ESSAYS IN ASSET PRICING AND INFORMATION ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Bryce Patrick Little

May 2017

ESSAYS IN ASSET PRICING AND INFORMATION ECONOMICS

Bryce Patrick Little, Ph.D.

Cornell University 2017

The first essay addresses how investors price risk in the stock market when they cannot observe the true long-run growth rate of their consumption and dividend endowments. The model presents a number of import insights into how the ability of long-run risks to explain the equity risk premium is directly related to the quality of investors information sets. The second essay addresses the welfare costs of ambiguity surrounding the probability distribution of shocks driving the growth rate of their consumption endowment. If an investor faces both long-run risk and rare disasters in their consumption edowment, then they would forgoe a large share of their lifetime consumption to absolve ambiguity surrounding disasters, but substantially less to remove ambiguity about long-run risk. The third essay presents a dynamic stochastic general equilibrium model of a small open economy (SOE) that faces time-varying volatility of news about their total factor productivity and real interest rate. News uncertainty shocks about the interest rate motivate the SOE to deleverage, however, the same class of shocks in total factor productivity have insubstantial effects.

BIOGRAPHICAL SKETCH

Prior to pursuing his Ph.D. in Economics at Cornell University, Bryce completed his undergraduate studies at the University of California, Los Angeles where he received a Bachelor of Science in Applied Mathematics. He also received a Master of Science in Statistics from Texas A&M University.

ACKNOWLEDGEMENTS

I would like to thank the members of my committee: Andrew Karolyi, Kristoffer Nimark, Karel Mertens, and Levon Barseghyan. Without their support and goodwill, this dissertation would not have been possible. In particular, I thank Andrew for his energy, enthusiasm, feedback, and willingness to invest in my research. I thank Kristoffer for his pateince, and mentoring me in the technical intricacies of information economics. I thank Karel for opening the door to macroeconomics, and providing me opportunities to follow in his footsteps. I thank Levon for his charismatic and intelligent feedback, and guidence.

Furthermore, I am thankful to my family and friends for their support and belief in me.

TABLE OF CONTENTS

Biographical Sketch . . . iii
Acknowledgements . . . iv
Table of Contents . . . v
List of Tables . . . vii
List of Figures . . . ix

1 Introduction **1**

2 Disagreement Over the Long-Run: The Effects of Higher Order Expectations on the Equity Risk Premium **5**
2.1 Introduction . . . 6
2.2 Model . . . 11
2.2.1 Information Sets and Higher Order Beliefs . . . 14
2.2.2 Solving the Model . . . 16
2.3 Calibration . . . 21
2.4 Discussions of Results . . . 24
2.4.1 Disagreement Shrinks the Risk Price for Long-Run Shocks . 27
2.4.2 The Risk Price for Long-Run Shocks is Countercyclical . . . 31
2.4.3 Three Factors Explain The Term Structure of Consumption Growth . . . 33
2.4.4 Disagreement Mutes Price Responses to Changes in State Variables . . . 36
2.4.5 Disagreement Increases During High Volatility Spells . . . 39
2.4.6 Consensus Forecast Errors are Priced Risk . . . 42
2.5 Conclusion . . . 43

3 Ambiguity Over the Long-Run: The Welfare Costs of Persistent Growth Shocks **48**
3.1 Introduction . . . 49
3.2 A Model of Consumption Growth . . . 51
3.3 Two Observationally Equivalent Agents . . . 52
3.3.1 Type-I Agent with EZ Preferences . . . 52
3.3.2 Type-II Agent with Multiplier Preferences . . . 53
3.3.3 The Worst-Case Probability Distribution . . . 55
3.4 Thought Experiments . . . 55
3.5 Conclussion . . . 60

4 News Uncertainty Shocks and Sovereign Debt **61**
4.1 Introduction . . . 62
4.2 Model . . . 65
4.2.1 Two Types of News Shocks . . . 69
4.2.2 Steady State . . . 70
4.3 Calibration . . . 71

4.4 Solution . . . 73
4.5 Discussion of Results . . . 74
4.6 Conclusion . . . 77

5 Conclusion **79**

A Appendix of Chapter 1 **82**
A.1 Law of Motion for Higher Order Beliefs . . . 83
A.2 Deriving the Stochastic Discount Factor . . . 85
A.3 Solving for Prices . . . 87
A.3.1 Approximating Returns . . . 87
A.3.2 Price-Consumption Ratio . . . 89
A.3.3 Price-Dividend Ratio . . . 90
A.3.4 Risk-Free Rate . . . 90
A.4 Conditional Variances . . . 91
A.5 Risk Premia . . . 95
A.5.1 Deriving Risk Prices . . . 95
A.5.2 Deriving Betas . . . 97
A.5.3 Equity Premium . . . 98
A.6 Numerical Algorithm . . . 98
A.7 An Estimation Procedure Utilizing Forecast Panel Data . . . 100
A.7.1 Conditional Cross-Sectional Covariance Matrix . . . 100
A.7.2 Likelihood . . . 101
A.7.3 An Adaptive Markov Chain Monte Carlo Estimation Procedure 102
A.8 Variance Decomposition of the Consumption Growth Term Structure 104

B Appendix of Chapter 2 **105**
B.1 Value Functions . . . 106
B.1.1 Type I & II Agents with $\gamma > 1$ or $\vartheta < +\infty$ . . . 106
B.1.2 Type I & II Agents with $\gamma = 1$ or $\vartheta = +\infty$ . . . 107
B.2 Derivations . . . 109
B.2.1 Derivation of the Distortion . . . 109
B.2.2 Worst-Case Distribution . . . 109
B.2.3 Model-Detection Error Probabilities . . . 110
B.2.4 Risk-Free Rate . . . 111

C Appendix of Chapter 3 **112**
C.1 Equilibrium Equations . . . 113
C.2 Competitive Equilibrium . . . 114
C.3 Impulse Response Functions . . . 115
C.3.1 Household's Problem with Recursive Preferences . . . 120
C.3.2 Stationarity . . . 121
C.4 Stochastic Discount Factor . . . 121

LIST OF TABLES

2.1 Parameter estimates for the preferences, consumption growth, and dividend growth categories from BKY. They estimate the model using the generalized method of moments with annual data beginning in 1930 and ending in 2008. I solve the model over a discrete grid of parameter values for noise φ_s. At the smallest value $\varphi_s = 0$, agents in the model have full information. At the largest value $\varphi_s = 4\varphi_x$, the risk price for long-run shocks is negligible. 25

2.2 Shares of variance in the empirical covariance matrix of X_t explained by the first three principal components (level, slope, and curvature). Values in the table are each principal component's associated eigenvalue divided by the sum of all eigenvalues. I estimate the empirical covariance matrix of X_t with a simulated sample of 1,000,000 observations under $\tau = 8$ and $\bar{k} = 12$. 36

2.3 Average interquartile range of real nondurable consumption growth forecasts from the Survey of Professional Forecasters conditioned on ex post NBER business cycle classifications. The top half of the table reports averages for the full sample from the third quarter of 1981 through the end of 2015. The bottom half reports averages from the second quarter of 1990 through 2015. The Federal Reserve Bank of Philadelphia took over the survey in the second quarter of 1990 and dramatically increased the breadth and quality of the sample. 42

3.1 Compensation for elimating different sources of physical risk and model uncertainty. The source column indicates what has been eliminated: risk, uncertainty, or risk and uncertainty simultaneously. The amount column refers to the cost of eliminating risk and/or uncertainty in terms of structural parameters. 58

3.2 Compensation for elimating different sources of physical risk and model uncertainty. The source column indicates what has been eliminated: risk, uncertainty, or risk and uncertainty simultaneously. The amount column refers to the cost of eliminating risk and/or uncertainty in terms of structural parameters. 59

4.1 Parameter values for the model calibration. All parameters are scaled to a quarterly frequency and are values for Argentina used in Fernández-Villaverde et al. [2011]. 72

A.1 Variance decomposition of higher order expectation in terms of their first three principal components (level, slope, and curvature). I compute the values in each row by projecting a simulated time series of higher order expectations onto the three time series of component scores for the level, slope and curvature factors. Each variance share is the square of the regression coefficient multiplied by the ratio of the factor variance to the variance of the higher order expectation. The length of the simulated sample is 1,000,000 observations with $\tau = 8$ and $\bar{k} = 12$. The t-statistics for the first stage regression are unanimously greater than one hundred, even after correcting the residuals for autocorrelation and heteroskedasticity, so I do not report them. 104

LIST OF FIGURES

2.1 The unconditional average risk prices for the long-run risk shock across different levels of noise. Full information indicates $\varphi_s = 0$. I calibrate positive values of noise to multiples of the time series volatility of the long-run risk shock. As investors' information sets grow more disparate, the average risk price for the long-run shock approaches zero. 30
2.2 The unconditional average risk prices and betas for the long-run risk shock across different levels of noise. Full information indicates $\varphi_s = 0$. I calibrate positive values of noise to multiples of the time series volatility of the long-run risk shock φ_x. As investors information sets grow more disparate, the average beta for the long-run shock approaches zero. 31
2.3 Risk prices for the long-run risk shock $\varepsilon_{x,t}$ over different volatility histories. I define "Smooth Sailing" as a τ-consecutive periods of low volatility, and oppositely, "Wild Times" as τ-consecutive periods of high volatility. The second bar immediately to the right of "Smooth Sailing" changes the most recent period from low to high volatility. The third bar does the same, but for the two most recent periods. The pattern continues until the last bar, where all states are high volatility. The scale of noise is calibrated $\varphi_s = 2\varphi_x$. 33
2.4 Plots of the coefficients from the first three principal components of the empirical covariance matrix of X_t with a simulated sample of 1,000,000 observations under $\tau = 8$ and $\bar{k} = 12$. X-axis labels refer to the order of expectation, and the y-axis are scalar magnitudes of loadings on principal components. 35
2.5 Impulse response functions for the price-consumption ratio from a two standard deviation shock to $\varepsilon_{x,t}$. Data plotted in the top graph was generated with consecutive periods of low volatility, and the bottom graph was generated with consecutive periods of high volatility. The level of noise is calibrated to $\varphi_s = 2\varphi_x$. 37
2.6 Impulse response functions for the price-dividend ratio from a two standard deviation shock to $\varepsilon_{x,t}$. Data plotted in the top graph was generated with consecutive periods of low volatility, and the bottom graph was generated with consecutive periods of high volatility. The level of noise is calibrated to $\varphi_s = 2\varphi_x$. 38

2.7 Impulse response functions for the hierarchy of expectations X_t from a two standard deviation shock to $\varepsilon_{x,t}$ (top), $\varepsilon_{c,t}$ (middle), and $\varepsilon_{d,t}$ (bottom). On the top graph, the thick dark blue line represents $x_t^{(0)}$, but this variable is absent from the other two graphs because it does not respond to consumption or dividend shocks. The increasingly lighter shades of turquoise and green correspond to higher orders of expectations. The level of noise is calibrated to $\varphi_s = 2\varphi_x$. 45
2.8 The top plot exhibits a simulation of the expectations hierarchy for $T = 100$ periods. The first fifty periods are low volatility states, and the last fifty periods are high volatility. A vertical dashed line separates the two regimes. The bottom plot shows the interquartile range of consumption growth forecasts. Each period, I randomly sample $N = 50$ independent draws to simulate a panel of forecasts. I then estimate the 25th and 75th percentiles and compute their difference. 46
2.9 Cross-sectional dispersion of annual real nondurable consumption growth forecasts from the Survey of Professional Forecasters. The forecast data are sampled at a quarterly frequency, beginning in the third quarter of 1981 and ending in the fourth quarter of 2015. I compute annual consumption growth for each forecaster in the sample by subtracting their forecast of the previous period revised log consumption level from the level forecasted four quarters ahead of the end of the measurement period. 47

3.1 The lines represent the share of consumption the representative agent is willing to forgo to absovle ambiguity surrounding a specific type of physical risk for a given level of γ. 57

4.1 Time series plot of first difference of 2 year forward swap rates on German sovereign debt with a 3 month tenor. Time observations are at a quarterly frequency. 63
4.2 Partially identified time series of news shocks to TFP in the United states at a quarterly frequency, 1959-2005. 64
4.3 Policy functions for bond holdings over a 100 period simulation. Line colors correspond to the order of the perturbation method used- red is accurate to the first order, green is second, and blue is third. The finely dotted horizontal line at 6 is the steady state level of bond, and the line with thicker dashes below is the ergodic mean. 75

4.4 Impulse response functions for sovereign bond holdings responding to a one standard error news uncertainty shock when the news is 1, 4, 8, and 12 periods in the future. The dark indigo line with the largest response corresponds to news 12 periods ahead, the blue line to news 8 periods ahead, and so on. 77

C.1 Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about the interest rate. . 116
C.2 Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about the interest rate. . 117
C.3 Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about TFP. 118
C.4 Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about TFP. 119

CHAPTER 1

INTRODUCTION

I present three essays which jointly highlight what happens to asset prices when investors information sets and/or beleifs are somehow flawed. They focus on three epsitemological situations faced by investors considering risk: can't see, don't know, and not sure. In the first essay, investors can't see the true growth rate, and therefore make their best inference given their information sets. This turns a popular result driving the equity premium upside down. In the second essay, investors don't know the true distribution of risk. They make a constrained-pessimistic best guess, and this reveals that the risk of a rare bad event is much more costly than a more probable (but smaller) bad event that persists for a long time. The third essay focuses on when investors are not sure about how much uncertainty there is about future events. This leads to substantial fluctations in economic aggregates in response to changing quantities of news risk. Taken together, these essays underline the importance of information in asset price formation.

The first essay is in Chapter 2, which presents a model of the equity risk premium. Investors price the stock market with imperfect information. The interesting trade-off in the model is how the quality of investors' information impacts the equilibrium price of risk. As the quality of information deteriorates, the price of risk for shocks impacting investors' marginal utility shrinks toward zero, diminishing the risk premium on the stock market. Further, the model matches features of time-varying disagreement observed in data from professional forecasters. As volatility of key economic aggregates changes over time, so does the cross-sectional dispersion of investors forecasts.

These results are best motivated at a more granular level. The economy is populated with many investors, each having identical risk preferences. Individual investors have a consumption endowment that grows at exactly the same rate as aggregate consumption. The stock market pays out the aggregate dividend in every period. Consumption and dividend growth are linked by a small, persistent process that jointly drives their conditional mean growth rates. Shocks to this process are known as long-run risks. If investors in the model had full information, long-run risks could generate a price of risk sufficient to explain the historical equity risk premium in the United States. But investors cannot directly observe the true growth rate, and instead have a noisy signal providing some information about the truth. There is no arbitrage in the economy, meaning each investor correctly prices the stock market in expectation. Further, it's common knowledge that investors form model consistent expectations. Taken together, these facts imply that average expectations of growth are what drive asset prices over time. These average expectations are otherwise known as higher order expectations.

Investors infer average expectations with the Kalman filter. They use their private signals in combination with publicly observable data on asset prices and economic aggregates to forecast the hidden values of average expectations. In equilibrium, the cross-sectional average of investors' forecasts of average expectations pins down the law of motion for average expectations. And these average expectations are the state variables that drive asset prices.

As the noise in investors' signals grows large, the equilibrium compensation for bearing long-run risk approaches zero. The equity risk premium shrinks. The Kalman filter estimates of average expectations are equivalent to investors' beliefs. And as input quality to the filter deteriorates, investors perception of physical risk is smoothed over time. Therefore the perceived quantity of risk falls, hence so do risk premia. The filter also matches the time-varying dispersion of forecasts observed in the data: during bad economic times, forecasts of consumption and dividend growth fan out. And in good times, forecasts condense. Investor beliefs are functions of volatility. When volatility is high, the cross-sectional dispersion of beliefs spread out.

Chapter 3 presents a model of a representative agent who faces long-run risks and rare disasters in their consumption endowment. However, the agent does not know the true probability distribution of these risks. Futher, they have concerns for robustness: whatever model the agent operationalizes should be reasonably pessimistic given they do not know the truth. Thus, the agent seeks to construct a model of consumption growth that simultaneously addresses both model uncertainty and physical risk. The agent solves this problem by choosing the worst case probability distribution for long-run risks and rare disasters subject to an entropy constraint: the resulting worst-cast distributions must be hard to reject given the observed data. With the solution in hand, I ask how much lifetime consumption would the agent give up to resolve ambiguity surrounding the risks in the endowment. For every level of risk aversion, the agent would pay a substantially larger share of their lifetime consumption to resolve ambiguity surrounding disasters relative to long-run risk.

Chapter 4 presents a model of a small open economy (SOE) facing time-varying uncertainty in news about their total factor productivity and real borrowing costs. An unanticipated change in the volatility of news is called a news uncertainty shock. When facing a positive news uncertainty shock about borrowing costs, the SOE responds by reducing its consumption and deleveraging external debt. When facing a similar shock for total factor productivity, the response of macroeconomic aggregates is small and negligible.

CHAPTER 2

DISAGREEMENT OVER THE LONG-RUN: THE EFFECTS OF HIGHER ORDER EXPECTATIONS ON THE EQUITY RISK PREMIUM

2.1 Introduction

The historically large return premium on risky stocks over safe bonds observed in the U.S. is an equilibrium outcome resulting from a large population of investors trading on their subjective views of asset prices and the state of the economy. Data suggests market participants hold consistently diverse beliefs about the anticipated trajectories of key economic aggregates important to asset prices. For instance, the Survey of Professional Forecasters exhibits large dispersion in expectations about the future path of real consumption. Yet the extant literature on the equity premium neglects targeting the observed level of disagreement in the data, or how disagreement evolves over the business cycle. In light of these facts, I contribute a novel consumption-based asset pricing model able to match import features of both the time series moments of asset prices in addition to the conditional cross-sectional moments of belief dispersion. In the model, investors have private information about the conditional mean growth rate of the economy. They use their private information in combination with publicly available data on prices and key aggregates to optimally form subjective forecasts of economic growth. In equilibrium, the model produces an important feature of belief dispersion observed in the data–that investors' forecasts tend to have greater dispersion during bad times for economic growth when uncertainty is highest. Further, the model shows that the size of compensation for risks tied to long-run economic growth is increasing in how precisely agents' discern the true state of the economy. If agents' largely disagree about economic growth, then the associated risk premia is small and inconsequential to asset returns. The model also exhibits endogenous countercyclical compensation for risks tied to long-run economic growth, matching the well established countercylicality of the equity premium borne out in the data.

I assume agents in the model cannot perfectly observe the conditional mean growth rate of their ex post identical consumption and dividend endowments. The stochastic processes driving the endowments are akin to the long-run risks model of Bansal and Yaron [2004] and Bansal et al. [2012]. The special feature of a long-run risks endowment is the conditional expected growth rates of both consumption and dividends are commonly driven by small, persistent, and mean-revering process. This feature portrays aggregate consumption and dividends as tending to weather sustained periods of above and below average growth throughout the business cycle. The shocks driving the conditional mean growth rate are what generate the large equity premium in a full-information long-run risks model. I refer to these shocks as long-run news. In my model, agents privately observe noisy signals obscuring the true conditional mean growth rate, in addition to publicly available price and economic data. In order for private information to play an important role in the the model, the time-series evolution of observable prices and economic aggregates cannot fully reveal the state of the conditional expected growth rate. Such a feature can be justified by the thinking of Grossman and Stiglitz [1980]. They posit that if information is costly to acquire, then prices, even if freely observed, cannot fully reveal all relevant information. Despite agents in the model not observing the true growth rate, rational expectations of its law of motion is common knowledge amongst agents. In the absence of arbitrage opportunities, agents are forced into a situation first described by Townsend [1983] and Sargent [1991] where they must, "...forecast the forecast of others..." in order to evaluate statistical expectations of future states in the economy. In equilibrium, the consequence of agents' private information, rational expectations, and no arbitrage is the emergence of higher order expectations. That is, the average forecast of the average forecast becomes a state variable in the model relevant to asset prices. I employ the technique for solving rational expectations models with private information devloped in Nimark [2011] to solve the model. Common knowledge of rational expectations of the true state is sufficient to pin down a stochastic process for higher order expectations. Agents use Bayesian methods to estimate the expectations hiearchy, which, in turn, also serve as the law of motion for the hierarchy. This realization underlies all of the interesting results.

The first main result is the model can generate a countercyclical risk price for long-run news. That is, the price agents would pay to hedge long-run shocks in their stochastic discount factors changes with economic conditions. This is a stark contrast to the full information model, where the risk price for long-run news in the full information model is constant and large. The result can best be understood by first noting the endowment process for all agents in the model is subject to common heteroskedastic shocks. If the scale of noise in agents' private signals is constant over the business cycle, then the precision of agents' signals increases when volatility is high. Therefore, in the aggregate, information about the long-run growth prospects of the economy improves during spells of high volatility–agents' effectively improve their collective forecast accuracy of the conditional mean growth rate. This mechanical feature of the model captures the intuition in the empirical work of Loh and Stulz [2015] that shows market participants' earnings forecasts are relatively more accurate during spells of high volatility. They argue this result in the data stems from agents' career concerns, which drive them to work harder during during bad times.

The second main result is the risk price for long-run news is decreasing in the scale of noise clouding investors' perceptions of the true state: the more agents' heterogeneous information sets disagree about growth, the lower the premium for long-run risk. The intuition behind a shrinking risk price is, in equilibrium, asset prices are less responsive to news about the state of the economy when noise inhibits agents' ability to percieve the true state. Risk prices directly inherit this property of actual asset prices. Further, when the scale of noise is large relative to the time series volatility of long-run news, about three to four times larger, then the risk price for long-run shocks is small and insubstantial. Likewise, asset returns also grow less sensitive to long-run news when noise is relatively high. In a nutshell, if agents' information sets are too disparate, even a strong distaste for shocks to long-run growth cannot explain the equity premium.

The model is also capable of generating time-varying cross-sectional disagreement in agents' consumption growth forecasts that matches features of the data. A simple exercise of partitioning the real consumption growth panel from the Survey of Professional Forecasters into recession and expansion regimes, periods when volatility is relatively low and high, shows the interquartile range of consumption growth forecasts is 1.28% during expansions relative to 1.70% in recessions. After a recent spell of high volatility, agents prior uncertainty about the state increases, which drives the larger cross-sectional dispersion in forecasts relative to times when volatility has been low or normal. This feature permits the model to flexibly target moments of data from real market participants that measure belief heterogeneity. Models of heterogeneous beliefs are notoriously difficult to calibrate, and the lion's share of such models therefore result in stylized theories with at most a handful of agents. A plausible (and exciting) next step in this research agenda is to incorporate measures of trading volume, and to therefore produce an asset pricing model that jointly explains conditional moments of the time series of returns, measures of heterogeneous beliefs, and trading activity.

Perhaps the theoretical foundation for heterogeneous beliefs in asset pricing stems sprouts from Harrison and Kreps [1978]. They formulate a dynamic economy populated by risk neutral traders set apart from one another by their differences in beliefs about the cash flows of a risky asset. If short sales of the risky asset are prohibited, speculative behavior arises: traders are willing to pay a premium above and beyond the price that would otherwise induce them to hold the risky asset forevermore. The premium exists because of other traders' willingness to pay higher prices, contingent, of course, on their relative optimism. Varian [1985] extends the application of heterogeneous beliefs to risk averse agents operating in a complete market. He shows, again supposing agents have homogeneous preferences, that asset prices tend to be lower when agents' beliefs are more dispersed. Using the theoretical chasis of Detemple and Murthy [1994], Zapatero [1998] poses a situation where two agents with logarithmic utility have heterogeneous beliefs about aggregate consumption growth. He finds that differences in beliefs largely increase the volatility of the real interest rate. Basak [2005] presents a continuous time economy where risk averse agents can disagree on the mean of consumption growth. If the agents' preferences are homogeneous, the relative optimism (pessimism) of a particular agent amplifies their degree of risk sharing. If the disagreement process is correlated with agents consumption, new risk premia terms enter the model alongside the traditional compensation for uncertain consumption flows. Recent work by Albagli et al. [2015] shows diverse beliefs held by Bayesian agents creates a wedge in the way asset prices respond to innovations in state variables. Barillas and Nimark [2016a] study an economy where agents' higher order expectations drive a speculative component of bond yields that is orthogonal to the traditional component of yields driven by common risk factors. Barillas and Nimark [2016b] etend their theory with an empirical affine term structure model with higher order expectations that is able to match the cross-sectional dispersion in forecast data of long-term bond yields.

The paper closest to this one is Bhandari [2016]. He reverse-engineers heterogeneous beliefs by recasting the incomplete markets asset pricing model of Constantinides and Duffie [1996] with idiosyncratic risk into a complete markets model with heterogeneous beliefs. The data generating process for idiosyncratic consumption is interpreted as a data generating process for likelihood ratios that distort individual agents' beliefs. A key difference from my paper is there is no private information in the model–agents' beliefs are driven by a carefully chosen exogenous process that satisfies desirable aggregation conditions.

Section 2 discusses the model in detail. Second 3 refers to how I calibrate the model. Section 4 discusses resuts, and section 5 concludes.

2.2 Model

Time is discrete, indexed by $t \in \mathbb{N}$. There is a continuum of islands $i \in (0, 1)$, and each island i is populated by one infinitely lived investor with identical preferences of Epstein and Zin [1989] and Weil [1989]. Investor i maximizes her lifetime utility over real consumption flows $C_{i,t}$ subject to a sequence of budget constraints:

$$V_{i,t} = \max_{\{C_{i,t}\}} \left((1-\beta) C_{i,t}^{1-1/\psi} + \beta \mathbb{E}_{i,t} \left[V_{i,t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \tag{2.1}$$

subject to,

$$W_{i,t+1} = R_{c,t+1} \left(W_{i,t} - C_{i,t} \right) \tag{2.2}$$

Investors accumulate real wealth $W_{i,t}$ and save with an asset that yields the return on the aggregate consumption claim $R_{c,t+1}$. The parameter $\beta \in (0, 1)$ models investors' time preference, $\psi > 0$ is the elasticity of intertemporal substituion (EIS), and $\gamma > 0$ is the coefficient of relative risk aversion. I define $\vartheta \equiv \frac{1-\gamma}{1-1/\psi}$ to simplify notation when solving the model. Importantly, when $\psi > 1$ investors prefer the early resolution of uncertainty, and when $\psi < 1$, investors would rather wait until uncertainty is resolved in the future.

Aggregate log consumption growth $\triangle c_{t+1}$ and log dividend growth $\triangle d_{t+1}$ follow exogenous stochastic processes:

$$x_{t+1} = \rho_x x_t + \varphi_x \sigma(s_{t+1}) \varepsilon_{x,t+1} \tag{2.3}$$

$$\triangle c_{t+1} = \mu_c + x_t + \varphi_c \sigma(s_{t+1}) \varepsilon_{c,t+1} \tag{2.4}$$

$$\triangle d_{t+1} = \mu_d + \rho_d x_t + \varphi_d \sigma(s_{t+1}) \varepsilon_{d,t+1} + \rho_{cd} \varphi_c \sigma(s_{t+1}) \varepsilon_{c,t+1} \tag{2.5}$$

where lowercase letters represent natural logarithms, $\triangle c_{t+1} = \log C_{t+1} - \log C_t$. The state variable x_t models a persistent, mean-reverting conditional growth rate of consumption and dividends as in Bansal and Yaron [2004]. Aggregate shocks are independent and identically distributed normal random variables:

$$\varepsilon_{x,t+1}, \varepsilon_{c,t+1}, \varepsilon_{d,t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0,1) \tag{2.6}$$

I allow for correlation between consumption and dividend growth by introducing the parameter $\rho_{cd} \in \mathbb{R}$ as in Bansal et al. [2012]. Aggregate shocks share a common stochastic volatility $\sigma(s_{t+1})$ that is a function of a binary state variable:

$$s_t \sim_{\text{i.i.d.}} \text{Bernoulli}(\pi_h) \tag{2.7}$$

When $s_{t+1} = 1$, the volatility of consumption and dividend growth $\sigma(s_{t+1} = 1) = \varsigma_h > 1$ is high relatively to normal times $\sigma(s_{t+1} = 0) = 1$. Histories of volatility states have superscripts $s^{t+1} \equiv \{\dots, s_{t-1}, s_t, s_{t+1}\}$.[1] Aggregate consumption is distributed equally to all investors in each period, hence there is no heterogeneity

[1] I deviate from the first order autoregression law of motion for a common stochastic variance

in ex post consumption growth. The difference between the aggregate consumption and dividend endowments defines agents' labor income process. For parsimony, I collect all parameters of the model in a vector Θ.

Investor i prices some asset j such that the consumption Euler equation derived from the optimality conditions to the problem posted by (2.1) and (2.2):

$$\mathbb{E}_{i,t}\left[\beta^{\vartheta}\left(\frac{C_{i,t+1}}{C_{i,t}}\right)^{-\frac{\vartheta}{\psi}} R_{c,t+1}^{\vartheta-1} R_{j,t+1}\right] = 1 \tag{2.8}$$

I assume consumption growth and returns for any assets in consideration are jointly conditionally lognormal. Then, investor i's log stochastic discount factor $m_{i,t+1}$ is a function of consumption growth and the return on the aggregate consumption claim $r_{c,t+1}$:

$$m_{i,t+1} \equiv \vartheta \log \beta - \frac{\vartheta}{\psi} \triangle c_{t+1} + \big(\vartheta - 1\big) r_{c,t+1} \tag{2.9}$$

and their consumption Euler equation may be conveniently expressed in logs:

$$\mathbb{E}_{i,t}\big[m_{i,t+1} + r_{j,t+1}\big] + 1/2 \mathrm{Var}_{i,t}\big(m_{i,t+1} + r_{j,t+1}\big) = 0 \tag{2.10}$$

There is no arbitrage, thus each investor's Euler equation pricing any asset j holds in expectation. For the sake of expositional clarity, I utilize the no arbitrage assumption to construct an "average agent" whose Euler equation satisfies the cross-sectional average of (2.10) for every period t:

$$\int \mathbb{E}_{i,t}\big[m_{i,t+1} + r_{j,t+1}\big] di + 1/2 \int \mathrm{Var}_{i,t}\big(m_{i,t+1} + r_{j,t+1}\big) di = 0 \tag{2.11}$$

While the true identity of the average agent is unknown, deliberately assuming the point of view of the average agent conveniently replaces the task of accounting for an arbitrarily chosen atomistic agent.

in Bansal and Yaron [2004] because solving the model necessitates an iterative procedure to find a fixed point over a countable number of volatility histories. A continuous support for volatility shocks paired with persistence makes the task of counting volatility histories, even under a discretization procedure, numerically unwieldy.

2.2.1 Information Sets and Higher Order Beliefs

Investors cannot directly observe the conditional mean of consumption growth, x_t. Rather, each investor i privately observes a noisy signal $x_{i,t}$ in each period t:

$$x_{i,t} = x_t + \varphi_s \epsilon_{i,t} \tag{2.12}$$

where $\epsilon_{i,t} \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$ is independent and identically distributed over time, as well as in the cross-section of investors. In any given period agents differ solely by their expectations of consumption growth. This is the heart of the model.

The consequence of assuming limited information, rational expectations, and no arbitrage is that agents' higher order beliefs enter the state space of the model. That is, agents recursively forecast the forecast of other agents. I motivate this idea with a simple example: consider the case of power utility when $\vartheta = 1$. After solving forward in time, the Euler equation for agent i pricing the dividend claim evaluates to:

$$\begin{aligned} pd_t = \delta + \left(1 - {}^1\!/\!_\psi\right)\mathbb{E}_{i,t}\left[x_t\right] di + \kappa_{1,m}\left(1 - {}^1\!/\!_\psi\right)\mathbb{E}_{i,t}\left[\mathbb{E}_{i,t+1}\left[x_{t+1}\right]\right] + \cdots \\ \cdots + \mathbb{E}_{i,t}\left[\mathbb{E}_{i,t+1}\left[\cdots \mathbb{E}_{i,t+\tau}\left[x_{t+\tau}\right]\cdots\right]\right] + \cdots + {}^1\!/\!_2 \sum_{\tau=0}^{\infty} \kappa_{1,m}^{\tau} \mathrm{Var}_{i,t+\tau}\left(m_{t+1+\tau} + r_{m,t+1+\tau}\right) \end{aligned} \tag{2.13}$$

where δ is a constant in terms of model parameters. Inside investor i's expectation operator conditional on their information in period t, I substitute out investor i's future conditional expectations and replace them with expectations of some agent $j \neq i$:

$$\mathbb{E}_{i,t}\left[\mathbb{E}_{i,t+1}\left[\cdots \mathbb{E}_{i,t+\tau}\left[x_{t+\tau}\right]\cdots\right]\right] = \mathbb{E}_{i,t}\left[\mathbb{E}_{j,t+1}\left[\cdots \mathbb{E}_{j,t+\tau}\left[x_{t+\tau}\right]\cdots\right]\right] \tag{2.14}$$

for $\tau > 0$. Investor i has no special information about future expectations of investor j. I assume investor i's average expectation is coincident with investor i's expection of investor j's expectation:

$$\mathbb{E}_{i,t}\big[\mathbb{E}_{i,t+1}\big[\cdots \mathbb{E}_{i,t+\tau}\big[x_{t+\tau}\big]\cdots\big]\big] = \mathbb{E}_{i,t}\Big[\int \mathbb{E}_{j,t+1}\big[\cdots \int \mathbb{E}_{j,t+\tau}\big[x_{t+\tau}\big]dj\cdots\big]dj\Big] \tag{2.15}$$

Applying the assumptions in equations (2.14) and (2.15), equation (2.13) can be modified for the average investor by aggregating over the cross-section of investors:

$$\begin{aligned} pd_t = \delta + \big(1 - {}^1\!/\!_\psi\big)\int \mathbb{E}_{i,t}\big[x_t\big]di + \kappa_{1,m}\big(1 - {}^1\!/\!_\psi\big)\int \mathbb{E}_{i,t}\Big[\int \mathbb{E}_{j,t+1}\big[x_{t+1}\big]dj\Big]di + \cdots \\ \cdots + \int \mathbb{E}_{i,t}\Big[\int \mathbb{E}_{j,t+1}\big[\cdots \int \mathbb{E}_{j,t+\tau}\big[x_{t+\tau}\big]dj\cdots\big]dj\Big]di + \cdots \\ \cdots + {}^1\!/\!_2\sum_{\tau=0}^{\infty}\kappa_{1,m}^{\tau}\int \mathrm{Var}_{i,t+\tau}\big(m_{t+1+\tau} + r_{m,t+1+\tau}\big)di \end{aligned} \tag{2.16}$$

From the perspective of the average agent in period t, the cross-sectional average expectation $\int \mathbb{E}_{i,t}\big[x_t\big]di$ is relevant for determining the price, which I call the first order expectation. In the model, it is common knowledge that investors form rational expectations. Hence investors' first order expectations are optimal forecasts of the true state, and the first order expectations behave as stochastic processes with publicly known characteristics. Investors use this knowledge when forming second order expectations. Since prices recursively discount the present value of consumption growth into the infinite future, so too does $\int \mathbb{E}_{i,t}\big[\int \mathbb{E}_{j,t+1}\big[x_{t+1}\big]dj\big]di$ and every higher order expectation in the future become relevant state variables. To tame notation, I follow the convention of Nimark [2011] and define each higher order expectation:

$$x^{(k)} \equiv \int \mathbb{E}_{i,t}\big[x_t^{(k-1)}\big]di \tag{2.17}$$

with $x_t^{(0)} \equiv x_t$ and the full hierarchy of expectations $X_t \in \mathbb{R}^\infty$:

$$X_t \equiv \begin{bmatrix} x_t \\ x_t^{(1)} \\ x_t^{(2)} \\ \vdots \end{bmatrix} \tag{2.18}$$

I refer to X_t as the *term structure of consensus growth.* To simplify the algebraic manipulations required to solve the model, I define the operator $\mathbf{H}$ which acts on X_t to annihilate the lowest order expectation and replaces it with the next highest:

$$\mathbf{H} = \begin{bmatrix} \underset{\infty\times 1}{\mathbf{0}} & \underset{\infty\times\infty}{\mathbf{I}_\infty} \end{bmatrix} \tag{2.19}$$

For example, first order expectations are simply the first element of $\mathbf{H}X_t$:

$$\int \mathbb{E}_{i,t}\big[x_t\big]di = \mathbf{e}_1'\mathbf{H}X_t \tag{2.20}$$

2.2.2 Solving the Model

Because I focus the model on explaining time series moments of the equity premium, I price three assets as in Beeler and Campbell [2012]: a claim on aggregate consumption, a claim on aggregate dividends, and the ex ante risk free rate. For analytical convenience, I log linearize the return on the consumption and dividend claims like Campbell and Shiller [1988]:

$$r_{c,t+1} = \kappa_0 + \kappa_1 pc_{t+1} - pc_t + \triangle c_{t+1} \tag{2.21}$$

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} pd_{t+1} - pd_t + \triangle d_{t+1} \tag{2.22}$$

The ex ante risk-free rate $r_{f,t}$ derives from the Euler equation pricing an asset that matures after one period with a certain a numeriare payoff equal to one:

$$r_{f,t} = -\int \mathbb{E}_{i,t}[m_{t+1}]di - {}^1\!/_2 \int \mathrm{Var}_{i,t}\big(m_{t+1}\big)di \tag{2.23}$$

I conjecture (and verify in the appendix) affine state-contingent functional forms for the three endogenous prices:

$$pc_t = g_0^{\mathrm{pc}}(s^t) + g_x^{\mathrm{pc}}(s^t)' X_t \tag{2.24}$$

$$pd_t = g_0^{\mathrm{pd}}(s^t) + g_x^{\mathrm{pd}}(s^t)' X_t \tag{2.25}$$

$$r_{f,t} = g_0^{\mathrm{rf}}(s^t) + g_x^{\mathrm{rf}}(s^t)' X_t \tag{2.26}$$

where the (s^t) notation indicates a functional dependence of the coefficients on the realized history of volatility states. Additionally, I conjecture X_t follows a first order vector autoregression with state-contingent coefficients:

$$X_t = \mathbf{A}(s^t) X_{t-1} + \mathbf{B}(s^t) \begin{bmatrix} \varepsilon_t \\ \epsilon_{i,t} \end{bmatrix} \tag{2.27}$$

where ε_t is a $n_\varepsilon \times 1$ vector of aggregate shocks. Investor i is aware of (2.27) but cannot directly observe X_t. Rather, they observe a $n_y \times 1$ vector $Y_{i,t}$ of aggregate growth rates and returns, in addition to their private noisy signal:

$$Y_{i,t} = \begin{bmatrix} x_{i,t} \\ \triangle c_t \\ \triangle d_t \\ r_{m,t} \\ r_{f,t-1} \end{bmatrix} \tag{2.28}$$

where the quantities in (2.28) evolve according to equations (2.3), (2.4), (2.5), (2.22), and (2.23). Taken together with (2.27), these equations form a state space system:

$$X_t = \mathbf{A}(s^t) X_{t-1} + \mathbf{B}(s^t) \begin{bmatrix} \varepsilon_t \\ \epsilon_{i,t} \end{bmatrix} \tag{2.29}$$

$$Y_{i,t} = \boldsymbol{\mu}_Y(s^t) + \mathbf{C}_1(s^t) X_t + \mathbf{C}_2(s^t) X_{t-1} + \mathbf{D}(s_t) \begin{bmatrix} \varepsilon_t \\ \epsilon_{i,t} \end{bmatrix} \tag{2.30}$$

I assume the time series observations of the market return $r_{m,t+1}$ and ex ante risk free rate $r_{f,t}$ are peppered with very small noise shocks $\varepsilon_{rm,t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$ and $\varepsilon_{rf,t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0,1)$ that are scaled by φ_{rm} and φ_{rf}, respectively. The noise shocks play an important role increasing the stochastic rank of the state space system. Without the noise shocks, the true state x_t is an invertible function of observables, effectively nullifying the role of private information in the model. Bansal et al. [2012] utilize invertibility, for example, to identify x_t by projecting consumption growth on the risk free rate and the price-dividend ratio. The state x_t is fully revealed to their representative agent in every period. Nonetheless, investors never learn the true state in my model. The best investor i can do is to endogenously form expectations of X_t that are optimal by some objective function, given their information set. A natural candidate for such expectations is the Kalman filter, which is optimal in the sense it is the minimum variance estimator of the state. Accordingly, investor i forecasts the state $\mathbb{E}_{i,t}\big[X_t\big] \equiv X_{i,t|t}$ by utilizing their knowledge (2.29) and (2.30) in combination with the Kalman filter:

$$X_{i,t|t} = \mathbf{A}(s^t)X_{i,t-1|t-1} + \mathbf{K}(s^t)\big(Y_{i,t} - \boldsymbol{\mu}_Y(s^t) - \big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)X_{i,t-1|t-1}\big) \tag{2.31}$$

The technical approach investors employ to form expectations is, from the econometrician's point of view, similar to Schorfheide et al. [2016], who use a Kalman filter along with a rich set of measurement errors to estimate long-run risks in the United States with data reaching back to the Great Depression. The Kalman filter is also used extensively in real world economic forecasting, so its application in the model is a realistic foundation for how practitioners can estimate consensus beliefs. Given the expectations (2.31) of each investor $i \in (0,1)$, the average agent estimates the state by a simple aggregation of individual agents expectations:

$$X_{t|t} \equiv \int_0^1 X_{i,t|t} di \tag{2.32}$$

With knowledge of the average agent's expectations, it is possible now to construct (2.27). The true long-run risk component x_t follows the publicly known law of motion given in (2.3). Higher order expectations are formed as endogenous functions of the known model parameters Θ and (2.32). I construct the sets of matrices $\mathbf{A}(s^t)$ and $\mathbf{B}(s^t)$ by carefully summing the implied dynamics of (2.3) and (2.32):

$$\mathbf{A}(s^t) = \begin{bmatrix} \underset{1\times 1}{\rho_x} & \underset{1\times \bar{k}}{\mathbf{0}} \\ \underset{\bar{k}\times 1}{\mathbf{0}} & \underset{\bar{k}\times\bar{k}}{\mathbf{0}} \end{bmatrix} + \begin{bmatrix} \underset{1\times 1}{0} & \underset{1\times \bar{k}}{\mathbf{0}} \\ \underset{\bar{k}\times 1}{\mathbf{0}} & \underset{\bar{k}\times\bar{k}}{\left[\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\right]_-} \end{bmatrix} + \begin{bmatrix} \underset{1\times\bar{k}+1}{\mathbf{0}} \\ \underset{\bar{k}\times\bar{k}+1}{\left[\mathbf{K}(s^t)(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t))\right]_-} \end{bmatrix} \tag{2.33}$$

$$\mathbf{B}(s^t) = \begin{bmatrix} \underset{1\times 1}{\varphi_x} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{n_\varepsilon\times 1}{\mathbf{0}} & \underset{n_\varepsilon\times n_\varepsilon}{\mathbf{0}} \end{bmatrix} + \begin{bmatrix} \underset{1\times n_\varepsilon+1}{\mathbf{0}} \\ \underset{1\times n_\varepsilon+1}{\left[\mathbf{K}(s^t)(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s_t))\right]_-} \end{bmatrix} \tag{2.34}$$

In the matrices above, the object $\mathbf{K}(s^t)$ is the Kalman gain and $\mathbf{D}_\varepsilon(s^t)$ is a submatrix of $\mathbf{D}(s^t)$ omitting the noise shock, which are both explicitly characterized in the appendix. In their native forms, (2.33) and (2.34) describe the infinite dimensional dynamics of the expectation hierarchy. Therefore any tractable numerical solution to the model necessitates a finite dimensional approximation to the expectations hierarchy X_t, truncating the size of $\mathbf{A}(s^t)$ and $\mathbf{B}(s^t)$. The minus subscript $[\cdot]_-$ indicates the approximation step required to set the order of $X_t \in \mathbb{R}^{\bar{k}+1}$ for some $0 < \bar{k} < \infty$. The choice of $\bar{k}$ induces approximation errors in the evolution of the state X_t, though Nimark [2011] shows such errors can become negligible for finite $\bar{k}$. His method demonstrates price dynamics can quickly converge by plotting impulse response functions for increasing values of $\bar{k}$.

Another numerical issue sprouts from the fact the matrices in (2.33) and (2.34) are functionally dependent on the full history of volatility states s^t. Given the model is structured upon an infinite time horizon, this implies the length of s^t grows one at a time with each successive period. Since I assume volatility can attain only two values, a solution to the model has 2^t distinct histories of volatility states, and after a handful of periods the model is more or less impossible to solve on a conventional computer. I follow Nimark [2014] by shortening the history of s^t to rolling window of a predetermined length τ. In this way, the model is solvable over the 2^τ histories of s^t. I denote the set of truncated volatility histories $\mathbb{S}(\tau) \equiv \{(s_{t-\tau+1}, s_{t-\tau+2}, \ldots, s_{t-1}, s_t)\}$. Because the number of histories grows exponentially in τ, so too does the computer time required to solve the model. But computational speed alone is not a sufficient criteria for choosing the number of states to track, as shorter histories induce larger approximation errors in the Kalman filtering equations. Therefore an ideal choice of τ is both numerically fast and yields small errors. With this in mind, I choose the value $\tau = 8$, which produces 256 distinct histories for every object in the model dependent upon s^t. With the state dynamics of X_t tamed by restricting attention to histories s^t of finite length, and considering only the first $\bar{k}$ orders of expectations, conditional expectations within the model must be defined explicitly in order to solve the model. For example, the expectation of the average agent in period t of a random variable in period $t+1$ requires use of truncated histories lead forward one period in time:

$$\int \mathbb{E}_{i,t}[g_x^{\mathrm{pc}}(s^{t+1})X_{t+1}]di = \left(\pi g_x^{\mathrm{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + \left(1-\pi\right)g_x^{\mathrm{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\right)\mathbf{H}X_t \tag{2.35}$$

where the histories s_1^{t+1} and s_0^{t+1} drop the most distant volatility state and add a realization of either a high or low volatility state in the next period, respectively. I present the full details of a fixed point algorithm for solving the model in the appendix.

Rounding up the model solution, I define a limited information equilibrium:

1. A limited information equilibrium is for every truncated volatility history $s^t \in \mathbb{S}(\tau)$:

1. Functions mapping the state into endogenous quantities (2.24), (2.25), and (2.26).
2. A law of motion for the state (2.29).
3. The aggregate resource constraints:

$$\int W_{i,t} di = W_{t+1} \tag{2.36}$$

$$\int C_{i,t} di = C_{t+1} \tag{2.37}$$

$$W_{t+1} = R_{c,t+1}\big(W_t - C_t\big) \tag{2.38}$$

Such that all investors $i \in (0, 1)$ satisfy the optimality conditions of their utility maximization problems (2.1) subject to their budget constraints (2.2).

2.3 Calibration

In total, the model has sixteen parameters:

$$\Theta \equiv \big\{\gamma, \psi, \beta, \mu_c, \mu_d, \rho_x, \rho_d, \rho_{cd}, \varphi_x, \varphi_c, \varphi_d, \varphi_{rm}, \varphi_{rf}, \varphi_s, \varsigma, \pi\big\} \tag{2.39}$$

Table 2.1 shows their calibrated values. I calibrate eleven of the sixteen model parameters in Θ relating to preferences, consumption growth, and dividend growth using the estimates from Bansal et al. [2012] (BKY). A number of papers estimate parameters in the long-run risks model, for example, Bansal and Yaron [2004] (BY), Beeler and Campbell [2012] (BC), and Schorfheide et al. [2016] (SSY). The commonality in these papers is that they all focus on U.S. data after 1930, but vary by their econometric techniques and model assumptions. BY and BKY employ the generalized method of moments to esitmate the model with annual data. SSY perform a Bayesian analysis using a sequential monte carlo method to estimate a nonlinear state-space model with mixed frequency data. Despite differences in econometric methodology and some minor theoretical variations in the models, all of the estimations manifest consistency in their best-fit parameter values. Risk aversion γ is at or near ten, which is commonly hailed as a success for the long-run risks model in light of the many vintages of asset pricing models prone to generating values of γ in excess of one hundred– a magnitude that is impossible to reconcile with the microeconomic evidence. The EIS parameter ψ is consistently greater than one, which is congruent with the microeconomic evidence of Gruber [2013], and provides a number of important implications within the model. For instance, given the level of risk aversion, investors prefer to resolve uncertainty surrounding future utility flows sooner rather than later. This drives up the equity premium when volatility is high, or when expected consumption growth is low. Hall [1988] and Campbell [1999] argue for a value of $\psi < 1$, close to zero. They estimate EIS by interpreting it as the slope from projecting time series data of consumption growth on the real risk free rate and price-dividend ratio. Bansal and Yaron [2004] show this estimate of EIS is biased downward by stochastic volatility in the process for consumption growth. All evidence notwithstanding, the debate over EIS remains an unsettled matter. The estimations performed by BY, BKY, and CB tend to fit the risk free rate poorly, a challenge unanimously shared by modern consumption based asset pricing models. SSY improve the model fit by adding a preference shock similar to Albuquerque et al. [2015]. However, the exogenous preference shock serves, more or less, as a reduced form stochastic process suited to match the risk free rate and lacks a firm microfoundation.

All of the estimations mentioned thus far produce a value of ρ_x near one, indicating the existence of a small, mean-reverting component to the conditional mean of consumption growth. In the full information setting, a large value of ρ_x assures the risk price for long-run risk shocks $\varepsilon_{x,t+1}$ is large, and is a necessary component to matching the equity premium. In a recent paper, Nakamura et al. [2016] estimate parameters in a consumption process with long-run risks using a panel of sixteen countries with over one hundred and twenty years of data. Their estimation yields strong evidence for long-run risks in consumption growth, and has the added benefit of excluding asset price data, immunizing their parameter estimates from the accusation that asset price data forces the large magnitude of ρ_x.

One aspect of my model that reduces its fit to the data is my choice to construct a stochastic volatility process that is independent and identically distributed over two discrete states. All variants of the full information long-run risks model discussed in this section have at least one first order autoregressive process driving the square of volatility. These models can outperform mine insofar as they offer a continuous support for $\sigma(s_{t+1})$ and time series persistence, features both supported by the data. Continuous support for $\sigma(s_{t+1})$ is problematic for my numerical procedure because it necessitates computing infinitely many fixed points over infinitely many volatility histories. Discretizing $\sigma(s_{t+1})$ helps by making the number of possible volatility histories $\mathbb{S}$ countably infinite. Truncating the set of histories to include only the past τ observations generates a finite set $\mathbb{S}(\tau)$ with precisely 2^τ elements, reducing the computational burden from infinitely many fixed points to 2^τ. Given τ, the probability of a high volatility state π and the level of a high volatility state ς must be chosen in a way that best preserves the time series fit of the model. This choice offers some flexibility since the parameters π and ς are not well identified–owing to the infinitely many ways to partition the time series data into two samples. Therefore I make a stylistic choice to reflect an intuitive notion of "bad times" that is consistent with the data by setting $\varsigma = 2$ and $\pi = 0.05$. This means, on average, a serious bout of high volatility occurs approximately once every twenty years. I came to these numbers by simulating ten million observations from the law of motion for variance using the parameters from BKY. If the value of variance veered into negative territory, I censored it with a small positive number. Then, I scaled each observation by the ergodic mean of variance, and took the square root to produce an average volatility multiplier. The parameter value $\pi = 0.05$ reflects the $1 - \pi$ empirical quantile of the ergodic distribution of volatility multipliers, and $\varsigma = 2$ is the sample mean of volatility multipliers in excess in excess of the $1 - \pi$ quantile.

2.4 Discussions of Results

The results of the model primary focus on risk premia of financial assets. What level of compensation do investors require to accept some non-zero weighting of a given return distribution in their portfolios? The answer to this question is settled by how the returns of a particular asset of interest covary with investors' stochastic discount factors. Assets that produce streams of returns with low covariance to the average investor's stochastic discount factor carry a small risk premium, becuase they are less likely to inconvenience the investor with bad payoffs when

Parameter	Calibrated Value
Preferences	
γ	10
ψ	1.5
β	0.9989
Consumption Growth	
μ_c	0.0015
φ_c	0.0072
ρ_x	0.975
φ_x	$0.038\varphi_c$
Dividend Growth	
μ_d	0.0015
ρ_d	2.5
ρ_{cd}	2.6
φ_d	$5.96\varphi_c$
Noise	
φ_s	$\{0, \varphi_x, 2\varphi_x, 3\varphi_x, 4\varphi_x\}$
φ_{rm}	0.02
φ_{rf}	0.02
Volatility	
ς	2.00
π	0.05

Table 2.1: Parameter estimates for the preferences, consumption growth, and dividend growth categories from BKY. They estimate the model using the generalized method of moments with annual data beginning in 1930 and ending in 2008. I solve the model over a discrete grid of parameter values for noise φ_s. At the smallest value $\varphi_s = 0$, agents in the model have full information. At the largest value $\varphi_s = 4\varphi_x$, the risk price for long-run shocks is negligible.

their marginal utility growth is highest. Oppositely, returns of assets with high covariance to the investor's marginal utility growth are a poor hedge for bad times and therefore require the highest compensation to hold. Because I price asssets in the model from the point of view of the average agent, it is this agent's stochastic discount factor I consider when thinking about risk premia.

Formally, the equity premium is the negative average covariance between the stochastic discount factor m_{t+1} and the market return $r_{m,t+1}$:

$$\int \mathbb{E}_{i,t}\big[r_{m,t+1} - r_{f,t}\big]di + \frac{1}{2}\int \mathrm{Var}_{i,t}\big(r_{m,t+1}\big)di = -\int \mathrm{Cov}_{i,t}(m_{t+1}, r_{m,t+1})di \quad (2.40)$$

where, provided all relevant quantites satisfy joint conditional log-normality, the average covariance is a simple expectation:

$$-\int \mathrm{Cov}_{i,t}(m_{t+1}, r_{m,t+1})di = -\int \mathbb{E}_{i,t}\big[\lambda_{\varepsilon}(s^{t+1})'\Sigma(s_{t+1})\varepsilon_{t+1}\varepsilon_{t+1}'\Sigma(s_{t+1})'\beta_{\varepsilon}(s^{t+1})\big]di \quad (2.41)$$

The fundamental difference between equations (2.40) and (2.41) and consumption-based models with representative agents are that the moments are cross-sectional averages. I organize objects on the right-hand side of equation (2.41) inside the expectaion such that they bare conventional economic interpretations. The $n_{\varepsilon} \times 1$ vector $\lambda_{\varepsilon}(s^{t+1})$ contains risk prices for the shocks in ε_{t+1}. A risk price is the premium an investor would pay to perfectly hedge a specific source of risk in their stochastic discount factor. For instance, if an asset provided a return exactly equal to the consumption shock every period, then an investor would require a premium equal to the risk price of the consumption shock to hold the asset. The derivation of $\lambda_{\varepsilon}(s^{t+1})$ simply requires solving for the innovation in the average agent's stochastic discount factor:

$$m_{t+1} - \int \mathbb{E}_{i,t}\big[m_{t+1}\big]di \equiv \lambda_{\varepsilon}(s^{t+1})'\Sigma(s_{t+1})\varepsilon_{t+1} + \lambda_{X}(s^{t+1})'X_t + \lambda_{\sigma}(s^{t+1}) \quad (2.42)$$

where the $n_\varepsilon \times n_\varepsilon$ matrix $\Sigma(s_{t+1})$ is a diagonal matrix that scales each shock in ε_{t+1} by their respective time series volatility. I refer to $\Sigma(s_{t+1})$ as the quantity of risk. Lastly, the $n_\varepsilon \times 1$ vector $\beta_\varepsilon(s^{t+1})$ are sensitivities of innovations in the market return $r_{m,t+1}$ to aggregate shocks ε_{t+1}:

$$r_{m,t+1} - \int \mathbb{E}_{i,t}\big[r_{m,t+1}\big]di = \beta_\varepsilon(s^{t+1})'\Sigma(s_{t+1})\varepsilon_{t+1} + \beta_X(s^{t+1})X_t + \beta_\sigma(s^{t+1}) \quad (2.43)$$

A detailed derivation of all of the objects in (2.41) is in the appendix.

2.4.1 Disagreement Shrinks the Risk Price for Long-Run Shocks

Figure 2.1 shows the unconditional mean risk price for the long-run shock is a decreasing function of noise. In fact, for a very small level of disagreement, the risk price for long-run shocks is negligible. When the level of noise is equal to the time series volatility of the long-run shocks, $\varphi_s = \varphi_x$, the risk price for long-run shocks is approximately cut in half relative to the full information model. And when noise is set to twice the time series volatility of long-run shocks, the reduction in the risk price relative to the full information model is more than eighty percent. The bottom line is investors are less willing to pay a substantive premium to hedge long-run risk shocks in their stochastic discount factors when the precisions of their signals is low relative to the time series volatility of the long-run risk shock.

To understand this result, consider the risk price for the long-run shock under full information $\varphi_s = 0$:

$$\lambda_{\varepsilon_x} = \frac{\kappa_1\big(\gamma - {}^1\!/\!_\psi\big)}{1 - \kappa_1\rho_x} \quad (2.44)$$

The large risk price for long-run shocks is driven by a highly persistent autoregressive process for the conditional mean x_t, in addition to the coefficient of relative risk aversion being markedly larger than the inverse EIS $\gamma - 1/\psi > 0$. Moreover, it is straightforward to see that under power utility, a special case of Epstein-Zin-Weil preferences where $\gamma = 1/\psi$, the risk price for long-run shocks is necessarily zero. The important takeaway is that under full information, equation (2.44) is large and does not change over time.

When the scale of noise is positive $\varphi_s > 0$, higher order expectations enter the state space of the model and the mathematical form of risk price for the long-run risk shock changes to:[2]

$$\lambda_{\varepsilon_x}(s^t) = \kappa_1\big(\vartheta - 1\big)g_x^{\text{pc}}(s^t)' \frac{\mathbf{B}(s^t)\mathbf{e}_x}{\varphi_x \sigma(s_t)} \tag{2.45}$$

Thus the risk price in equation (2.45) is the product of model parameters, the endogenous coefficients mapping higher order expectations to the price-consumption ratio, and the scaled first column of the instantaneous impact matrix. The linearization constant $\kappa_1 \in (0,1)$ is a function of the steady state price-consumption ratio, and is very close to one in value regardless of the scale of noise I employ, hence the latter two terms must drive the negative relationship between noise and the risk price.

The motivation for this effect is found in the expectations of the average agent (2.32). In each period, the average agent's higher order expectations are a weighted sum of their a priori estimate of the state and a noisy innovation in observed data. When the scale of noise is zero, $\varphi_s = 0$, the average agent perfectly observes the state in every time period, therefore the Kalman gain allocates zero weight to the observed evolution of macroeconomic time series and places unity weight

[2]Despite the visible differences between equations (2.44) and (2.45), the numerical solution to the model perfectly equates these two quantities when $\varphi_s = 0$.

on investors' perfectly accurate signals of the conditional mean of consumption growth. Accordingly, higher order expectations are identical $x_t^{(0)} = x_t^{(k)}$ for all $k > 0$ and $t \in \mathbb{N}$. As the scale of noise increases $\varphi_s \to \infty$ and the quality of investors' information sets deteriorates, the Kalman gain shifts weight away from the noisy signal toward the observed macroeconomic aggregates and prices, as these alternate data series become valuable sources of information regarding investors' conditional expectations of consumption growth. Importantly, the magnitude of the shift away from the noisy signal and toward data innovations increases when signals are less precise. This effects the price dynamics $g_x^{\mathrm{pc}}(s^t)$ and the scaled impact matrix ${}^{\mathbf{B}(s^t)\mathbf{e}_x}/_{\varphi_x \sigma(s_t)}$ in a way that attenuates the magnitude of the risk price for the long-run shock. That is, the elements of $g_x^{\mathrm{pc}}(s^t)$ and ${}^{\mathbf{B}(s^t)\mathbf{e}_x}/_{\varphi_x \sigma(s_t)}$ are decreasing in the level of noise φ_s, so their product in equation (2.45) naturally leads to a smaller risk price. This is driven by the first column of the Kalman gain $\mathbf{K}(s^t)$ in the construction of $g_x^{\mathrm{pc}}(s^t)$ and ${}^{\mathbf{B}(s^t)\mathbf{e}_x}/_{\varphi_x \sigma(s_t)}$ through equations (A.43) and (2.34).

Figure 2 shows the market return's average beta for the long-run risk shock is decreasing in noise. The market return is less sensitive to long-run risks as investors' information sets disperse. The intuition for this negative relationship is similar to the negative relationship between risk prices and noise. In the full-information benchmark, the market return beta for long-run risk is:

$$\beta_{\varepsilon_x} = \kappa_{1,m} \frac{\rho_d - {}^1/\psi}{1 - \kappa_{1,m}\rho_x} \tag{2.46}$$

Equation (2.46) shows the market return is more sensitive to long-run risk shocks as financial leverage ρ_d increases, when EIS ψ increases, when long-run consumption growth is more persistent, and for a higher steady state price-dividend ratio. When

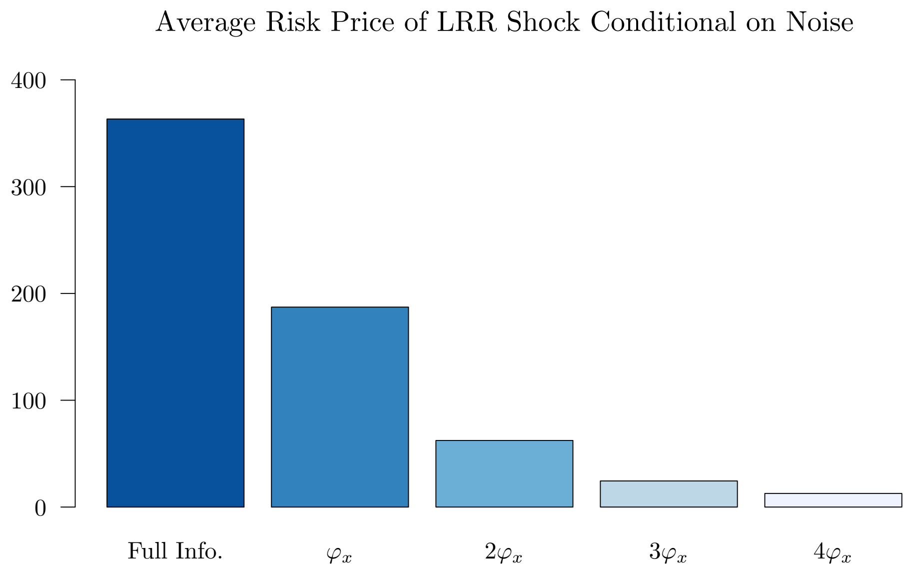

Figure 2.1: The unconditional average risk prices for the long-run risk shock across different levels of noise. Full information indicates $\varphi_s = 0$. I calibrate positive values of noise to multiples of the time series volatility of the long-run risk shock. As investors' information sets grow more disparate, the average risk price for the long-run shock approaches zero.

the scale of noise $\varphi_s > 0$, the beta for the long-run risk shock is given by:

$$\beta_{\varepsilon_x}(s^t) = \kappa_{1,m} g_x^{pd}(s^t)' \frac{\mathbf{B}(s^t)\mathbf{e}_x}{\varphi_x \sigma(s_t)} \tag{2.47}$$

What drives the shrinking $\beta_{\varepsilon_x}(s^t)$ as $\varphi_s \to \infty$? It cannot be $\kappa_{1,m}$, because the steady state price-dividend ratio does not change much with φ_s. The correct attribution focuses on the shrinking elements in $g_x^{pd}(s^t)$ and ${}^{\mathbf{B}(s^t)\mathbf{e}_x}/_{\varphi_x\sigma(s_t)}$ as noise φ_s increases. Again, the Kalman gain $\mathbf{K}(s^t)$ is the source of these changes. For ${}^{\mathbf{B}(s^t)\mathbf{e}_x}/_{\varphi_x\sigma(s_t)}$, I refer to the explanation given earlier in this section, and logic similar to $g_x^{\mathrm{pc}}(s^t)$ holds for $g_x^{pd}(s^t)$ through equation (A.47). As a result, innovations in the market return are less sensitive to long-run shocks when investors' have heterogeneous beliefs surrounding expected consumption growth.

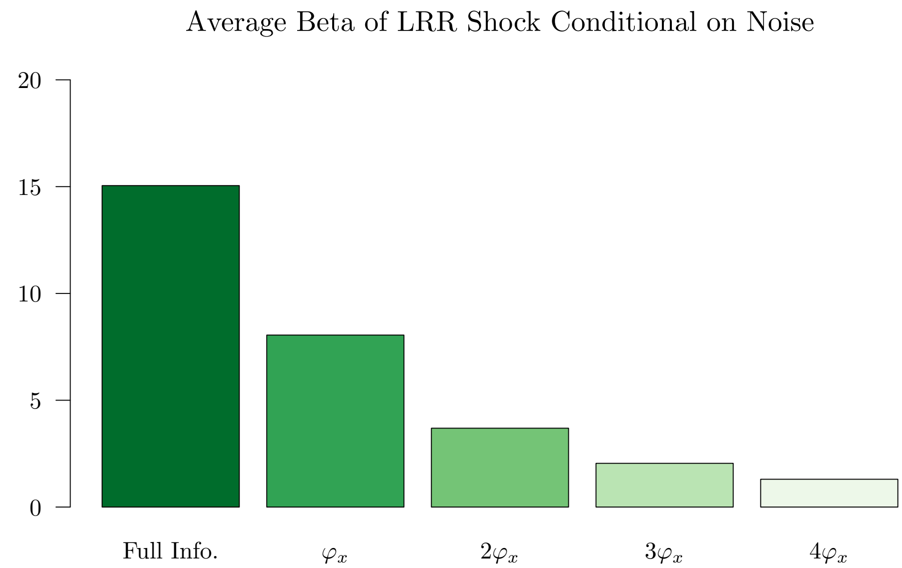

Figure 2.2: The unconditional average risk prices and betas for the long-run risk shock across different levels of noise. Full information indicates $\varphi_s = 0$. I calibrate positive values of noise to multiples of the time series volatility of the long-run risk shock φ_x. As investors information sets grow more disparate, the average beta for the long-run shock approaches zero.

2.4.2 The Risk Price for Long-Run Shocks is Countercyclical

Figure 3 shows the risk price for long-run shocks is increasing in recent spells of high volatility. The leftmost bar indicates the risk price for the long-run risk shock is slightly larger than fifty after τ periods of consecutive low volatility, a history of states I call "Smooth Sailing". The next bar immediately to the right of "Smooth Sailing" has the same history of low volatility states, except the most recent volatility state is high. In this case, the risk price nearly doubles to a value of about one hundred. Moving to the the next bar to the right, the history is of low volatility states, except the most recent two periods are high volatility. Here, the risk price is over one hundred and fifteen. This pattern continues, monotonically adding another high volatility state until the full history is uniformly high, which I call "Wild Times". The takeaway is the risk price for the long-run risk shocks increases at a decreasing rate in the local history of time series volatility, yielding a value slightly larger than one hundred and twenty three during "Wild Times". The time-varying behavior of the risk price in equation (2.45) is in stark contrast to the time-invariant risk price (2.44) of Bansal and Yaron [2004]. While Bansal and Yaron [2004] do achieve time-varying risk *premia*, this is the consequence of scaling equation (2.44) by an exogenous law of motion for the quantity of risk. The model presented in this paper achieves a similar effect because the quantity of risk $\Sigma(s_{t+1})$ scales risk prices in (2.41), however, the sources of time-varying risk premia are twofold because $\lambda_\varepsilon(s^{t+1})$ changes over time as well.

So why is the risk-price for long-run shocks positively related to time series volatility? The law of motion for x_t, equation (2.3), evolves according to a heteroskedastic shock. Investors observe a signal (2.12) that obscures x_t with homoskedastic noise. The result is the precision of the signal (2.12) is higher during high volatility states of nature because the volatility of the process driving the true state $\varphi_x \sigma(s_{t+1})$ increases relative to the scale of noise φ_s:

$$\int \mathrm{Var}_{i,t}(x_{i,t+1}) = \big(\varphi_x \sigma(s_{t+1})\big)^2 + \varphi_s^2 \tag{2.48}$$

Then, as shown earlier in Section 2.4.1, the risk price for long-run shocks approaches its full information limit as the scale of noise approaches zero. I argue for this feature in the model, that investors have more precise signals in high volatility states of nature, with the recent empirical work of Loh and Stulz [2015]. They show in "bad times", or when ex ante uncertainty is high, analyst forecasts of earnings are more accurate than in "good times" when uncertainty is low. Loh and Stulz argue increasing analyst accuracy during bad times is the result of intensified efforts arising from career concerns. While my model abstracts from the specifics of agents' effort or career worries, it accurately captures the same signal-to-noise dynamics with a homoskedastic noise shock clouding investor information and a hetroskedastic shock driving the state.

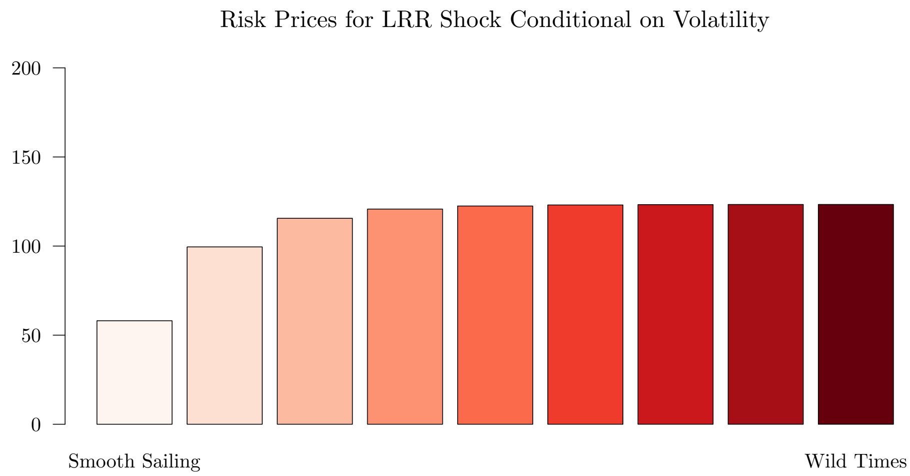

Figure 2.3: Risk prices for the long-run risk shock $\varepsilon_{x,t}$ over different volatility histories. I define "Smooth Sailing" as a τ-consecutive periods of low volatility, and oppositely, "Wild Times" as τ-consecutive periods of high volatility. The second bar immediately to the right of "Smooth Sailing" changes the most recent period from low to high volatility. The third bar does the same, but for the two most recent periods. The pattern continues until the last bar, where all states are high volatility. The scale of noise is calibrated $\varphi_s = 2\varphi_x$.

2.4.3 Three Factors Explain The Term Structure of Consumption Growth

Litterman and Scheinkman [1991] show three statistical factors summarize nearly all of the time series variation in treasury bond yields. They construct factors by computing the first three principal components from the empirical covariance matrix of a panel of bond yields. The first factor equally weights all bond yields, a level factor: all bond yields rise and fall together. The second factor puts high weight on short maturity yields and low weight on long maturity yields, a slope factor, which proxies a steepening or flattening of the yield curve. The third factor is a curvature factor, where long and short duration yields move opposite intermediate maturity yields. The success of Litterman and Scheinkman is not limited to explaining the comovement of bond yields. Research using their procedure reveals interesting level, slope, and curvature factors in other asset pricing domains. For example, Lustig et al. [2011] find a slope factor explains most of the common variation in the cross-section of currency returns. And in a recent paper, Clarke [2016] discovers a level, slope, and curvature factor in portfolios of stocks sorted by their expected returns. He finds, "A standard linearized Consumption Capital Asset Pricing Model explains almost all of the spread in average returns across portfolios. Since the Level, Slope, and Curvature model explains almost all of the variance in the expected return sorted portfolios, another interpretation is that the model represents high frequency factors mimicking changes in expected future consumption growth."

Figure 2.4 shows a level, slope, and curvature factor for the expectations hierarchy. I construct the three factors by simulating one million observations of the expectations hierarchy and then computing principal components from the empirical covariance matrix. The first factor equally weights higher order expectations, a level factor. The second factor puts high weight on lower order expectations, and low weight on higher order expectations, a slope factor. The third factor puts high weight on lower and higher order expectations, an low weight on intermediate order expectations, a curvature factor. Table 2.2 decomposes the variance explained by the three factors. The first factor accounts for over ninety eight percent of the vari-

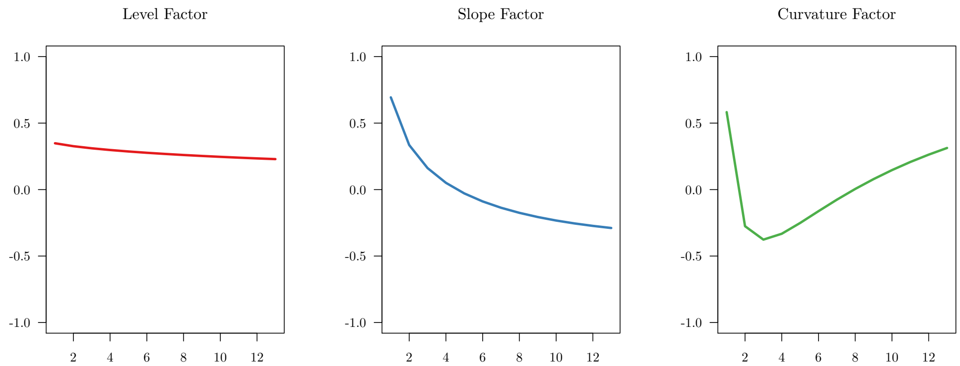

Figure 2.4: Plots of the coefficients from the first three principal components of the empirical covariance matrix of X_t with a simulated sample of 1,000,000 observations under $\tau = 8$ and $\bar{k} = 12$. X-axis labels refer to the order of expectation, and the y-axis are scalar magnitudes of loadings on principal components.

ation in higher order expectations, indicating the strong comovement in consensus expectations of consumption growth. The second factor explains considerably less than the first, nearly two percent of the variation. And the curvature factor explains merely twenty basis points. Taken together, the three factors account for over ninety nine percent of the time series variation in higher order expectations. Table A.1 decomposes the variance of the individual higher order expectations. The level factor explains over ninety percent of the variation for every order of expectation. The slope and curvature factors have the most explanatory power in the lowest and highest order expectations, and very little in intermediate order expectations.

	Level	Slope	Curvature
Share of Explained Variance	0.982	0.016	0.002

Table 2.2: Shares of variance in the empirical covariance matrix of X_t explained by the first three principal components (level, slope, and curvature). Values in the table are each principal component's associated eigenvalue divided by the sum of all eigenvalues. I estimate the empirical covariance matrix of X_t with a simulated sample of 1,000,000 observations under $\tau = 8$ and $\bar{k} = 12$.

2.4.4 Disagreement Mutes Price Responses to Changes in State Variables

Figure 2.5 shows impulse response of the price-consumption ratio to a long-run risk shock. The blue line represents the response under full information, and the green line represents the response under limited information. The relative response of the full information price-consumption ratio is over fifty percent larger than the price-consumption ratio under limited information. Why does the price-consumption ratio in the limited information economy react to an important shock with inertia? The answer lies in the signal extraction problem faced by investors. Under full information, investors observe the long-run shock, and given their perfect knowledge of the state, prices accurately represent the evolving state of the economy in future periods. Under limited information, an arbitrary investor observes the state with noise, and then solves a signal extraction problem to devise an operative best-estimate of the state in each future period via the Kalman filter. The hierarchy of expectations X_t is a construction of investors' beliefs surrounding the future path of x_t, representing investors' consensus expectations of their future average consumption growth. The Kalman filter uses information from *all* aggregate shocks to optimally estimate the hidden state, thus the noise present in investors' signals shifts weight away from the long-run shock and toward the aggregate consumption and dividend shocks. This creates meaningful inertia in investors' expectations, and consequently in the price-consumption ratio.

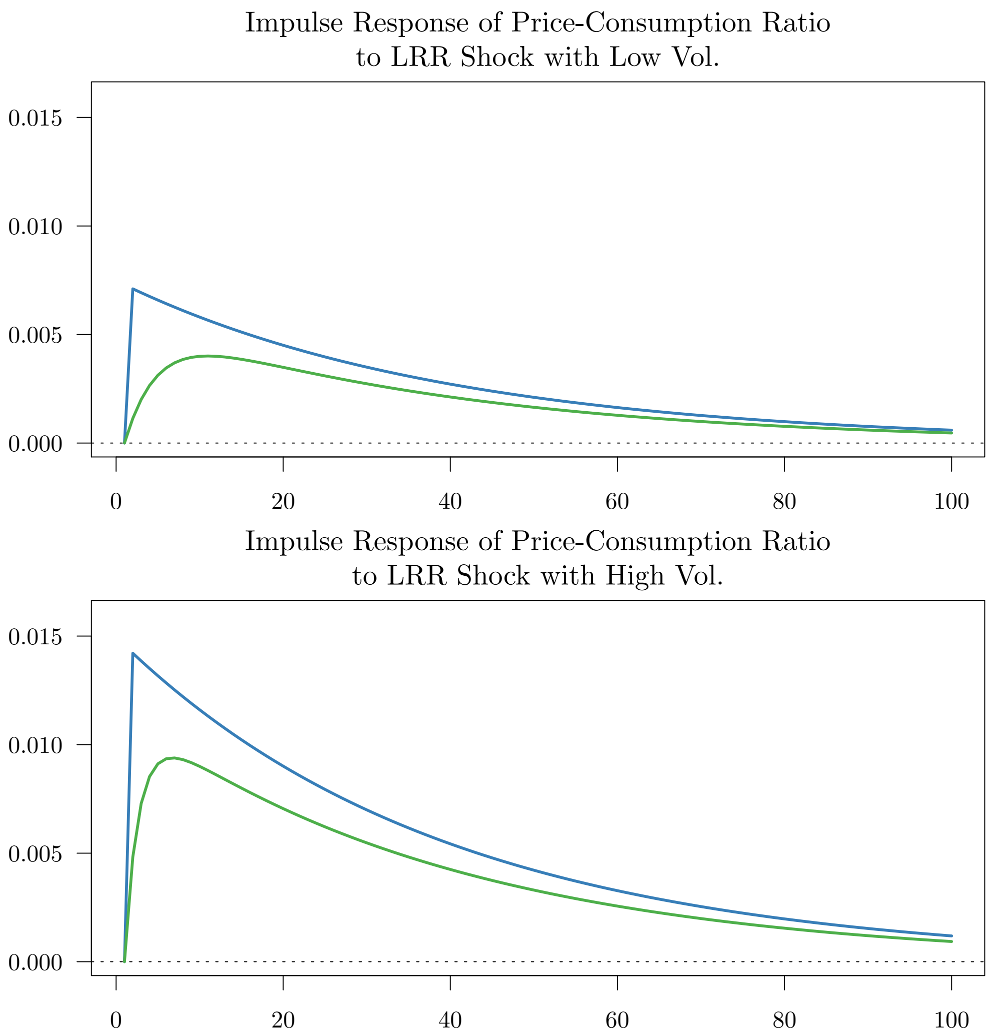

Figure 2.5: Impulse response functions for the price-consumption ratio from a two standard deviation shock to $\varepsilon_{x,t}$. Data plotted in the top graph was generated with consecutive periods of low volatility, and the bottom graph was generated with consecutive periods of high volatility. The level of noise is calibrated to $\varphi_s = 2\varphi_x$.

Figure 2.7 shows impulse response functions for the hierarchy of expectations X_t to a long-run risk shock, consumption shock, and dividend shock. In the top plot, the expectations hierarchy responds positively to a positive long-run risk shock. The first-order expectation $x_t^{(1)}$ nearly mimics the behavior of the true state x_t, the

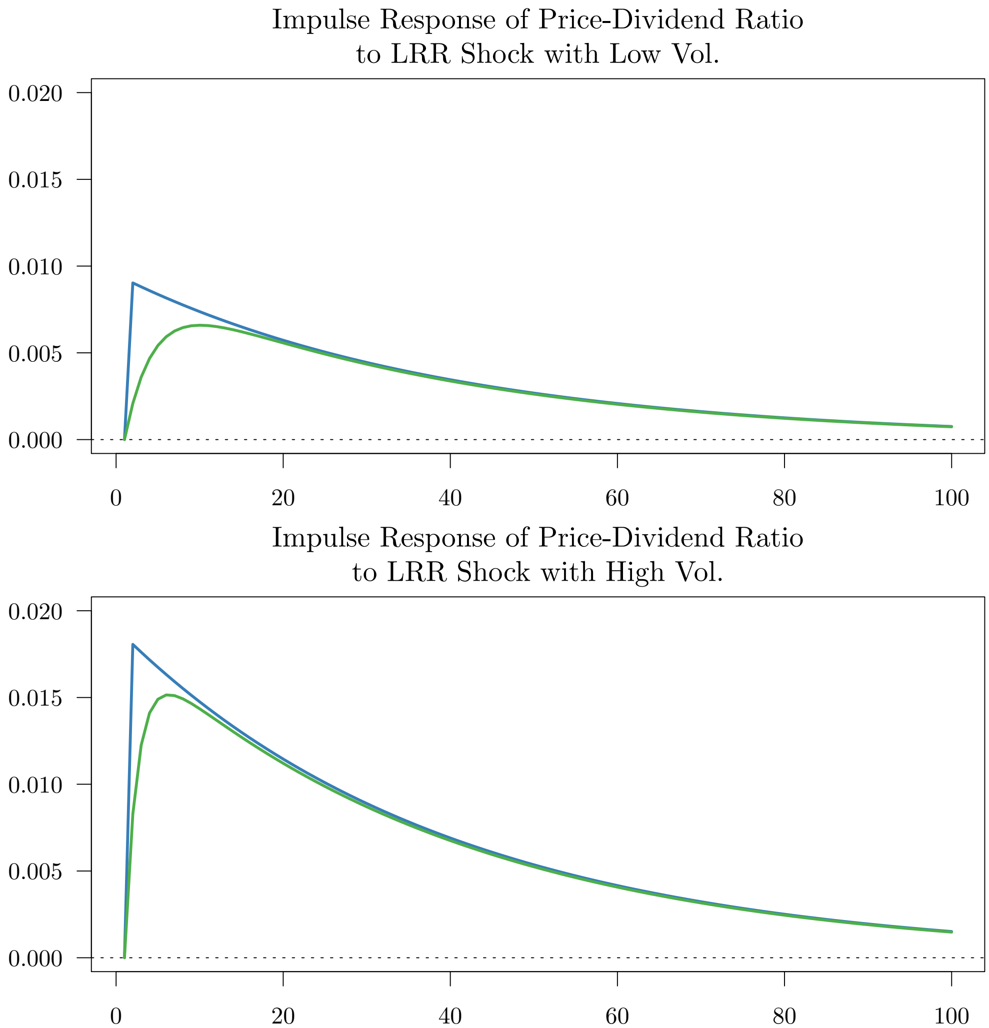

Figure 2.6: Impulse response functions for the price-dividend ratio from a two standard deviation shock to $\varepsilon_{x,t}$. Data plotted in the top graph was generated with consecutive periods of low volatility, and the bottom graph was generated with consecutive periods of high volatility. The level of noise is calibrated to $\varphi_s = 2\varphi_x$.

second order expectation $x_t^{(2)}$ nearly mimics the first order expectation $x_t^{(1)}$, and so on. Nimark [2011] shows the comovement of the expectations hierarchy X_t is a decreasing function of the scale of the idiosyncratic noise shocks φ_s. As agents' information sets converge to full information, the expectations hierarchy X_t becomes perfectly correlated with the true state x_t, which is equivalent to the original setting of Bansal and Yaron [2004]. Further, in the full information economy, the law of motion for x_t is purely exogenous, therefore under no circumstances would x_t ever respond to a consumption shock or dividend shock. The expectations hierarchy X_t, on the other hand, *does* respond to consumption and dividend shocks under limited information. The bottom two plots of Figure 2.7 exhibit this interesting behavior. In response to a positive consumption shock, the largest orders of expectations have a pronounced response relative to the lowest orders. Under the metaphor that X_t is a term structure of expectations, a positive consumption shock is tantamount to a positive shock to a steepening factor. Intuitively, the sign makes sense because a positive consumption shock impacts both consumption and dividend growth positively, in the same way a positive shock to x_t would. The opposite dynamics hold true for the dividend shock–the expectations hierarchy X_t responds negatively, and most dramatically in the largest orders of expectations.

2.4.5 Disagreement Increases During High Volatility Spells

Figure 2.8 presents a plot of the expectations hierarchy simulated for one hundred periods. Below it is a plot of the interquartile range of consumption growth forecasts. The first fifty periods of the simulation have uniformly low volatility. Then, at the halfway point, the remaining fifty periods have high volatility. Evidently the higher volatility of long-run shocks leads to larger fluctuations in the expectations hierarchy throughout the second half of the sample. The interquartile range of consumption growth forecasts nearly doubles after transitioning into the high volatility regime from the low volatility regime. Time series volatility leads to high levels of disagreement in consumption growth forecasts through the Kalman updating equation (2.31). To see why, notice each forecast of consumption growth $f_{i,t+1|t}$ is treated as an independent and identical draw form a conditional distribution of forecasts:

$$f_{i,t+1|t} \sim_{\text{i.i.d.}} \mathcal{N}\big(\mu_c + \mathbf{e}_1'\mathbf{H}X_t,\ \mathbf{e}_1'\mathbf{H}\mathbf{P}_i(s^t)\mathbf{H}'\mathbf{e}_1\big) \tag{2.49}$$

where the $\mathbf{P}_i(s^t)$ is the conditional cross-sectional covariance matrix:

$$\mathbf{P}_i(s^t) = \mathbb{E}_{i,t}\left[\Big(\mathbb{E}_{i,t}\big[X_t\big] - \textstyle\int \mathbb{E}_{i,t}\big[X_t\big]dj\Big)\Big(\mathbb{E}_{i,t}\big[X_t\big] - \textstyle\int \mathbb{E}_{i,t}\big[X_t\big]dj\Big)'\right] \tag{2.50}$$

Recall the precision of investors signals $x_{i,t}$ varies over time given the heteroskedasticity in long-run shocks, thus increasing the precision of signals during bad states of nature with high volatility. The consequence of greater precision is the first column of the Kalman gain places larger weight on $x_{i,t}$ and less weight on other observed sources of information. By subtracting equation (2.32) from (2.31), it is clear to see how the time-variation in the Kalman gain directly influences the role investor-specific noise plays in consensus forecast dispersion:

$$\begin{aligned}
&\mathbb{E}_{i,t}\big[X_t\big] - \int \mathbb{E}_{i,t}\big[X_t\big]dj = \\
&\big(\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big)\big(\mathbb{E}_{i,t-1}\big[X_{t-1}\big] - \textstyle\int \mathbb{E}_{i,t-1}\big[X_{t-1}\big]dj\big) + \boxed{\mathbf{K}(s^t)\mathbf{D}_\epsilon \epsilon_{i,t}}
\end{aligned} \tag{2.51}$$

where the matrix $\mathbf{D}_\epsilon$ simply contains the scale parameters for investors' signals:

$$\mathbf{D}_\epsilon \equiv \begin{bmatrix} \varphi_s \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.52}$$

Thus cross-sectional forecasts of consumption growth will tend to disperse when time the time series volatility of aggregate shocks is high. This feature of the model is broadly consistent with the data. Figure 2.9 plots the cross-sectional dispersion of real nondurable consumption growth forecasts from the Survey of Professional

Forecasters (SPF). Table 2.3 conditions time series averages of the interquartile range of consumption growth forecasts on ex post business cycle classifications provided by the National Bureau of Economic Research (NBER). In recession states of the business cycle, the average interquartile range is about 1.70% relative to 1.28% during expansion states, indicating greater dispersion in forecasts during bad times. While the model matches the direction of the relationship between cross-sectional disagreement and time series volatility, it does not match the scale. The survey of professional forecasters data yields estimates of the interquartile range measured in percent, while forecast disagreement generated by the model is measured in mere basis points. There are a number of possible explanations for this fact. First, Clements [2014] provides evidence of anti-herding in SPF data for inflation and output. Rulke et al. [2012] find similar results on a data set including professional forecasters from around the globe. Forecasters working in the private sector may have career incentives to enhance the marketability of their forecasts by straying from the pack, increasing forecast dispersion as a result. For instance, if customers rank forecasters by their recent relative performance, and not their career spanning mean-squared errors, forecasters can capture a large share of new customers by providing extreme forecasts. An example of this is providing forecasts of a harsh recession when economic indicators begin to show signs of slowing growth, and then claiming valuable foresight ex post if the extreme prediction pans out. A second possibility is I have misspecified the information sets of investors by focusing disagreement on a state variable with a relatively small volatility. Simply expanding the state-space of the model to include a state with a volatility similar to the time series volatility of aggregate consumption growth would increase the spread in consumption growth forecasts to a scale seen in the data. Third, it may be that professional forecasters are not privy to long-run risks in the way utilized in the asset pricing literature, and requiring forecasting agents of to think of the world as an endowment economy is not a fruitful econometric exercise.

Cycle	Interquartile Range
1983 - 2015	
Expansion	0.0128
Recession	0.0170
1990 - 2015	
Expansion	0.0095
Recession	0.0143

Table 2.3: Average interquartile range of real nondurable consumption growth forecasts from the Survey of Professional Forecasters conditioned on ex post NBER business cycle classifications. The top half of the table reports averages for the full sample from the third quarter of 1981 through the end of 2015. The bottom half reports averages from the second quarter of 1990 through 2015. The Federal Reserve Bank of Philadelphia took over the survey in the second quarter of 1990 and dramatically increased the breadth and quality of the sample.

2.4.6 Consensus Forecast Errors are Priced Risk

Equation (2.53) shows the risk price for the expectations hierarchy:

$$\lambda_X(s^t)' = -\gamma \mathbf{e}_1'\big(\mathbf{I} - \mathbf{H}\big) + \big(\vartheta - 1\big)\kappa_1\Big(g_x^{\mathrm{pc}}(s^t)'\mathbf{A}(s^t) - \textstyle\int \mathbb{E}_{i,t-1}\big[g_x^{\mathrm{pc}}(s^t)'\mathbf{A}(s^t)\mathbf{H}\big]\,di\Big) \tag{2.53}$$

The first term in (2.53) indicates the difference between the true state and the first order expectation carries a risk price of γ. Therefore investors are willing to pay a premium to hedge their first order expectations deviation from the true conditional mean growth rate. The second term in (2.53) prices a forecast error in the evolution of the price consumption ratio. However, the second term is small and insubstantial, and may be ignored for that reason. In the full information economy, the risk price for aggregate consumption shocks is simply the coefficient of relative risk aversion $\lambda_{\varepsilon_c} = \gamma$, and the risk price in the limited information economy has approximately the same magnitude $\lambda_X(s^t) \approx \gamma$. Intuitively, this observation implies investors dislike consensus forecast errors between their first order expectations and the conditional mean of consumption growth equally as much as transient shocks to aggregate consumption. The tendency for first order expectations to deviate from the true conditional mean is modulated by the noise parameter. Section 2.4.4 shows that the in the expectations hierarchy is increasing in the scale of noise φ_s. Therefore the quantitative importance of consensus forecast errors grows as investor information sets diverge.

2.5 Conclusion

Consumption based asset pricing links investors' marginal utility of consumption to discount rates that price assets. If investors have Epstein-Zin-Weil preferences, and there is a small, persistent component in expected consumption growth, then news about the long-run economic growth carries a large risk price. This result revolves around full information: at every point in time, the state of the economy is precisely known to a representative investor. I extend the long-run risks economy to include a large number of heterogeneous investors with imperfect knowledge of conditional expected consumption and dividend growth. Investors use their knowledge of deep parameters in the economy and the laws of motion of driving the unobserved state variables to estimate the true position of the economy via a signal extraction exercise. Consensus beliefs about economic growth compose the state space of the model, altering price dynamics relative to the full information benchmark. Beliefs about expected economic growth respond more sluggishly to shocks in the limited information economy because noise obscures the true conditional growth rate. When investors' information sets about the state diverge, the risk price for long-run news about economic growth becomes small and inconsequential. The implication is, at most, only a handful of basis points can separate the most optimistic and pessimistic investors' private beliefs about economic growth in order for long-run risks to serve as a source of risk premium. Consumption growth forecasts from professional forecasters exhibit a level of cross-sectional dispersion that is counterfactually large relative to the model. As investors information sets converge toward the truth, the risk price for long-run shocks increases toward the full information limit. Risk prices exhibit large countercylicalility as investor beliefs grow more accurate. This happens because the precision of investors' private signals increases during periods of high volatility in economic aggregates and asset prices. This paper attempts to take heterogeneous beliefs seriously in an economy with long-run risks, and shows disagreement is a challenge for this literature, opening new empirical targets for the model.

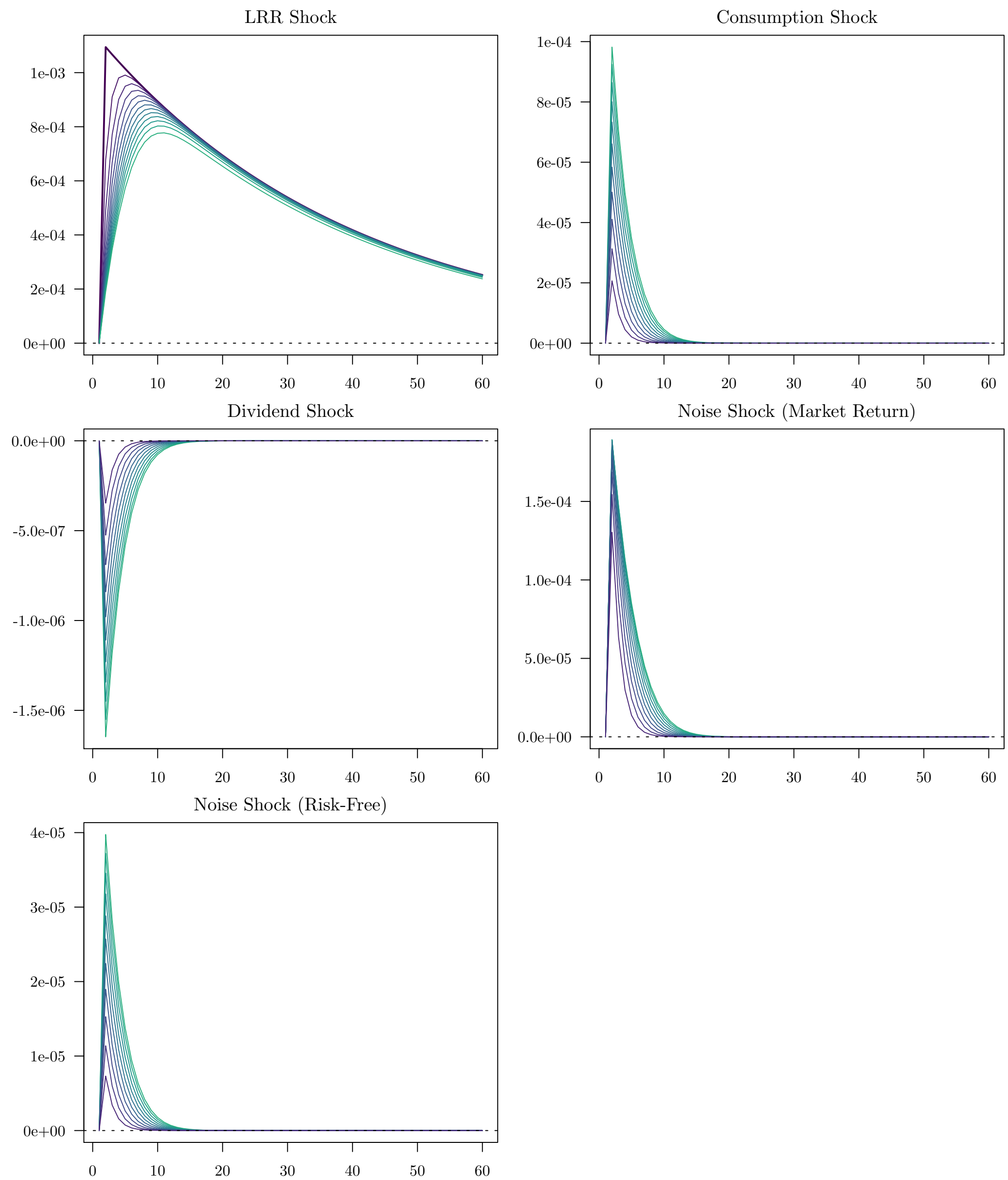

Figure 2.7: Impulse response functions for the hierarchy of expectations X_t from a two standard deviation shock to $\varepsilon_{x,t}$ (top), $\varepsilon_{c,t}$ (middle), and $\varepsilon_{d,t}$ (bottom). On the top graph, the thick dark blue line represents $x_t^{(0)}$, but this variable is absent from the other two graphs because it does not respond to consumption or dividend shocks. The increasingly lighter shades of turquoise and green correspond to higher orders of expectations. The level of noise is calibrated to $\varphi_s = 2\varphi_x$.

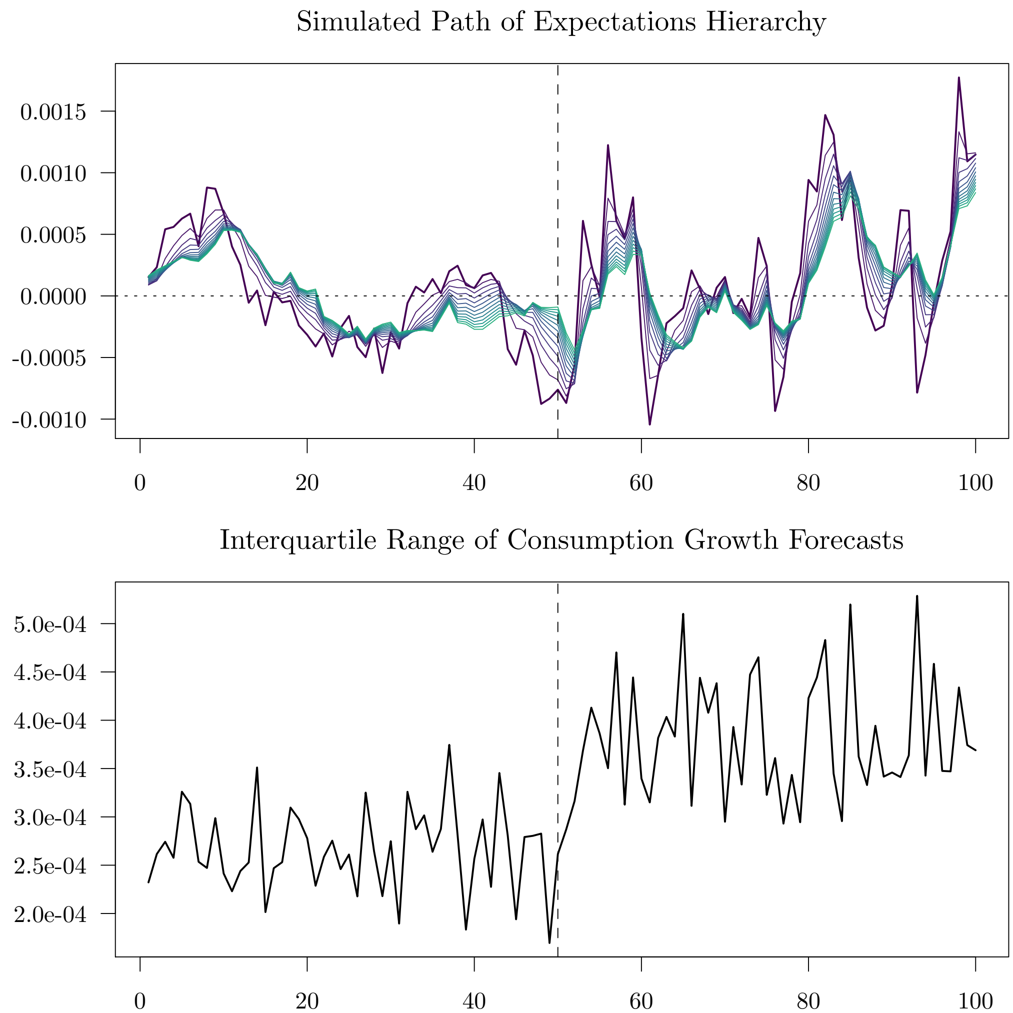

Figure 2.8: The top plot exhibits a simulation of the expectations hierarchy for $T = 100$ periods. The first fifty periods are low volatility states, and the last fifty periods are high volatility. A vertical dashed line separates the two regimes. The bottom plot shows the interquartile range of consumption growth forecasts. Each period, I randomly sample $N = 50$ independent draws to simulate a panel of forecasts. I then estimate the 25th and 75th percentiles and compute their difference.

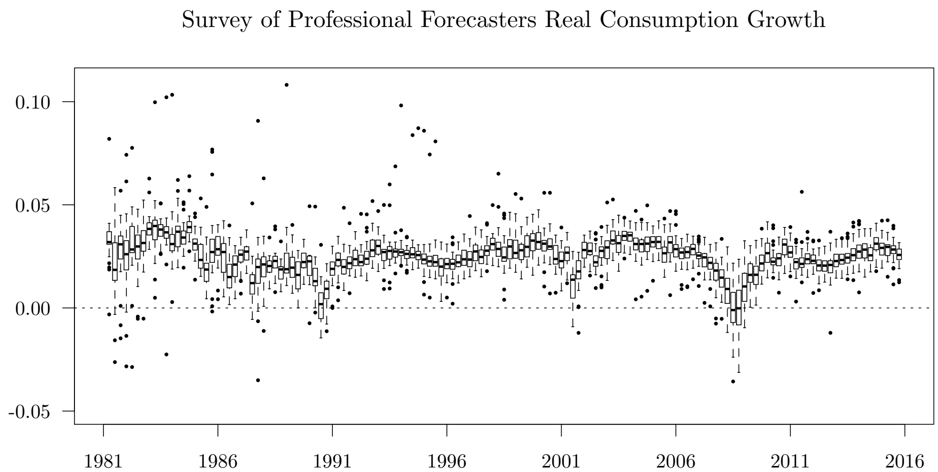

Figure 2.9: Cross-sectional dispersion of annual real nondurable consumption growth forecasts from the Survey of Professional Forecasters. The forecast data are sampled at a quarterly frequency, beginning in the third quarter of 1981 and ending in the fourth quarter of 2015. I compute annual consumption growth for each forecaster in the sample by subtracting their forecast of the previous period revised log consumption level from the level forecasted four quarters ahead of the end of the measurement period.

CHAPTER 3

AMBIGUITY OVER THE LONG-RUN: THE WELFARE COSTS OF PERSISTENT GROWTH SHOCKS

3.1 Introduction

Mehra and Prescott [1985] demonstrate a simple model of a representative agent with power utility cannot replicate the size of the equity premium. A subsequent literature seeks to "solve" the equity premium puzzle by enriching economic models to match the statistical moments of asset price data. Generally, this literature focuses on two paths to refining such a model. First, one can equip a representative agent with exotic preferences, with the goal in mind of increasing the flexibility to the stochastic discount factor to better price risky streams of payoffs. Ideally, the preference specification and parameter values are corroborated by an economic theory and microeconomic evidence. Second, and typically in combination with an exotic preference specification, one can augment the exogenous stochastic process driving consumption growth with different types of risks. The model of consumption growth should be supported by a basic econometric calibration with a time series of observed aggregate consumption data. Despite the variety of preferences and risks in structural asset pricing models, most share a common assumption that agents inhabit in a world with rational expectations. That is, the representative agent knows the correct risk model of consumption growth, and prices risky payoffs with the correct model.

Using the framework Barillas et al. [2009], I take a different route, a route where the representative agent has the correct risk model of consumption growth in mind as an approximating model, but is uncertain of the model's validity. To surmount model uncertainty, she employs a robust risk model by choosing the worst-case probability distribution of consumption growth subject to an entropy constraint. The entropy constraint dictates the worst-case model cannot be "too far" from the approximating model in a statistical sense–one cannot arbitrarily expect tomorrow to be armaggedon with probability one if a simple statistical test suggests such a model is near the realm impossibility. The result of this exercise is an ability to robustly price risky assets under a pessimistic view of consumption risk. If assets are priced in this way, what share of the risk premium on common stocks can be attributed to model uncertainty? How much would an economic agent pay to absolve themself from model uncertainty? In particular, I seek answers to these questions with an updated model of consumption growth having both long-run risks (LRR) of Bansal and Yaron [2004] and rare disasters of Barro [2006].

The LRR literature conjectures consumption growth has both a persistent component and stochastic volatility. When paired with the preferences of Epstein and Zin [1989] (EZ), a LRR model of consumption growth does a good job matching key statistical moments of asset price data. A caveat is Bansal and Yaron rely on a value of the intertemporal elasticity of substitution (IES) to be larger than one, which conflicts with some earlier estimates, like Hall [1988], who estimates IES to be approximately one half. The IES is a theoretical parameter that regulates economic agents willingness to substitute consumption over time, and subsequently drives savings behavior. The implication for asset pricing being that, via a standard Euler equation relationship, fluctuations in the risk-free rate are amplified fluctuations in consumption growth by a factor greater than one.

The rare disasters literature begins with Rietz [1988], who shows that the infrequent market crashes are capable of generating a large equity premium and a low risk-free rate. Barro [2006] expands this literature by constructing a prolific panel data set of global consumption growth that reaches back as far as 1790. With a subsample of 35 countries over 100 years, Barro finds the probability of aggregate consumption declining by 15 percent or more in a given year is 1.7 percent per year. Both Barro and Rietz rely on a simple asset pricing model where the representative agent has time-separable power utility, and is subject to iid shocks. More recent work in Barro [2009], Barro and Jin [2011], and Nakamura et al. [2013] discard time-separable utility in favor of EZ preferences. Asset pricing models with disasters and EZ pereferences, similar to the LRR literature, require a value of IES to be greater than one.[1]

The paper that is most similar to this one is Bidder and Smith [2013]. However, they focus on a consumption model with stochastic volatility, and without long-run risks and rare disasters.

3.2 A Model of Consumption Growth

Consumption growth is indexed by discrete time $t = 0, 1, 2, \ldots$ and has the form,

$$\triangle c_{t+1} = \mu + x_t + \varepsilon^c_{t+1} + \varepsilon^d_{t+1} \tag{3.1}$$

$$x_{t+1} = \rho x_t + \varepsilon^x_{t+1} \tag{3.2}$$

$$\varepsilon^c_{t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma_c^2) \tag{3.3}$$

$$\varepsilon^x_{t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma_x^2) \tag{3.4}$$

$$\varepsilon^d_{t+1} \sim_{\text{i.i.d.}} \text{Bernoulli}(\pi_d) \cdot \log(1 - b) \tag{3.5}$$

where the notation $\triangle c_{t+1} = c_{t+1} - c_t$ is the first difference of logs and μ is the trend growth rate. An uppercase letter $C_t = \exp(c_t)$ is the exponent of the log value. Rare disasters take two distinct values $\varepsilon^d_{t+1} \in \{0, \log(1 - b)\}$, and follow a

[1] In an unpublished set of presentation slides, Viktor Tsyrennikov is to my knowledge the first to apply the framework of Barillas et al. [2009] to a model of rare disasters.

two-state Markov chain with a stationary distribution,

$$f^d(\varepsilon^d_{t+1}) = \pi_d^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}(1-\pi_d)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}}.$$

Taken together, the shocks are also represented as a vector ε_{t+1},

$$\varepsilon_{t+1} \equiv \begin{pmatrix} \varepsilon^d_{t+1} \\ \varepsilon^c_{t+1} \\ \varepsilon^x_{t+1} \end{pmatrix} \sim f^\varepsilon(\varepsilon_{t+1}), \tag{3.6}$$

where $f^\varepsilon(\cdot)$ is the joint distribution function and $\varepsilon^{t+1} = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_t, \varepsilon_{t+1})$.

3.3 Two Observationally Equivalent Agents

3.3.1 Type-I Agent with EZ Preferences

The type-I agent faces consumption growth (3.1)-(3.2) and has peferences of Epstein and Zin [1989] and Tallarini [2000],

$$\mathbf{V}^{\mathrm{I}}(s_t) = c_t - \beta\vartheta \log \mathbb{E}_t[\exp(-\mathbf{V}^{\mathrm{I}}(s_{t+1})/\vartheta)], \tag{3.7}$$

where the utility paramter ϑ can be expressed as,

$$\vartheta = -\frac{1}{(1-\beta)(1-\gamma)}. \tag{3.8}$$

The parameter $\beta \in (0,1)$ is a one-period time discount rate, $\gamma > 0$ is atemporal risk aversion, and EIS is fixed equal to one. When $\gamma >$ EIS, the type-I agent prefers the early resolution of uncertainty. The state vector s_t consists of consumption c_t and the time-varying component of the growth rate x_t,

$$s_t = \begin{pmatrix} c_t \\ x_t \end{pmatrix}. \tag{3.9}$$

By standard arguments, the stochastic discount factor of the type-I is,

$$\Lambda_{t,t+1} = \beta \frac{C_t}{C_{t+1}} \frac{\exp(-\mathbf{V}^{\mathrm{I}}(s_{t+1})/\vartheta)}{\mathbb{E}_t[\exp(-\mathbf{V}^{\mathrm{I}}(s_{t+1})/\vartheta)]}, \tag{3.10}$$

and can be used to price any risky payoff. An advantages of fixing EIS equal to one is that the value function has a closed-form solution. Simply guess the affine form,

$$\mathbf{V}^{\mathrm{I}}(s_t) = \alpha_0 + \alpha_c c_t + \alpha_x x_t, \tag{3.11}$$

and after some algebra, arrive at the solution,

$$\alpha_0 = \frac{\beta}{1-\beta}\Big(\frac{\mu}{1-\beta} - \Big(\frac{\beta}{(1-\beta\rho)(1-\beta)}\Big)^2 \frac{\sigma_x^2}{2\vartheta} - \Big(\frac{1}{1-\beta}\Big)^2 \frac{\sigma_c^2}{2\vartheta} - \vartheta \log\Big(1 - \pi_d + \pi_d (1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)\Big) \tag{3.12}$$

$$\alpha_c = \frac{1}{1-\beta} \tag{3.13}$$

$$\alpha_x = \frac{\beta}{(1-\beta\rho)(1-\beta)}. \tag{3.14}$$

3.3.2 Type-II Agent with Multiplier Preferences

The type-II agent faces consumption growth (3.1)-(3.2) and has multiplier preferences as in Hansen and Sargent [2001] and Barillas et al. [2009],

$$\mathbf{V}^{\mathrm{II}}(s_0) = \min_{m_{t+1}} \sum_{t=0}^{\infty} \mathbb{E}\Big[\beta^t M_t \Big(c_t + \beta\vartheta \mathbb{E}\Big[m_{t+1} \ln(m_{t+1}) | s^t\Big]\Big) | s^t\Big] \tag{3.15}$$

where,

$$M_{t+1} = m_{t+1} M_t \tag{3.16}$$

$$\mathbb{E}[m_{t+1} | \varepsilon^t, s_0] = 1 \tag{3.17}$$

$$m_{t+1} \geq 0 \tag{3.18}$$

$$M_0 = 1, \tag{3.19}$$

with s_0 given. This agent does not know the true probability distribution $f^{\varepsilon}(\cdot)$ of shocks. To deal with this concern, the type-II agent chooses a scalar m_{t+1} that is Borel measurable with respect to (ε^{t+1}, s_0). The choice of m_{t+1} minimizes the agent's lifetime expected utility, subject to an entropy constraint parameterized by ϑ. Intuitively, this agent chooses a probability distribution for risks to consumption growth that is the worst *indistinguishable* case. That is, it is very difficult for a discerning econometrician to separate the worst-case distribution from the actual data generating process with a relatively short time series of data. The recursively defined sequence $\{M_t\}_{t=0}^{\infty}$ of likelihood ratios is a non-negative martingale such that,

$$\mathbb{E}[M_{t+1}|\varepsilon^t, s_0] = M_t. \tag{3.20}$$

Therefore m_{t+1} is also known as the minimizing martingale increment. The likelihood ratios $\{M_t\}_{t=0}^{\infty}$ represent the distance between the type-II agent's constrained worst-case model and the approximating model (3.1)-(3.2).

After scaling by M, the same problem restated recursively is,

$$\mathbf{V}^{\text{II}}(s) = c + \min_{m(\varepsilon)\geq 0}\Big(\beta \int \big(m(\varepsilon)\mathbf{V}^{\text{II}}(s') + \vartheta m(\varepsilon)\log(m(\varepsilon))\big) f^{\varepsilon}(\varepsilon) d\varepsilon\Big) \tag{3.21}$$

where,

$$\mathbb{E}[m(\varepsilon)] = 1. \tag{3.22}$$

Solving for the minimizer $\hat{m}_{t+1}(s_{t+1})$ and substituting it back into the Bellman equation (3.21) yields,

$$\mathbf{V}^{\text{II}}(s_t) = c_t - \beta\vartheta \log \mathbb{E}_t[\exp(-\mathbf{V}^{\text{II}}(s_{t+1})/\vartheta)], \tag{3.23}$$

with the minimizing martingale increment,

$$\hat{m}_{t+1}(s_{t+1}) = \frac{\exp(-\mathbf{V}^{\text{II}}(s_{t+1})/\vartheta)}{\mathbb{E}_t[\exp(-\mathbf{V}^{\text{II}}(s_{t+1})/\vartheta)]}. \tag{3.24}$$

3.3.3 The Worst-Case Probability Distribution

The minimizing martingale increment (3.24) is proportional to,

$$\hat{m}_{t+1}(\varepsilon_{t+1}) \propto \exp\Big(-\frac{1}{\vartheta(1-\beta)}\varepsilon^c_{t+1} - \frac{1}{\vartheta(1-\beta)}\varepsilon^d_{t+1} - \frac{\beta}{\vartheta(1-\beta\rho)(1-\beta)}\varepsilon^x_{t+1}\Big). \quad (3.25)$$

To arrive at the worst-case distribution $\tilde{f}^\varepsilon(\cdot)$, the joint distribution of the approximating model $f^\varepsilon(\cdot)$ is simply multiplied by the distortion,

$$\tilde{f}^\varepsilon(\varepsilon_{t+1}) \propto \hat{m}_{t+1}(\varepsilon_{t+1}) f^\varepsilon(\varepsilon_{t+1}). \quad (3.26)$$

Under the worst-case distribution, the locations of the transient consumption shock and the persistent shock shift downward,

$$\varepsilon^c_{t+1} \sim \mathcal{N}\Big(-\frac{\sigma_c^2}{\vartheta(1-\beta)}, \sigma_c^2\Big) \quad (3.27)$$

$$\varepsilon^x_{t+1} \sim \mathcal{N}\Big(-\frac{\beta\sigma_x^2}{\vartheta(1-\beta\rho)(1-\beta)}, \sigma_x^2\Big). \quad (3.28)$$

And, the disaster risk probability is slightly higher,

$$f^d(\varepsilon^d_{t+1}) \propto \Big(e^{-\frac{\log(1-b)}{\vartheta(1-\beta)}}\pi_d\Big)^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}\Big(1-\pi_d\Big)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}}. \quad (3.29)$$

Evidently, the type-II agent's pessism is reflected in their choice of $\hat{m}_{t+1}$.

3.4 Thought Experiments

How much consumption would the type-I agent forgo each period to eliminate different components of physical risk from (3.1)-(3.2)? To answer this question, one needs the value function (3.11) for the type-I agent under (3.1)-(3.2), in addition to a value function solved under a less-risky process for consumption growth. Equating these two value functions permits one to solve for consumption under the less-risky process. The difference in these two consumption allocations is how much the type-I agent would pay to eliminate physical risk. To perform this analysis, consider the following four processes for consumption growth that eliminate different sources of physical risk, and are all equal in expectation to (3.1)-(3.2).

No Short-Run Risks:

$$\triangle c_{t+1} = \mu + \frac{1}{2}\sigma_c^2 + x_t + \varepsilon_{t+1}^d \tag{3.30}$$

$$x_{t+1} = \rho_x x_t + \varepsilon_{t+1}^x \tag{3.31}$$

No Long-Run Risks:

$$\triangle c_{t+1} = \mu + \varepsilon_{t+1}^c + \varepsilon_{t+1}^d \tag{3.32}$$

No Disasters:

$$\triangle c_{t+1} = \mu + \pi_d \log(1-b) + x_t + \varepsilon_{t+1}^c \tag{3.33}$$

$$x_{t+1} = \rho_x x_t + \varepsilon_{t+1}^x \tag{3.34}$$

No Risk:

$$\triangle c_{t+1} = \mu + \frac{1}{2}\sigma_c^2 + \pi_d \log(1-b) \tag{3.35}$$

Let $j \in \{\text{c,x,d}\}$ index the type of physical risk, and the value function with risk j eliminated is $\mathbf{V}_j^{\text{I}}(\cdot)$. Then the compensating variation $c_0 - c_{0,j}^{\text{I}}$ for a type-I agent eliminating risk j can be found by solving,

$$\mathbf{V}^{\text{I}}(c_0) = \mathbf{V}_j^{\text{I}}(c_{0,j}^{\text{I}}). \tag{3.36}$$

A type-II agent has the same value function as a type-I agent, so $c_{0,j}^{\text{I}} = c_{0,j}^{\text{II}}$. For a type-I agent, $c_0 - c_{0,j}^{\text{I}}$ is compensation for risk, while for a type-II agent, $c_0 - c_{0,j}^{\text{II}}$

is compensation for both risk and uncertainty. The principle difference between these two quantities is their expression in terms of β and γ for a type-I agent or ϑ for a type-II agent.

A type-II agent who does not fear model misspecification has $\vartheta = +\infty$ (or $\gamma = 1$). The compensating variation for this type of agent after eliminating physical risk is,

$$c_0 - c_{0,j}^{\text{II}}(r) = \lim_{\vartheta \to \infty} (c_0 - c_{0,j}^{\text{II}}). \tag{3.37}$$

The compensating variation for model uncertainty $c_{0,j}^{\text{II}}(r) - c_0^{\text{II}}$ can then be solved for with the following decomposition,

$$c_0 - c_{0,j}^{\text{II}} = (c_0 - c_{0,j}^{\text{II}}(r)) + (c_{0,j}^{\text{II}}(r) - c_{0,j}^{\text{II}}). \tag{3.38}$$

As Table 1 and Table 2 show, the representative agent is willing to pay more to resolve ambiguity surrounding rare disasters relative to long-run risks. Figure 1 drives this point home best:

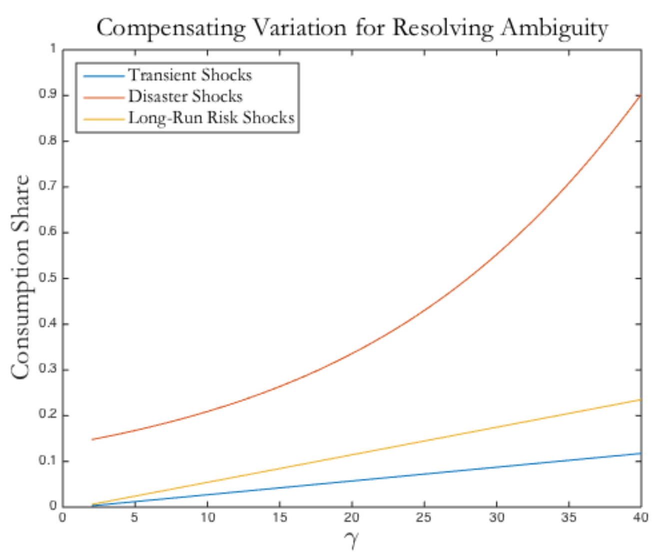

Figure 3.1: The lines represent the share of consumption the representative agent is willing to forgo to absovle ambiguity surrounding a specific type of physical risk for a given level of γ.

Compensation	Amount	Source
$c_0 - c^{\mathrm{I}}_{0,\mathrm{c}}$	$\frac{\beta\sigma_c^2\gamma}{2(1-\beta)}$	risk
$c_0 - c^{\mathrm{I}}_{0,\mathrm{x}}$	$-\frac{\beta^3}{(1-\beta\rho)^2}\frac{(1-\gamma)}{(1-\beta)}\frac{\sigma_x^2}{2\vartheta} - \frac{\beta}{1-\beta\rho}x_0$	risk
$c_0 - c^{\mathrm{I}}_{0,\mathrm{d}}$	$\frac{\beta}{1-\beta}\pi_d \log(1-b) - \frac{1}{(1-\beta)(1-\gamma)}\log\Big(1-\pi_d+\pi_d(1-b)^{1-\gamma}\Big)$	risk
$c_0 - c^{\mathrm{I}}_{0,\mathrm{cxd}}$	$\beta\left(\frac{\frac{1}{2}\sigma_c^2+\pi_d\log(1-b)}{1-\beta} - \frac{\beta^3}{(1-\beta\rho)^2}\frac{(1-\gamma)}{(1-\beta)}\frac{\sigma_x^2}{2\vartheta} - \frac{(1-\gamma)}{(1-\beta)}\frac{\sigma_c^2}{2\vartheta} - \frac{1}{(1-\beta)(1-\gamma)}\log\Big(1-\pi_d+\pi_d(1-b)^{1-\gamma}\Big)\right)$	risk
$c_0 - c^{\mathrm{II}}_{0,\mathrm{c}}$	$\frac{\beta\sigma_c^2}{2(1-\beta)}\left(1+\frac{1}{(1-\beta)\vartheta}\right)$	risk + uncertainty
$c_0 - c^{\mathrm{II}}_{0,\mathrm{x}}$	$\beta\left(\frac{\beta}{(1-\beta\rho)(1-\beta)}\right)^2\frac{\sigma_x^2}{2\vartheta} - \frac{\beta}{1-\beta\rho}x_0$	risk + uncertainty
$c_0 - c^{\mathrm{II}}_{0,\mathrm{d}}$	$\frac{\beta}{1-\beta}\pi_d\log(1-b) + \vartheta\log\Big(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)$	risk + uncertainty
$c_0 - c^{\mathrm{II}}_{0,\mathrm{cxd}}$	$\beta\left(\frac{\frac{1}{2}\sigma_c^2+\pi_d\log(1-b)}{1-\beta} + \left(\frac{\beta}{(1-\beta\rho)(1-\beta)}\right)^2\frac{\sigma_x^2}{2\vartheta} + \left(\frac{1}{1-\beta}\right)^2\frac{\sigma_c^2}{2\vartheta} + \vartheta\log\Big(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)\right)$	risk + uncertainty

Table 3.1: Compensation for elimating different sources of physical risk and model uncertainty. The source column indicates what has been eliminated: risk, uncertainty, or risk and uncertainty simultaneously. The amount column refers to the cost of eliminating risk and/or uncertainty in terms of structural parameters.

Compensation	Amount	Source
$c_0 - c^{\text{II}}_{0,\text{c}}(r)$	$\frac{\beta\sigma_c^2}{2(1-\beta)}$	risk
$c_0 - c^{\text{II}}_{0,\text{x}}(r)$	$-\frac{\beta}{1-\beta\rho}x_0$	risk
$c_0 - c^{\text{II}}_{0,\text{d}}(r)$	$\frac{\beta}{1-\beta}\pi_d \log(1-b)$	risk
$c_0 - c^{\text{II}}_{0,\text{cxd}}(r)$	$\frac{\beta}{1-\beta}\left(\frac{1}{2}\sigma_c^2 + \pi_d \log(1-b)\right)$	risk
$c^{\text{II}}_{0,\text{c}}(r) - c^{\text{II}}_{0,\text{c}}$	$\frac{\beta\sigma_c^2}{2(1-\beta)}\left(\frac{1}{(1-\beta)\vartheta}\right)$	uncertainty
$c^{\text{II}}_{0,\text{x}}(r) - c^{\text{II}}_{0,\text{x}}$	$\beta\left(\frac{\beta}{(1-\beta\rho)(1-\beta)}\right)^2\frac{\sigma_x^2}{2\vartheta}$	uncertainty
$c^{\text{II}}_{0,\text{d}}(r) - c^{\text{II}}_{0,\text{d}}$	$\vartheta\log\left(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\right)$	uncertainty
$c^{\text{II}}_{0,\text{cxd}}(r) - c^{\text{II}}_{0,\text{cxd}}$	$\beta\left(\left(\frac{\beta}{(1-\beta\rho)(1-\beta)}\right)^2\frac{\sigma_x^2}{2\vartheta} + \left(\frac{1}{1-\beta}\right)^2\frac{\sigma_c^2}{2\vartheta} + \vartheta\log\left(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\right)\right)$	uncertainty

Table 3.2: Compensation for elimating different sources of physical risk and model uncertainty. The source column indicates what has been eliminated: risk, uncertainty, or risk and uncertainty simultaneously. The amount column refers to the cost of eliminating risk and/or uncertainty in terms of structural parameters.

3.5 Conclussion

The structural asset pricing literature has long sought to explain statistical moments of aggregate asset price fluctuations with models of full information rational expectations. While these models are critically important benchmarks for measurment, in reality the agents who set the prices on risky assets do not all share a common risk model for a fixed set of underlying physical risks. A more realistic approach acknowledges that the true underyling risks are ambiguous to economic agent, and that agents must somehow make the best possible decisions nonetheless. Two models of consumption growth have come to dominate the recent literature of structural explanations of the equity premium, long-run risks and rare disasters. I cast both of these models in a theoretical world where the representative agent does not know the true stochastic processes for these two sources of risk, but behaves optimally by assuming a worst-case joint distribution that would be difficult to reject with real data. A subsequent welfare analysis reveals that the representative agent would be willing to forgo a relatively larger share of their consumption each period for the rest of their infinite life to remove ambiguity surrounding rare disasters relative to ambiguity surrounding long-run risks.

CHAPTER 4

NEWS UNCERTAINTY SHOCKS AND SOVEREIGN DEBT

4.1 Introduction

The budding literature on news shocks shows economic agents' response to anticipated changes in an economy accounts for a substnatial share of macroeconomic fluctuations. I contribute to the conversation on news shocks with a model where news uncertainty, defined as the second moment of the stochastic process generating news shocks, is allowed to randomly fluctuate over time. Intuitively, this models exogenous, time-varying changes in agents' level of certainty of news about future outcomes. In this paper, I focus on news uncertainty shocks to total factor productivity (TFP) and interest rates in a small open economy. I show that the magnitude of an economy's response to news uncertainty shocks is positively related to the length of time before news shocks realization in the level, and that an economy's response to news uncertainty can eclipse the response to contemporaneous volatility shocks.[1]

Why are news uncertainty shocks a valid modelling decision? First, consider the data. A news shock to interest rates in the context of an open economy can be interpreted as a shock to the level of a forward rate. Thus, time-varying news uncertainty is equivalent to conditional heteroskedasticity in forward rates. Figure 4.1 displays the first differences of the two year forward rate on German Bunds. I use this particular time series because in an open economy DSGE, a shock to the level of interest rates eight quarters in the future is equivalent to a shock to the two year forward rate. As the plot shows, throughout the financial crisis, innovations to the forward rate are large, though they substantially attenuate beginning around 2012.

[1]There are many different interesting applications for news uncertainty shocks not studied in this paper, in particular, fiscal policy.

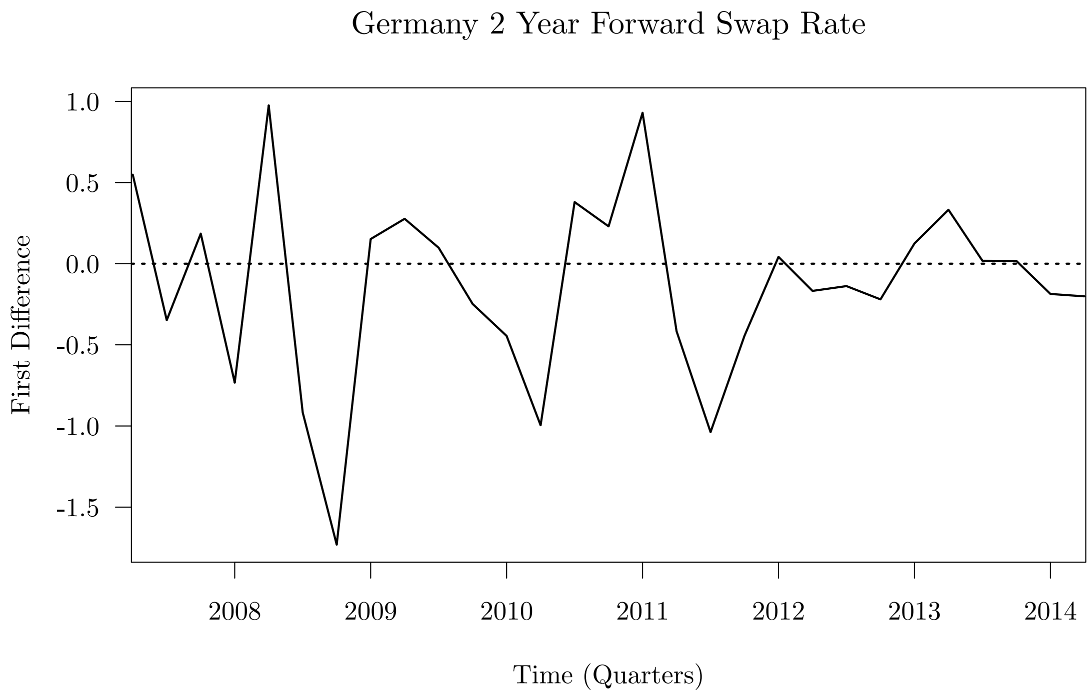

Figure 4.1: Time series plot of first difference of 2 year forward swap rates on German sovereign debt with a 3 month tenor. Time observations are at a quarterly frequency.

Figure 4.2 displays the estimated time series of TFP news shocks using the identification scheme of Barsky and Sims [2011] and the matlab code and data of Kurmann and Otrok [2013]. Here, a TFP news shock is defined to be a shock orthogonal to a conventional TFP shock that accounts for the largest share of variation in future TFP. News about TFP is anemic in the 1960's, as well as the 1990's and 2000s. The variance of TFP news is noticeable larger throughout the 1970's and 1980's. A Ljung-Box test of the squared TFP news shock time series yields strong evidence of conditional heteroskedisticity.[2]

[2]The Ljung-Box test is a portmanteau test, meaning it does not provide a sharp definition of an alternative hypothesis. Further, conditional heteroskedasticity may disappear under future identification schemes for TFP news shocks. I take the TFP argument from the empirical angle with a grain of salt, and wish to point out that homoskedasticity was a natural assumption to initiate studies of economic models with news shocks, though it may be more difficult to justify against robust alternatives in the wake of advancing numerical methods that are evermore easier

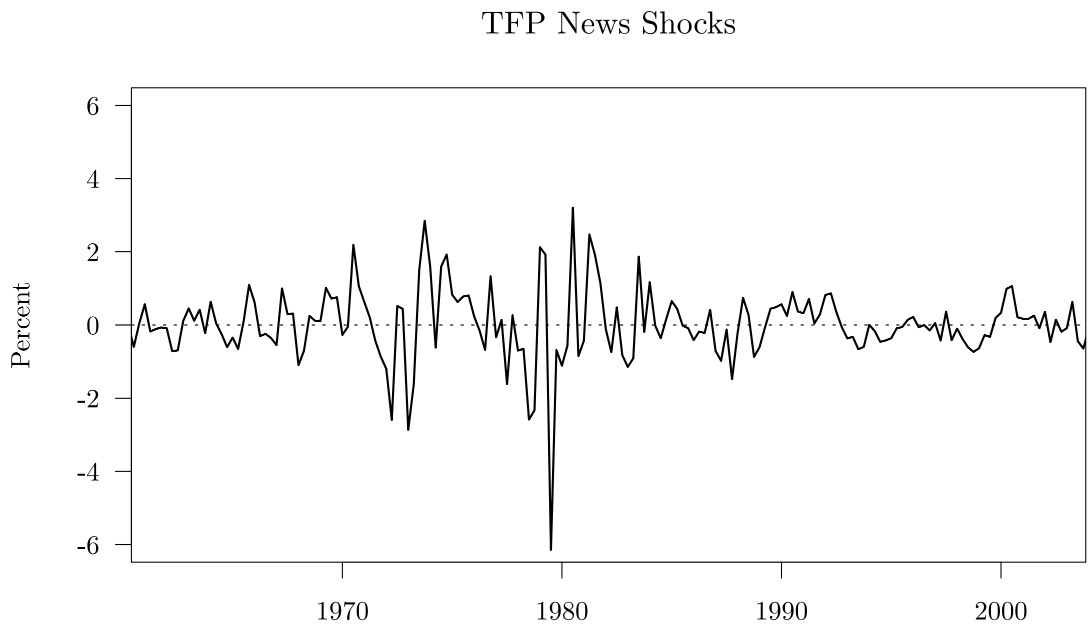

Figure 4.2: Partially identified time series of news shocks to TFP in the United states at a quarterly frequency, 1959-2005.

The recent literature on news shocks begins with Cochrane [1994] arguing that a bundle of standard contemporaneous shocks to money and TFP are unable to account for a majority of economic fluctuations, and that econometric models incorporating shocks to anticipated levels of fundamentals may serve as a plausible candidate for increasing statistical explanatory power. Beaudry and Portier [2006] create an identification scheme for news shocks to TFP for a VAR with U.S. data, and show that news shocks may account for approximately 50% of the fluctuations in consumption, investment, and hours worked at business cycle frequencies. Jaimovich and Rebelo [2009] create a real business cycle model where households observe news about TFP and investment technology, which can replicate recession dynamics resulting from lackluster news, and not negative TFP shocks. Barsky

to implement.

and Sims [2011] further construct a novel identification scheme for news shocks, and similarly show that news shocks are accountable for around 50% of fluctuations in key economic aggregates.

Section 2 of the paper describes the model in detail. Sections 3 and 4 are the calibration and solution technique. Section 5 is a discussion of the results, and section 6 concludes.

4.2 Model

There is a small open economy populated by a representative household with recursive preferences of Epstein and Zin [1989] and Weil [1990], with flow utility of Greenwood et al. [1988] (GHH),

$$\mathbf{V}_t = \left(\left(C_t - \Gamma_{t-1}\frac{1}{\nu}h_t^{\nu}\right)^{1-1/\psi} + \beta\mathbf{R}_t^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \tag{4.1}$$

$$\mathbf{R}_t^{1-\gamma} = \mathbb{E}_t\Big[\mathbf{V}_{t+1}^{1-\gamma}\Big], \tag{4.2}$$

where $\mathbf{R}_t$ is the continuation value. Time is discrete, indexed by the non-negative integers. The household procures utility with consumption C_t of a homogenous good, and forfeits utility by working hours h_t. The parameter $\nu > 0$ determines the elasticity of the household's labor supply to wages. Three parameters govern the household's attitudes toward time and uncertainty: ψ is the inter temporal elasticity of substitution, γ is risk aversion, and $\beta \in (0, 1)$ is a one period time discount rate. [3]

As in Aguiar and Gopinath [2007], the embedded GHH flow utility function is

[3]The notation $\mathbf{V}_t$ represents the household's value function at time t, which is clarified further in the appendix.

augmented with a stochastic process for cumulative growth,

$$\Gamma_t = e^{g_t}\Gamma_{t-1} \tag{4.3}$$

with growth shocks following an AR(1) process,

$$g_t = (1-\rho_g)\mu_g + \rho_g g_{t-1} + \eta^g \varepsilon_t^g, \qquad \varepsilon_t^g \sim \mathcal{N}(0,1). \tag{4.4}$$

I induce stationarity in the model by dividing quantities in time t by the lagged value of cumulative growth, Γ_{t-1}. For example, the stationary transformation of the utility function is,

$$V_t = \frac{\mathbf{V}_t}{\Gamma_{t-1}} \tag{4.5}$$

$$\mathcal{R}_t = \frac{\mathbf{R}_t}{\Gamma_{t-1}}, \tag{4.6}$$

which yields,

$$V_t = \left(\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)}\mathcal{R}_t^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \tag{4.7}$$

$$\mathcal{R}_t^{1-\gamma} = \mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right]. \tag{4.8}$$

I write stationary variables lowercase, with the exeption of the two utility variables in this example. Hereafter, all equations I introduce are stationary.

The household's objective is to maximize lifetime expected utility subject to a flow budget constraint,

$$c_t + i_t + b_t - \frac{e^{g_t}b_{t+1}}{1+e^{r_t}} + \frac{\varphi_{ba}}{2}\left(e^{g_t}b_{t+1} - b\right)^2 \leq w_t h_t + q_t k_t + d_t. \tag{4.9}$$

In every period t, it may choose consumption c_t, hours worked h_t, investment i_t, and holdings of international risk-free bonds b_{t+1}. Deviations of bond holdings from steady state b incur quadratic adjustment costs, scaled by the parameter φ_{ba}.

The household participates in a competitive labor market, and earns a wage w_t per each unit of labor it sells to the representative firm. It also participates in a competitive capital market, and earns a rental rate q_t per each unit of capital lent. Any profits earned by the firm are transferred back to the household as a dividend d_t in the same period.

The investment good and consumption good are equivalent, and may be costlessly and instantaneously transformed into one another. Capital is durable and may be stored for future use, though it depreciates at a constant rate of δ per period and is subject to quadratic adjustment costs,

$$e^{g_t} k_{t+1} = (1-\delta)k_t + i_t - \frac{\varphi_{ka}}{2}\left(\frac{e^{g_t} k_{t+1}}{k_t} - e^{\mu_g}\right)^2 k_t, \tag{4.10}$$

scaled by the parameter φ_{ka}.

Financial markets are incomplete– the household may borrow or lend with a single risk free bond at the exogenous interest rate $1+e^{r_t}$, which I assume is driven by stochastic processes,

$$\epsilon_t^r = \rho_r \varepsilon_{t-1}^r + e^{\sigma_t^r} \varepsilon_t^r, \qquad \varepsilon_t^r \sim \mathcal{N}(0,1) \tag{4.11}$$

$$\sigma_t^r = (1-\rho_{\sigma^r})\mu^{\sigma^r} + \rho_{\sigma^r}\sigma_{t-1}^r + \eta^{\sigma^r}\varepsilon_t^{\sigma^r}, \qquad \varepsilon_t^{\sigma^r} \sim \mathcal{N}(0,1) \tag{4.12}$$

$$e^{r_t} = e^{\bar{r}} + \epsilon_t^r. \tag{4.13}$$

The first order necessary conditions of consumption and labor choices in the household's utility maximization problem give rise to two equations,

$$\lambda_t = V_t^{1/\psi}\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{-1/\psi} \tag{4.14}$$

$$\lambda_t w_t = V_t^{1/\psi} h_t^{\nu-1}\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{-1/\psi} \tag{4.15}$$

where λ_t is a Lagrange multiplier for the household's flow budget constraint.

In period t, the household values payoffs in period $t+1$ with a stochastic discount factor,

$$\Lambda_{t,t+1} = \beta e^{-g_t/\psi} \frac{\lambda_{t+1}}{\lambda_t} \left(\frac{\mathcal{R}_t}{V_{t+1}} \right)^{\gamma - 1/\psi}. \tag{4.16}$$

First order conditions for bond holdings and investment give rise to two Euler equations,

$$1 = \mathbb{E}_t \left[\Lambda_{t,t+1} \left(\frac{1}{\varphi_b(b_{t+1} - b) + 1/1+e^{r_t}} \right) \right] \tag{4.17}$$

$$1 = \mathbb{E}_t \left[\Lambda_{t,t+1} \left(\varphi_k \frac{y_{t+1}}{k_{t+1}} + 1 - \delta + \frac{\varphi_{ka}}{2} \left(\left(\frac{e^{g_t} k_{t+2}}{k_{t+1}} \right)^2 - e^{2\mu_g} \right) \right) - \varphi_k \left(\frac{e^{g_t} k_{t+1}}{k_t} - e^{\mu_g} \right) \right] \tag{4.18}$$

The private sector is populated by a profit maximizing representative firm, which purchases labor and rents capital from the household in order to produce the consumption good. Output of the representative firm follows a Cobb-Douglas type production function,

$$y_t = e^{z_t} k_t^{\varphi_k} (e^{g_t} h_t)^{\varphi_h}, \tag{4.19}$$

where I restrict $\varphi_k + \varphi_h = 1$ to ensure constant returns to scale. The level of total factor productivity (TFP) is given by a stationary stochastic process e^{z_t}, where,

$$z_t = \rho_z z_{t-1} + e^{\sigma_t^z} \varepsilon_t^z, \qquad \varepsilon_t^z \sim \mathcal{N}(0,1). \tag{4.20}$$

$$\sigma_t^z = (1 - \rho_{\sigma^z}) \mu^{\sigma^z} + \rho_{\sigma^z} \sigma_{t-1}^z + \eta^{\sigma^z} \varepsilon_t^{\sigma^z}, \qquad \varepsilon_t^{\sigma^z} \sim \mathcal{N}(0,1) \tag{4.21}$$

Both markets for labor and capital are competitive, hence the first order necessary conditions of profit maximization also price these factors, respectively,

$$w_t = \varphi_h \frac{y_t}{h_t} \tag{4.22}$$

$$q_t = \varphi_k \frac{y_t}{k_t}. \tag{4.23}$$

I assume two transversality conditions for capital and bonds, which taken together with the Euler equations, are sufficient conditions for the existence of a maximum to the household's problem:

$$\lim_{t\to\infty} \beta^t \lambda_{t+1} b_{t+1} = 0 \tag{4.24}$$

$$\lim_{t\to\infty} \beta^t \lambda_{t+1} k_{t+1} = 0 \tag{4.25}$$

4.2.1 Two Types of News Shocks

In period t, the household observes a news shock $[\text{news}]_t^{t+\tau}$ that effects the level of an exogenous stochastic process in period $t+\tau$,

$$\sigma_t^n = (1-\rho_{\sigma^n})\mu^{\sigma^n} + \rho_{\sigma^n}\sigma_{t-1}^n + \eta^{\sigma^n}\varepsilon_t^{\sigma^n}, \qquad \varepsilon_t^{\sigma^n} \sim \mathcal{N}(0,1) \tag{4.26}$$

$$[\text{news}]_t^{t+\tau} = e^{\sigma_t^n}\varepsilon_t^n \qquad \varepsilon_t^n \sim \mathcal{N}(0,1) \tag{4.27}$$

where the volatility of news shocks follows an AR(1) process. The first moment of the news shock process follows Jaimovich and Rebelo [2009], equation (4.26) is novel component. I write the vector containing all observed news shocks at time t as $news_t$, and more explicitly as equation (C.3).

News about the Interest Rate

In the first model, I add news shocks to the stochastic process for interest rates at time horizons 4 quarters and 8 quarters prior to realization in the level,

$$\epsilon_t^r = \rho_r \varepsilon_{t-1}^r + e^{\sigma_t^r}\varepsilon_t^r + [\text{news}]_t^{t+4} + [\text{news}]_t^{t+8}, \qquad \varepsilon_t^r \sim \mathcal{N}(0,1). \tag{4.28}$$

Thus in every period t, the household observes two news shocks: a distant shock two years away, and a revision shock to a news shock which was initially observed one year before. Both shocks occurring in period t share a common distribution, so a volatility shock to news jointly increases the scale of both the distant and revision shocks.

News about TFP

In the second model, I add news shocks at the same horizons to the process for TFP,

$$z_t = \rho_z z_{t-1} + e^{\sigma_t^z} \varepsilon_t^z + [\text{news}]_t^{t+4} + [\text{news}]_t^{t+8}, \qquad \varepsilon_t^z \sim \mathcal{N}(0,1). \tag{4.29}$$

4.2.2 Steady State

I label steady state quantities by omitting the time subscript. To compute the steady state, first, I solve for capital as a function of labor by combining the household first order conditions for labor and consumption with the firm's first order condition for labor. Then, I substitute out labor in the Euler equation for capital, which yields the steady state level of capital in terms of given parameters,

$$k = \left(\left(\Lambda^{-1} - 1 - \delta \right) \left(\varphi_h e^{\mu_g \varphi_h} \left(\varphi_k e^{\mu_g \varphi_h} \right)^{\frac{\varphi_h}{\nu - \varphi_h}} \right)^{-1} \right)^{\frac{1}{\varphi_k - 1 + \frac{\varphi_k \varphi_h}{\nu - \varphi_h}}}. \tag{4.30}$$

The steady state level of capital is then used to solve for labor with the previously established relationship,

$$h = \left(\varphi_h e^{\mu_g \varphi_h} k^{\varphi_k} \right)^{\frac{1}{\nu - \varphi_h}}. \tag{4.31}$$

Having the steady state levels of capital and labor in hand, the solutions for the remaining variables are straightforward:

$$y = k^{\varphi_k}\big(e^{\mu_g}h\big)^{\varphi_h} \tag{4.32}$$

$$\bar{r} = \ln\big(\Lambda^{-1} - 1\big) \tag{4.33}$$

$$i = \big(e^{\mu_g} - 1 + \delta\big)k \tag{4.34}$$

$$c = y - i - b + \frac{b}{1 + e^{\bar{r}}} \tag{4.35}$$

$$V = \frac{c - \frac{1}{\nu}h^{\nu}}{\big(1 - \beta^{\mu_g(1-1/\psi)}\big)^{\frac{1}{1-1/\psi}}} \tag{4.36}$$

$$\mathcal{R} = V \tag{4.37}$$

$$\lambda = V^{1/\psi}\big(c - \frac{1}{\nu}h^{\nu}\big)^{1-1/\psi} \tag{4.38}$$

4.3 Calibration

Table 4.1 presents the calibration of the model economy, which closely resembles Argentina in Fernández-Villaverde et al. [2011]. Parameters are scaled such that one period in the model corresponds to one quarter in calendar time. I set the labor elasticity constant to $\nu = 1000$, so the household's labor response to an interest rate shock is small and negative. The sign of $\gamma - 1/\psi$ determines the household's attitude toward the resolution of uncertainty over expected utility outcomes. In the case $\gamma - 1/\psi > 0$, the household prefers an early resolution to uncertainty. If $\gamma = 1/\psi$, equations (4.1) and (4.2) are equivalent to time separable GHH preferences. And lastly, if $\gamma < 1/\psi$, the household prefers the resolution of uncertainty later, rather than sooner. Because $\gamma - 1/\psi = 6 > 0$, the household is averse to increased uncertainty in expected future utility outcomes.

Description	Parameter	Value
Discount Rate	β	0.980
Depreciation	δ	0.014
Risk Aversion	γ	8
Labor Elasticity Constant	ν	1000
Intertemporal Elasticity of Substitution	ψ	0.500
Steady State Bond Holdings	b	6
Average Real Interest Rate	$\bar{r}$	0.019
Capital Share	φ_k	0.320
Labor Share	φ_h	0.680
Bond Adjustment Cost Scale	φ_{ba}	1.75e-3
Capital Adjustment Cost Scale	φ_{ka}	6
TFP Persistence	ρ_z	0.950
TFP Volatility	μ^{σ^z}	ln(0.0075)
TFP Volatility Persistence	ρ_{σ^z}	0.950
TFP Volatility of Volatility	η^{σ^z}	0.300
Long Run Growth Rate	μ_g	ln(1.0048)
Growth Persistence	ρ_g	0.200
Growth Volatility	η^g	0.015
Interest Rate Persistence	ρ_r	0.950
Interest Rate Volatility of Volatility	η^{σ^r}	0.150
Interest Rate Volatility	μ_{σ^r}	-6
Interest Rate Volatility Persistence	ρ_{σ^r}	0.950
News Volatility	μ^{σ^n}	
News about r_t		ln(0.0075)
News about z_t		-6
News Volatility Persistence	ρ_{σ^n}	0.950
News Volatility of Volatility	η^{σ^n}	
News about r_t		0.100
News about z_t		0.150

Table 4.1: Parameter values for the model calibration. All parameters are scaled to a quarterly frequency and are values for Argentina used in Fernández-Villaverde et al. [2011].

4.4 Solution

Models of dynamic economies with stochastic volatility necessitate a solution accurate to the third order or higher. First order approximations are certainty equivalent, hence the resulting policy functions are invariant to disturbances in second moments of the shocks driving the model. Second order solutions capture only interaction terms between shocks. Therefore, I use a perturbation technique to estimate policy functions accurate to the third order.

Impulse response functions generated by higher order approximations require additional care beyond what is typically desired for first order solutions. Simply iterating the policy functions is liable to generate explosive paths for the economy. This happens because, given the recursive structure of the model, it is possible for simulated paths to depend upon orders of state variables higher than three. And these higher order dependencies bias the model, though pruning algorithms are capable of removing such nuisance terms and restoring policy stability. I use the pruning algorithm of Andreasen et al. [2013].

The logic for utilizing the deterministic steady state to initiate analyses of impulse response functions holds for first order solutions, owing to certainty equivalence. If optimal solutions are equivalent in both the stochastic and deterministic cases, simply initiate the stochastic economy from the deterministic equilibrium. Fernández-Villaverde et al. [2011] argue the loss of certainty equivalence that comes with higher order approximations is likely to drive simulated paths of an economy away from steady state in the long run, and thus propose an alternative technique addressing this fact. I use their algorithm to generate impulse responses, which goes as follows: the first step is to simulate the economy for 2096 periods beginning from the steady state. I discard the first 2000 observations as burn-in, and then compute the ergodic mean for the economy with the remaining 96 observations. Next, beginning at the ergodic mean, for each shock I separately initiate a one standard deviation impulse and iterate the economy forward 36 quarters.

Figure 4.3 displays plots of the bond policy functions approximated to three distinct orders of accuracy, simulated for 100 periods beginning at the ergodic mean. First, observe the ergodic mean of debt is indeed lower than the deterministic steady state. The second and third order accurate policies move in lockstep, however, notice that by the end of the 100 periods, the first order accurate policy deviates to a level more than 50% above the other two policies.

4.5 Discussion of Results

Figures C.1 and C.2 display impulse response functions for the model with news uncertainty shocks about the interest rate. Rows are indexed by endogenous variables, and columns by shocks. The vertical axes are scaled in percentage deviations from the ergodic mean, and horizontal axes measure time in quarters. For each displayed endogenous variable, a one standard deviation shock is realized in the first period. Recall a news shocks is first observed 8 quarters prior to its realization in the level, and then revision shocks subsequently occur 4 and 8 periods later. The column for news shocks has horizontal dotted lines at 4 and 8 quarters to highlight these points in time. The initial impulse to the news shock may be interpreted as the household learning about an uncertain anticipated change to the level occurring 8 periods ahead. Revision shocks after 4 and 8 periods are zero, thus what the household initially anticipates comes to pass. Like GHH preferences, the labor decision arising from GHH-EZ pereferences is immunized from wealth

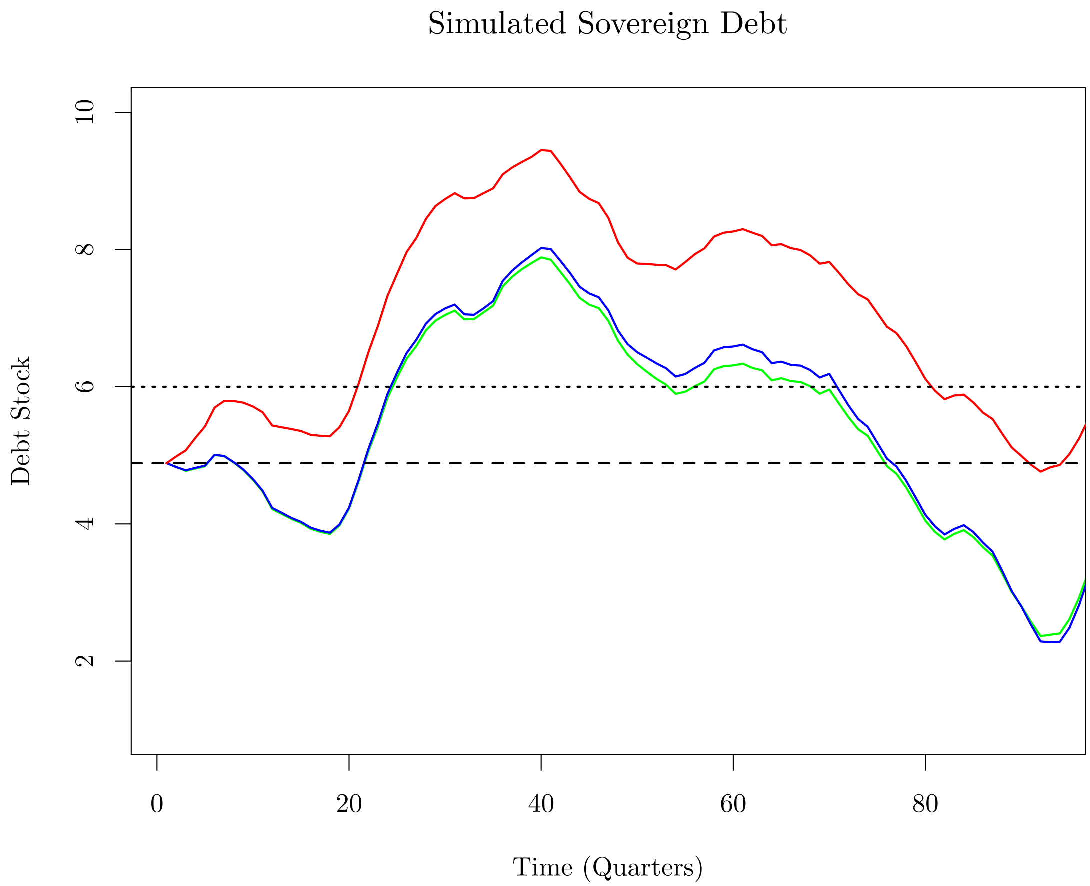

Figure 4.3: Policy functions for bond holdings over a 100 period simulation. Line colors correspond to the order of the perturbation method used- red is accurate to the first order, green is second, and blue is third. The finely dotted horizontal line at 6 is the steady state level of bond, and the line with thicker dashes below is the ergodic mean.

effects driven by changes in the interest rates. In both models with news shocks to interest rates and TFP, the labor decisions are non-factors in explaining the model dynamics.

For a shock to news uncertainty, output falls by a mere 2 basis points. Consumption falls by nearly 20 basis points, which is offset by declines in both investment and bond holdings. The response patterns endogenous variables exhibit to a news uncertainty shock are generally mirrored by shocks to interest rate volatility, although the magnitudes of declines are smaller for consumption and bond holdings. Both news shocks and contemporaneous level shocks additively feed the process for interest rates, and it stands to reason that shocks to these processes' second moments generate similar dynamics. So, what then drives the difference in the magnitudes? Figure 4.4 displays how sovereign bond responds to news uncertainty shocks about the interest rate at different time horizons $\tau \in \{1, 4, 8, 12\}$,

$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + e^{\sigma_t^r} \varepsilon_t^r + [\text{news}]_t^{t+\tau}. \tag{4.39}$$

As the plot shows, the magnitude of response to a news uncertainty shock is increasing in the time horizon of news.

Adding more intermediate revision shocks also increases the magnitude of the economy's response, because this increases the volatility of the underlying stochastic process. For instance, suppose news shocks arrive eight periods in the future, news uncertainty is highly persistent, and there are seven intermediate revision shocks to news in period before the incorporation of news in the level. A large positive shock to news uncertainty forebodes a high degree of variation in the subsequent seven revisions.

Figures C.3 and C.4 display impulse response functions for the model with news uncertainty shocks to TFP. Output rises by not even a basis point. Consumption falls 6 basis points and bond holdings fall 15 basis points, and the savings are used to fund increased investment. Despite the small magnitudes, the response of news uncertainty relative to a TFP volatility shock is threefold for consumption and bond holdings.

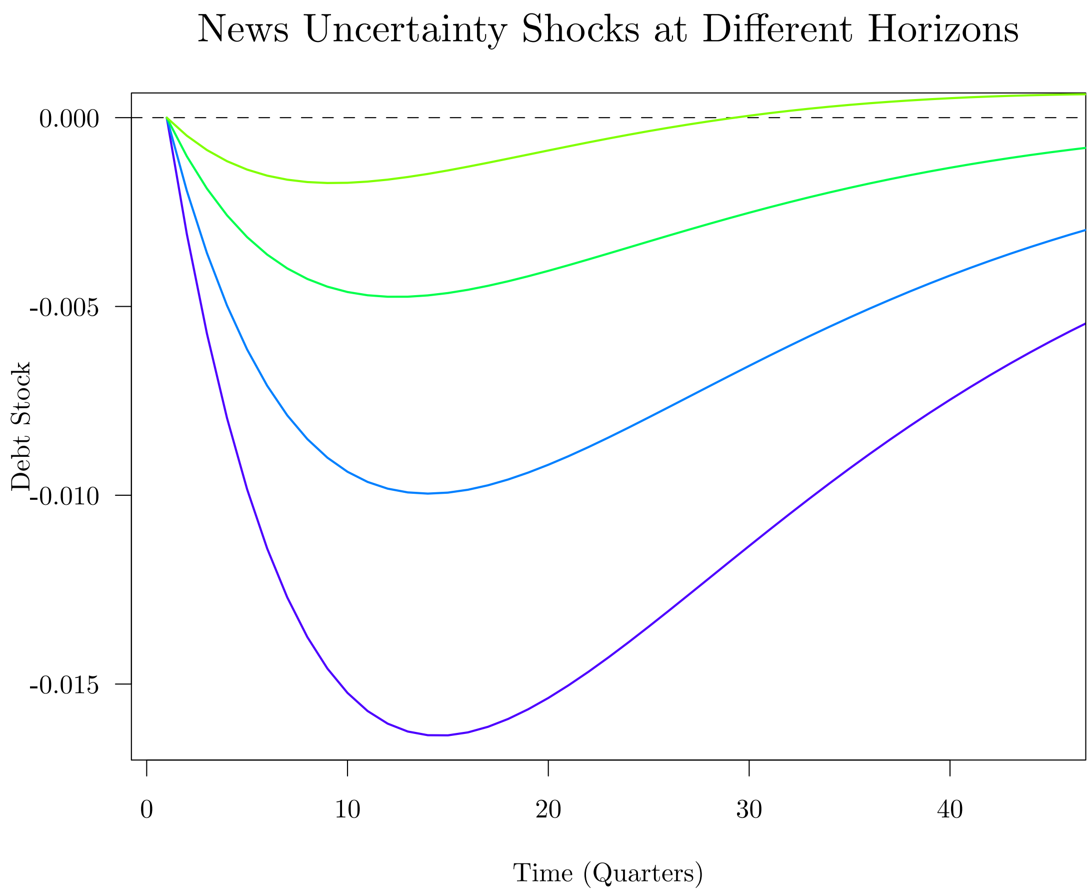

Figure 4.4: Impulse response functions for sovereign bond holdings responding to a one standard error news uncertainty shock when the news is 1, 4, 8, and 12 periods in the future. The dark indigo line with the largest response corresponds to news 12 periods ahead, the blue line to news 8 periods ahead, and so on.

4.6 Conclusion

News uncertainty shocks are important. In a small open economy, shocks to news uncertainty about the intererst rate generate large responses in bond holdings, investment, and consumption. News uncertainty about TFP plays a lesser role, generating meager responses in bond holdings and consumption. One glaring omission from this paper is the application of news uncertainty shocks to fiscal policy, which might be an interesting avenue for future research. Further, are news uncertainty shocks positively correlated with contemporaneous volatility shocks? If yes, to what degree are responses amplified? And how do news uncertainty shocks effect asset pricing in a closed economy?

CHAPTER 5

CONCLUSION

Chapter 2 shows the quality of investors' information sets shapes their beliefs about risk, and, in equilibrium, the price of that risk. A popular and well-established literature argues a small, mean-reverting stochastic process that drives average consumption and average dividend growth motivates the large historical equity premium in the United States. That argument rests on investors being able to perfectly observe the state of the growth rate at every point in time. In reality, investors cannot, and I show that when their information sets are realistically challenged with noise, then the compensation for long-run risks is reduced to an extent that it cannot explain the equity risk premium. I also present a novel mechanism in the consumption-based asset pricing literature that can explain features of the time-varying dispersion in consumption growth forecasts. Chapter 3 illustrates a related situation where a representative investor faces long-run risk and rare disasters in their consumption endowment, but does not know the actual probability distribution of these shocks. The agent estimates the worst-case distribution for these risks subject to a constraint that their pessimistic beliefs must be difficult to reject given the data. I find that the agent would give up a much larger share of their lifetime consumption to remove ambiguity around rare disasters relative to long-run risks. Chapter 4 presents a novel theoretical source of risk in a dynamic stochastic general equilibrium model: news uncertainty shocks. This is time-variation in the volatility of news shocks. When a small open economy is subjected to a positive news uncertainty shock in their borrowing costs, that is, their future interest rate becomes less certain, they will making meaningful adjustments in aggregate consumption and deleverage.

To conclude, asset pricing models offer important insights to finance professional and everyday investors. Yet if they are founded upon the assumption that agents have perfect information sets and form rational expectations consistent with nature, then asset pricing models risk generating results that match the data yet fail to match reality. Economists are therefore tasked with creating realistic belief structures for economic agents that can match both the data and our fundamental understanding of how human beings metabolise risk and return. This is challenging work, and will have to continue to push methodological limits to create suitable models for explaining the fundamental behavior of asset prices.

APPENDIX A

APPENDIX OF CHAPTER 1

A.1 Law of Motion for Higher Order Beliefs

Investor i encounters the state-space system:

$$X_t = \mathbf{A}(s^t)X_{t-1} + \mathbf{B}(s^t)\begin{bmatrix}\varepsilon_t \\ \epsilon_{i,t}\end{bmatrix} \tag{A.1}$$

$$Y_{i,t} = \boldsymbol{\mu}_Y(s^t) + \mathbf{C}_1(s^t)X_t + \mathbf{C}_2(s^t)X_{t-1} + \mathbf{D}(s_t)\begin{bmatrix}\varepsilon_t \\ \epsilon_{i,t}\end{bmatrix} \tag{A.2}$$

Define the conditional expectation of the state $X_{i,t|t} \equiv \mathbb{E}_{i,t}\big[X_t\big]$. Applying the Kalman filter of Nimark [2015] for a state space system with normal shocks and a lagged state vector in the measurement equation, the conditional expectation may be expressed:

$$X_{i,t|t} = \mathbf{A}(s^t)X_{i,t-1|t-1} + \mathbf{K}(s^t)\big(Y_{i,t} - \boldsymbol{\mu}_Y(s^t) - \big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)X_{i,t-1|t-1}\big) \tag{A.3}$$

Define the matrices $\mathbf{D}_\varepsilon(s^t)$ and $\mathbf{D}_\epsilon$:

$$\mathbf{D}_\varepsilon(s^t) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_c\sigma(s_{t+1}) & 0 & 0 & 0 \\ 0 & \rho_{cd}\varphi_c\sigma(s_{t+1}) & \varphi_d\sigma(s_{t+1}) & 0 & 0 \\ 0 & 0 & 0 & \varphi_{rm} & 0 \\ 0 & 0 & 0 & 0 & \varphi_{rf} \end{bmatrix} \tag{A.4}$$

$$\mathbf{D}_\epsilon = \begin{bmatrix} \varphi_s \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.5}$$

Let the subscript 1 denote the posterior, and 0 the prior. Then posterior covariance matrix may be computed:

$$\begin{aligned}\mathbf{P}_1(s^t) = \mathbf{P}_0(s^t) - \\ \mathbf{K}(s^t)\Big(\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\mathbf{P}_1(s^{t-1})\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)' + \\ \big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s^t)\big)\big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s^t)\big)' + \mathbf{D}_\epsilon\mathbf{D}_\epsilon'\Big)\mathbf{K}(s^t)'\end{aligned} \tag{A.6}$$

And the prior covariance matrix:

$$\mathbf{P}_0(s^t) = \mathbf{A}(s^t)\mathbf{P}_1(s^{t-1})\mathbf{A}(s^t)' + \mathbf{B}(s^t)\mathbf{B}(s^t)' \tag{A.7}$$

Lastly, the Kalman gain:

$$\begin{aligned}\mathbf{K}(s^t) = \\ \Big(\mathbf{A}(s^t)\mathbf{P}_1(s^{t-1})\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)' + \mathbf{B}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s^t)\big)'\Big) \\ \Big(\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\mathbf{P}_1(s^{t-1})\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)' + \\ \big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s^t)\big)\big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s^t)\big)' + \mathbf{D}_\epsilon\mathbf{D}_\epsilon'\Big)^{-1}\end{aligned} \tag{A.8}$$

Integrating the posterior estimate of the state over the cross section of agents:

$$X_{t|t} \equiv \int_0^1 X_{i,t|t} di \tag{A.9}$$

Which yields the infinite dimensional state transition equation for higher order beliefs:

$$\begin{aligned}X_{t|t} = \big(\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big)X_{t-1|t-1} + \\ \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)X_{t-1} + \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s_t)\big)\varepsilon_t\end{aligned} \tag{A.10}$$

The state transition equation can be written as:

$$\begin{bmatrix} x_t \\ X_{t|t} \end{bmatrix} = \mathbf{A}(s^t) \begin{bmatrix} x_{t-1} \\ X_{t-1|t-1} \end{bmatrix} + \mathbf{B}(s^t)\varepsilon_t \tag{A.11}$$

To make the problem numerically tractable, I approximate the infinite dimensional law of motion by choosing the first $\bar{k}$ orders of expectations. I construct the matrices $\mathbf{A}(s^t)$ and $\mathbf{B}(s^t)$ as follows:

$$\mathbf{A}(s^t) = \begin{bmatrix} \underset{1\times 1}{\rho_x} & \underset{1\times \bar{k}}{\mathbf{0}} \\ \underset{\bar{k}\times 1}{\mathbf{0}} & \underset{\bar{k}\times\bar{k}}{\mathbf{0}} \end{bmatrix} + \begin{bmatrix} \underset{1\times 1}{0} & \underset{1\times \bar{k}}{\mathbf{0}} \\ \underset{\bar{k}\times 1}{\mathbf{0}} & \underset{\bar{k}\times\bar{k}}{\big[\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big]_-} \end{bmatrix} + \begin{bmatrix} \underset{1\times\bar{k}+1}{\mathbf{0}} \\ \underset{\bar{k}\times\bar{k}+1}{\big[\mathbf{K}(s^t)(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t))\big]_-} \end{bmatrix} \tag{A.12}$$

$$\mathbf{B}(s^t) = \begin{bmatrix} \underset{1\times 1}{\varphi_x} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{n_\varepsilon\times 1}{\mathbf{0}} & \underset{n_\varepsilon\times n_\varepsilon}{\mathbf{0}} \end{bmatrix} + \begin{bmatrix} \underset{1\times n_\varepsilon+1}{\mathbf{0}} \\ \underset{1\times n_\varepsilon+1}{\big[\mathbf{K}(s^t)(\mathbf{C}_1(s^t)\mathbf{B}(s^t) + \mathbf{D}_\varepsilon(s_t))\big]_-} \end{bmatrix} \tag{A.13}$$

where a minus subscript indicates orders of expectations greater than $\bar{k}$ have been dropped.

A.2 Deriving the Stochastic Discount Factor

Investor i solves the following utility maximization problem:

$$V_{i,t} = \max_{\{C_{i,t}\}} \left((1-\beta) C_{i,t}^{1-1/\psi} + \beta \mathbb{E}_{i,t}\big[V_{i,t+1}^{1-\gamma}\big]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \tag{A.14}$$

subject to,

$$W_{i,t+1} = R_{c,t+1}\big(W_{i,t} - C_{i,t}\big) \tag{A.15}$$

I conjecture $V_{i,t} = \varpi_{i,t} W_{i,t}$ and re-express the value function accordingly:

$$\varpi_{i,t} W_{i,t} = \max_{\{C_{i,t}\}} \left((1-\beta) C_{i,t}^{1-1/\psi} + \beta \big(W_{i,t} - C_{i,t}\big)^{1-1/\psi} \mathbb{E}_{i,t}\Big[\big(\varpi_{i,t+1} R_{c,t+1}\big)^{1-\gamma}\Big]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \tag{A.16}$$

Re-arranging:

$$\varpi_{i,t}^{1-1/\psi} = (1-\beta)\big(\upsilon_{i,t}\big)^{1-1/\psi} + \beta\big(1-\upsilon_{i,t}\big)^{1-1/\psi} \mathbb{E}_{i,t}\Big[\big(\varpi_{i,t+1} R_{c,t+1}\big)^{1-\gamma}\Big]^{\frac{1-1/\psi}{1-\gamma}} \tag{A.17}$$

where $\upsilon_{i,t} = {}^{C_{i,t}}/_{W_{i,t}}$. Taking the first order condition and solving for the conditional expectation term:

$$\mathbb{E}_{i,t}\Big[\big(\varpi_{i,t+1} R_{c,t+1}\big)^{1-\gamma}\Big]^{\frac{1-1/\psi}{1-\gamma}} = \frac{1-\beta}{\beta}\left(\frac{1-\upsilon_{i,t}}{\upsilon_{i,t}}\right)^{1/\psi} \tag{A.18}$$

Substituing the first order condition (A.18) back into the objective function leads to:

$$\varpi_{i,t} = \big(1-\beta\big)^{\frac{1}{1-1/\psi}} \big(\upsilon_{i,t}\big)^{\frac{-1/\psi}{1-1/\psi}} \tag{A.19}$$

Then I re-express the term inside the conditional expectation in terms of the consumption-wealth ratio and the return on the consumption claim:

$$\varpi_{i,t+1} R_{c,t+1} = \big(1-\beta\big)^{\frac{1}{1-1/\psi}} \big(\upsilon_{i,t+1}\big)^{\frac{-1/\psi}{1-1/\psi}} R_{c,t+1} \tag{A.20}$$

I substitute the budget constraint in period t into the consumption-wealth ratio in period $t+1$:

$$\upsilon_{i,t+1} = \frac{\triangle C_{i,t+1}}{R_{c,t+1}} \frac{\upsilon_{i,t}}{1-\upsilon_{i,t}} \tag{A.21}$$

where $\triangle C_{i,t+1} = {}^{C_{i,t+1}}/_{C_{i,t}}$. Putting this back into the term inside the conditional expectation:

$$\varpi_{i,t+1} R_{c,t+1} = \big(1-\beta\big)^{\frac{1}{1-1/\psi}} \left(\frac{\triangle C_{i,t+1}}{R_{c,t+1}} \frac{\upsilon_{i,t}}{1-\upsilon_{i,t}}\right)^{\frac{-1/\psi}{1-1/\psi}} R_{c,t+1} \tag{A.22}$$

Now putting this back into the first order condition (A.18):

$$\mathbb{E}_{i,t}\left[\left(1-\beta\right)^{\vartheta}\left(\frac{\triangle C_{i,t+1}}{R_{c,t+1}}\frac{\upsilon_{i,t}}{1-\upsilon_{i,t}}\right)^{-\vartheta/\psi}R_{c,t+1}^{1-\gamma}\right]=\left(\frac{1-\beta}{\beta}\left(\frac{1-\upsilon_{i,t}}{\upsilon_{i,t}}\right)^{1/\psi}\right)^{\vartheta} \tag{A.23}$$

Simplifying provides:

$$\mathbb{E}_{i,t}\left[\beta^{\vartheta}\triangle C_{i,t+1}^{-\vartheta/\psi}R_{c,t+1}^{\vartheta-1}\right]=1 \tag{A.24}$$

Next, define the log stochastic discount factor for investor i as:

$$m_{i,t+1}\equiv\vartheta\log\beta-\vartheta/\psi\triangle c_{i,t+1}+\left(\vartheta-1\right)r_{c,t+1} \tag{A.25}$$

The stochastic discount factor for the average agent aggregates over the cross-section:

$$m_{t+1}\equiv\int m_{i,t+1}di \tag{A.26}$$

A.3 Solving for Prices

A.3.1 Approximating Returns

I approximate the return on the aggregate consumption claim $R_{c,t+1}$ with by log-linearizing the return function about a deterministic steady state $pc\equiv\log P-\log C$. The one period real return of the consumtion claim is:

$$R_{c,t+1}=\frac{P_{c,t+1}+C_t}{P_{c,t}} \tag{A.27}$$

Let $r_{c,t+1}\equiv\log R_{c,t+1}$. Taking logs of both sides:

$$r_{c,t+1}=\log\left(\frac{P_{c,t+1}+C_t}{P_{c,t}}\right) \tag{A.28}$$

After some intermediate algebraic manipulations:

$$r_{c,t+1} = \log\left(\frac{P_{t+1}}{C_t} + 1\right) - \log\left(\frac{P_t}{C_{t-1}}\right) + \log\left(\frac{C_t}{C_{t-1}}\right) \tag{A.29}$$

Denote $\triangle c_{t+1} \equiv \log C_{t+1} - \log C_t$ and $pc_t \equiv \log P_t - \log C_{t-1}$. Then exponentiating and taking logs yields:

$$r_{c,t+1} = \log\big(\exp(pc_{t+1}) + 1\big) - pc_t + \triangle c_{t+1} \tag{A.30}$$

Taking the first order Taylor series expansion of the first term about pc:

$$\log\big(\exp(pc_{t+1}) + 1\big) \approx \log\big(\exp(pc) + 1\big) + \frac{\exp(pc)}{\exp(pc) + 1}\big(pc_t - pc\big) \tag{A.31}$$

Define the linearization constants:

$$\kappa_0 \equiv \log\big(\exp(pc) + 1\big) - \frac{pc\exp(pc)}{\exp(pc) + 1} \tag{A.32}$$

$$\kappa_1 \equiv \frac{\exp(pc)}{\exp(pc) + 1} \tag{A.33}$$

Then the log return on the aggregate consumption claim is:

$$r_{c,t+1} = \kappa_0 + \kappa_1 pc_{t+1} - pc_t + \triangle c_{t+1} \tag{A.34}$$

By the same line of reasoning, the terms for log return on the aggregate dividend claim:

$$\kappa_{0,m} \equiv \log\big(\exp(pd) + 1\big) - \frac{pd\exp(pd)}{\exp(pd) + 1} \tag{A.35}$$

$$\kappa_{1,m} \equiv \frac{\exp(pd)}{\exp(pd) + 1} \tag{A.36}$$

And the log return on the aggregate dividend claim:

$$r_{m,t+1} = \kappa_0 + \kappa_1 pd_{t+1} - pd_t + \triangle d_{t+1} \tag{A.37}$$

A.3.2 Price-Consumption Ratio

I solve for the price-consumption ratio with the method of undetermined coefficients. Begin with the Euler equation pricing the aggregate consumption claim:

$$0 = \vartheta \log \beta + \vartheta \kappa_0 + \vartheta\big(1 - {}^1\!/\!_\psi\big)\mu_c + \vartheta\big(1 - {}^1\!/\!_\psi\big) \textstyle\int \mathbb{E}_{i,t}[x_t]di +$$
$$\vartheta\big(\kappa_1 \textstyle\int \mathbb{E}_{i,t}[pc_{t+1}]di - pc_t\big) + {}^1\!/\!_2 \textstyle\int \mathrm{Var}_{i,t}\big(m_{t+1} + r_{c,t+1}\big)di \quad \text{(A.38)}$$

Substitute in the affine form of the price consumption ratio:

$$pc_t = g_0^{\mathrm{pc}}(s^t) + g_x^{\mathrm{pc}}(s^t)'X_t \quad \text{(A.39)}$$

This leads to:

$$0 = \vartheta \log \beta + \vartheta \kappa_0 + \vartheta\big(1 - {}^1\!/\!_\psi\big)\mu_c + \vartheta\big(\kappa_1 \textstyle\int \mathbb{E}_{i,t}[g_0^{\mathrm{pc}}(s^{t+1})]di - g_0^{\mathrm{pc}}(s^t)\big) +$$
$$\vartheta\big(1 - {}^1\!/\!_\psi\big) \textstyle\int \mathbb{E}_{i,t}[x_t]di + \vartheta\big(\kappa_1 \textstyle\int \mathbb{E}_{i,t}[g_x^{\mathrm{pc}}(s^{t+1})'X_{t+1}]di - g_x^{\mathrm{pc}}(s^t)'X_t\big)$$
$$+ {}^1\!/\!_2 \textstyle\int \mathrm{Var}_{i,t}\big(m_{t+1} + r_{c,t+1}\big)di \quad \text{(A.40)}$$

Evaluating expectations:

$$\textstyle\int \mathbb{E}_{i,t}[g_x^{\mathrm{pc}}(s^{t+1})X_{t+1}]di = \big(\pi g_x^{\mathrm{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + \big(1 - \pi\big) g_x^{\mathrm{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\big)\mathbf{H}X_t \quad \text{(A.41)}$$

$$\textstyle\int \mathbb{E}_{i,t}[g_0^{\mathrm{pc}}(s^{t+1})]di = \pi g_0^{\mathrm{pc}}(s_1^{t+1}) + (1 - \pi) g_0^{\mathrm{pc}}(s_0^{t+1}) \quad \text{(A.42)}$$

Matching coefficients reveals:

$$g_x^{\mathrm{pc}}(s^t)' = \big(1 - {}^1\!/\!_\psi\big)\mathbf{e}_1'\mathbf{H} + \kappa_1\big(\pi g_x^{\mathrm{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + \big(1 - \pi\big) g_x^{\mathrm{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\big)\mathbf{H} \quad \text{(A.43)}$$

And the intercept term is therefore:

$$g_0^{\mathrm{pc}}(s^t) = \log \beta + \kappa_0 + \big(1 - {}^1\!/\!_\psi\big)\mu_c + \kappa_1\big(\pi g_0^{\mathrm{pc}}(s_1^{t+1}) + (1 - \pi) g_0^{\mathrm{pc}}(s_0^{t+1})\big) +$$
$${}^\pi\!/\!_{2\vartheta} \textstyle\int \mathrm{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s_1^{t+1})'Z_{t+1}\big)di + {}^{(1-\pi)}\!/\!_{2\vartheta} \textstyle\int \mathrm{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s_0^{t+1})'Z_{t+1}\big)di \quad \text{(A.44)}$$

A.3.3 Price-Dividend Ratio

The Euler equation for the market return:

$$
\begin{aligned}
0 = \vartheta \log \beta + \big(\vartheta - 1 - {}^{\vartheta}/_{\psi}\big)\mu_c + \big(\vartheta - 1\big)\kappa_0 + \kappa_{0,m} + \mu_d + \\
\big(\vartheta - 1 - {}^{\vartheta}/_{\psi} + \rho_d\big)\textstyle\int \mathbb{E}_{i,t}\big[x_t\big]di + \big(\vartheta - 1\big)\big(\kappa_1 \textstyle\int \mathbb{E}_{i,t}\big[pc_{t+1}\big]di - pc_t\big) + \\
\big(\kappa_{1,m}\textstyle\int \mathbb{E}_{i,t}\big[pd_{t+1}\big]di - pd_t\big) + {}^1/_2 \textstyle\int \mathrm{Var}_{i,t}\big(m_{t+1} + r_{m,t+1}\big)di \quad \text{(A.45)}
\end{aligned}
$$

I conjecture an affine form for the price-dividend ratio:

$$
pd_t = g_0^{\text{pd}}(s^t) + g_x^{\text{pd}}(s^t)' X_t \quad \text{(A.46)}
$$

Then after taking expectations and substituting, the following equation must hold:

$$
\begin{aligned}
g_x^{\text{pd}}(s^t)' = \big(\vartheta - 1 - {}^{\vartheta}/_{\psi} + \rho_d\big)\mathbf{e}_1'\mathbf{H} + \\
\big(\vartheta - 1\big)\Big(\kappa_1\Big(\pi g_x^{\text{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + \big(1-\pi\big)g_x^{\text{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\Big)\mathbf{H} - g_x^{\text{pc}}(s^t)'\Big) + \\
\kappa_{1,m}\Big(\pi g_x^{\text{pd}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + \big(1-\pi\big)g_x^{\text{pd}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\Big)\mathbf{H} \quad \text{(A.47)}
\end{aligned}
$$

With the constant term:

$$
\begin{aligned}
g_0^{pd}(s^t) = \vartheta \log \beta + \big(\vartheta - 1 - {}^{\vartheta}/_{\psi}\big)\mu_c + \big(\vartheta - 1\big)\kappa_0 + \kappa_{0,m} + \mu_d + \\
\kappa_{1,m}\Big(\pi g_0^{\text{pd}}(s_1^{t+1}) + (1-\pi)g_0^{\text{pd}}(s_0^{t+1})\Big) + \big(\vartheta - 1\big)\Big(\kappa_1\Big(\pi g_0^{\text{pc}}(s_1^{t+1}) + (1-\pi)g_0^{\text{pc}}(s_0^{t+1})\Big) - g_0^{\text{pc}}(s^t)\Big) + \\
\pi \textstyle\int \mathrm{Var}_{i,t}\big(\boldsymbol{\gamma}^{pd}(s_1^{t+1})'Z_{t+1}\big)di + (1-\pi)\textstyle\int \mathrm{Var}_{i,t}\big(\boldsymbol{\gamma}^{pd}(s_0^{t+1})'Z_{t+1}\big)di \quad \text{(A.48)}
\end{aligned}
$$

A.3.4 Risk-Free Rate

We can solve for the risk-free rate with:

$$
r_{f,t} = -\textstyle\int \mathbb{E}_{i,t}[m_{t+1}]di - {}^1/_2 \textstyle\int \mathrm{Var}_{i,t}\big(m_{t+1}\big)di \quad \text{(A.49)}
$$

I conjecture:

$$r_{f,t} = g_0^{\text{rf}}(s^t) + g_x^{\text{rf}}(s^t)' X_t \tag{A.50}$$

Solving for $r_{f,t}$ yields:

$$r_{f,t} = -\big(\vartheta \log \beta + (\vartheta - 1)\kappa_0 + (\vartheta - \vartheta/\psi - 1)\mu_c\big) - \vartheta\big(1 - 1/\psi - 1/\vartheta\big)\textstyle\int \mathbb{E}_{i,t}[x_t]di - \\ (\vartheta - 1)\big(\kappa_1 \textstyle\int \mathbb{E}_{i,t}[pc_{t+1}]di - pc_t\big) - 1/2\textstyle\int \text{Var}_{i,t}\big(m_{t+1}\big)di \tag{A.51}$$

The coefficient on higher order beliefs:

$$g_x^{\text{rf}}(s^t)' = -\vartheta\big(1 - 1/\psi - 1/\vartheta\big)\mathbf{e}_1'\mathbf{H} - \\ (\vartheta - 1)\Big(\kappa_1\big(\pi g_x^{\text{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + (1-\pi)g_x^{\text{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\big)\mathbf{H} - g_x^{\text{pc}}(s^t)'\Big) \tag{A.52}$$

Which leaves us with an intercept term:

$$g_0^{\text{rf}}(s^t) = -\vartheta \log \beta - (\vartheta - 1)\kappa_0 - (\vartheta - \vartheta/\psi - 1)\mu_c - \\ (\vartheta - 1)\Big(\kappa_1\big(\pi g_0^{\text{pc}}(s_1^{t+1}) + (1-\pi)g_0^{\text{pc}}(s_0^{t+1})\big) - g_0^{\text{pc}}(s^t)\Big) - \\ \pi \textstyle\int \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{\text{rf}}(s_1^{t+1})' Z_{t+1}\big)di - (1-\pi)\textstyle\int \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{\text{rf}}(s_0^{t+1})' Z_{t+1}\big)di \tag{A.53}$$

A.4 Conditional Variances

In order to solve the equations for pd_t, pc_t, and $r_{f,t}$, I need to solve for three conditional variances. This may be accomplished by recasting the state space system to include aggregate shocks in the state vector, and then utilizing the Kalman filter to compute the state conditional covariance matrix. Define the new state vector Z_{t+1} as:

$$Z_{t+1} \equiv \begin{bmatrix} X_{t+1} \\ \varepsilon_{t+1} \end{bmatrix} \tag{A.54}$$

The new state space system can be described by the state transition equation:

$$\begin{bmatrix} X_{t+1} \\ \varepsilon_{t+1} \end{bmatrix} = \begin{bmatrix} \underset{\bar{k}+1\times\bar{k}+1}{\mathbf{A}(s^{t+1})} & \underset{\bar{k}+1\times n_\varepsilon}{\mathbf{0}} \\ \underset{n_\varepsilon\times\bar{k}+1}{\mathbf{0}} & \underset{n_\varepsilon\times n_\varepsilon}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} X_t \\ \varepsilon_t \end{bmatrix} + \begin{bmatrix} \underset{\bar{k}+1\times n_\varepsilon}{\mathbf{B}(s^{t+1})} & \underset{\bar{k}+1\times 1}{\mathbf{0}} \\ \underset{n_\varepsilon\times n_\varepsilon}{\mathbf{I}} & \underset{n_\varepsilon\times 1}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1} \\ \epsilon_{i,t+1} \end{bmatrix} \tag{A.55}$$

and the measurement equation:

$$\begin{bmatrix} x_{i,t+1} \\ \triangle c_{t+1} \\ \triangle d_{t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix} = \boldsymbol{\mu}_Y(s^{t+1}) + \begin{bmatrix} \underset{1\times\bar{k}+1}{\mathbf{e}_1'} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{1\times\bar{k}+1}{\mathbf{0}} & \underset{1\times n_\varepsilon}{\varphi_c\sigma(s_{t+1})\mathbf{e}_c'} \\ \underset{1\times\bar{k}+1}{\mathbf{0}} & \underset{1\times n_\varepsilon}{\big(\varphi_d\mathbf{e}_d' + \rho_{cd}\varphi_c\mathbf{e}_c'\big)\sigma(s_{t+1})} \\ \underset{1\times\bar{k}+1}{\kappa_{1,m}g_x^{\text{pd}}(s^{t+1})'} & \underset{1\times n_\varepsilon}{\big(\varphi_d\mathbf{e}_d' + \rho_{cd}\varphi_c\mathbf{e}_c'\big)\sigma(s_{t+1}) + \varphi_{rm}\mathbf{e}_{rm}'} \\ \underset{1\times\bar{k}+1}{\mathbf{0}} & \underset{1\times n_\varepsilon}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} X_{t+1} \\ \varepsilon_{t+1} \end{bmatrix} + \begin{bmatrix} \underset{1\times\bar{k}+1}{\mathbf{0}} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{1\times\bar{k}+1}{\mathbf{e}_1'} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{1\times\bar{k}+1}{\rho_d\mathbf{e}_1'} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{1\times\bar{k}+1}{\rho_d\mathbf{e}_1' - g_x^{\text{pd}}(s^t)'} & \underset{1\times n_\varepsilon}{\mathbf{0}} \\ \underset{1\times\bar{k}+1}{g_x^{\text{rf}}(s^t)'} & \underset{1\times n_\varepsilon}{\varphi_{rf}\mathbf{e}_{rf}'} \end{bmatrix} \begin{bmatrix} X_t \\ \varepsilon_t \end{bmatrix} + \begin{bmatrix} \underset{n_y\times n_\varepsilon}{\mathbf{0}} & \underset{n_y\times 1}{\varphi_s\mathbf{e}_1'} \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1} \\ \epsilon_{i,t+1} \end{bmatrix} \tag{A.56}$$

where:

$$\boldsymbol{\mu}_Y(s^{t+1}) \equiv \begin{bmatrix} 0 \\ \mu_c \\ \mu_d \\ \kappa_{0,m} + \kappa_{1,m} g_0^{\text{pd}}(s^{t+1}) - g_0^{\text{pd}}(s^t) \\ g_0^{\text{rf}}(s^t) \end{bmatrix} \tag{A.57}$$

The conditional covariance matrix of the state in this system is:

$$\mathbf{P}_Z(s^{t+1}) \equiv \mathbb{E}_{i,t}\Big[\big(Z_{t+1} - Z_{t+1|t+1}\big)\big(Z_{t+1} - Z_{t+1|t+1}\big)'\Big] \tag{A.58}$$

The conditional variances for the price-consumption ratio, price-dividend ratio, and risk free rate are can be expressed as functions of Z_{t+1}:

$$\text{Var}_{i,t}\big(m_{t+1} + r_{c,t+1}\big) = \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s^{t+1})' Z_{t+1}\big) \tag{A.59}$$

$$\text{Var}_{i,t}\big(m_{t+1} + r_{m,t+1}\big) = \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pd}(s^{t+1})' Z_{t+1}\big) \tag{A.60}$$

$$\text{Var}_{i,t}\big(m_{t+1}\big) = \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{\text{rf}}(s^{t+1})' Z_{t+1}\big) \tag{A.61}$$

where the vector for the price-consumption ratio is defined:

$$\boldsymbol{\gamma}^{pc}(s^{t+1}) \equiv \begin{bmatrix} \vartheta \kappa_1 g_x^{\text{pc}}(s^{t+1}) \\ 0 \\ \vartheta\big(1 - {}^1\!/\!_\psi\big)\varphi_c \sigma(s_{t+1}) \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{A.62}$$

and the price-dividend ratio:

$$\boldsymbol{\gamma}^{pd}(s^{t+1}) \equiv \begin{bmatrix} (\vartheta - 1)\kappa_1 g_x^{\text{pc}}(s^{t+1}) + \kappa_{1,m} g_x^{\text{pd}}(s^{t+1}) \\ 0 \\ (\vartheta - \vartheta/\psi - 1 + \rho_{cd})\varphi_c \sigma(s_{t+1}) \\ \varphi_d \sigma(s_{t+1}) \\ \varphi_{rm} \\ 0 \end{bmatrix} \tag{A.63}$$

and lastly, the risk free rate:

$$\boldsymbol{\gamma}^{\text{rf}}(s^{t+1}) \equiv \begin{bmatrix} (\vartheta - 1)\kappa_1 g_x^{\text{pc}}(s^{t+1}) \\ 0 \\ (\vartheta - \vartheta/\psi - 1)\varphi_c \sigma(s_{t+1}) \\ 0 \\ 0 \\ \varphi_{rf} \end{bmatrix} \tag{A.64}$$

This variance operator conditions on the information set in period t, therefore the next period volatility state s_{t+1} is a random variable. Further, the conditional variance is a mixture of normal distributions with constant mixing weights π and $1-\pi$. By the properties of mixtures of normal distributions with constant mixing weights, the conditional variance is a simple weighted average:

$$\text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s^{t+1})' Z_{t+1}\big) \approx \\ \pi \text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s_1^{t+1})' Z_{t+1}\big) + (1-\pi)\text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s_0^{t+1})' Z_{t+1}\big) \tag{A.65}$$

I approximate the variance by dropping the small mean correction terms in (A.65). The impact of approximating conditional variances is negligible on the model, because risk prices for volatility are close to zero with or without the cumbersome mean corrections that necessitate solving the model on a large numerical grid over

X_t. Using the approximation from (A.65) and the state covariance matrix (A.58), the conditional covariance matrix for the price-consumption ratio is:

$$\begin{aligned}&\text{Var}_{i,t}\big(\boldsymbol{\gamma}^{pc}(s^{t+1})'Z_{t+1}\big) = \\ &\qquad \pi\boldsymbol{\gamma}^{pc}(s_1^{t+1})'\mathbf{P}_Z(s_1^{t+1})\boldsymbol{\gamma}^{pc}(s_1^{t+1}) + \big(1-\pi\big)\boldsymbol{\gamma}^{pc}(s_0^{t+1})'\mathbf{P}_Z(s_0^{t+1})\boldsymbol{\gamma}^{pc}(s_0^{t+1}) \quad \text{(A.66)}\end{aligned}$$

Similar experession hold for the conditional variances for the price-dividend ratio and risk free rate.

A.5 Risk Premia

A.5.1 Deriving Risk Prices

To compute risk prices, begin with the stochastic discount factor in innovations form:

$$m_{t+1} - \textstyle\int \mathbb{E}_{i,t}\big[m_{t+1}\big]di \tag{A.67}$$

Define the expected value of the slope and intercepts of the price function:

$$\textstyle\int \mathbb{E}_{i,t}\big[g_0^{\text{pc}}(s^{t+1})\big]di = \pi g_0^{\text{pc}}(s_1^{t+1}) + (1-\pi)g_0^{\text{pc}}(s_0^{t+1}) \tag{A.68}$$

$$\textstyle\int \mathbb{E}_{i,t}\big[g_x^{\text{pc}}(s^{t+1})'X_{t+1}\big]di = \big(\pi g_x^{\text{pc}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + (1-\pi)g_x^{\text{pc}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\big)\mathbf{H}X_t \tag{A.69}$$

Dropping constants and substituting leads to:

$$
\begin{aligned}
m_{t+1} - \textstyle\int \mathbb{E}_{i,t}\big[m_{t+1}\big]di = &\\
&\big(\vartheta - 1\big)\kappa_1\big(g_0^{\text{pc}}(s^{t+1}) - \textstyle\int \mathbb{E}_{i,t}\big[g_0^{\text{pc}}(s^{t+1})\big]di\big) + \\
&\big(-\gamma\varphi_c\sigma(s_{t+1})\mathbf{e}_c' + \big(\vartheta - 1\big)\kappa_1 g_x^{\text{pc}}(s^{t+1})'\mathbf{B}(s^{t+1})\big)\varepsilon_{t+1} + \\
\big(-\gamma\mathbf{e}_1'\big(\mathbf{I} - \mathbf{H}\big) + \big(\vartheta - 1\big)\kappa_1\big(g_x^{\text{pc}}(s^{t+1})'\mathbf{A}(s^{t+1}) &- \textstyle\int \mathbb{E}_{i,t}\big[g_x^{\text{pc}}(s^{t+1})'\mathbf{A}(s^{t+1})\mathbf{H}\big]di\big)\big)X_t
\end{aligned}
\tag{A.70}
$$

This results in risk-prices for:

$$
\lambda_\varepsilon(s^{t+1})' \equiv \big(-\gamma\varphi_c\sigma(s_{t+1})\mathbf{e}_c' + \big(\vartheta - 1\big)\kappa_1 g_x^{\text{pc}}(s^{t+1})'\mathbf{B}(s^{t+1})\big)\Sigma(s_{t+1})^{-1} \tag{A.71}
$$

$$
\lambda_\sigma(s^{t+1}) \equiv \big(\vartheta - 1\big)\kappa_1\big(g_0^{\text{pc}}(s^{t+1}) - \textstyle\int \mathbb{E}_{i,t}\big[g_0^{\text{pc}}(s^{t+1})\big]di\big) \tag{A.72}
$$

$$
\begin{aligned}
\lambda_X(s^{t+1})' \equiv -\gamma\mathbf{e}_1'\big(\mathbf{I} - \mathbf{H}\big) + \big(\vartheta - 1\big)\kappa_1\big(g_x^{\text{pc}}(s^{t+1})'\mathbf{A}(s^{t+1}) - &\\
\textstyle\int \mathbb{E}_{i,t}\big[g_x^{\text{pc}}(s^{t+1})'\mathbf{A}(s^{t+1})\mathbf{H}\big]di\big)&
\end{aligned}
\tag{A.73}
$$

where I define the matrix $\Sigma(s_{t+1})$ such that:

$$
\Sigma(s_{t+1}) \equiv \begin{bmatrix}
\varphi_x\sigma(s_{t+1}) & 0 & 0 & 0 & 0 \\
0 & \varphi_c\sigma(s_{t+1}) & 0 & 0 & 0 \\
0 & 0 & \varphi_d\sigma(s_{t+1}) & 0 & 0 \\
0 & 0 & 0 & \varphi_{rm} & 0 \\
0 & 0 & 0 & 0 & \varphi_{rf}
\end{bmatrix} \tag{A.74}
$$

The result is innovations to the stochastic discount factor may be written:

$$
m_{t+1} - \textstyle\int \mathbb{E}_{i,t}\big[m_{t+1}\big]di \equiv \lambda_\varepsilon(s^{t+1})'\Sigma(s_{t+1})\varepsilon_{t+1} + \lambda_X(s^{t+1})'X_t + \lambda_\sigma(s^{t+1}) \tag{A.75}
$$

A.5.2 Deriving Betas

To compute betas for the market return by exploiting the equation:

$$r_{m,t+1} - \int \mathbb{E}_{i,t}\big[r_{m,t+1}\big]di = \kappa_{1,m}\Big(pd_{t+1} - \int \mathbb{E}_{i,t}\big[pd_{t+1}\big]di\Big) + \Big(\triangle d_{t+1} - \int \mathbb{E}_{i,t}\big[\triangle d_{t+1}\big]di\Big) \quad \text{(A.76)}$$

Define the expected value of the slope and intercepts of the price function:

$$\int \mathbb{E}_{i,t}\big[g_0^{\text{pd}}(s^{t+1})\big]di = \pi g_0^{\text{pd}}(s_1^{t+1}) + (1-\pi) g_0^{\text{pd}}(s_0^{t+1}) \quad \text{(A.77)}$$

$$\int \mathbb{E}_{i,t}\big[g_x^{\text{pd}}(s^{t+1})' X_{t+1}\big]di = \Big(\pi g_x^{\text{pd}}(s_1^{t+1})'\mathbf{A}(s_1^{t+1}) + (1-\pi) g_x^{\text{pd}}(s_0^{t+1})'\mathbf{A}(s_0^{t+1})\Big)\mathbf{H}X_t \quad \text{(A.78)}$$

Then the market return innovation can be expressed as:

$$\begin{aligned} r_{m,t+1} - \int \mathbb{E}_{i,t}\big[r_{m,t+1}\big]di = {} & \Big(\kappa_{1,m} g_x^{pd}(s^{t+1})'\mathbf{B}(s^{t+1}) + \varphi_d \sigma(s_{t+1})\mathbf{e}_d'\Big)\varepsilon_{t+1} + \\ & + \Big(\rho_d \mathbf{e}_1'\big(\mathbf{I}-\mathbf{H}\big) + \kappa_{1,m}\Big(g_x^{pd}(s^{t+1})'\mathbf{A}(s^{t+1}) - \int \mathbb{E}_{i,t}\big[g_x^{\text{pd}}(s^{t+1})'\mathbf{A}(s^{t+1})\mathbf{H}\big]di\Big)\Big)X_t + \\ & + \kappa_{1,m}\Big(g_0^{pd}(s^{t+1}) - \int \mathbb{E}_{i,t}\big[g_0^{\text{pd}}(s^{t+1})\big]di\Big) \end{aligned} \quad \text{(A.79)}$$

Which can be written in terms of betas:

$$r_{m,t+1} - \int \mathbb{E}_{i,t}\big[r_{m,t+1}\big]di = \beta_\varepsilon(s^{t+1})\Sigma(s_{t+1})\varepsilon_{t+1} + \beta_X(s^{t+1})X_t + \beta_\sigma(s^{t+1}) \quad \text{(A.80)}$$

and are defined as:

$$\beta_\varepsilon(s^{t+1}) \equiv \Big(\varphi_d \sigma(s_{t+1})\mathbf{e}_d' + \kappa_{1,m} g_x^{pd}(s^{t+1})'\mathbf{B}(s^{t+1})\Big)\Sigma(s_{t+1})^{-1} \quad \text{(A.81)}$$

$$\beta_x(s^{t+1}) \equiv \rho_d \mathbf{e}_1'\big(\mathbf{I}-\mathbf{H}\big) + \kappa_{1,m}\Big(g_x^{pd}(s^{t+1})'\mathbf{A}(s^{t+1}) - \int \mathbb{E}_{i,t}\big[g_x^{\text{pd}}(s^{t+1})'\mathbf{A}(s^{t+1})\mathbf{H}\big]di\Big) \quad \text{(A.82)}$$

$$\beta_\sigma(s^{t+1}) \equiv \kappa_{1,m}\Big(g_0^{pd}(s^{t+1}) - \int \mathbb{E}_{i,t}\big[g_0^{\text{pd}}(s^{t+1})\big]di\Big) \quad \text{(A.83)}$$

A.5.3 Equity Premium

The equity premium is the covariance between the stochastic discount factor and the market return:

$$\int \mathbb{E}_{i,t}\big[r_{m,t+1} - r_{f,t}\big]di + \frac{1}{2}\int \mathrm{Var}_{i,t}\big(r_{m,t+1}\big)di = -\int \mathrm{Cov}_{i,t}(m_{t+1}, r_{m,t+1})di \tag{A.84}$$

Using equations (A.75) and (A.80), the covariance is straightforward:

$$\int \mathrm{Cov}_{i,t}(m_{t+1}, r_{m,t+1})di = \int \mathbb{E}_{i,t}\big[\lambda_\varepsilon(s^{t+1})'\Sigma(s_{t+1})\varepsilon_{t+1}\varepsilon'_{t+1}\Sigma(s_{t+1})'\beta_\varepsilon(s^{t+1})\big]di \tag{A.85}$$

which implies risk premia may be computed by:

$$-\int \mathrm{Cov}_{i,t}(m_{t+1}, r_{m,t+1})di = -\pi\lambda_\varepsilon(s_1^{t+1})'\Sigma(s_1)\Sigma(s_1)'\beta_\varepsilon(s_1^{t+1}) - (1-\pi)\lambda_\varepsilon(s_0^{t+1})'\Sigma(s_0)\Sigma(s_0)'\beta_\varepsilon(s_0^{t+1}) \tag{A.86}$$

A.6 Numerical Algorithm

For a given steady state (pd, pc) and truncated volatility history $s^t \in \mathbb{S}(\tau)$, the model solution is a fixed point to the equations:

- State space system: (2.29), (2.30)
- Prices: (A.43), (A.44), (A.47), (A.48), (A.52), (A.53)
- Conditional covariance matrices: (A.58)

The procedure is as follows:

1. Fix a choice for $\bar{k}$, the order of expectations, and τ, the number of lagged states to use.
   (a) Construct a set of indexes $\mathcal{J} : \mathbb{S}(\tau) \to \mathbb{N}$ for every truncated history in $s^t \in \mathbb{S}(\tau)$.
2. Choose a small error tolerance $\xi > 0$, preferably close to 1×10^{-12}.
3. Given Θ, compute the steady states $\big(pc, pd\big)$ by solving the equations:

$$\mathbb{E}\big[pc - g_0^{pc}(s^t)\big] = 0 \tag{A.87}$$

$$\mathbb{E}\big[pd - g_0^{pd}(s^t)\big] = 0 \tag{A.88}$$

   such that $\max\big\{\big\|pc - g_0^{pc}(s^t)\big\|, \big\|pd - g_0^{pd}(s^t)\big\|\big\} < \xi$ is satisfied. If $|\mathcal{J}|$ is large, programming the objective function for the two steady state equations can be cumbersome. Substituting the full information steady state is useful in this condition if the approximation errors are small.
4. Iterate the Kalman filter over each history $j \in \mathcal{J}$ to find a fixed point for (2.29), (2.30), (A.43), (A.47), (A.52), using the error tolerance ξ to check for convergence of the matrices as in step 3. Convergence is faster if the Kalman filter begins with a well educated guess.
5. Conditional on (2.29), (2.30), (A.43), (A.47), (A.52), iterate the Kalman filter over each history $j \in \mathcal{J}$ to find a fixed point for (A.58) using the error tolerance ξ to check for convergence.
6. Conditional on (2.29), (2.30), (A.43), (A.47), (A.52), iterate equations (A.44), (A.48), (A.53) over each history $j \in \mathcal{J}$ using the error tolerance ξ to check for convergence.

A.7 An Estimation Procedure Utilizing Forecast Panel Data

A.7.1 Conditional Cross-Sectional Covariance Matrix

The conditional cross-sectional covariance matrix is defined:

$$\mathbf{P}_i(s^t) \equiv \mathbb{E}_{i,t}\left[\left(\mathbb{E}_{i,t}\big[X_t\big] - \textstyle\int \mathbb{E}_{j,t}\big[X_t\big]dj\right)\left(\mathbb{E}_{i,t}\big[X_t\big] - \textstyle\int \mathbb{E}_{j,t}\big[X_t\big]dj\right)'\right] \tag{A.89}$$

Expanding the deviations of investor i's forecast from the aggregate:

$$X_{i,t|t} - X_{t|t} = \\ \big(\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big)\big(X_{i,t-1|t-1} - X_{t-1|t-1}\big) + \mathbf{K}(s^t)\mathbf{D}_\epsilon \epsilon_{i,t} \tag{A.90}$$

The covariance matrices $\mathbf{P}_i(s^t)$ for all 2^τ histories may be computed as a fixed point of the following system of equations:

$$\mathbf{P}_i(s^t) = \\ \big(\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big)\mathbf{P}_i(s^{t-1})\cdot \\ \big(\mathbf{A}(s^t) - \mathbf{K}(s^t)\big(\mathbf{C}_1(s^t)\mathbf{A}(s^t) + \mathbf{C}_2(s^t)\big)\big)' + \\ \mathbf{K}(s^t)\mathbf{D}_\epsilon\mathbf{D}_\epsilon'\mathbf{K}_t(s^t)' \tag{A.91}$$

A.7.2 Likelihood

The vector of measured variables Z_{t+1} is defined:

$$Z_{t+1} \equiv \begin{bmatrix} \triangle c_{t+1} \\ \triangle d_{t+1} \\ r_{m,t+1} \\ r_{f,t} \\ \underset{n_{\mathbf{f}}(t+1)\times 1}{\mathbf{f}_{t+2|t+1}} \end{bmatrix} \tag{A.92}$$

Fix a finite sample of observations $Z^T = \{Z_1, Z_2, \ldots, Z_{T-1}, Z_T\}$ for estimation of the model. The dimension of the forecast vector $n_{\mathbf{f}}(t+1)$ is a function of time due to the changing number of respondents in the SPF data each period. Let $\eta_{t+1} \sim_{\text{i.i.d.}} \mathcal{N}(0, \mathbf{I}_{n_{\mathbf{f}}(t+1)})$ be a vector of shocks driving individual forecasts:

$$\mathbf{f}_{t+2|t+1} = \left(\mu_c + \mathbf{e}_1'\mathbf{H}X_{t+1}\right) \cdot \mathbf{1}_{n_{\mathbf{f}}(t+1)} + \sqrt{\mathbf{e}_1'\mathbf{H}\mathbf{P}_i(s^{t+1})\mathbf{H}'\mathbf{e}_1} \cdot \mathbf{I}_{n_{\mathbf{f}}(t+1)}\eta_{t+1} \tag{A.93}$$

Then Z_{t+1} evolves according to the state-space system:

$$X_{t+1} = \mathbf{A}(s^{t+1})X_t + \tilde{\mathbf{B}}(s^{t+1}) \begin{bmatrix} \varepsilon_{t+1} \\ \eta_{t+1} \end{bmatrix} \tag{A.94}$$

$$Z_{t+1} = \boldsymbol{\mu}_Z(s^{t+1}) + \mathbf{Q}_1(s^{t+1})X_{t+1} + \mathbf{Q}_2(s^{t+1})X_t + \mathbf{R}(s^{t+1}) \begin{bmatrix} \varepsilon_{t+1} \\ \eta_{t+1} \end{bmatrix} \tag{A.95}$$

where the matrix $\mathbf{R}(s^{t+1})$ is a combination of aggregate and forecast shock impact matrices and $\tilde{\mathbf{B}}(s^{t+1})$ is modified to accompany the forecast shocks:

$$\mathbf{R}(s^{t+1}) \equiv \begin{bmatrix} \mathbf{D}_\varepsilon(s^{t+1})_- & \underset{n_Y - 1\times n_{\mathbf{f}}(t+1)}{\mathbf{0}} \\ \underset{n_{\mathbf{f}}(t+1)\times n_\varepsilon}{\mathbf{0}} & \sqrt{\mathbf{e}_1'\mathbf{H}\mathbf{P}_i(s^{t+1})\mathbf{H}'\mathbf{e}_1} \cdot \mathbf{I}_{n_{\mathbf{f}}(t+1)} \end{bmatrix} \tag{A.96}$$

$$\tilde{\mathbf{B}}(s^{t+1}) \equiv \begin{bmatrix} \mathbf{B}(s^{t+1}) & \underset{\bar{k}+1\times n_{\mathbf{f}}(t+1)}{\mathbf{0}} \end{bmatrix} \tag{A.97}$$

The likelihood function can be evaluated with a prediction error decomposition:

$$\log \mathcal{L}\big(\Theta, s^T|Z^T\big) = -\frac{1}{2}\sum_{t=1}^{T}\big(2\pi \cdot \dim(Z_t) + \log|\Omega(s^t)| + \tilde{Z}_t'\Omega(s^t)^{-1}\tilde{Z}_t\big) \tag{A.98}$$

where $\Omega(s^t) \equiv \mathbb{E}\big[\tilde{Z}_t\tilde{Z}_t'\big]$ is the covariance matrix associated with the innovation vector $\tilde{Z}_t$:

$$\Omega(s^t) = \big(\mathbf{Q}_1(s^t)\mathbf{A}(s^t) + \mathbf{Q}_2(s^t)\big)\tilde{\mathbf{P}}_0(s^t)\big(\mathbf{Q}_1(s^t)\mathbf{A}(s^t) + \mathbf{Q}_2(s^t)\big)' + \big(\mathbf{Q}_1(s^t)\tilde{\mathbf{B}}(s^t) + \mathbf{R}(s^t)\big)\big(\mathbf{Q}_1(s^t)\tilde{\mathbf{B}}(s^t) + \mathbf{R}(s^t)\big)' \tag{A.99}$$

The innovation $\tilde{Z}_t \equiv Z_t - Z_{t|t-1}$ may be computed with the law of iterated projections:

$$\tilde{Z}_t = Z_t - \boldsymbol{\mu}_Z(s^t) - \big(\mathbf{Q}_1(s^t)\mathbf{A}(s^t) + \mathbf{Q}_2(s^t)\big)X_{t-1|t-1} \tag{A.100}$$

A.7.3 An Adaptive Markov Chain Monte Carlo Estimation Procedure

- Choose an initial $\Theta_{(0)}$ with prior $\mathcal{P}_\Theta(\cdot)$ and a history $s^T_{(0)}$ with prior $\mathcal{P}_s(\cdot)$.
- Solve the model with $\Theta_{(0)}$.
- For $\tau = 0, 1, 2, \ldots, \mathcal{T}$:
  - Randomly partition $\Theta_{(\tau)}$ into two blocks with equal numbers of parameters. For each block $b = 1, 2$:
    * Draw a candidate vector $\tilde{\Theta}_{(\tau),b} \sim \mathcal{N}\big(\Theta_{(\tau-1),b}, \mathbf{V}_{\Theta,b}\big)$.
    * Solve the model for $\tilde{\Theta}_{(\tau),b}$.

* With probability

$$\pi_{\Theta,b} = \max\left\{\min\left\{\frac{\mathcal{L}\big(\tilde{\Theta}_{(\tau),b}, s^T_{(\tau-1)}|Z^T\big)\mathcal{P}_\Theta\big(\tilde{\Theta}_{(\tau-1),b}\big)\mathcal{P}_s\big(s^T_{(\tau-1)}|\tilde{\Theta}_{(\tau),b}\big)}{\mathcal{L}\big(\Theta_{(\tau-1)}, s^T_{(\tau-1)}|Z^T\big)\mathcal{P}_\Theta\big(\Theta_{(\tau-1)}\big)\mathcal{P}_s\big(s^T_{(\tau-1)}|\Theta_{(\tau-1)}\big)}, 1\right\}, 0\right\} \tag{A.101}$$

· Set $\Theta_{(\tau)} = \tilde{\Theta}_{(\tau),b}$

· Otherwise $\Theta_{(\tau)} = \Theta_{(\tau-1)}$

– Draw candidate $\tilde{s}^T_{(\tau)} \sim Q_s\big(s^T_{(\tau-1)}|\Theta_{(\tau)}, Z^T\big)$

* With probability

$$\pi_s = \max\left\{\min\left\{\frac{\mathcal{L}\big(\Theta_{(\tau)}, \tilde{s}^T_{(\tau)}|Z^T\big)\mathcal{P}_s\big(\tilde{s}^T_{(\tau)}|\Theta_{(\tau)}\big)}{\mathcal{L}\big(\Theta_{(\tau)}, s^T_{(\tau-1)}|Z^T\big)\mathcal{P}_s\big(s^T_{(\tau-1)}|\Theta_{(\tau)}\big)}, 1\right\}, 0\right\} \tag{A.102}$$

· Set $s^T_{(\tau)} = \tilde{s}^T_{(\tau)}$

· Otherwise $s^T_{(\tau)} = s^T_{(\tau-1)}$

As in Nimark [2014], the proposal density Q_s works by randomly changing each observation in $s^T_{(\tau)}$ from one to zero (or zero to one) with a given probability $\alpha = 0.05$. The proposal covariance matrix $\mathbf{V}_\Theta$ can be computed as the empirical covariance matrix of candidate parameters $\big\{\Theta_{(0)}, \Theta_{(1)}, \ldots, \Theta_{(\tau-2)}, \Theta_{(\tau-1)}\big\}$ multiplied by a small scaling constant $a \in \mathbb{R}_+$:

$$\mathbf{V}_\Theta = a\widehat{\text{Cov}}(\Theta, \Theta) \tag{A.103}$$

A.8 Variance Decomposition of the Consumption Growth Term Structure

Dependent Variable	Independent Variable		
	Level	Slope	Curvature
$x_t^{(0)}$	0.922	0.073	0.005
$x_t^{(1)}$	0.975	0.024	0.001
$x_t^{(2)}$	0.990	0.007	0.002
$x_t^{(3)}$	0.996	0.001	0.002
$x_t^{(4)}$	0.998	0.000	0.001
$x_t^{(5)}$	0.998	0.001	0.001
$x_t^{(6)}$	0.997	0.003	0.000
$x_t^{(7)}$	0.994	0.006	0.000
$x_t^{(8)}$	0.991	0.009	0.000
$x_t^{(9)}$	0.987	0.012	0.000
$x_t^{(10)}$	0.983	0.016	0.001
$x_t^{(11)}$	0.979	0.019	0.002
$x_t^{(12)}$	0.975	0.022	0.003

Table A.1: Variance decomposition of higher order expectation in terms of their first three principal components (level, slope, and curvature). I compute the values in each row by projecting a simulated time series of higher order expectations onto the three time series of component scores for the level, slope and curvature factors. Each variance share is the square of the regression coefficient multiplied by the ratio of the factor variance to the variance of the higher order expectation. The length of the simulated sample is 1,000,000 observations with $\tau = 8$ and $\bar{k} = 12$. The t-statistics for the first stage regression are unanimously greater than one hundred, even after correcting the residuals for autocorrelation and heteroskedasticity, so I do not report them.

APPENDIX B

APPENDIX OF CHAPTER 2

B.1 Value Functions

B.1.1 Type I & II Agents with $\gamma > 1$ or $\vartheta < +\infty$

Full Policy Functions

$$\begin{aligned}
\alpha_0 &= \frac{\beta}{1-\beta}\Bigg(\frac{\mu}{1-\beta} - \Big(\frac{\beta}{(1-\beta\rho)(1-\beta)}\Big)^2 \frac{\sigma_x^2}{2\vartheta} - \\
&\qquad \Big(\frac{1}{1-\beta}\Big)^2 \frac{\sigma_c^2}{2\vartheta} - \vartheta \log\Big(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)\Bigg) \\
\alpha_c &= \frac{1}{1-\beta} \\
\alpha_x &= \frac{\beta}{(1-\beta\rho)(1-\beta)}
\end{aligned}$$

No Short-Run Risk

$$\begin{aligned}
\alpha_0 &= \frac{\beta}{1-\beta}\Bigg(\frac{\mu+\frac{1}{2}\sigma_c^2}{1-\beta} - \Big(\frac{\beta}{(1-\beta\rho)(1-\beta)}\Big)^2 \frac{\sigma_x^2}{2\vartheta} - \\
&\qquad \vartheta \log\Big(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)\Bigg) \\
\alpha_c &= \frac{1}{1-\beta} \\
\alpha_x &= \frac{\beta}{(1-\beta\rho)(1-\beta)}
\end{aligned}$$

No Long-Run Risk

$$\begin{aligned}
\alpha_0 &= \frac{\beta}{1-\beta}\Bigg(\frac{\mu}{1-\beta} - \Big(\frac{1}{1-\beta}\Big)^2 \frac{\sigma_c^2}{2\vartheta} - \vartheta \log\Big(1-\pi_d+\pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\Big)\Bigg) \\
\alpha_c &= \frac{1}{1-\beta}
\end{aligned}$$

No Disasters

$$\alpha_0 = \frac{\beta}{1-\beta}\left(\frac{\mu + \pi_d \log(1-b)}{1-\beta} - \left(\frac{\beta}{(1-\beta\rho)(1-\beta)}\right)^2 \frac{\sigma_x^2}{2\vartheta} - \left(\frac{1}{1-\beta}\right)^2 \frac{\sigma_c^2}{2\vartheta}\right)$$
$$\alpha_c = \frac{1}{1-\beta}$$
$$\alpha_x = \frac{\beta}{(1-\beta\rho)(1-\beta)}$$

No Risk

$$\alpha_0 = \frac{\beta}{(1-\beta)^2}\left(\mu + \frac{1}{2}\sigma_c^2 + \pi_d \log(1-b)\right)$$
$$\alpha_c = \frac{1}{1-\beta}$$

B.1.2 Type I & II Agents with $\gamma = 1$ or $\vartheta = +\infty$

No Short-Run Risk

$$\alpha_0 = \frac{\beta}{1-\beta}\left(\frac{\mu + \varphi_c}{1-\beta} + \lim_{\vartheta \to \infty} \log\left(\left(1 - \pi_d + \pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\right)^{-\vartheta}\right)\right)$$
$$\alpha_c = \frac{1}{1-\beta}$$
$$\alpha_x = \frac{\beta}{(1-\beta\rho)(1-\beta)}$$

No Long-Run Risks

$$\alpha_0 = \frac{\beta}{1-\beta}\left(\frac{\mu + \varphi_x}{1-\beta} + \lim_{\vartheta \to \infty} \log\left(\left(1 - \pi_d + \pi_d(1-b)^{-\frac{1}{(1-\beta)\vartheta}}\right)^{-\vartheta}\right)\right)$$
$$\alpha_c = \frac{1}{1-\beta}$$

No Disasters

$$
\begin{aligned}
\alpha_0 &= \frac{\beta(\mu + \pi_d \log(1-b))}{(1-\beta)^2} \\
\alpha_c &= \frac{1}{1-\beta} \\
\alpha_x &= \frac{\beta}{(1-\beta\rho)(1-\beta)}
\end{aligned}
$$

No Risk

$$
\begin{aligned}
\alpha_0 &= \frac{\beta(\mu + \varphi)}{(1-\beta)^2} \\
\alpha_c &= \frac{1}{1-\beta}
\end{aligned}
$$

B.2 Derivations

B.2.1 Derivation of the Distortion

$$
\begin{aligned}
\hat{m}_{t+1}(s_{t+1}) &= \frac{\exp(-\mathbf{V}^{\text{II}}(s_{t+1})/\vartheta)}{\mathbb{E}_t[\exp(-\mathbf{V}^{\text{II}}(s_{t+1})/\vartheta)]} \\
&= \frac{\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}c_{t+1} - \frac{\alpha_x}{\vartheta}x_{t+1})}{\mathbb{E}_t[\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}c_{t+1} - \frac{\alpha_x}{\vartheta}x_{t+1})]} \\
&= \frac{\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}(\mu + c_t + x_t + \varepsilon^c_{t+1} + \varepsilon^d_{t+1}) - \frac{\alpha_x}{\vartheta}(\rho x_t + \varepsilon^x_{t+1}))}{\mathbb{E}_t[\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}(\mu + c_t + x_t + \varepsilon^c_{t+1} + \varepsilon^d_{t+1}) - \frac{\alpha_x}{\vartheta}(\rho x_t + \varepsilon^x_{t+1}))]} \\
&= \frac{\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}\mu - \frac{\alpha_c}{\vartheta}c_t - (\frac{\alpha_c}{\vartheta} + \frac{\alpha_x \rho}{\vartheta})x_t - \frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1}))}{\mathbb{E}_t[\exp(-\frac{\alpha_0}{\vartheta} - \frac{\alpha_c}{\vartheta}\mu - \frac{\alpha_c}{\vartheta}c_t - (\frac{\alpha_c}{\vartheta} + \frac{\alpha_x \rho}{\vartheta})x_t - \frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1}))]} \\
&= \frac{\exp(-\frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1}))}{\mathbb{E}_t[\exp(-\frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1}))]} \\
&= \frac{\exp(-\frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1})}{\exp(\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2})(1 - \pi_d + \pi_d(1-b)^{-\frac{\alpha_c}{\vartheta}})\exp(\frac{\alpha_x^2\sigma_x^2}{2\vartheta^2})}
\end{aligned}
$$

B.2.2 Worst-Case Distribution

Individually, the probability distributions of shocks are,

$$
\begin{aligned}
f^d(\varepsilon^d_{t+1}) &= \pi_d^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}(1-\pi_d)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}} \\
f^c(\varepsilon^c_{t+1}) &= \frac{1}{\sqrt{2\pi}\sigma_c}e^{-\frac{(\varepsilon^c_{t+1})^2}{2\sigma_c^2}} \\
f^x(\varepsilon^x_{t+1}) &= \frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{(\varepsilon^x_{t+1})^2}{2\sigma_x^2}}.
\end{aligned}
$$

By independence, the joint distribution is the product of the marginals,

$$
\begin{aligned}
f^\varepsilon(\varepsilon_{t+1}) &= f^d(\varepsilon^d_{t+1})f^c(\varepsilon^c_{t+1})f^x(\varepsilon^x_{t+1}) \\
&= \pi_d^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}(1-\pi_d)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}}\frac{1}{2\pi\sigma_c\sigma_x}e^{-\frac{(\varepsilon^c_{t+1})^2}{2\sigma_c^2}-\frac{(\varepsilon^x_{t+1})^2}{2\sigma_x^2}}.
\end{aligned}
$$

The distortion,

$$\hat{m}_{t+1}(\varepsilon_{t+1}) = \frac{\exp\Big(-\frac{1}{\vartheta(1-\beta)}\varepsilon^c_{t+1} - \frac{1}{\vartheta(1-\beta)}\varepsilon^d_{t+1} - \frac{\beta}{\vartheta(1-\beta\rho)(1-\beta)}\varepsilon^x_{t+1}\Big)}{\exp\Big(\frac{1}{(1-\beta)^2}\frac{\sigma_c^2}{2\vartheta}\Big)\Big(\pi_d + (1-\pi_d)(1-b)^{-\frac{1}{\vartheta(1-\beta)}}\Big)\exp\Big(\Big(\frac{\beta}{(1-\beta\rho)(1-\beta)}\Big)^2\frac{\sigma_x^2}{2\vartheta}\Big)}$$
$$\propto \exp\Big(-\frac{1}{\vartheta(1-\beta)}\varepsilon^c_{t+1} - \frac{1}{\vartheta(1-\beta)}\varepsilon^d_{t+1} - \frac{\beta}{\vartheta(1-\beta\rho)(1-\beta)}\varepsilon^x_{t+1}\Big)$$

The worst-case distribution,

$$\tilde{f}^\varepsilon(\varepsilon_{t+1}) \propto m_{t+1}(\varepsilon_{t+1}) f^\varepsilon(\varepsilon_{t+1}).$$

This evaluates to,

$$\tilde{f}^\varepsilon(\varepsilon_{t+1}) \propto \Big(e^{-\frac{\log(1-b)}{\vartheta(1-\beta)}}\pi_d\Big)^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}\Big(1-\pi_d\Big)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}}\frac{1}{2\pi\sigma_c\sigma_x}e^{-\frac{\left(\varepsilon^c_{t+1}+\frac{\sigma_c^2}{\vartheta(1-\beta)}\right)^2}{2\sigma_c^2}-\frac{\left(\varepsilon^x_{t+1}+\frac{\beta\sigma_x^2}{\vartheta(1-\beta\rho)(1-\beta)}\right)^2}{2\sigma_x^2}}$$
$$= (\tilde{\pi}_d)^{\frac{\varepsilon^d_{t+1}}{\log(1-b)}}(1-\tilde{\pi}_d)^{1-\frac{\varepsilon^d_{t+1}}{\log(1-b)}}\frac{1}{2\pi\sigma_c\sigma_x}e^{-\frac{\left(\varepsilon^c_{t+1}+\frac{\sigma_c^2}{\vartheta(1-\beta)}\right)^2}{2\sigma_c^2}-\frac{\left(\varepsilon^x_{t+1}+\frac{\beta\sigma_x^2}{\vartheta(1-\beta\rho)(1-\beta)}\right)^2}{2\sigma_x^2}}$$

B.2.3 Model-Detection Error Probabilities

The approximating model (3.1)-(3.2) can be thought of as a vector autoregression (VAR),

$$\begin{pmatrix} c_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & \rho \end{pmatrix}\begin{pmatrix} c_t \\ x_t \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \varepsilon^d_{t+1} \\ \varepsilon^c_{t+1} \\ \varepsilon^x_{t+1} \end{pmatrix}.$$

The worst-case distribution is,

$$\begin{pmatrix} c_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & \rho \end{pmatrix}\begin{pmatrix} c_t \\ x_t \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\left(\begin{pmatrix} \tilde{\varepsilon}^d_{t+1} \\ \tilde{\varepsilon}^c_{t+1} \\ \varepsilon^x_{t+1} \end{pmatrix} + \begin{pmatrix} w^d_{t+1} \\ w^c_{t+1} \\ w^x_{t+1} \end{pmatrix}\right),$$

where $\tilde{\varepsilon}_{t+1}$ is a shock with the same distribution as ε_{t+1} and w_{t+1} the mean of the worst-case distribution. This simplifies to,

$$\begin{pmatrix} c_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} \mu + \tilde{\pi}\log(1-b) - \frac{\sigma_c^2}{\vartheta(1-\beta)} \\ -\frac{\beta\sigma_x^2}{\vartheta(1-\beta\rho)(1-\beta)} \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} c_t \\ x_t \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\varepsilon}^d_{t+1} \\ \tilde{\varepsilon}^c_{t+1} \\ \tilde{\varepsilon}^x_{t+1} \end{pmatrix}.$$

Now, we can re-think of this as a Markov-switching VAR with Normal shocks and a hidden disaster state,

$$\begin{pmatrix} c_{t+1} \\ x_{t+1} \end{pmatrix} = \tilde{\mu}(\varepsilon^d_{t+1}) + \begin{pmatrix} 1 & 1 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} c_t \\ x_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\varepsilon}^c_{t+1} \\ \tilde{\varepsilon}^x_{t+1} \end{pmatrix},$$

where,

$$\tilde{\mu}(\varepsilon^d_{t+1}) = \begin{pmatrix} \mu + \tilde{\pi}\log(1-b) - \frac{\sigma_c^2}{\vartheta(1-\beta)} + \varepsilon^d_{t+1} \\ -\frac{\beta\sigma_x^2}{\vartheta(1-\beta\rho)(1-\beta)} \end{pmatrix}.$$

B.2.4 Risk-Free Rate

$$\begin{aligned}
\frac{1}{r_t^f} &= \mathbb{E}_t[\Lambda_{t,t+1}] \\
&= \mathbb{E}_t\left[\beta\frac{C_t}{C_{t+1}}\frac{\exp(-\mathbf{V}^{\text{I}}(s_{t+1})/\vartheta)}{\mathbb{E}_t[\exp(-\mathbf{V}^{\text{I}}(s_{t+1})/\vartheta)]}\right] \\
&= \mathbb{E}_t\left[\beta\frac{\exp(c_t)}{\exp(c_{t+1})}\frac{\exp(-\frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1})}{\exp(\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2})(1-\pi_d+\pi_d(1-b)^{-\frac{\alpha_c}{\vartheta}})\exp(\frac{\alpha_x^2\sigma_x^2}{2\vartheta^2})}\right] \\
&= \mathbb{E}_t\left[\frac{\beta}{\exp(\mu + x_t + \varepsilon^d_{t+1} + \varepsilon^c_{t+1})}\frac{\exp(-\frac{\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1})}{\exp(\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2})(1-\pi_d+\pi_d(1-b)^{-\frac{\alpha_c}{\vartheta}})\exp(\frac{\alpha_x^2\sigma_x^2}{2\vartheta^2})}\right] \\
&= \mathbb{E}_t\left[\frac{\beta}{\exp(\mu + x_t)}\frac{\exp(-\frac{(1+\vartheta)\alpha_c}{\vartheta}\varepsilon^c_{t+1} - \frac{(1+\vartheta)\alpha_c}{\vartheta}\varepsilon^d_{t+1} - \frac{\alpha_x}{\vartheta}\varepsilon^x_{t+1})}{\exp(\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2})(1-\pi_d+\pi_d(1-b)^{-\frac{\alpha_c}{\vartheta}})\exp(\frac{\alpha_x^2\sigma_x^2}{2\vartheta^2})}\right] \\
&= \frac{\beta}{\exp(\mu + x_t)}\exp\left(\left((1+\vartheta)^2 - 1\right)\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2}\right)\frac{(1-\pi_d+\pi_d(1-b)^{-\frac{(1+\vartheta)\alpha_c}{\vartheta}})}{(1-\pi_d+\pi_d(1-b)^{-\frac{\alpha_c}{\vartheta}})} \\
&= \frac{\beta}{\exp(\mu + x_t)}\exp\left(\left((1+\vartheta)^2 - 1\right)\frac{\alpha_c^2\sigma_c^2}{2\vartheta^2}\right)\varphi_d
\end{aligned}$$

APPENDIX C

APPENDIX OF CHAPTER 3

C.1 Equilibrium Equations

$$
\begin{aligned}
V_t &= \left(\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)}\mathcal{R}_t^{1-1/\psi}\right)^{\frac{1}{1-1/\psi}} \\
\mathcal{R}_t^{1-\gamma} &= \mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right] \\
y_t &= c_t + i_t + b_t - \frac{e^{g_t}b_{t+1}}{1+e^{r_t}} + \frac{\varphi_b}{2}\left(b_{t+1} - b\right)^2 \\
e^{g_t}k_{t+1} &= (1-\delta)k_t + i_t - \frac{\varphi_k}{2}\left(\frac{e^{g_t}k_{t+1}}{k_t} - e^{\mu_g}\right)^2 k_t \\
\Lambda_{t,t+1} &= \beta e^{-g_t/\psi}\frac{\lambda_{t+1}}{\lambda_t}\cdot\left(\frac{\mathcal{R}_t}{V_{t+1}}\right)^{\gamma-1/\psi} \\
\lambda_t &= V_t^{1/\psi}\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{-1/\psi} \\
\lambda_t w_t &= V_t^{1/\psi}h_t^{\nu-1}\left(c_t - \frac{1}{\nu}h_t^\nu\right)^{-1/\psi} \\
1 &= \mathbb{E}_t\left[\Lambda_{t,t+1}\left(\frac{1}{\varphi_b(b_{t+1}-b) + e^{g_t}/1+e^{r_t}}\right)\right] \\
1 &= \mathbb{E}_t\left[\Lambda_{t,t+1}\left(\varphi_k\frac{y_{t+1}}{k_{t+1}} + 1 - \delta + \frac{\varphi_k}{2}\left(\left(\frac{e^{g_{t+1}}k_{t+2}}{k_{t+1}}\right)^2 - e^{2\mu_g}\right)\right) - \varphi_k\left(\frac{e^{g_t}k_{t+1}}{k_t} - e^{\mu_g}\right)\right] \\
y_t &= e^{z_t}k_t^{\varphi_k}(e^{g_t}h_t)^{\varphi_h} \\
z_t &= \rho_z z_{t-1} + \varepsilon_t^z + [\text{news}]_t^{t+\tau} \\
\sigma_t^z &= (1-\rho_{\sigma^z})\mu^{\sigma^z} + \rho_{\sigma^z}\sigma_{t-1}^z + \eta^{\sigma^z}\varepsilon_t^{\sigma^z}, \\
g_t &= (1-\rho_g)\mu_g + \rho_g g_{t-1} + \eta^g\varepsilon_t^g \\
\epsilon_t^r &= \rho_r\varepsilon_{t-1}^r + e^{\sigma_t^r}\varepsilon_t^r + [\text{news}]_t^{t+\tau} \\
\sigma_t^r &= (1-\rho_{\sigma^r})\mu^{\sigma^r} + \rho_{\sigma^r}\sigma_{t-1}^r + \eta^{\sigma^r}\varepsilon_t^{\sigma^r} \\
e^{r_t} &= e^{\bar{r}} + \varepsilon_t^r \\
\sigma_t^n &= (1-\rho_{\sigma^n})\mu^{\sigma^n} + \rho_{\sigma^n}\sigma_{t-1}^n + \eta^{\sigma^n}\varepsilon_t^{\sigma^n} \\
[\text{news}]_t^{t+\tau} &= e^{\sigma_t^n}\varepsilon_t^n
\end{aligned}
$$

C.2 Competitive Equilibrium

A competitive equilibrium is a sequence $\{V_t, \mathcal{R}_t, c_t, h_t, i_t, k_t, b_t\}_{t=0}^{\infty}$ and prices $\{r_t\}_{t=0}^{\infty}$ such that in every period:

1. Given prices, the representative household solves its utility maximization problem.
2. Given prices, the representative firm maximizes profits.
3. All resource constraints and laws of motions bind.

C.3 Impulse Response Functions

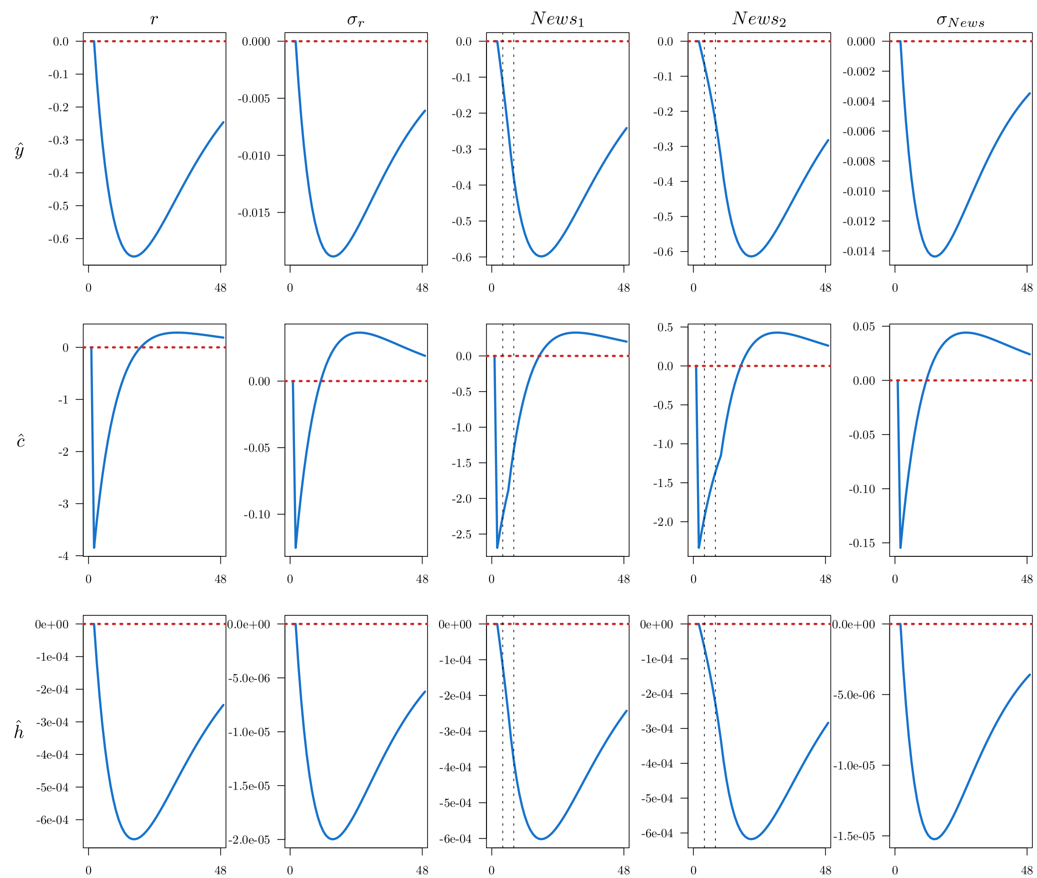

Figure C.1: Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about the interest rate.

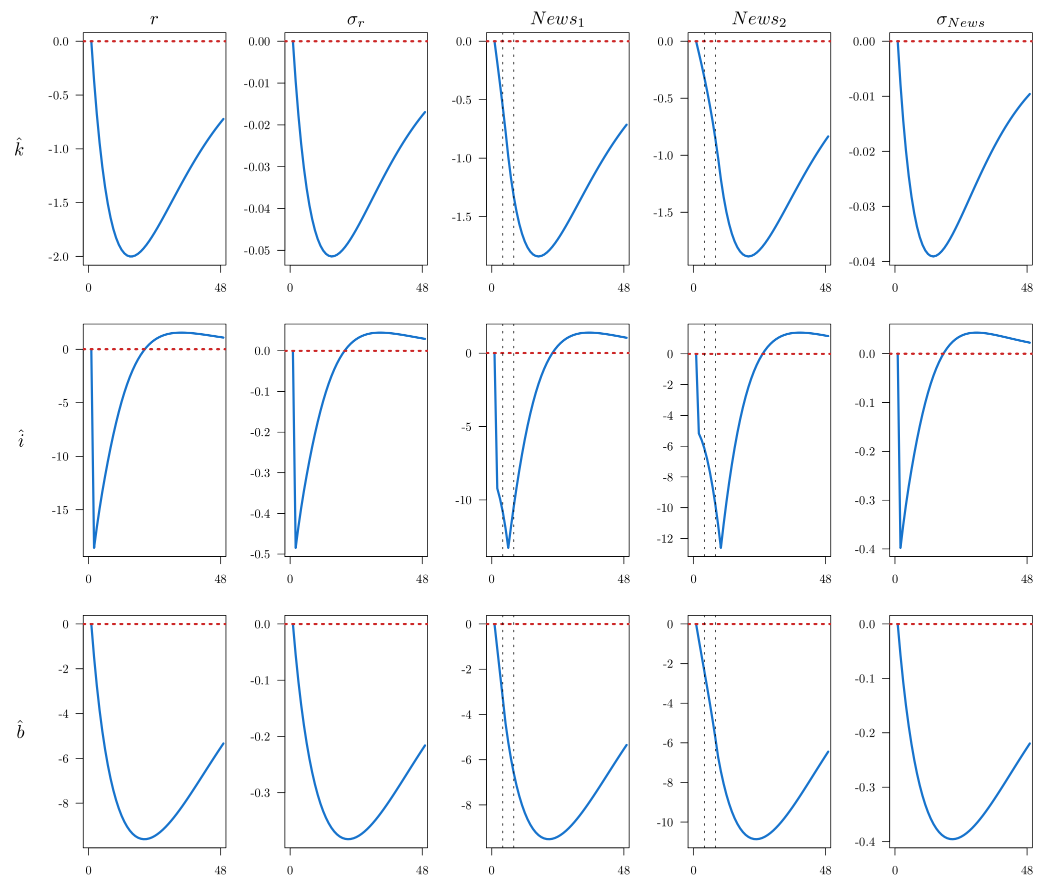

Figure C.2: Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about the interest rate.

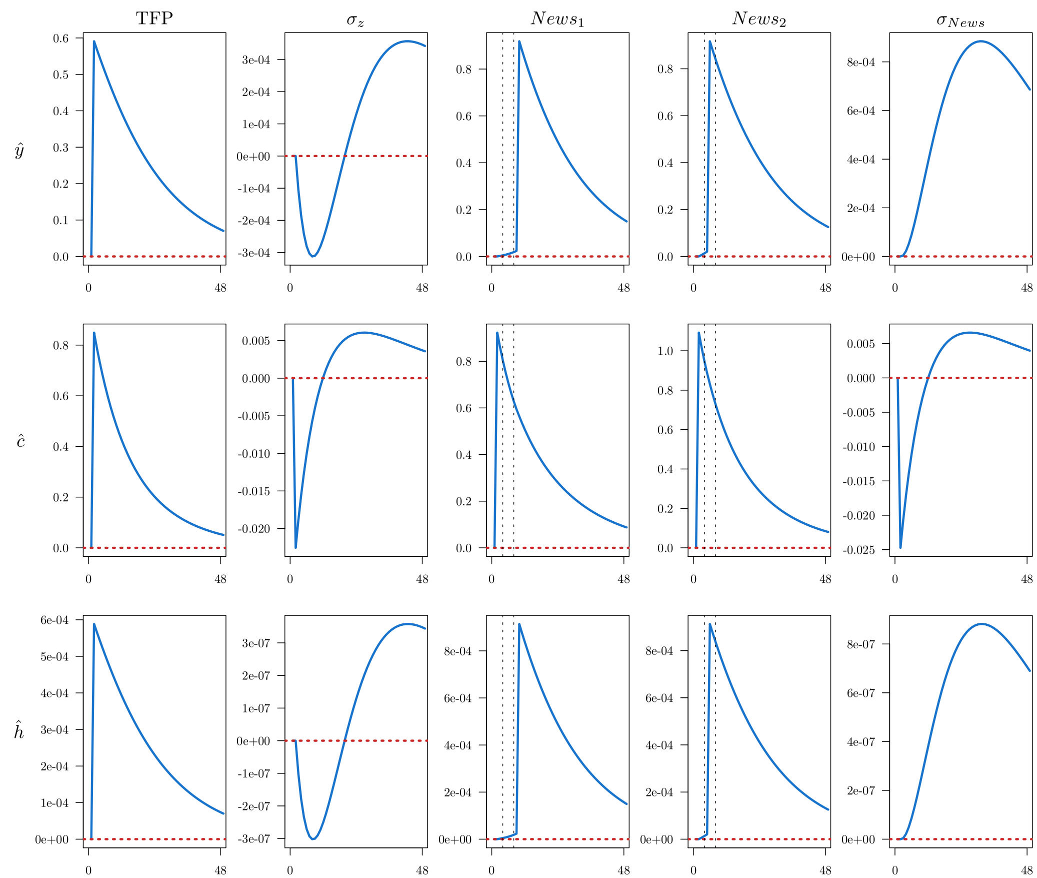

Figure C.3: Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about TFP.

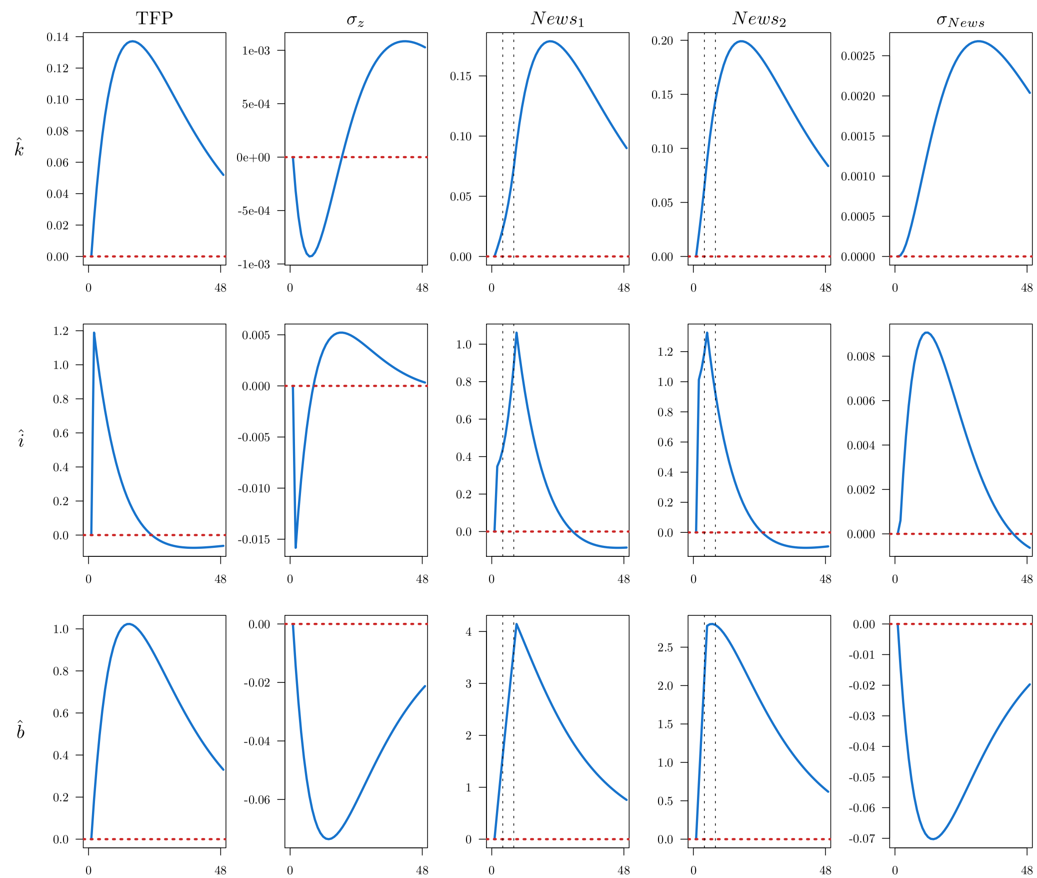

Figure C.4: Plots display reposes of endogenous variables (rows) to a one standard deviation impulse of an exogenous shock (columns). The vertical axis is scaled in percentage points, and the horizontal time axis is in quarters. News in this model is about TFP.

C.3.1 Household's Problem with Recursive Preferences

The notation $V_t = V(k_t, b_t, \varepsilon_t, news_t)$ represents the following optimization problem,

$$V(k_t, b_t, news_t) = \tag{C.1}$$

$$\max_{c_t, h_t, i_t, b_{t+1}} \left(\left(c_t - \frac{1}{\nu} h_t^{\nu} \right)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)} \mathbb{E}_t \left[V(k_{t+1}, b_{t+1}, news_{t+1})^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}}$$

subject to,

$$y_t = c_t + i_t + b_t - \frac{e^{g_t} b_{t+1}}{1+e^{r_t}} + \frac{\varphi_b}{2}\left(b_{t+1} - b \right)^2$$

with contemporaneous shocks ε_t and $news_t$,

$$\varepsilon_t = \begin{pmatrix} \varepsilon_t^g \\ \varepsilon_t^z \\ \varepsilon_t^{\sigma^z} \\ \varepsilon_t^r \\ \varepsilon_t^{\sigma^r} \end{pmatrix} \tag{C.2}$$

$$news_t = \begin{pmatrix} e^{\sigma_{t+8}^n} \varepsilon_{t+8}^n \\ e^{\sigma_{t+7}^n} \varepsilon_{t+7}^n \\ e^{\sigma_{t+6}^n} \varepsilon_{t+6}^n \\ e^{\sigma_{t+5}^n} \varepsilon_{t+5}^n \\ e^{\sigma_{t+4}^n} \varepsilon_{t+4}^n \\ e^{\sigma_{t+3}^n} \varepsilon_{t+3}^n \\ e^{\sigma_{t+2}^n} \varepsilon_{t+2}^n \\ e^{\sigma_{t+1}^n} \varepsilon_{t+1}^n \\ e^{\sigma_t^n} \varepsilon_t^n \end{pmatrix}, \tag{C.3}$$

given all other laws of motion and the firm's production function.

C.3.2 Stationarity

Define the function V_t as the stationary utility function,

$$V_t = \frac{1}{\Gamma_{t-1}} \mathbf{V}_t. \tag{C.4}$$

The analytic form of V_t follows,

$$\begin{aligned}
\mathbf{V}_t &= \left(\left(C_t - \Gamma_{t-1} \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta \left[\mathbf{V}_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \\
&= \left(\left(\Gamma_{t-1} c_t - \Gamma_{t-1} \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta \mathbb{E}_t \left[\mathbf{V}_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \\
&= \left(\Gamma_{t-1}^{1-1/\psi} \left(c_t - \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta \mathbb{E}_t \left[\Gamma_t^{1-\gamma} V_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \\
&= \left(\Gamma_{t-1}^{1-1/\psi} \left(c_t - \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta \Gamma_t^{1-1/\psi} \mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \\
&= \Gamma_{t-1} \left(\left(c_t - \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)} \mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}} \\
&= \Gamma_{t-1} V_t.
\end{aligned}$$

Hence,

$$V_t = \left(\left(c_t - \frac{1}{\nu} h_t^\nu \right)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)} \mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right]^{\frac{1-1/\psi}{1-\gamma}} \right)^{\frac{1}{1-1/\psi}}. \tag{C.5}$$

C.4 Stochastic Discount Factor

Define a generalized stochastic discount factor as,

$$\Lambda_{t,t+1} = \frac{\partial V_t / \partial c_{t+1}}{\partial V_t / \partial c_t}. \tag{C.6}$$

Then,

$$\begin{aligned}\frac{\partial V_t}{\partial c_t} &= \left(u(c_t,h_t)^{1-1/\psi} + \beta e^{g_t(1-1/\psi)}\mathcal{R}_t^{1-1/\psi}\right)^{\frac{1/\psi}{1-1/\psi}} u(c_t,h_t)^{-1/\psi}u_c(c_t,h_t)\\ &= V_t^{1/\psi}u(c_t,h_t)^{-1/\psi}u_c(c_t,h_t)\\ \frac{\partial V_t}{\partial c_{t+1}} &= \beta e^{g_t(1-1/\psi)}V_t^{1/\psi}\mathcal{R}_t^{\gamma-1/\psi}V_{t+1}^{1/\psi-\gamma}u(c_{t+1},h_{t+1})^{-1/\psi}u_c(c_{t+1},h_{t+1})\end{aligned}$$

And dividing the two resulting equations yields,

$$\begin{aligned}\Lambda_{t,t+1} &= \frac{\partial V_t/\partial c_{t+1}}{\partial V_t/\partial c_t} \qquad \text{(C.7)}\\ &= \beta e^{g_t(1-1/\psi)}\frac{V_t^{1/\psi}\mathcal{R}_t^{\gamma-1/\psi}V_{t+1}^{1/\psi-\gamma}u(c_{t+1},h_{t+1})^{-1/\psi}u_c(c_{t+1},h_{t+1})}{V_t^{1/\psi}u(c_t,h_t)^{-1/\psi}u_c(c_t,h_t)}\\ &= \beta e^{g_t(1-1/\psi)}\left(\frac{u(c_{t+1},h_{t+1})}{u(c_t,h_t)}\right)^{-1/\psi}\left(\frac{u_c(c_{t+1},h_{t+1})}{u_c(c_t,h_t)}\right)\left(\frac{\mathcal{R}_t}{V_{t+1}}\right)^{\gamma-1/\psi} \qquad \text{(C.8)}\end{aligned}$$

Substituting in the analytic form of preferences,

$$\Lambda_{t,t+1} = \beta e^{g_t(1-1/\psi)}\left(\frac{c_{t+1}-\frac{1}{\nu}h_{t+1}^{\nu}}{c_t-\frac{1}{\nu}h_t^{\nu}}\right)^{-1/\psi}\left(\frac{\mathcal{R}_t}{V_{t+1}}\right)^{\gamma-1/\psi}.$$

Do note that in the paper the stochastic discount factor is scaled by e^{g_t} because this simplifies the presentation of the Euler equations.

BIBLIOGRAPHY

Mark Aguiar and Gita Gopinath. Emerging Market Business Cycles: The Cycle is the Trend. *Journal of Political Economy*, 115:69–102, 2007.

Elias Albagli, Christian Hellwig, and Aleh Tsyvinski. A Theory of Asset Prices Based on Heterogeneous Information. *Working Paper*, 2015.

Rui Albuquerque, Martin Eichenbaum, Victor Luo, and Sergio Rebelo. Valuation Risk and Asset Pricing. *Working Paper*, 2015.

Martin M. Andreasen, Jesús Fernández-Villaverde, and Juan Rubio-Ramírez. The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Application. *NBER Working Paper*, (18983), April 2013.

Ravi Bansal and Amir Yaron. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance*, 59(4):1481–1509, 2004.

Ravi Bansal, Dana Kiku, and Amir Yaron. An Empirical Evaluation of the Long-Run Risks Model for Asset Prices. *Critical Finance Review*, 1:183–221, 2012.

Francisco Barillas and Kristoffer Nimark. Speculation and the Term Structure of Interest Rates. *Working Paper*, 2016a.

Francisco Barillas and Kristoffer Nimark. Speculation and the Bond Market: An Empirical No-Arbitrage Framework. *Working Paper*, 2016b.

Francisco Barillas, Lars Peter Hansen, and Thomas J. Sargent. Doubts or variability? *Journal of Economic Theory*, 144(6):2388–2418, November 2009.

Robert J. Barro. Rare Disasters and Asset Markets in the Twentieth Century. *The Quarterly Journal of Economics*, 121(3):823–866, 2006.

Robert J. Barro. Rare Disasters, Asset Prices, and Welfare Costs. *American Economic Review*, 99(1):243–264, 2009.

Robert J. Barro and Tao Jin. On the Size Distribution of Macroeconomic Disasters. *Econometrica*, 79(5):1567–1589, September 2011.

Robert B. Barsky and Eric R. Sims. News Shocks and Business Cycles. *Journal of Monetary Economics*, 58(3):273–289, 2011.

Suleyman Basak. Asset Pricing with Heterogeneous Beliefs. *Journal of Banking and Finance*, 29:2849–2881, 2005.

Paul Beaudry and Franck Portier. Stock Prices, News, and Economic Fluctuations. *American Economic Review*, 96(4):1293–1307, 2006.

Jason Beeler and John Y. Campbell. The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment. *Critical Finance Review*, 1:141–182, 2012.

Anmol Bhandari. Reverse Engineering Heterogeneous Beliefs. *Working Paper*, March 2016.

Rhys Bidder and Matthew E. Smith. Doubts and Variability: A Robust Perspective on Exotic Consumption Series. *Federal Reserve Bank of San Francisco Working Paper*, September 2013.

John Y. Campbell. Asset Prices, Consumption and the Business Cycle. In John B. Taylor and Michael Woodford, editors, *Handbook of Macroeconomics*, volume 1. Elsevier, 1999.

John Y. Campbell and Robert J. Shiller. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance*, 43(3):661–676, 1988.

Charles Clarke. The Level, Slope, and Curve Factor Model for Stocks. *Working Paper*, 2016.

Michael P Clements. Do US Macroeconomic Forecasters Exaggerate Their Differences? *Working Paper*, 2014.

John Cochrane. Shocks. *Carnegie-Rochester Conference Series on Public Policy*, 41:295–364, December 1994.

Geroge M. Constantinides and Darrell Duffie. Asset Pricing with Heterogeneous Consumers. *Journal of Political Economy*, 104:219–240, 1996.

Jerome Detemple and Shashidhar Murthy. Intertemporal Asset Pricing with Heterogeneous Beleifs. *Journal of Economic Theory*, 62:294–320, 1994.

Larry G. Epstein and Stanley E. Zin. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 4:937–969, 1989.

Jesús Fernández-Villaverde, Pablo Guerrón-Quintana, and Martın Uribe. Risk matters: The real effects of volatility shocks. *American Economic Review*, 101: 2530–2561, 2011.

Jeremy Greenwood, Zvi Hercowitz, and Gregory W. Huffman. Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review*, 78 (3):402–417, June 1988.

Sanford Grossman and Joseph Stiglitz. On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 70(3):393–408, 1980.

Jonathan Gruber. A Tax Based Estimate of the Elasticity of Iintertemporal Substitution. *Quarterly Journal of Finance*, 3(1), 2013.

Robert Hall. Iintertemporal Substitution in Consumption. *Journal of Political Economy*, 90:339–357, 1988.

Lars Peter Hansen and Thomas J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91:60–66, 2001.

Michael Harrison and David Kreps. Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *Quarterly Journal of Economics*, 92: 323–336, 1978.

Nir Jaimovich and Sergio Rebelo. Can News about the Future Drive the Business Cycle? *American Economic Review*, 99(4):1097–1118, 2009.

Andre Kurmann and Christopher Otrok. News Shocks and the Slope of the Term Structure of Interest Rates. *American Economic Review*, 103(6):2612–32, 2013.

Robert Litterman and Jose Scheinkman. Common Factors Affecting Bond Returns. *The Journal of Fixed Income*, 1991.

Roger Loh and Rene M. Stulz. Is Sell-Side Research More Valuable in Bad Times? *Working Paper*, 2015.

Hanno Lustig, Nikolai Roussanov, and Adrien Verdelhan. Common Risk Factors in Currency Markets. *Review of Financial Studies*, 24(11), 2011.

Rajnish Mehra and Edward C. Prescott. The Equity Premium: A Puzzle. *Journal of Monetary*, 15:145–161, 1985.

Emi Nakamura, Jon Steinsson, Robert J. Barro, and Jose Ursua. Crises and Recoveries in an Empirical Model of Consumption Disasters. *American Economic Journal: Macroeconomics*, 5(3):35–75, 2013.

Emi Nakamura, Dmitriy Sergeyev, and Jon Steinsson. Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence. *Working Paper*, 2016.

Kristoffer Nimark. Dynamic Higher Order Expectations. *Working Paper*, 2011.

Kristoffer Nimark. Man-Bites-Dog Business Cycles. *American Economic Review*, 104(8):2320–2367, 2014.

Kristoffer Nimark. A Low Dimensional Kalman Filter for Systems with Lagged States in the Measurement Equation. *Economics Letters*, 127:10–13, 2015.

Thomas A. Rietz. The Equity Risk Premium: A Solution. *Journal of Monetary Economics*, 22:117–131, 1988.

Jan-Christoph Rulke, Maria Silgoner, and Jula Worz. Herding Behavior of Business Cycle Forecasters in Times of Economic Crises. *Working Paper*, 2012.

Thomas J. Sargent. Equilibrium with Signal Extraction from Endogenous Variables. *Journal of Economic Dynamics and Control*, 15(2):245–273, 1991.

Frank Schorfheide, Dongho Song, and Amir Yaron. Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach. *Working Paper*, 2016.

Thomas D. Tallarini. Risk-Sensitive Real Business Cycles. *Journal of Monetary Economics*, 45(3):507–532, 2000.

Robert M. Townsend. Forecasting the Forecast of Others. *Journal of Political Economy*, 91(4):545–588, 1983.

Hal Varian. Divergence of Opinion in Complete Markets. *Journal of Finance*, 40: 309–317, 1985.

Philippe Weil. The Equity Premium and the Risk-Free Rate Puzzle. *Journal of Monetary Economics*, 24(3):401–421, 1989.

Philippe Weil. Nonexpected Utility in Macroeconomics. *The Quarterly Journal of Economics*, 105(1):29–42, 1990.

Fernando Zapatero. Effects of Financial Innovations on Market Volatility When Beliefs are Heterogeneous. *Journal of Economic Dynamics and Control*, 22: 597–626, 1998.