(t) - \lambda^2(t)| < \eta_1$ for all t and $\lim_{t\to\infty} \lambda^1(t) - \lambda^2(t) = \epsilon_1$.

If $\lambda^1(t) - \lambda^2(t)$ is bounded, so is $\log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t)$. It is therefore natural to assume that $\left|\log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t)\right| < \eta_2$ for all t and $\lim_{t\to\infty} \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t) = \epsilon_2$.

For all $\epsilon > 0$, there exists t_1 such that $|\lambda^1(t) - \lambda^2(t) - \epsilon_1| < \epsilon$ if $t > t_1$.

For all $\epsilon > 0$, there exists t_2 such that $\left|\log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t) - \epsilon_2\right| < \epsilon$ if $t > t_2$.

Let $\epsilon < \frac{1}{2}|\epsilon_1 + \epsilon_2|$ and $t_0 = \max\{t_1, t_2\}$, we have

$$(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t) < \epsilon_1 + \epsilon_2 + 2\epsilon < 0 \quad t > t_0 \tag{B.21}$$

Since $\lim_{t\to\infty} \frac{1}{t-t_0} \int_{t_0}^t (\epsilon_1 + \epsilon_2 + 2\epsilon)ds = \epsilon_1 + \epsilon_2 + 2\epsilon$, we can conclude that

$$\lim_{t\to\infty} \frac{1}{t-t_0} \int_{t_0}^t (\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)ds \leq \epsilon_1 + \epsilon_2 + 2\epsilon \tag{B.22}$$

If $\epsilon_1 \neq 0$, then $\epsilon_2 \neq 0$

$$\lim_{t\to\infty} k(t) = 0 \quad \text{a.s.}$$

If an agent cannot learn the truth, his extinction is guaranteed.

Part 2. We assume that $\left|(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t)\right| < \frac{1}{t^\alpha}$ for all $t > t_0$.

We know that $(\lambda^1(t) - \lambda^2(t)) + \log\left(\frac{\lambda^2(t)}{\lambda^1(t)}\right)\lambda^1(t) < 0$. We have

$$\int_{t_0}^{t} \left((\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)\right)ds = -\int_{t_0}^{t} \left|(\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)\right|ds \tag{B.23}$$

and

$$\int_{t_0}^{t} \left|(\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)\right|ds < \int_{t_0}^{t} \frac{1}{s^\alpha}ds \tag{B.24}$$

we need $\alpha > 1$ to make sure that

$$\lim_{t\to\infty} \int_{t_0}^{t} \left|(\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)\right|ds < \infty \tag{B.25}$$

and

$$\lim_{t\to\infty} \int_{t_0}^{t} \left((\lambda^1(s) - \lambda^2(s)) + \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\lambda^1(s)\right)ds < \infty \tag{B.26}$$

The variance of $\int_0^t \log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)d\tilde{N}(s)$ is $E\int_0^t \left(\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\right)^2 \lambda^1(s)ds$. In order to make sure that the variance is finite, we need $\int_0^t \left(\log\left(\frac{\lambda^2(s)}{\lambda^1(s)}\right)\right)^2 ds < \infty$.

If the agent with inaccurate beliefs can learn the truth quickly enough, both agents survive as t goes to infinity.

Proof of Proposition 2.4.5.

Proof. Part 1. Under incomplete markets with learning, the wealth ratio at time t is

$$k(t) = k(0)exp\Bigg(\int_0^t \left(\frac{\lambda^2_{RN}(s))\lambda^1(s) - \lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)} - \frac{1}{2}(\theta^2(s) - \theta^1(s))^2\right)ds$$
$$+ \int_0^t (\theta^2(s) - \theta^1(s))dW(s) + \int_0^t \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)dN(s)\Bigg) \tag{B.27}$$

According to proposition 2.4.2, $\int_0^t \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)d\tilde{N}(s)$ is a martingale and $\lim_{t\to\infty} \frac{\int_0^t \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)d\tilde{N}(s)}{t}$ is zero almost surely.

We assume that $\left|\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)}\right| < \eta_3$ for all t and $\lim_{t\to\infty} \frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} = \epsilon_3$.

If $\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)}$ is bounded, so is $\log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t)$. It is therefore natural to assume that $\left|\log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t)\right| < \eta_4$ for all t and $\lim_{t\to\infty}\log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t) = \epsilon_4$.

For all $\epsilon > 0$, there exists t_1 such that $|\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} - \epsilon_3| < \epsilon$ if $t > t_1$.

For all $\epsilon > 0$, there exists t_2 such that $\left|\log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t) - \epsilon_4\right| < \epsilon$ if $t > t_2$.

Let $\epsilon < \frac{1}{2}|\epsilon_1 + \epsilon_2|$ and $t_0 = \max\{t_1, t_2\}$, we have

$$\frac{\lambda^2_{RN}(t))\lambda^1(t) - \lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t) < \epsilon_3 + \epsilon_4 + 2\epsilon < 0 \quad t > t_0 \tag{B.28}$$

Since $\lim_{t\to\infty}\frac{1}{t-t_0}\int_{t_0}^t(\epsilon_3+\epsilon_4+2\epsilon)ds = \epsilon_3+\epsilon_4+2\epsilon$, we can conclude that

$$\lim_{t\to\infty}\frac{1}{t-t_0}\int_{t_0}^t \frac{\lambda^2_{RN}(s))\lambda^1(s) - \lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)} + \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)ds \le \epsilon_3 + \epsilon_4 + 2\epsilon \tag{B.29}$$

If $\epsilon_3 \neq 0$, then $\epsilon_4 \neq 0$

$$\lim_{t\to\infty} k(t) = 0 \quad \text{a.s.}$$

If an agent cannot learn the truth, his extinction is guaranteed.

Part 2. We assume that $\left|\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t)\right| < \frac{1}{t^\alpha}$ for all $t > t_0$.

We know that $\frac{\lambda^2_{RN}(t))\lambda^1(t)-\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^2_{RN}(t)} + \log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\lambda^1(t) < 0$. We have

$$\begin{aligned}\int_{t_0}^t\left(\frac{\lambda^2_{RN}(s))\lambda^1(s) - \lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)} + \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)\right)ds = \\ -\int_{t_0}^t\left|\frac{\lambda^2_{RN}(s))\lambda^1(s) - \lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)} + \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)\right|ds\end{aligned} \tag{B.30}$$

and

$$\int_{t_0}^t\left|\frac{\lambda^2_{RN}(s))\lambda^1(s) - \lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)} + \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)\right|ds < \int_{t_0}^t \frac{1}{s^\alpha}ds \tag{B.31}$$

we need $\alpha > 1$ to make sure that

$$\lim_{t\to\infty}\int_{t_0}^{t}\left|\frac{\lambda^2_{RN}(s))\lambda^1(s)-\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)}+\log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)\right|ds<\infty \tag{B.32}$$

and

$$\lim_{t\to\infty}\int_{t_0}^{t}\left(\frac{\lambda^2_{RN}(s))\lambda^1(s)-\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^2_{RN}(s)}+\log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\lambda^1(s)\right)ds<\infty \tag{B.33}$$

The variance of $\int_0^t \log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)d\tilde{N}(s)$ is $E\int_0^t\left(\log\left(\frac{\lambda^1_{RN}(s)\lambda^2(s)}{\lambda^1(s)\lambda^2_{RN}(s)}\right)\right)^2\lambda^1(s)ds$. In order to make sure that the variance is finite, we need $\int_0^t\left(\log\left(\frac{\lambda^1_{RN}(t)\lambda^2(t)}{\lambda^1(t)\lambda^2_{RN}(t)}\right)\right)^2 ds<\infty$.

We also need to worry about $(\theta^2(s)-\theta^1(s))^2$. In order to make sure that both $\int_0^t -\frac{1}{2}(\theta^2(s)-\theta^1(s))^2ds$ and $\int_0^t(\theta^2(s)-\theta^1(s))dW(s)$ are finite. We need

$$\int_0^t(\theta^2(s)-\theta^1(s))^2ds<\infty \tag{B.34}$$

If the agent with inaccurate beliefs can learn the truth quickly enough, both agents survive as t goes to infinity.

Proof of Proposition 2.5.1.

Proof.

$$Pr(\lambda(t+h)=\Lambda_k|N(t+h)-N(t)=0)=\frac{Pr(\lambda(t+h)=\Lambda_k,N(t+h)-N(t)=0)}{Pr(N(t+h)-N(t)=0)} \tag{B.35}$$

$$Pr(\lambda(t+h)=\Lambda_k,N(t+h)-N(t)=0)=\sum_{j\neq k}q_{jk}p^1_j(t)h+p^1_k(t)(1-q_{kk}h-\Lambda_k h)+o(h) \tag{B.36}$$

Sum the above equation over all possible states, we get

$$Pr(N(t+h)-N(t)=0)=1-\sum_k p^1_k(t)\Lambda_k h+o(h) \tag{B.37}$$

If no jumps happen in this interval, then

$$p^1_k(t+h)=Pr(\lambda(t+h)=\Lambda_k|N(t+h)-N(t)=0) \tag{B.38}$$

$$\lim_{h\to 0}\frac{p^1_k(t+h)-p^1_k(t)}{h}=\sum_j q_{jk}p^1_j(t)-p^1_k(t)(\Lambda_k-\lambda^1(t)) \tag{B.39}$$

$$dp^1_k(t)=\sum_j q_{jk}p^1_j(t)dt-p^1_k(t)(\Lambda_k-\lambda^1(t))dt \tag{B.40}$$

If one jump happens in this interval, then

$$p_k^1(t+h) = Pr(\lambda(t+h) = \Lambda_k | N(t+h) - N(t) = 1) = \frac{p_k^1(t)\Lambda_k h}{\lambda^1(t)h} = \frac{p_k^1(t)\Lambda_k}{\lambda^1(t)} \tag{B.41}$$

$$p_k^1(t+h) - p_k^1(t) = p_k^1(t)\frac{\Lambda_k - \lambda^1(t)}{\lambda^1(t)} \tag{B.42}$$

If we let $h \to 0$, then right after the moment when jump happens, we add the increment in (B.42) to (B.40) to obtain (72).

Proof of Proposition 2.5.3.

Proof.

$$\begin{aligned} E[l(x,\hat{x}) - l(x, E(x|\mathcal{G}))] &= E[x\log(\frac{x}{\hat{x}}) - x + \hat{x}] - E[x\log(\frac{x}{E(x|\mathcal{G})}) - x + E(x|\mathcal{G})] \\ &= E[x\log(\frac{E(x|\mathcal{G})}{\hat{x}}) - E(x|\mathcal{G}) + \hat{x}] \\ &= E[E[x\log(\frac{E(x|\mathcal{G})}{\hat{x}}) - E(x|\mathcal{G}) + \hat{x}|\mathcal{G}]] \\ &= E[E(x|\mathcal{G})\log(\frac{E(x|\mathcal{G})}{\hat{x}}) - E(x|\mathcal{G}) + \hat{x}] = E[l(E(x|\mathcal{G}),\hat{x})] \geq 0 \end{aligned} \tag{B.43}$$

$l(E(x|\mathcal{G}),\hat{x}) = 0$ when $\hat{x} = E(x|\mathcal{G})$, $E(x|\mathcal{G})$ is the unique estimator minimizing the mean loss under l.

$$\begin{aligned} E[l_{SE}(x,\hat{x}) - l_{SE}(x, E(x|\mathcal{G}))] &= E[(x-\hat{x})^2] - E[(x - E(x|\mathcal{G}))^2] \\ &= E[(E(x|\mathcal{G}) - \hat{x})(2x - \hat{x} - E(x|\mathcal{G}))] \\ &= E[E[(E(x|\mathcal{G}) - \hat{x})(2x - \hat{x} - E(x|\mathcal{G}))|\mathcal{G}]] \\ &= E[(E(x|\mathcal{G}) - \hat{x})E[(2x - \hat{x} - E(x|\mathcal{G}))|\mathcal{G}]] = E[l_{SE}(E(x|\mathcal{G}),\hat{x})] \geq 0 \end{aligned} \tag{B.44}$$

$l_{SE}(E(x|\mathcal{G}),\hat{x}) = 0$ when $\hat{x} = E(x|\mathcal{G})$, $E(x|\mathcal{G})$ is the unique estimator minimizing the mean loss under l_{SE}.

Appendix C

Portfolio Choice and Intergenerational Inequality in China: Theory and Evidence

Proof of Lemma 3.3.2. See Dieckmann (2011) for the proof of eqn.(3.15) and eqn.(3.16). The derivation of $\xi_{s,t}$ process follows first by applying the Girsanov theorem for the jump process, s.t:

$$dN_{s,t} - \bar{\lambda}_{s,t}dt = dN_t(\bar{\lambda}_t) - \bar{\lambda}_t dt \tag{C.1}$$

With the change of measure, we can rewrite eqn.(3.15) into

$$\frac{d\xi_{s,t}}{\xi_{s,t^-}} = \left(\bar{\lambda}_{s,t} - \lambda^N_{s,t} - r_t + (\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1)(\lambda_{s,t} - \bar{\lambda}_t)\right) dt - \theta_{s,t} dZ_t + \left(\frac{\lambda^N_{s,t}}{\bar{\lambda}_{s,t}} - 1\right) dN_t(\bar{\lambda}_t) \tag{C.2}$$

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

$$\frac{d\eta_{s,t}}{\eta_{s,t}} = \left(\frac{1}{1+\bar{\kappa}}\lambda_{s,t} - \lambda^N_t\right) dt + \left[\frac{1+\bar{\kappa}}{\bar{\kappa}}\left(-\frac{2\lambda^N_t}{\lambda_t} - 1\right) - 1\right] dN(\bar{\lambda}_t) \tag{C.3}$$

Proof of proposition 3.3.2. To get the coefficient of the stock price, we can write down the formula for stock prices, i.e.,

$$\begin{aligned} S_t &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty \xi_u D_u du\right] \\ &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty e^{-(\rho+\delta(1-\beta))u}\eta_u du\right] \\ &= \frac{1}{\xi_t}\eta_t\int_t^\infty e^{-(\rho+\delta(1-\beta))u}du \\ &= \frac{1}{\rho+\delta(1-\beta)}Y_t \end{aligned} \tag{C.4}$$

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

$$\frac{dS_t}{S_{t^-}} = \frac{dY_t}{Y_{t^-}} \tag{C.5}$$

Recall that the compounded stock market value follows the following process

$$\frac{dS_t + D_t dt}{S_{t^-}} = \mu_t^S dt + \sigma^S dZ_t + \kappa_t^S dN_t(\lambda_t) \tag{C.6}$$

Matching coefficients, one get

$$\mu^S = \mu + \rho + \delta(1-\beta); \sigma^S = \sigma; \kappa_t^S = \kappa_t \tag{C.7}$$

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

$$\mu_t^P = -\kappa_t^P \lambda_t^N + r_t = -\frac{\kappa_t}{1+\bar{\kappa}}\mathbb{E}_{s,t}(\bar{\lambda}_{s,t}) + r_t \tag{C.8}$$

C.1 Proof of Proposition 3.3.3

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

$$\frac{dw_{s,t}}{w_{s,t^-}} = \hat{\mu}_{s,t}dt + \hat{\sigma}_{s,t}dZ_t + \hat{\kappa}_{s,t}dN_t \tag{C.9}$$

where dZ_t and dN_t represent aggregate Brownian motion and jump shocks. To simplify notation, we 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

$$\begin{aligned} A(t+dt) &= \int_{-\infty}^{\infty} f(w)n_{t+dt}dw \\ &= \int_{-\infty}^{\infty} (f(w)+df(w))n(w) - \delta f(w)n(w)dw \\ &= \int_{-\infty}^{\infty} f(w)(1-\delta)n(w)dw + \int_{-\infty}^{\infty} df(w)n(w)dw \end{aligned} \tag{C.10}$$

We then have

$$d(A(t)) = -\int_{-\infty}^{\infty} \delta n(w) f(w) dw + \int_{-\infty}^{\infty} df(w) n(w) dw. \tag{C.11}$$

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

$$df(w) = f'(w)[\hat{\mu} w dt + \hat{\sigma} w dZ] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt + [f(w(1+\hat{\kappa})) - f(w)] dN \tag{C.12}$$

Using integration by parts, we have

$$\begin{aligned}
\int_{-\infty}^{\infty} df(w) n(w) dw &= \int_{-\infty}^{\infty} \left[f'(w) \left[\hat{\mu} w dt + \hat{\sigma} w dZ \right] + \frac{1}{2} f''(w) \hat{\sigma}^2 w^2 dt \right] n(w) dw \\
&+ \int_{-\infty}^{\infty} \left[f(w(1+\hat{\kappa})) - f(w) \right] n(w) dN dw \\
&= \int_{-\infty}^{\infty} f(w) \left[-\frac{\partial}{\partial w} \left(n(w) \hat{\mu} w dt + n(w) \hat{\sigma} w dZ_t \right) + \frac{1}{2} f(w) \frac{\partial^2}{\partial w^2} \left(n(w) \hat{\sigma}^2 w^2 \right) dt \right] \\
&+ \int_{-\infty}^{\infty} \left[n(w(1+\hat{\kappa})) - n(w) \right] f(w) dN dw
\end{aligned} \tag{C.13}$$

Notice that the way we write down changes in $A(t)$ in (C.11) fixes the density of w in the state space and calculate with Ito's Lemma how $f(w)$ will change. One can also approximate $d(A(t))$ by linearly extrapolating the density at each point, that is,

$$d(A(t)) = \int_{-\infty}^{\infty} f(w) \frac{\partial n}{\partial t} dt dw = \int_{-\infty}^{\infty} df(w) n(w) dw \tag{C.14}$$

Plugging in the expression in eqn. (C.13), and equating the integrands, we get

$$dn = -\frac{\partial}{\partial w} (n \hat{\mu} w dt + n \hat{\sigma} w dZ) + \frac{1}{2} \frac{\partial^2}{\partial w^2} (n \hat{\sigma}^2 w^2) dt - \delta n + [n(w(1+\hat{\kappa}), t)) - n(w,t)] dN \tag{C.15}$$

As one can see, the distribution of this variable is stochastic, and that there is no closed form solution in general. However, we can still ask the question, what is the long-run stationary distribution of this variable in this economy? That is, what is the solution of $dp(w) = \mathbb{E}_t (dn(w)) = 0$? [1] By averaging out the KFP equation, we then have

$$-\frac{\partial}{\partial w} \left(\mathbb{E}(\hat{\mu}) w p(w) \right) + \frac{\partial^2}{\partial w^2} \left(\frac{\mathbb{E}(\hat{\sigma}^2)}{2} w^2 p(w) \right) - \delta p(w) + \lambda (p^J - p) = 0 \tag{C.16}$$

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

$$\tilde{w}_{s,t} = \frac{w_{s,t}}{\omega Y_t} \tag{C.17}$$

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

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

$$\frac{dw_{s,t}}{w_{s,t^-}} = \left(\sigma^2 + r - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right)\right)dt + \sigma dZ + \left(\frac{\bar{\lambda}_{s,t}}{\lambda_t^N} - 1\right)dN_t \tag{C.18}$$

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

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \left(\sigma^2 + r - \bar{\lambda}_{s,t} + \lambda_t^N + \delta + (\lambda_{s,t} - \bar{\lambda}_t^0)\left(\frac{\lambda_{s,t}}{\lambda_t^N} - 1\right) - \mu\right)dt + \left(\frac{\lambda_{s,t}}{\mathbb{E}(\lambda_{s,t})}(1+\kappa_t) - 1\right)dN_t \tag{C.19}$$

which in short-hand can be written as

$$\frac{d\tilde{w}_{s,t}}{\tilde{w}_{s,t^-}} = \hat{\mu}(\lambda_{s,t})dt + \hat{\kappa}(\lambda_{s,t})dN_t \tag{C.20}$$

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

$$dx = \hat{\mu}dt + \log(1+\hat{\kappa})dN_t \tag{C.21}$$

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

$$p(x) = \int_0^\infty n(x,\lambda)d\lambda \tag{C.22}$$

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.,

$$n(x,\lambda) = n_1(x|\lambda)n_2(\lambda) \tag{C.23}$$

From eqn. (C.21), we can write down the dynamics of $n_1(x|\lambda)$, i.e.,

$$0 = -\frac{\partial n_1}{\partial x}\hat{\mu} + \lambda^0\left(n_1(\log(1+\hat{\kappa}) + x) - n_1\right) - \delta n_1 \tag{C.24}$$

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

$$\zeta \approx \zeta_0 + (\lambda - \lambda^0)\zeta_1 \tag{C.25}$$

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

$$d\lambda_{s,t} = (\lambda_{s,t^-} - \lambda_l)(\lambda_{s,t} - \lambda_h)dt - (\lambda_{s,t^-} - \lambda_h)(\lambda_{s,t^-} - \lambda_l)\frac{(1 + \lambda_{s,t^-})}{\lambda_{s,t^-}}dN_t \tag{C.26}$$

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

$$0 = -\frac{\partial n_2}{\partial \lambda}(\lambda - \lambda_h)(\lambda - \lambda_l) - n_2(2\lambda - \lambda_l - \lambda_h + \delta) + \lambda^0(n_2^J - n_2) \tag{C.27}$$

We can guess and verify the following approximate exponential solution

$$n_2(\lambda) \approx e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} \tag{C.28}$$

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

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

$$\begin{aligned} p(x) &= G_0 e^{(\zeta_0 - \lambda^0\zeta_1)x} \int_{\lambda_l}^{\lambda_h} e^{\lambda\zeta_1 x} e^{g_0 + g_1\lambda + \frac{g_2}{2}\lambda^2} d\lambda \\ &= \underbrace{Ge^{\zeta_0 x}}_{RE} \underbrace{[\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}]}_{Learning} \end{aligned} \tag{C.29}$$

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

$$\begin{aligned} \lim_{x\to\infty} \frac{p(x)}{p^{RE}(x)} &= \lim_{x\to\infty} [\zeta_1 x + g_1]^{-1}[e^{(\lambda_h - \lambda^0)\zeta_1 x} - e^{(\lambda_l - \lambda^0)\zeta_1 x}] \\ &= \lim_{x\to\infty} \zeta_1^{-1}\left[-(\lambda_l - \lambda^0)\zeta_1 e^{(\lambda_l - \lambda^0)\zeta_1 x}\right] \end{aligned} \tag{C.30}$$

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

$$\lim_{x\to\infty} p(x) > \lim_{x\to\infty} p^{RE}(x) \tag{C.31}$$

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

C.2 Verification of Newborn Consumption Share

We start by defining β_t, i.e.,

$$\beta_t = \frac{c_{t,t}}{Y_t} = \frac{(\rho+\delta)w_{t,t}}{Y_t} \tag{C.32}$$

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

$$\begin{aligned} W_{t,t} &= \frac{1}{\xi_t}\mathbb{E}_t\left[\int_t^\infty e^{-\delta(u-t)\xi_u}\omega Y_u du\right] \\ &= \omega Y_t \mathbb{E}_t\left[\int_t^\infty e^{-(\rho+\delta+\delta(1-\beta))(u-t)}\frac{\bar{\eta}_u}{\bar{\eta}_t}du\right] \\ &= \frac{\omega Y_t}{\rho+\delta+\delta(1-\beta)} \end{aligned} \tag{C.33}$$

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

$$\beta = \frac{1}{\rho+\delta+\delta(1-\beta)} \tag{C.34}$$

This renders the two solutions

$$\beta_{1,2} = \frac{\rho+2\delta}{2\delta} \pm \frac{\sqrt{\rho^2+4(\rho+\delta)\delta(1-\omega)}}{2\delta} \tag{C.35}$$

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

$$\beta = \frac{\rho+2\delta}{2\delta} - \frac{\sqrt{\rho^2+4(\rho+\delta)\delta(1-\omega)}}{2\delta} \tag{C.36}$$