CERGE
Center for Economics Research and Graduate Education
Charles University

Essays on Asset Pricing

Mykola Babiak

Dissertation

Prague, October 2019

Mykola Babiak

Essays on Asset Pricing

Dissertation

Prague, October 2019

Dissertation Committee

Michal Kejak (CERGE-EI; Chair)

Ctirad Slavik (CERGE-EI)

Sergey Slobodyan (CERGE-EI)

Roman Kozhan (Warwick Business School)

Referees

Daniele Bianchi (Queen Mary University of London)

Anmol Bhandari (University of Minnesota)

Dedication

In memory of my grandfather, Mykola Babiak Sr.

Table of Contents

Abstract ix

Acknowledgments xi

Introduction 1

1 Generalized Disappointment Aversion and the Variance Term Structure 5

1.1 Introduction 6
1.2 Variance and Skewness Risk 11
1.3 Model Setup 15
1.3.1 Generalized Disappointment Aversion Risk Preferences 15
1.3.2 Endowments and Inference Problem 16
1.4 Asset Prices 18
1.4.1 Equilibrium and Pricing Kernel 18
1.4.2 Model Solution 19
1.4.3 Prices and Returns of Variance Swaps 19
1.4.4 Variance and Skew Risk Premiums 20
1.4.5 Option Prices and Implied Volatilities 21
1.5 Calibration and Quantitative Results 21
1.5.1 Calibrated Parameters 22
1.5.2 Endowments and Equity Returns 24
1.5.3 The Price of Variance Risk 25
1.5.4 Variance and Skew Risk Premiums 29
1.5.5 The Term Structure of Implied Volatilities 32
1.6 Sensitivity Analysis 33
1.6.1 The Variance Term Structure 34
1.6.2 Equity Returns and Moment Risk Premiums 35
1.6.3 Implied Volatilities 38
1.7 Conclusion 41

A Appendix **43**
A.1 Data . . . 43
A.1.1 Consumption, Dividends, and Market Returns . . . 43
A.1.2 The Variance Premium Data . . . 44
A.1.3 Option Data for the Skew Premium and Implied Volatility Skew . 45
A.2 Representative Agent's Maximization Problem . . . 46
A.3 The Numerical Solution . . . 49
A.3.1 The Projection Method . . . 50
A.3.2 Implementation in Matlab . . . 52
A.3.3 Accuracy of the Numerical Methods . . . 54

2 Parameter Learning in Production Economies **57**
2.1 Introduction . . . 58
2.2 The Model . . . 63
2.2.1 The Representative Household . . . 64
2.2.2 The Representative Firm . . . 65
2.2.3 Technology . . . 65
2.2.4 Equilibrium Asset Prices . . . 66
2.3 Results . . . 68
2.3.1 Parameter Values . . . 68
2.3.2 Parameter Uncertainty . . . 70
2.3.3 Pricing a Claim to Firm's Levered Dividends . . . 71
2.3.4 Pricing a Claim to Calibrated Dividends . . . 77
2.3.5 Adding Costly Reversibility . . . 81
2.4 Sensitivity Analysis . . . 84
2.5 Conclusion . . . 88

B Appendix **89**
B.1 Numerical Algorithm: Anticipated Utility . . . 89
B.1.1 All Known Parameters . . . 90
B.2 Numerical Algorithm: Priced Parameter Uncertainty . . . 91
B.2.1 Unknown Transition Probabilities . . . 92
B.2.2 Unknown Transition Probabilities and Unknown Mean Growth Rates 97
B.2.3 Existence of Equilibrium . . . 100
B.3 Impulse Responses . . . 100

3 Option Prices and Learning about Productivity Dynamics **103**
3.1 Introduction . . . 104
3.2 The Model . . . 108
3.2.1 The Representative Household . . . 108
3.2.2 The Representative Firm . . . 109
3.2.3 Technology . . . 110
3.2.4 Equilibrium Asset Prices . . . 110
3.3 Quantitative Analysis . . . 113
3.3.1 Calibration . . . 114
3.3.2 Unconditional Moments . . . 116

3.3.3 Conditional Moments . . . 122
3.4 The Impact of Priced Parameter Uncertainty and Volatility Risks . . . 124
3.5 Conclusion . . . 128

C Appendix **131**
C.1 Numerical Algorithm: All Known Parameters . . . 131
C.2 Numerical Algorithm: Anticipated Utility . . . 133
C.3 Numerical Algorithm: Priced Parameter Uncertainty . . . 134
C.3.1 Unknown Transition Probabilities, Unknown Mean Growth Rates, and Unknown Volatilities . . . 134
C.3.2 Existence of Equilibrium . . . 140

Bibliography **141**

Abstract

In the first chapter, I analyze how investor asymmetric risk attitude can rationalize the variance term structure in the aggregate stock market. Contrary to leading asset pricing theories, recent empirical evidence indicates that it was costless to hedge long-term volatility in aggregate stock market returns over the last two decades, whereas investors paid large premia for insurance against the unexpected realized variance. I offer a generalized disappointment aversion explanation that can also account for the variance and skew risk premiums in equity returns and the implied volatility skew of index options. The proposed model also captures other features of the data including the low risk-free rate, the high equity premium, and excess stock market volatility.

In the second chapter, we examine how parameter learning amplifies the impact of macroeconomic shocks on equity prices and quantities in a standard production economy where a representative agent has Epstein-Zin preferences. An investor observes technology shocks that follow a regime-switching process, but does not know the underlying model parameters governing the short-term and long-run perspectives of economic growth. We show that rational parameter learning endogenously generates long-run productivity and consumption risks that help explain a wide array of dynamic pricing phenomena. The asset pricing implications of subjective long-run risks crucially depend on the introduction of a procyclical dividend process consistent with the data.

In the third chapter, we demonstrate that incorporating time-varying productivity volatility and priced parameter uncertainty in a production economy can explain index option prices, equity returns, the risk-free rate, and macroeconomic quantities. A Bayesian investor learns about the true parameters governing mean, persistence, and volatility of productivity growth. Rational parameter learning amplifies the conditional risk premium and volatility especially at the onset of recessions. We estimate the model based on post-war U.S. data and find that it can capture the implied volatility surface and the variance premium. Intuitively, the agent pays a large premium for index options because they hedge future belief revisions.

Abstrakt

Současné empirické důkazy navzdory předním teoriím o oceňování aktiv naznačují, že finanční trhy kompenzují krátkodobé riziko volatility akcií. Rovnovážný model s rizikovými preferencemi se zobecněnou averzí vůči riziku a výjimečnými událostmi vytváří časovou strukturu variability cen swapů a akciových výnosů konzistentní s daty. Kalibrace navíc vysvětluje variabilitu a zešikmení prémie za riziko u akciových výnosů a zešikmení implikované volatility opcí na akciové indexy. Souběžně zachycuje nejdůležitější momenty základních faktorů, akciových výnosů a bezrizikové úrokové míry. Klíčem k intuitivnímu pochopení je skutečnost, že výsledky pramení z endogenní variability pravděpodobnosti výjimečných událostí vedoucích ke zklamání v růstu spotřeby.

Zkoumáme, jakým způsobem učení parametrů posiluje dopad makroekonomic-kých šoků na ceny akcií a jiné veličiny ve standardní produkční ekonomice, kde má reprezentativní agent Epstein-Zin preference. Investor pozoruje technologické šoky, jejichž dynamika je dána procesem s proměnlivým režimem, ale nezná skryté parametry modelu, které řídí krátkodobé a dlouhodobé vyhlídky ekonomického růstu. Ukazujeme, že racionální učení parametrů endogenně generuje dlouhodobé riziko v ekonomickém růstu a ve spotřebě, což pomáhá vysvětlit širokou škálu feno-ménů dynamického oceňování aktiv. Implikace dlouhodobých subjektivních rizik pro oceňování aktiv zásadně závisí na zavedení procyklického procesu dividend, který je konzistentní s daty.

Ukazujeme, že zahrnutí v čase proměnlivé volatility produktivity a oceňování s parametrem nejistoty v produkční ekonomice může vysvětlit ceny opcí na akciové indexy, akciové výnosy, bezrizikový výnos a makroekonomické veličiny. Bayesiánský investor se učí o skutečných parametrech určujících průměr, persistenci a volatilitu růstu produktivity. Racionální parametr učení zvyšuje podmíněné prémie za riziko a volatilitu, obzvláště pak na počátku recese. Provádíme odhad modelu založeného na poválečných datech z USA a zjišťujeme, že je schopen zachytit plochu implikovaných volatilit a prémii za rozptyl. Intuitivně lze chápat výsledky tak, že agent platí vysoké prémie za opce na akciové indexy, protože hedgují budoucí revize názorů.

Acknowledgments

I am very grateful to my advisors Michal Pakos and Michal Kejak for their guidance, encouragement, and the generosity they have shown me with support. I would also like to thank all members of my dissertation committee: Ctirad Slavik, Sergey Slobodyan, and Roman Kozhan. I have benefited greatly from their helpful comments, suggestions, and our many hours of conversations.

I acknowledge the special role of CERGE-EI's academic community in shaping my ideas. I am also indebted to the members of the Study Affairs Office, the Academic Skills Center, the Finance Department, and the Secretariat. Their assistance facilitated a smoother dissertation process.

I wrote a larger part of my dissertation during my research mobility stay at Warwick Business School (WBS). I thank the faculty and staff members at WBS who provided me with valuable comments on my drafts and presentations. I am particularly grateful to Roman Kozhan and Daniele Bianchi for their exceptional contribution to my development as a researcher.

I send a special thank you to Olenka for her support and to my entire family for giving me their love and support. Finally, and most importantly, I would like to thank my mother Nadiya and my grandmother Olga for their continuous love and encouragement through every step of my life. I owe all my accomplishments to my mother and grandmother who are everything to me.

All errors remaining in this text are entirely my own.

Czech Republic, Prague
October 2019

Mykola Babiak

Introduction

All models are wrong, but some are useful.

George E. P. Box

My dissertation consists of three essays investigating how investor learning and macroeconomic uncertainty affect asset prices and the real economy in general equilibrium settings. The models studied in this dissertation deviate from the traditional frameworks in various ways, but they all examine different aspects of the same question: How does incomplete agent information affect the decision-making of agents, and what are the resulting implications for financial markets?

The first chapter analyzes how investor asymmetric attitude towards risk can explain the term structure of variance risk in the aggregate stock market. Recent empirical evidence indicates that it was costless to hedge long-term volatility over the last two decades, whereas investors paid large premia for insurance against the unexpected realized variance (Dew-Becker et al. 2017). Yet leading asset pricing models fail to explain this stylized fact by overpricing news to future volatility (Du 2011; Drechsler and Yaron 2011; Wachter 2013). This chapter provides an explanation of the variance term structure.

I account for the empirical term structure of variance swap prices and returns in a consumption-based exchange economy with generalized disappointment aversion risk preferences of Routledge and Zin (2010) and rare events in the spirit of Rietz (1988) and Barro (2006). I model consumption growth via a hidden Markov chain with two regimes: an "expansion" state and a rare "depression" state. The negative news to consumption growth implies that the probability of being in the expansion state partially falls, and so does the equity price. The combination of more pessimistic beliefs and low consumption growth raises the agent's marginal utility. Crucially, GDA preferences penalize *disappointing* belief revisions that correspond to continuation utilities below a scaled certainty equivalent. I show that this asymmetric risk attitude towards downside consumption shocks generates a sizeable crash risk in the short term and mean reversion in variance swap prices in the long term. This helps explain the observed term structure of Sharpe ratios on variance claims, which is steep and negative for maturities shorter than three months and becomes flat and slightly positive from the three-month to 12-month horizons.

The second chapter examines how rational learning about parameters governing the short-term and long-run perspectives of economic growth amplifies the impact of macroeconomic shocks on equity prices and quantities in a standard production economy. The common assumption of prior learning models is that economic agents learn about unknown parameters over time but treat their current beliefs as true parameter values when computing asset prices (Weitzman 2007; Cogley and Sargent 2008; Pastor and Veronesi 2009). In a joint work with Dr Roman Kozhan, we depart from the extant literature by exploring the implications of priced parameter uncertainty, which incorporates future revisions of parameter beliefs into the decision-making process. The results indicate that rational learning about unknown parameters together with recursive preferences give rise to subjective long-lasting macroeconomic risks. These risks are priced under investor preference for early resolution of uncertainty and hence they help reproduce salient features of equity returns and comovements of macroeconomic variables. Time variation in beliefs leads to fluctuations in the equity risk premium and generates long-term predictability of excess returns consistent with the data. The asset pricing implications of subjective long-run risks crucially depend on the introduction of a procyclical dividend process in a production economy. We provide an extension of the standard model with investment frictions to account for this feature.

The third chapter is concerned with the analysis of learning about time-varying macroeconomic volatility and its effects on index option prices. A large strand of the literature emphasizes the importance of fluctuating macroeconomic uncertainty (see, for example, Justiniano and Primiceri (2008), Bloom (2009), Fernandez-Villaverde et al. (2011), Born and Pfeifer (2014), Christiano, Motto, and Rostagno (2014), Gilchrist, Sim, and Zakrajsek (2014), Liu and Miao (2014) and more recent studies by Leduc and Liu (2016), Basu and Bundick (2017), Bloom et al. (2018)). This chapter contributes to the existing literature by investigating the links between derivative prices and investor rational learning about volatility risk. It proposes a simple extension of the production economy of the second chapter of my dissertation to learning about regime-switching volatility to illustrate how rationally accounting for structural uncertainty can explain large premiums embedded in index option prices.

In the model, the investor faces uncertainty about the persistence of business cycles, mean growth of the economy, and time-varying productivity volatility. The key mechanism of the framework is as follows. First, in the presence of parameter uncertainty, learning generates time-variation in posterior estimates of unknown parameters creating an additional channel by which shocks to productivity growth introduce extra fluctuations in the investor's marginal utility. Second, rational pricing of beliefs amplifies the impact of parameter uncertainty on the stochastic discount factor, conditional moments of returns, and asset prices. The agent is concerned about future revisions, especially those in response to negative news to technology growth, and hence he is willing to pay a large premium for insurance against pessimistic updates. The deep out-of-the-money put options on the aggregate stock market index provide such insurance and hence bear high parameter uncertainty premiums. We show that this mechanism generates a steep implied volatility skew, which closely replicates the shape observed in the data. Furthermore, the conditional volatility of equity return variance is amplified, thus, raising the investor's concerns about the high realized variance in stock returns. In order to hedge his concerns, the agent is willing to pay large prices for variance swaps, which would provide a high payoff in states of high return volatility. In contrast, anticipated utility pricing, which ignores parameter uncertainty in decision-making, and full knowledge cannot reproduce the size of risk premiums. Quantitatively speaking, the models with anticipated utility or full information generate an average variance premium close to zero and flat implied volatility curves approximately equal to the annualized stock market volatility.

Chapter 1

Generalized Disappointment Aversion and the Variance Term Structure

Mykola Babiak[1]

Abstract

Contrary to leading asset pricing theories, recent empirical evidence indicates that financial markets compensate short-term equity volatility risk. An equilibrium model with generalized disappointment aversion risk preferences and rare events reconciles the term structure of variance swap prices and returns, consistent with the data. In addition, a calibration explains the variance and skew risk premiums in equity returns and the implied volatility skew of index options while capturing salient moments of fundamentals, equity returns, and the risk-free rate. The key intuition for the results stems from endogenous variation in the probability of disappointing events in consumption growth.

[1] I am especially indebted to Roman Kozhan for constructive suggestions that greatly improved the paper. I appreciate helpful comments from Daniele Bianchi, Jaroslav Borovicka, Byeongju Jeong, Michael Johannes, Keneth L. Judd, Marek Kapicka, Michal Kejak, Michal Pakoš, David Schreindorfer, Veronika Selezneva, Ctirad Slavik, Sergey Slobodyan, Stijn Van Nieuwerburgh, Ansgar Walther, conference participants at the 2016 EEA-ESEM Meeting, the 2016 Meeting of the Society for Computational Economics, the 2016 Zurich Initiative for Computational Economics, the 2016 Annual Conference of the Swiss Society for Financial Market Research, the 2018 RES Annual Conference, the 2018 RES Symposium of Junior Researchers, the 2018 Spanish Economic Association Meeting, and seminar participants at Columbia Business School, Warwick Business School, Lancaster University Management School, Collegio Carlo Alberto, Durham University Business School, the University of Gothenburg, and the University of Groningen. The financial support from the Charles University Grant Agency (GAUK No. 151016) and the Czech Science Foundation project No. P402/12/G097 (DYME Dynamic Models in Economics) is gratefully acknowledged.

1.1 Introduction

The consumption-based asset pricing literature has recently been revived by generalized models of long-run risks (Bansal and Yaron 2004) and rare disasters (Rietz 1988; Barro 2006) to capture many characteristics of the equity and derivatives markets. Nevertheless, leading theories fail to explain the timing of volatility risk (Dew-Becker et al. 2017).[1] The most successful asset pricing models commonly employ Epstein and Zin (1989) preferences, which, coupled with non-standard shocks to fundamentals, can generate sizeable equity risk premiums. In standard calibrations, investors are assumed to have a preference for early resolution of uncertainty and therefore they price long-term equity volatility strongly. Contrary to these predictions of well-known structural models, Dew-Becker et al. (2017) show that financial markets compensate short-term volatility risk.[2] The goal of this paper is to reconcile the observed variance term structure, while capturing salient properties of the risk-free rate, equity and equity index option prices.

In this paper, I account for the empirical term structure of variance swap prices and returns in a consumption-based exchange economy with generalized disappointment aversion (GDA) risk preferences of Routledge and Zin (2010) and rare events in the spirit of Rietz (1988) and Barro (2006). I model consumption growth via a hidden Markov chain with two regimes: an "expansion" state and a rare "depression" state. The negative news to consumption growth implies that the probability of being in the expansion state partially falls, and so does the equity price. The combination of more pessimistic beliefs and low consumption growth raises the agent's marginal utility. Crucially, GDA preferences penalize *disappointing* belief revisions that correspond to continuation utilities below a scaled certainty equivalent. I show that this asymmetric risk attitude towards downside consumption shocks generates a sizeable crash risk in the short term and mean reversion in variance swap prices in the long term. This helps explain the observed term structure of Sharpe ratios on variance claims, which is steep and negative for maturities shorter than three months and becomes flat and slightly positive from the three-month

[1] Also see van Binsbergen, Brandt, and Koijen (2012) and van Binsbergen et al. (2013) who document a downward sloping term structure of equity risk premia and equity return volatility, which is at odds in leading asset pricing models.

[2] Analyzing the portfolios across 19 different markets, Dew-Becker, Giglio, and Kelly (2019) also conclude that, over the last three decades, it was highly costly for investors to hedge realized volatility but not forward-looking uncertainty. Not only was it costless to hedge news about future variance but Berger, Dew-Becker, and Giglio (2019) provide new empirical evidence that shocks to future uncertainty have no significant effect on the economy. Furthermore, Dew-Becker and Giglio (2019) construct a novel measure of cross-sectional uncertainty and find that investors also do not view shocks to cross-sectional uncertainty as bad.

to 12-month horizons.

I show the importance of generalized disappointment aversion by comparing the model with GDA preferences to calibrations with alternative preference specifications such as a disappointment aversion utility function (Gul 1991) and Epstein-Zin preferences (Epstein and Zin 1989). First, a disappointment-averse agent similarly puts more weight on disappointing utility outcomes, defined as being below the certainty equivalent. Compared to Routledge and Zin (2010)'s definition of disappointment, Gul (1991)'s preferences increase the disappointment threshold, and hence overweight more outcomes. I show that this generates an upward sloping term structure of variance swap prices and negative Sharpe ratios on variance claims for all horizons, which is inconsistent with the empirical evidence. The reason is that, with a too high threshold, disappointment aversion magnifies the effect of small consumption shocks on the pricing kernel, increasing an insurance premium against low realized variance. Second, in a model with Epstein-Zin preferences, the average forward variance prices also remain markedly increasing, thereby producing negative average returns on variance claims. The intuition is that, under standard calibrations of Epstein-Zin preferences, expectations about long-term fluctuations in cash flows are the underlying drivers of asset prices. This implies a high insurance premium against shocks to future volatility.

Delving further into the origins of such results, I look at conditional dynamics of the term structures over various business cycle conditions. I define normal times as periods when the investor holds a median belief. Assuming the model initially stays in normal times, I study the impact of one positive and three negative consumption growth innovations. In the upside scenario (good times), consumption growth is a 1.0 standard deviation above mean growth in an expansion. In the three downside scenarios (bad times), consumption growth turns negative and is 1.5, 2.5 and 3.0 standard deviations, respectively, below mean growth in an expansion.

Several results are noteworthy. First, at the one-month maturity, the average Sharpe ratios in the GDA economy are pro-cyclical (less negative in good times and more negative in bad times), consistent with the empirical evidence (Ait-Sahalia, Karaman, and Mancini 2019). The reason is that GDA preferences generate a beliefs-dependent pricing kernel with higher marginal utility in bad times (and especially disappointing states) of the economy, increasing an insurance premium against high realized variance associated with low-utility states.[3] Furthermore, at longer maturities, the GDA model predicts that the

[3] Routledge and Zin (2010) and Bonomo et al. (2011) provide a similar analysis of the stochastic

Sharpe ratios remain insignificantly different from zero when the economy is hit by small shocks, whereas they become upward sloping and positive in response to large negative news. Hence, it is large consumption drops in an expansion state that, coupled with generalized disappointment aversion, lead to the inversion in variance swap prices and positive returns on variance claims at longer horizons. The reason is that, in the GDA model, small shocks are not priced, due to a low disappointment threshold. In contrast, lower-tail shocks lead to disappointment and high realized variance. After extreme jumps in realized variance, the economy is more likely to experience higher growth rates in an expansion state and the investor expects lower volatility in later periods. Hence, future volatility following those brief volatility spikes is priced less in the model.

Second, in the disappointment aversion model, the variance term structure is the same across normal and good times, being negative and flat at all horizons, and it experiences a similar parallel shift in the three bad scenarios. Intuitively, the positive news are fairly uninformative as they only diminish the already low likelihood of consumption depressions. A piece of even small bad news, however, has a disproportionately large impact on marginal utility due to high disappointment aversion to downside shocks, which leads to overpricing volatility risks at all horizons. Third, in the Epstein-Zin model, the impact of different shocks on the variance term structure is determined by the sign of consumption innovations and is proportional to their magnitude. However, for all economic conditions, Epstein-Zin preferences price shocks to future expected volatility strongly. Therefore, Sharpe ratios are on average negative, being larger in absolute terms in bad times than in good and normal times.

The magnitude of average Sharpe ratios on one-month variance claims indicates a large variance risk premium in the aggregate stock market.[4] Using the data over the last two decades, I further document the existence of a skew risk premium. I measure it as a ratio between the physical and option-implied expectations of equity return skewness over a monthly horizon. One can interpret this quantity as the return on a skew swap, a contract paying the realized skewness of stock returns. In addition to these novel measures of equity risk premiums, the literature has been long concerned with the puzzling implied volatility skew. I examine whether the empirical evidence concerning the moment risk premiums and option prices can be understood through the lens of different preference

discount factor of GDA preferences in the settings with different consumption processes. Also, the beliefs-dependent effective risk aversion of my paper echoes the mechanism of Berrada, Detemple, and Rindisbacher (2018) with learning and a beliefs-dependent utility function.

[4] Also, see Choi, Mueller, and Vedolin (2017) for the variance risk premium in fixed income markets.

specifications.

I find that the model with GDA preferences can reasonably capture the size of both risk premiums, whereas the Epstein-Zin framework generates about half of the average premiums implied by the GDA model. The disappointment-averse specification performs even worse, generating the smallest size and volatility of the variance premium and incorrectly predicting a positive skew premium. Furthermore, I show that generalized disappointment aversion helps generate a steep volatility skew implied by one-month index options and replicate the term structure of implied volatilities. In contrast, risk aversion alone produces too low implied volatilities, and the disappointment aversion framework predicts basically flat volatility curves that are approximately equal to the annualized stock market volatility. Mechanically, disappointment and risk aversion stand between the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ probability measures through the Radon-Nikodym derivative. The risk aversion smoothly distorts the $\mathbb{Q}$-density towards the left tail through the pricing kernel, while disappointment aversion overweights the outcomes strictly below some reference point. Since GDA preferences enable control of the disappointment threshold, they become instrumental in generating a fatter left tail of a $\mathbb{Q}$-measure compared to smooth preferences. In contrast, disappointment aversion preferences penalize too many outcomes due to a high disappointment threshold, and hence cannot generate a negatively skewed risk-neutral distribution of returns.

In a thorough comparative analysis, I show that my results are robust to different calibrations of key parameter values. Specifically, I demonstrate that one cannot match the data in the model with Gul preferences by recalibrating a disappointment aversion parameter in the range $[0.45, 0.75]$, or in the model with Epstein-Zin preferences by changing a coefficient of relative risk aversion in the range $[4.5, 7.5]$. Following Pohl, Schmedders, and Wilms (2018), I check that global projection methods provide highly accurate solutions by generating very small numerical errors.

Related literature. This paper is related to at least three main strands of the literature. First, it contributes to the new and growing literature on the term structures of equity and variance claims (van Binsbergen, Brandt, and Koijen 2012; van Binsbergen et al. 2013; Dew-Becker et al. 2017). The researchers propose a different rationale for the observed downward sloping term structure of equity risk premia and equity return volatility such as labour frictions (Favilukis and Lin 2015; Márfe 2017), financial leverage (Belo, Collin-Dufresne, and R. 2015), disaster recoveries (Hasler and Márfe 2016), and learning

(Croce, Lettau, and Ludvigson 2014; Ai et al. 2018; Hasler, Khapko, and Márfe 2019).[5] I complement the findings of these papers by explaining the term structure of variance swap prices and returns and emphasizing the crucial role of generalized disappointment aversion in the pricing of variance risk.

Second, this study builds on the large literature exploring the asset pricing implications of the asymmetric preferences of Routledge and Zin (2010). Bonomo et al. (2011, Bonomo et al. (2015) construct a long-run risk model with GDA preferences to explain equity return moments and the risk-return trade-offs. Liu and Miao (2014) focus on production-based implications of GDA preferences. Augustin and Tedongap (2016) shed light on their role in explaining sovereign credit spreads. Dahlquist, Farago, and Tédongap (2016) study the portfolio choice of an investor with GDA preferences. Farago and Tédongap (2018) use generalized disappointment aversion to explain the cross-section of expected returns, whereas Delikouras (2017) and Delikouras and Kostakis (2019) similarly study cross-sectional implications of disappointment aversion. This paper is also related to Schreindorfer (2018), who shows that US consumption and dividend growth rates are more correlated in bad times than in good times. The author introduces this feature into a model with GDA preferences to explain the equity premium and some features of index options. None of the aforementioned papers considers investor learning, which my paper treats as a central driver of asset prices, and none examines the variance term structure. My paper is, to my knowledge, the first to reconcile the term structure of variance swaps. Furthermore, it does so while jointly capturing salient properties of equity returns, variance and skew risk premiums, and option prices. Finally, the extant literature mainly studies GDA preferences in the setting with long-run risks, while my paper focuses on the role of GDA preferences in a rare event model with learning.

Third, this paper is also related to leading asset pricing theories advocating that habit formation, rare disasters, and long-run risks in consumption provide explanations for equity returns and option prices. In the context of habits, Du (2011) shows that an extension of the model with habit formation (Campbell and Cochrane 1999) to include rare disasters can explain the observed implied volatility skew. Under the rare disasters umbrella, the implied volatility surface can be explained with extensions to model uncertainty about rare events (Liu, Pan, and Wang 2005), rare jumps in persistence (Benzoni, Collin-Dufresne, and Goldstein 2011), or stochastic probability of disasters (Seo and

[5] See van Binsbergen and Koijen (2017) for a comprehensive overview of the literature on the term structure of equity claims.

Wachter 2018). The long-run risks literature further introduces jump risks (Eraker and Shaliastovich 2008; Shaliastovich 2015) to rationalize option prices. The long-run risk models with transient non-Gaussian shocks to fundamentals (Bollerslev, Tauchen, and Zhou 2009; Drechsler and Yaron 2011) and multiple volatility risks (Zhou and Zhu 2014) prove to be successful in explaining the variance premium. Drechsler (2013) constructs the long-run risks framework with model uncertainty to explain both the variance premium and the implied volatility skew. The mechanism of my paper is distinct from the existing literature, since it points out the importance of the investor's generalized disappointment aversion for asset prices. Also, unlike other models, the approach in this work additionally explains the variance term structure and the skew risk premium.

The remainder of the paper is organized as follows. Section 1.2 reports the empirical evidence. Section 1.3 describes the economy. Section 1.4 derives asset prices in the model. Section 1.5 provides asset pricing results of the models with GDA, disappointment aversion and Epstein-Zin preferences. Section 2.5 concludes. Sections A.1, A.2, and A.3 of Appendix present the data, solve the representative agent's maximization problem, and discuss the application and accuracy of numerical methods.

1.2 Variance and Skewness Risk

This section describes several methods to quantify risk premia embedded in the variance and skewness of equity returns. First, following the discussion of Dew-Becker et al. (2017), I describe salient features of the term structure of variance claims in the equity index market and discuss the failure of leading asset pricing models to account for variance forward prices and returns. Second, I define and measure the one-month variance and skew risk premiums in aggregate stock returns. Third, I construct the volatility surface by extrapolating historical volatilities implied by equity index options.

Recent empirical studies focus on the term structure of dividend and variance claims. In particular, Dew-Becker et al. (2017) discover new prominent facts about the price of variance risk, which are at odds in well-known asset pricing theories. Their analysis is based on the pricing of volatility-linked assets, primarily variance swaps, and it yields two main results. First, they show that, over the period from 1996 to 2014, news about future volatility at horizons ranging from one quarter to 14 years is unpriced. Second, risk exposure to unexpected realized variance is significantly priced in the data. This leads to the conclusion that it was almost costless to hedge future variance over the last

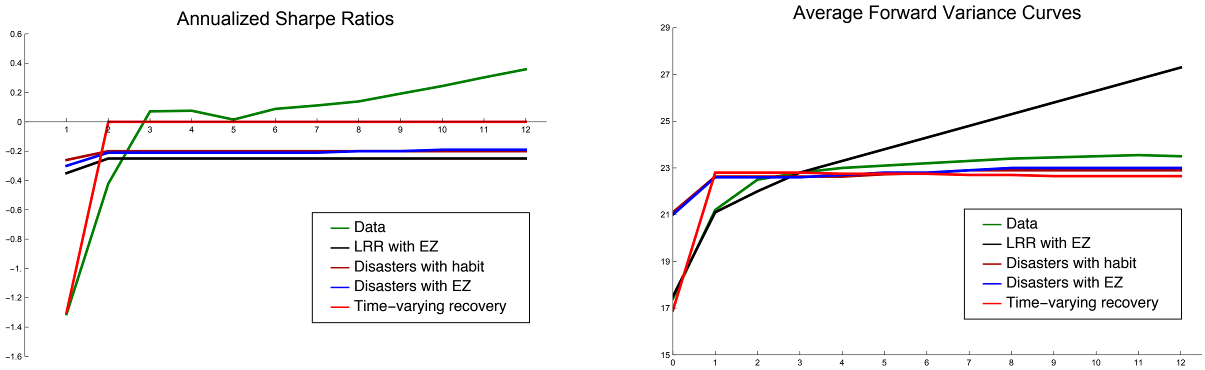

Figure 1.1: Average prices and annualized Sharpe ratios for forward variance claims. The left and right panels plot annualized Sharpe ratios and average prices for forward variance claims for the data and different models. The prices are reported in annualized volatility terms, $100 \times \sqrt{12 \times F_t^n}$. The empirical lines correspond to the US data from 1996 to 2013.

two decades, whereas investors paid a substantial premium for protection against extreme realized volatility.

Figure 1.1 is a reproduction of Figure 10 in Dew-Becker et al. (2017) and illustrates their main results. The left graph of Figure 1.1 compares annualized Sharpe ratios for forward variance claims in the data and in different models: a long-run risk model of Drechsler and Yaron (2011) with Epstein-Zin preferences, a disaster risk model of Du (2011) with habit formation, a time-varying disaster risk model of Wachter (2013) with Epstein-Zin preferences, and a rare disaster model of Gabaix (2012) with time-varying recovery rates. The right plot of Figure 1.1 compares empirical variance claim prices with those predicted by theoretical frameworks. As shown in the figure, the empirical Sharpe ratios are significantly negative for short maturities, especially for one-month variance forwards, whereas they become slightly positive for horizons from three to 12 months. In turn, the term structure of variance forwards is on average upward sloping in the data and significantly flattens with the horizon. The figure also shows that the long-run risk model and rare disaster frameworks with recursive preferences or habit formation fail to capture these prominent facts. Most notably, the three models generate almost flat Sharpe ratios on variance forwards, which are respectively far too small and too large for short and long maturities when compared to the data. A model with disasters and time-varying recovery rates does a better job of capturing the empirical patterns, though it cannot fully explain the upward trend in Sharpe ratios for longer periods.

Closely related to the variance term structure is the risk premium in the second and third moments of returns. A large strand of the literature has focused on the variance premium, while the skew premium has received little attention, especially from the theoretical research. This paper aims to explain both phenomena simultaneously. The variance premium can be defined as the difference between expectations of stock market return variance under the risk-neutral $\mathbb{Q}$ and actual physical $\mathbb{P}$ probability measures for a given horizon.[6] Formally, a τ-month variance premium at time t is

$$vp_t = \mathbb{E}_t^{\mathbb{Q}}\big[\text{Return Variation}(t, t+\tau)\big] - \mathbb{E}_t^{\mathbb{P}}\big[\text{Return Variation}(t, t+\tau)\big],$$

in which the total return variation is calculated over the period t to $t+\tau$. The quantity vp_t corresponds to the expected profits of a variance swap, which pays the equity's realized variance over the term of the contract. Britten-Jones and Neuberger (2000) and Carr and Wu (2009) show that this payoff can be replicated by a portfolio of European options. Like the variance swap, Kozhan, Neuberger, and Schneider (2013) consider a skew swap with a payoff equal to the equity's realized skewness. Bakshi, Kapadia, and Madan (2003) show that a skew contract can be replicated by a trading portfolio involving long OTM calls and short OTM puts. I follow Kozhan, Neuberger, and Schneider (2013) and define a τ-month skew risk premium at time t as

$$sp_t = \frac{\mathbb{E}_t^{\mathbb{P}}\big[\text{Return Skewness}(t, t+\tau)\big]}{\mathbb{E}_t^{\mathbb{Q}}\big[\text{Return Skewness}(t, t+\tau)\big]} - 1,$$

in which the total return skewness is calculated from t to $t+\tau$. In this paper, I focus on the one-month variance and skew risk premiums consistent with the literature. For the empirical analysis of the variance premium, I use the VIX index, S&P 500 index futures, and the S&P 500 index from the Chicago Board of Options Exchange (CBOE). The options data used to construct the skew premium is from OptionMetrics. The data sets employed in the analysis of the variance and skew measures cover the periods January 1990 to December 2016 and January 1996 to December 2016, respectively. I provide a description of the empirical strategy in Appendix.

Table 1.1 shows summary statistics for variance and skew risk premiums. A positive variance premium and a negative skew premium are consistent with the literature.[7] Since the prices of variance and skew swaps are on average greater than their corresponding payoffs, the average profits from writing these contracts are interpreted as insurance

[6] Consistent with the definitions in Bollerslev, Tauchen, and Zhou (2009), Bollerslev, Gibson, and Zhou (2011), and Drechsler and Yaron (2011).

[7] See Bakshi, Kapadia, and Madan (2003), Bollerslev, Tauchen, and Zhou (2009), and Kozhan, Neuberger, and Schneider (2013), among others.

Table 1.1
Summary statistics: variance and skew risk premiums.

	vp_t	sp_t
Mean	10.24	−42.12
Median	7.50	−68.11
SD	10.49	82.11
Max	83.70	447.37
Skewness	2.62	3.57
Kurtosis	14.15	16.26

This table reports monthly descriptive statistics for the conditional variance vp_t and skew sp_t premiums. Mean, Median, SD, Max, Skewness, and Kurtosis report the sample average, median, standard deviation, maximum, skewness, and kurtosis, respectively. The empirical statistics of the variance and skew risk premiums are for the US data from January 1990 to December 2016 and from January 1996 to December 2016, respectively.

premiums associated with higher moments of equity returns. Table 1.1 also shows that both premiums have large volatility, positive skewness, and a kurtosis coefficient much larger than three. The latter two characteristics indicate fat tails in the distributions of quantities.

The risk-neutral expectations of the second and third moments of equity returns are related to the level and the slope of the implied volatility surface, which remains a challenge for equilibrium asset pricing models. I construct the empirical implied volatility surface by performing a polynomial extrapolation of volatilities in the maturity time and strike prices. I use the option data from OptionMetrics for January 1996 to December 2016. I present the empirical methodology in Appendix.

The left plot in Figure 1.2 shows the implied volatility curve for 1-month maturity as a function of moneyness (a ratio of strike to spot price). The implied volatilities are downward sloping in moneyness and decline from 28% to slightly above 20% for a range of moneyness from 0.90 to 1.05. This shape is known in the literature as the implied volatility skew. The right plot in Figure 1.2 provides the implied volatility curve for ATM and 0.90 OTM put options as functions of maturity. The graph suggests that ATM volatilities increase slightly in the horizon and are around 22% for for 1, 3, and 6 month maturities, while OTM volatilities decline slightly in the horizon. Furthermore, the plot confirms that OTM volatilities are strictly higher than ATM volatilities for all times to expiration. Note that the implied volatilities are significantly above the annualized stock market volatility. Hence, it is difficult to rationalize the level and slope of the implied volatility curves given the historical stock market volatility.

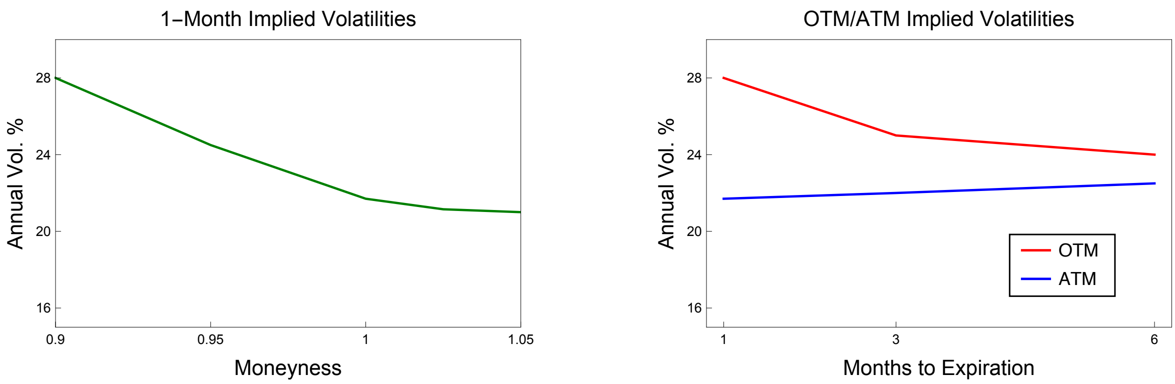

Figure 1.2: Implied volatilities. The left panel plots the empirical 1-month implied volatility curve as a function of moneyness. The right panel plots the empirical implied volatility curves for ATM and OTM options as functions of the time to maturity (in months). All curves are for the US data from January 1996 to December 2016.

1.3 Model Setup

This section presents the economy. In particular, it provides details of the agent's preferences and cash-flow processes for consumption and dividends.

1.3.1 Generalized Disappointment Aversion Risk Preferences

The environment is an infinite-horizon, discrete-time exchange economy with a representative agent receiving utility from a consumption stream. Following the recursive utility framework of Epstein and Zin (1989), the agent's utility V_t in period t is defined recursively as

$$V_t = \left[(1-\beta)C_t^{\rho} + \beta\mu_t^{\rho}\right]^{1/\rho}, \quad 0 < \beta < 1, \quad \rho \leq 1, \tag{1.1}$$

in which C_t is agent's consumption, β is the subjective discount factor, $1/(1-\rho)$ is the intertemporal elasticity of substitution (IES), and $\mu_t = \mu_t(V_{t+1})$ is the certainty equivalent of random future utility V_{t+1}.

The certainty equivalent captures the generalized disappointment aversion (GDA) risk attitude as defined by Routledge and Zin (2010). GDA preferences allocate more weight on the "disappointing" events compared to the expected utility, similarly to disappointment aversion risk preferences of Gul (1991). For Gul's disappointment aversion model, however, an outcome is viewed as disappointing when it is below the certainty equivalent, whereas for Routledge and Zin's generalized disappointment aversion specification a disappointing outcome is below a constant fraction of the implicit certainty equivalent.

Formally, the certainty equivalent of GDA preferences is defined as

$$\frac{[\mu_t(V_{t+1})]^\alpha}{\alpha} = \mathbb{E}_t\left[\frac{V_{t+1}^\alpha}{\alpha}\right] - \theta\mathbb{E}_t\left[\mathbb{I}\left(\frac{V_{t+1}}{\mu_t(V_{t+1})} \leqslant \delta\right)\left(\frac{[\delta\mu_t(V_{t+1})]^\alpha}{\alpha} - \frac{V_{t+1}^\alpha}{\alpha}\right)\right] \tag{1.2}$$

or equivalently

$$\mu_t(V_{t+1}) = \left(\mathbb{E}_t\left[V_{t+1}^\alpha \cdot \frac{1+\theta\mathbb{I}(V_{t+1} \leqslant \delta\mu_t(V_{t+1}))}{1+\theta\delta^\alpha\mathbb{E}_t\Big[\mathbb{I}(V_{t+1} \leqslant \delta\mu_t(V_{t+1}))\Big]}\right]\right)^{1/\alpha},$$

in which $\mathbb{I}(\cdot)$ denotes the indicator function, $1-\alpha > 0$ is the relative risk aversion, $\delta \leq 1$ and $\theta \geq 0$ represent a disappointment threshold and a disappointment aversion parameter, respectively. GDA preferences enable one control for a disappointment threshold by changing δ. Routledge and Zin (2010) preferences defined by (1.1) and (1.2) nest two preference specifications. The expected utility of Epstein and Zin (1989) can be obtained by setting $\theta = 0$. Assuming $\theta \neq 0$ and $\delta = 1$, GDA preferences reduce to the disappointment aversion utility of Gul (1991).

1.3.2 Endowments and Inference Problem

A popular paradigm in the asset pricing literature is the application of a regime-switching framework for modeling aggregate consumption growth.[8] I follow this tradition and subject log consumption growth to hidden shifts in the growth rate:

$$\Delta c_{t+1} = \mu_{s_{t+1}} + \sigma\varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0,1).$$

The consumption volatility σ is constant, whereas the mean growth rate $\mu_{s_{t+1}}$ is driven by a hidden two-state Markov chain s_{t+1} with a state space $\mathcal{S} = \{1,\ 2\}$ and a transition matrix

$$\mathcal{P} = \begin{pmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{pmatrix},$$

[8] Since Mehra and Prescott (1985) and Hamilton (1989), researchers have used these models to embed business cycle fluctuations in the mean and volatility of consumption growth (Cecchetti, Lam, and Mark 1990; Veronesi 1999; Ju and Miao 2012; Johannes, Lochstoer, and Mou 2016; Collin-Dufresne, Johannes, and Lochstoer 2016). By changing the number of states and parameters controlling the persistence and conditional distribution of regimes, these models can also embed the "peso problem" in the mean (Rietz 1988; Barro 2006; Backus, Chernov, and Martin 2011; Gabaix 2012) or persistence (Gillman, Kejak, and Pakos 2015) of consumption growth. Additionally, a particular calibration of a regime-switching model can also generate long-run risks (Bonomo et al. 2011; Bonomo et al. 2015) or economic recoveries (Hasler and Márfe 2016) in consumption and dividends.

in which $0 < \pi_{11} < 1$ and $0 < \pi_{22} < 1$ are transition probabilities. I further assume $\mu_2 < \mu_1$ to identify $s_{t+1} = 1$ and $s_{t+1} = 2$ as expansion and recession, respectively.

The reason for calibrating the model with two regimes is twofold. First, I want to maintain parsimony for the sake of convenient interpretation of results. Second, I do not introduce additional risks in consumption growth to isolate the impact of learning and GDA preferences. Of course, a model with time-varying expected growth (Bansal and Yaron 2004), more regimes (Bonomo et al. 2011; Bonomo et al. 2015), economic recoveries (Hasler and Márfe 2016), or a multidimensional learning problem (Johannes, Lochstoer, and Mou 2016) could enrich conditional dynamics and improve the performance of the model. However, I show that the model with Bayesian learning about a latent state of the economy and GDA preferences can successfully reproduce the variance term structure along with a wide array of asset pricing phenomena observed in the equity and derivatives markets.

I seek to price the equity as a levered consumption claim with monthly log dividend growth defined as follows:

$$\Delta d_{t+1} = g_d + \lambda \Delta c_{t+1} + \sigma_d e_{t+1}, \quad e_{t+1} \sim N(0,1), \tag{1.3}$$

in which λ is a leverage ratio on expected consumption growth. I use g_d to equalize long-run dividend and consumption growth rates, and σ_d to match the empirical dividend growth volatility. In addition, the chosen value of λ allows me to match the observed correlation between annual consumption and dividend growth rates.

The investor knows the true parameters and distribution of shocks in the model but does not observe the state s_{t+1} of the economy. Consequently, he updates a posterior belief about the hidden state s_{t+1}, conditional on the observable history of consumption and dividend growth rates at time t :

$$\mathcal{F}_t = \Big\{ (\Delta c_\tau, \Delta d_\tau) : 0 \leq \tau \leq t \Big\}.$$

The inference problem is to derive the evolution of $\pi_t = \mathbb{P}(s_{t+1} = 1 | \mathcal{F}_t)$ given the initial belief π_0 (the stationary prior). In this paper, I consider a Bayesian agent who updates his belief through Bayes' rule:

$$\pi_{t+1} = \frac{\pi_{11} f(\Delta c_{t+1}|1)\pi_t + (1-\pi_{22}) f(\Delta c_{t+1}|2)(1-\pi_t)}{f(\Delta c_{t+1}|1)\pi_t + f(\Delta c_{t+1}|2)(1-\pi_t)}, \tag{1.4}$$

in which

$$f(\Delta c_{t+1}|i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\Delta c_{t+1}-\mu_i)^2}{2\sigma^2}}, \quad i = 1, 2.$$

1.4 Asset Prices

I now characterize equilibrium conditions, discuss the impact of generalized disappointment aversion on the stochastic discount factor, and outline a sketch of a numerical solution. Then, I describe equilibrium asset prices in the economy.

1.4.1 Equilibrium and Pricing Kernel

Following Routledge and Zin (2010), I show (see Appendix A.2) that the gross return $R_{i,t+1}$ on the i-th traded asset satisfies the condition

$$\mathbb{E}_t \left[M_{t+1} R_{i,t+1} \right] = 1, \tag{1.5}$$

in which M_{t+1} is the pricing kernel of the economy defined as

$$M_{t+1} = \underbrace{\beta \left(\frac{C_{t+1}}{C_t} \right)^{\rho - 1}}_{M_{t+1}^{CRRA}} \cdot \underbrace{\left(\frac{V_{t+1}}{\mu_t(V_{t+1})} \right)^{\alpha - \rho}}_{M_{t+1}^{EZ}} \cdot \underbrace{\left(\frac{1 + \theta \mathbb{I}(V_{t+1} \leqslant \delta \mu_t(V_{t+1}))}{1 + \theta \delta^{\alpha} \mathbb{E}_t \Big[\mathbb{I}(V_{t+1} \leqslant \delta \mu_t(V_{t+1})) \Big]} \right)}_{M_{t+1}^{GDA}}. \tag{1.6}$$

There are different components in the pricing kernel. The first part M_{t+1}^{CRRA} is the stochastic discount factor of the time-separable power utility. The second multiplier M_{t+1}^{EZ} is the adjustment of Epstein-Zin preferences, which allow a separation between the coefficient of risk aversion and elasticity of intertemporal substitution. The third component M_{t+1}^{GDA} represents the generalized disappointment aversion adjustment. When the agent's utility is below a predefined fraction of the certainty equivalent, more weight is attached to the pricing kernel, magnifying the countercyclical dynamics of the pricing kernel. For a better understanding of the key role of generalized disappointment aversion, I consider the calibration of preference parameters where $\alpha = \rho$. Hence, the pricing kernel simplifies to

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{\rho - 1} \cdot \left(\frac{1 + \theta \mathbb{I}(V_{t+1} \leqslant \delta \mu_t(V_{t+1}))}{1 + \theta \delta^{\alpha} \mathbb{E}_t \Big[\mathbb{I}(V_{t+1} \leqslant \delta \mu_t(V_{t+1})) \Big]} \right).$$

1.4.2 Model Solution

Recently, Pohl, Schmedders, and Wilms (2018) show that the latest asset pricing models with long-run risks generate significant nonlinearities, which, coupled with the log-linearization of equilibrium quantities, can generate economically significant numerical errors. Hence, I solve the model numerically using global solution methods to accurately capture the nonlinear nature of the model under consideration.

I first need to solve for the return on the wealth portfolio R^{ω}_{t+1} (the return on the aggregate consumption claim) and then the equity return $R_{e,t+1}$ (the return on the aggregate dividend claim), which are implicitly defined by equation (1.5). Denoting the investor's wealth and equity price by W_t and P^e_t, the returns on the wealth portfolio and equity can be rewritten as

$$R^{\omega}_{t+1} = \frac{W_{t+1}}{W_t - C_t} = \frac{\frac{W_{t+1}}{C_{t+1}}}{\frac{W_t}{C_t} - 1} \cdot e^{\Delta c_{t+1}} \quad \wedge \quad R^e_{t+1} = \frac{P^e_{t+1} + D_{t+1}}{P^e_t} = \frac{\frac{P^e_{t+1}}{D_{t+1}} + 1}{\frac{P^e_t}{D_t}} \cdot e^{\Delta d_{t+1}}.$$

I conjecture that the wealth-consumption ratio $\frac{W_t}{C_t} = G(\pi_t)$ and the price-dividend ratio $\frac{P^e_t}{D_t} = H(\pi_t)$ are functions of the state belief π_t. I substitute R^{ω}_{t+1} and R^e_{t+1} into (1.5) and apply the projection method (Judd 1992) to approximate $G(\pi_t)$ and $H(\pi_t)$ by a basis of complete Chebyshev polynomials. The numerical solution and its accuracy in the asset pricing models of this paper are discussed in details in Appendix A.3.

Having solved for wealth-consumption and price-dividend ratios, I can simulate asset pricing moments associated with the risk-free rate, equity returns, and the price-dividend ratio. Further, I can numerically calculate variance swap prices and returns, quantities in the variance and skew risk premiums, and option prices.

1.4.3 Prices and Returns of Variance Swaps

Consider an n-month variance swap, a claim to realized variance over months $t+1$ to $t+n$. Given the discrete nature of the model, total variance of the return is equal to the sum of conditional variances RV_{t+i} in each subperiod. Following Dew-Becker et al. (2017), the price of an n-month variance swap is

$$VS^n_t = \mathbb{E}^Q_t \left[\sum_{i=1}^{n} RV_{t+i} \right].$$

In turn, the price of a zero coupon forward claim on realized variance is

$$F_t^n = \mathbb{E}_t^Q \left[RV_{t+n} \right].$$

Thus, F_t^n is equal to the risk-neutral expectation of return variance during the n-th month from the current period. F_t^0 is naturally defined as the realized variance in the current period. Next, I define the return on the n-month variance forward as a return on the trading strategy in which investors buy the n-month forward at time t and sell it in the next period as a forward claim with maturity $n-1$. The proceeds from selling the forward are then used to purchase a new n-month variance at price F_{t+1}^n. Formally, the excess return of an n-period variance forward is

$$R_{t+1}^n = \frac{F_{t+1}^{n-1} - F_t^n}{F_t^n}. \tag{1.7}$$

1.4.4 Variance and Skew Risk Premiums

The focus of this paper is on the monthly variance and skew risk premiums associated with equity returns. Since I calibrate the economy at the monthly frequency, the t-time monthly variance premium vp_t is defined as the difference between risk-neutral and physical expectations of the total return variance between t and $t+1$. As in Drechsler and Yaron (2011), the variance premium equals

$$vp_t = \mathbb{E}_t^{\mathbb{Q}}(var_{t+1}^{\mathbb{Q}}(r_{e,t+2})) - \mathbb{E}_t^{\mathbb{P}}(var_{t+1}^{\mathbb{P}}(r_{e,t+2})), \tag{1.8}$$

in which $var_{t+1}^{\mathbb{Q}}(r_{e,t+2})$ and $var_{t+1}^{\mathbb{P}}(r_{e,t+2})$ are $(t+1)$-period conditional variances of the log return $r_{e,t+2} = \ln(R_{e,t+2})$ under the risk-neutral $\mathbb{Q}$ and physical $\mathbb{P}$ probability measures, respectively. The t-time monthly skew premium is defined as a return on a skew swap, a contract paying the realized skew of the return between time t and $t+1$. As in Kozhan, Neuberger, and Schneider (2013), the skew premium equals

$$sk_t = \frac{\mathbb{E}_t^{\mathbb{P}}(skew_{t+1}^{\mathbb{P}}(r_{e,t+2}))}{\mathbb{E}_t^{\mathbb{Q}}(skew_{t+1}^{\mathbb{Q}}(r_{e,t+2}))} - 1,$$

in which $skew_{t+1}^{\mathbb{Q}}(r_{e,t+2})$ and $skew_{t+1}^{\mathbb{P}}(r_{e,t+2})$ are $(t+1)$-period conditional skewness of the log return $r_{e,t+2} = \ln(R_{e,t+2})$ under the risk-neutral $\mathbb{Q}$ and physical $\mathbb{P}$ probability measures, respectively.

1.4.5 Option Prices and Implied Volatilities

I now describe how I compute model-based option prices and solve for their Black-Scholes implied volatilities. Consider a European put option written on the price of the equity that is traded in the economy. Note that the equity price should not include dividend payments; that is, options are written on the ex-dividend stock price index. Using the Euler condition (1.5), the relative price $\mathcal{O}_t(\pi_t, \tau, K) = \frac{P_t^o(\pi_t,\tau,K)}{P_t^e(\pi_t)}$ of the τ-period European put option with the strike price K, expressed as a ratio to the initial price of the equity P_t^e, should satisfy

$$\mathcal{O}_t(\pi_t, \tau, K) = \mathbb{E}_t\left[\prod_{k=1}^{\tau} M_{t+k} \cdot \max\left(K - \frac{P_{t+\tau}^e}{P_t^e}, 0\right)\right]. \tag{1.9}$$

It is worth noting that a put price P_t^o depends on the equity price P_t^e, whereas the normalized price $\mathcal{O}_t$ does not. One can express the ratio $\frac{P_{t+\tau}^e}{P_t^e}$ in terms of dividend growth rates and price-dividend ratios on the equity and hence the state belief π_t provides sufficient information for the calculation of the option prices. Specifically, I compute model-based European put prices $\mathcal{O}_t = \mathcal{O}_t(\pi_t, \tau, K)$ via Monte Carlo simulations. I convert them into Black-Scholes implied volatilities with a properly annualized continuous interest rate $r_t = r_t(\pi_t)$ and dividend yield $q_t = q_t(\pi_t)$. Thus, given the time to maturity τ, the strike price K, the risk-free rate r_t, and dividend yield q_t, the implied volatility $\sigma_t = \sigma_t^{\text{BS}}(\pi_t, \tau, K)$ solves the equation:

$$\mathcal{O}_t = e^{-r_t\tau} \cdot K \cdot N(-d_2) - e^{-q_t\tau} \cdot N(-d_1), \tag{1.10}$$

in which

$$d_{1,2} = \left[-\ln(K) + \tau\left(r_t - q_t \pm \sigma_t^2/2\right)\right] / \left[\sigma_t\sqrt{\tau}\right].$$

1.5 Calibration and Quantitative Results

In this section, I first calibrate the cash-flow processes for consumption and dividends consistent with the historical US data from January 1930 to December 2016. To better understand the role of generalized disappointment aversion in the consumption-based asset pricing economy of this paper, I consider three specifications of preference parameters: a model with generalized disappointment aversion preferences (GDA), an economy with disappointment aversion preferences (DA), and a framework with Epstein-Zin preferences (EZ). The comparison of GDA and DA isolates the contribution of disappointment aversion, while the comparison of GDA and EZ illustrates the impact of the representative agent's preference for early resolution of uncertainty. Having solved the model numerically, I generate 10,000 simulations of each calibration and compare model-based statistics of cash flows and asset prices with corresponding empirical counterparts. Consistent with the historical data, the model-generated moments of returns and cash flows are based on the simulations with depressions, while the model-based statistics of variance forwards, higher-moment risk premiums, and option prices correspond to the simulations without depressions. The key results are robust to the inclusion of rare events, which are excluded to eliminate the impact of large consumption declines and to highlight the importance of learning and generalized disappointment aversion.

1.5.1 Calibrated Parameters

The top panel in Table 2.1 provides the parameter values of cash-flow processes for consumption and dividends. I begin with the parameters of a regime-switching process for aggregate consumption growth. As in Bansal and Yaron (2004), I make the model's time-averaged consumption statistics consistent with observed annual log consumption growth from 1930 to 2016. Following Collin-Dufresne, Johannes, and Lochstoer (2016), I calibrate a parsimonious model of monthly consumption growth with the recession state mimicking a consumption decline in the US during the Great Depression. Specifically, I set $\pi_{11} = 1151/1152$ and $\pi_{22} = 47/48$. These numbers imply an average duration of the high-growth state of about $(1-\pi_{11})^{-1} = 96$ years and the depression state of about $(1-\pi_{22})^{-1} = 4$ years. The unconditional probability of being in expansion $\overline{\pi}_{11} = (1-\pi_{22})/(2-\pi_{11}-\pi_{22})$ results in $\overline{\pi}_{11} = 0.96$ and hence the economy experiences one four-year depression per century, consistent with the historical data. For the mean growth rate, consumption tends to grow on average at the annual rates of about $\mu_1 \times 12 = 2.06\%$ and $\mu_2 \times 12 = -4.6\%$ in the expansion and depression states, respectively. The annualized consumption drop in the depression state is equal to an average annual decline in the real, per capita log consumption growth during the Great Depression and it is less severe than rare disasters, defined as a drop in annual consumption growth larger than ten percent (Rietz 1988; Barro 2006). I calibrate σ to match the empirical volatility of consumption growth.

Table 1.2
Parameter values.

Parameter	Description	Value		
π_{11}	Transition probability from expansion to expansion	1151/1152		
π_{22}	Transition probability from recession to recession	47/48		
$\mu_1 \times 12$	Consumption growth in expansion	2.06		
$\mu_2 \times 12$	Consumption growth in recession	-4.6		
$g_d \times 12$	Mean adjustment of dividend growth	-2.87		
$\sigma \times \sqrt{12}$	Std. deviation of consumption growth shock	2.6		
$\sigma_d \times \sqrt{12}$	Std. deviation of dividend growth shock	11.41		
λ	Leverage ratio	2.6		
		GDA	DA	EZ
β^{12}	Discount factor	0.99	0.99	0.99
$1-\alpha$	Risk aversion	1/1.5	1/1.5	6.0
$1/(1-\rho)$	EIS	1.5	1.5	1.5
θ	Disappointment aversion	8.41	0.6	0
δ	Disappointment threshold	0.930	1	

This table reports parameter values in the cash-flow processes and the three models: GDA, DA, and EZ.

My strategy to calibrate a rare bad state to the US Great Depression experience is identical to Collin-Dufresne, Johannes, and Lochstoer (2016), who study rational parameter learning in a model with rare events. In the context of the US history, Nakamura et al. (2013) identify two disaster episodes (1914-1922 and 1929-1933) during the twentieth century. Since the Great Depression is the only example of a consumption disaster in the US for the period considered in my paper, I naturally calibrate the recession state to this specific observation. Furthermore, Nakamura et al. (2013) note that rare disasters tend to unfold over multiple years. Thus, instead of assuming extreme instantaneous consumption disasters, I choose the milder depression state with an average duration corresponding to four years of the Great Depression, consistent with the empirical evidence.

I now turn to calibrating parameters in the dividend process. I regress the annual dividends on the annual consumption covering the period 1930-2016 and find the leverage ratio is around 2.5, a conservative number within an interval of plausible values from 1.5 to 4.5. The leverage ratio is an important parameter for two reasons. First, it controls the volatility of dividends in normal times. Second, it determines the decline of dividends in the depression state. Consequently, a larger leverage parameter would increase the payoff of put options, conditional on the depression realization. To compare my results to prior studies, particularly the disaster literature, I set the leverage ratio $\lambda = 2.6$, the value used in Seo and Wachter (2018). I further follow the literature and set g_d to equalize the long-run dividend and consumption growth. The standard deviation of the dividend process σ_d is used to generate large annual dividend volatility observed in the data.

The bottom panel in Table 2.1 summarizes the values of GDA, DA, and EZ specifications. I keep the subjective discount factor $\beta^{12} = 0.99$ and the EIS $1/(1-\rho) = 1.5$ the same for all preference specifications. For the GDA model, the coefficient of relative risk aversion is $1 - \alpha = 1/1.5$. This eliminates one degree of freedom caused by extra parameters in GDA preferences. I jointly choose the disappointment aversion parameter $\theta = 8.41$ and the disappointment threshold $\delta = 0.930$ to generate the high equity premium. The degree of disappointment aversion is consistent with the empirical literature, which reports a range of values from 3.29 to 8.41 (Delikouras 2017). Note that the variance term structure, the variance and skew premiums, and the implied volatility surface are not directly targeted in the model calibration.

For the DA model, I shut down the generalized disappointment aversion channel by setting $\delta = 1$. This inevitably generates larger risk aversion in good times due to an increased number of disappointing events, significantly distorting equity moments in the DA model. Thus, I decrease the disappointment aversion parameter $\theta = 0.6$ to match the observed equity premium. The remaining parameters are fixed at the initial values. For the EZ model, I turn off all (generalized) disappointment aversion by setting $\theta = 0$. The model operates only through the risk aversion channel in the recursive preferences of Epstein and Zin (1989) with the coefficient of relative risk aversion of $1 - \alpha = 6$. In this case, the representative agent exhibits a preference for early resolution of uncertainty, a popular workhorse in the asset pricing literature. Other parameters correspond to those in the GDA specification.

1.5.2 Endowments and Equity Returns

Before discussing the asset pricing implications of GDA, DA, and EZ economies, I look at the cash-flow dynamics predicted by a two-state regime-switching process. Panel A in Table 1.3 compares the annualized consumption and dividends moments of the data with those implied by the calibration in this paper. The model-based medians of the mean and volatility of consumption and dividend growth come out close to their empirical counterparts, although mean dividend growth is slightly higher in the simulations. The autocorrelation of cash flows is also in line with the empirical estimates. The leverage parameter captures the observed correlation between consumption and dividends. Overall, one can see that a cash-flow model of consumption and dividend growth matches the key empirical statistics well.

Table 1.3
Cash flows and stock market returns.

	Data	GDA			DA			EZ		
		5%	50%	95%	5%	50%	95%	5%	50%	95%
Panel A: Cash flows										
$E(\Delta c)$	1.83	0.91	1.85	2.40	0.91	1.85	2.40	0.91	1.85	2.40
$\sigma(\Delta c)$	2.22	1.90	2.28	3.19	1.90	2.28	3.19	1.90	2.28	3.19
$ac1(\Delta c)$	0.50	0.09	0.30	0.62	0.09	0.30	0.62	0.09	0.30	0.62
$E(\Delta d)$	1.44	−1.10	1.91	4.44	−1.10	1.91	4.44	−1.10	1.91	4.44
$\sigma(\Delta d)$	11.04	9.51	11.05	12.97	9.51	11.05	12.97	9.51	11.05	12.97
$ac1(\Delta d)$	0.19	0.09	0.27	0.46	0.09	0.27	0.46	0.09	0.27	0.46
$corr(\Delta c, \Delta d)$	0.55	0.38	0.55	0.71	0.38	0.55	0.71	0.38	0.55	0.71
Panel B: Returns										
$E(r_f)$	0.81	−0.13	0.86	1.49	0.68	1.14	1.20	0.22	1.03	1.41
$\sigma(r_f)$	1.87	1.48	2.52	3.51	0.04	0.25	1.22	0.73	1.50	2.34
$E(r_e - r_f)$	5.22	3.67	6.10	8.35	3.43	6.04	8.47	3.50	5.89	8.19
$\sigma(r_e - r_f)$	19.77	15.58	19.22	23.11	13.03	16.02	20.34	14.64	18.69	23.49
$E(pd)$	3.11	2.96	3.03	3.05	2.90	2.97	2.98	2.95	3.04	3.06
$\sigma(pd)$	0.33	0.04	0.08	0.18	0.01	0.05	0.18	0.03	0.08	0.22

Panel A reports moments of consumption and dividend growth denoted by Δc and Δd. Panel B reports moments of the log risk-free rate r_f, the excess log equity returns $r_e - r_f$, and the log price-dividend ratio pd. The entries are annualized statistics. The empirical moments are for the US data from January 1930 to December 2016. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report percentiles of sample statistics based on these artificial series. I use the common notations for the sample mean E, standard deviation σ, autocorrelation $ac1$, and correlation $corr$.

Panel B in Table 1.3 reports the key annualized moments of the risk-free rate, equity returns, and the price-dividend ratio for the three specifications. All three models do a good job of accounting for the salient features of equity returns, as all predict the low risk-free rate, the large equity premium and volatility of excess returns. Also, the volatility of the risk-free rate and the level of the log price-dividend ratio correspond well to the empirical estimates under all specifications. The main shortcoming of the three models is too low volatility of the log price-dividend ratio.

1.5.3 The Price of Variance Risk

Figure 1.3 compares the empirical and model-based term structures of Sharpe ratios and prices for forward variance claims. These graphs assess how well different preferences can explain the patterns in the data. The left plot in Figure 1.3 shows that the GDA model does a good job of matching the overall shape of annualized Sharpe ratios. In particular, it generates a curve that is very steep for the one-month returns and then has a small and positive slope for the longer horizons. The figure also shows that both EZ and DA specifications fail to reconcile the concave and upward shape of the term

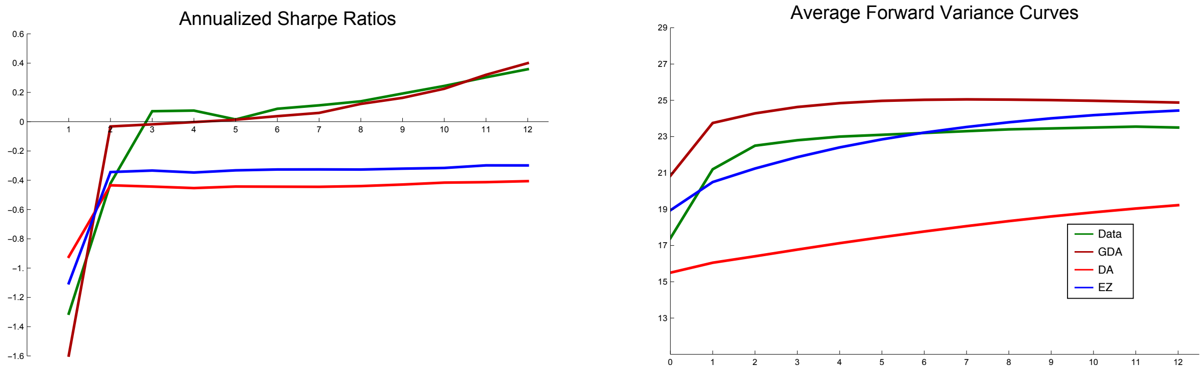

Figure 1.3: Sharpe ratios and forward variance claim prices. The left and right panels plot annualized Sharpe ratios and average prices for forward variance claims for the data and the three models: GDA, DA, and EZ. The prices are reported in annualized volatility terms, $100 \times \sqrt{12 \times F_t^n}$. The empirical lines correspond to the US data from 1996 to 2013. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report medians of sample statistics based on these artificial series.

structure. Consistent with the findings of Dew-Becker et al. (2017), the calibration with Epstein-Zin preferences underprices volatility risk in the short term and overprices future variance in the long term. The results for the DA model show that disappointment aversion generates even higher Sharpe ratios for the longer maturities, while the one-month forwards are underpriced compared to the data as well as to the GDA and EZ models. The right panel in Figure 1.3 plots the average variance swap prices of different maturities in the data and the three models. The empirical curve has an upward and concave shape for the horizons from one to 12 months and it flattens significantly at the longer end. In contrast to the empirical evidence, the DA and EZ specifications predict strongly upward sloping term structures of variance forwards. Although the GDA model generates slightly higher prices of variance swaps, it successfully captures the concave shape and especially the flatness of the curve at longer maturities.

The interpretation of our results in the EZ economy is similar to the intuition provided by Dew-Becker et al. (2017). The risk-averse investor considers the states with high expected future volatility as periods of low lifetime utility. With Epstein-Zin preferences, low utility increases the pricing kernel and, thus, the agent with a strong preference for early resolution of uncertainty requires a large compensation for future consumption volatility. The economic intuition for the results in the DA model is similar. In this case, even though the coefficient of relative risk aversion is very low, the investor is still extremely averse to expected future volatility due to high disappointment aversion. Because even small negative news about consumption growth leads to an investor's disappointment and high volatility, he is willing to pay higher prices for forward variance swaps

Table 1.4
Model tests using annualized Sharpe ratios for forward variance claims.

	p-value		
	GDA	DA	EZ
Simulated 1mo/SR $\leq$ empirical SR	0.91	0.10	0.26
Simulated 3mo/SR $\geq$ empirical SR	0.37	< 0.01	0.03
Simulated 12mo/SR $\geq$ empirical SR	0.51	< 0.01	< 0.01
Joint test: 1mo/SR $\leq$ data $\wedge$ 3mo/SR $\geq$ data $\wedge$ 12mo/SR $\geq$ data	0.32	< 0.01	< 0.01

For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to the length of the variance swap data. In each simulation, I calculate average annualized Sharpe ratios for forward variance claims with one-, three-, and 12-month maturities. For each model, the first row shows the fractions of samples in which the simulated Sharpe ratios are at least as small as the empirical one-month estimates. The second and third rows present the fraction of samples in which the simulated Sharpe ratios are at least as large as the empirical three-month and 12-month estimates, respectively. The entries of the bottom row are the fraction of samples in which all three conditions are satisfied simultaneously.

with longer maturities, which hedge his concerns about future disappointing events.

The generalized disappointment-averse investor also has an asymmetric risk attitude towards downside risk in consumption growth, but unlike the disappointment-averse agent, he places more weight on outcomes that are sufficiently deep in the left tail of the consumption distribution. Therefore, the generalized disappointment-averse agent perceives unexpected realized volatility in the short-term more risky than uncertainty in the distant future. This generates the term structure of forward variance prices that is very steep for maturities less than three months but remains flat for longer horizons. One can also generate higher volatility of the pricing kernel by controlling the disappointment threshold. Consequently, a more volatile stochastic discount factor in the GDA model leads to higher Sharpe ratios for the one-month forward variance swaps, relative to the DA and EZ economies.

Table 1.4 augments the results in Figure 1.3 by reporting the p-values of annualized Sharpe ratios with respect to their finite-sample distribution in all three specifications. For each model, it shows the fraction of samples across 10,000 artificial simulations of the corresponding economy that satisfy one of the four conditions. For the first three conditions, simulated average Sharpe ratios for one-, three-, and 12-month horizons should be respectively smaller, larger, and larger than the empirical estimates. One can interpret these fractions as the p-values for a one-sided test of the model generating as negative or as positive average Sharpe ratios for a particular maturity as in the data. For the last condition, simulated statistics should jointly satisfy the first three requirements. This number corresponds to the p-value for a test of the model matching the observed upward sloping shape of the term structure.

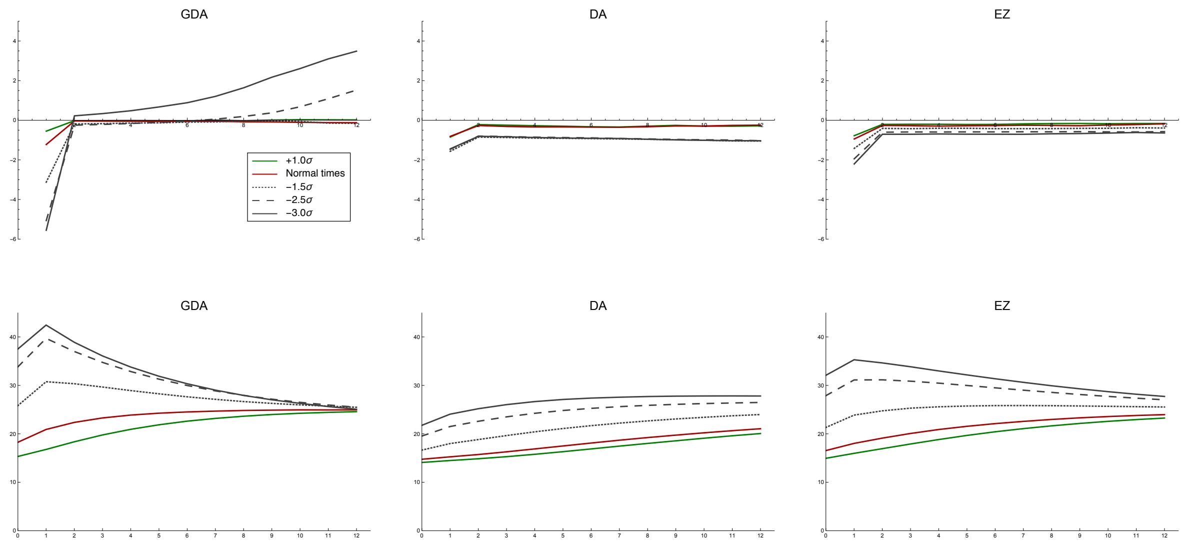

Figure 1.4: Sharpe ratios and forward variance claim prices versus consumption growth shocks. The figure plots annualized Sharpe ratios (top panels) and average prices (bottom panels) for forward variance claims for the three models: GDA, DA, and EZ. For each model, the respective panel shows the term structures in good, normal, and bad times. The economy is initially in normal times, corresponding to a median posterior belief. In good (bad) times denoted by "$+1.0\sigma$" ("-1.5σ", "-2.5σ", and "-3.0σ"), consumption growth is a 1.0 (1.5, 2.5, and 3.0) standard deviation(s) above (below) an average growth in an expansion state.

Table 1.4 shows that it is difficult to reject any of the three models based on the one-month variance swap returns only. Specifically, one would expect to see as small average one-month Sharpe ratios as observed empirically in 91%, 10%, and 26% of the time in the GDA, DA, and EZ specifications, respectively. At longer maturities, however, one can reject at the 5% level the null hypothesis that the DA or EZ frameworks generate the variance swap data. The GDA model instead generates large p-values for all tests and cannot be rejected. In particular, the models with disappointment aversion or Epstein-Zin preferences would predict as positive Sharpe ratios at longer maturities as in the data in fewer than 3% of simulations, while the likelihood of replicating the overall shape is less than 0.1%. This is in stark contrast to the GDA model, which captures negative Sharpe ratios at the short end and positive ones at the long end in 32% of the time.

As an additional exercise, Figure 1.4 provides impulse response analysis of the conditional term structure of Sharpe ratios and average prices for forward variance claims. I assume initially the economy is in normal times, when the investor holds a median belief. I study conditional dynamics of the term structures in the next period when consumption growth is a 1.0 standard deviation above and 1.5, 2.5, and 3.0 standard deviations below an average growth in an expansion state (called good and bad times, respectively). The top panels in Figure 1.4 show that the model-generated Sharpe ratios for the DA and

EZ models are negative for all economic conditions and hence the two preference specifications cannot capture slightly positive average Sharpe ratios at longer maturities. The economy with GDA preferences generates a procyclical and steep curve for short-term claims, consistent with Ait-Sahalia, Karaman, and Mancini (2019) who study variance swap prices in different economic conditions. The term structure in the GDA model is insignificant for maturities longer than two months in good and normal times as well as bad but not depression-like states, whereas it becomes steeper and positive in response to large consumption declines. The latter feature of the GDA model enables to match the data.

To better understand the differences in conditional Sharpe ratios predicted by different preferences, the bottom panels in Figure 1.4 plot impulse responses of variance swap prices. The average curve for the DA model remains upward sloping in good, normal, and bad times. This result explains negative average returns on a variance swap contract. For the EZ economy, the term structure of variance forwards switches from strongly increasing in normal and good times to slightly increasing in bad (but not severe) times, and it even becomes weakly downward sloping in very bad times. Nevertheless, this amplification of short-term prices is too weak to generate on average positive returns on holding a variance forward. In contrast, the generalized disappointment aversion channel in the GDA model inverts the term structure in all bad scenarios considered, and the degree of inversion is substantially stronger compared to the economy with Epstein-Zin preferences. This suggests that GDA preferences provide a key amplification mechanism for the pricing of short-term variance risk in bad times that enables to replicate the empirically observed patters.

1.5.4 Variance and Skew Risk Premiums

Panel A in Table 1.5 collects moments of the variance premium and conditional variances of the market return under the actual and risk-neutral probability measures in the data and the three models. It shows that the GDA model is able to generate a large and volatile variance premium. It is well-known that the variance premium distribution is fat-tailed as characterized by the positive skewness and excess kurtosis. The GDA model qualitatively respects the non-normality of the distribution, although the sample skewness and kurtosis statistics are smaller relative to the data. Generalized disappointment aversion allows me to successfully account for the first and second moments of the variance premium with empirically consistent conditional return variances under both probability measures. The model predicts that the total return variance is more volatile under the risk-neutral distribution relative to the physical distribution and that both volatilities are persistent, as they are in the data.

Empirical literature further documents return predictability by the variance premium. To study this predictive relation, I regress the one-, three-, and six-month cumulative excess log returns, which are expressed in percentages, on the lagged monthly variance premium. Consistent with the existing literature, the "Data" column of Panel B in Table 1.5 indicates a positive impact as measured by positive and slightly decreasing regression coefficients. Also, there is an increasing predictive power as measured by increasing R^2s over longer horizons. The GDA model replicates these empirical findings by matching the magnitude of coefficients and R^2 statistics.

Table 1.5 shows that the model with disappointment aversion preferences produces a mean and volatility of the variance premium that are more than five times smaller than with the generalized utility function. Turning off the generalized disappointment aversion channel also leads to a significant reduction in the volatility of return variance in the DA model. As the variance premium decreases, its predictive power for the excess log returns also suffers. This is manifested in the lower R^2s and empirically inconsistent regression coefficients. Next, I turn off any source of (generalized) disappointment aversion and consider a representative agent with Epstein-Zin preferences. As shown in Table 1.5, the EZ model leads to around a two-fold increase in the mean and volatility of the variance premium relative to the DA model, but sample statistics are less than half of the numbers in the GDA model. The table suggests that a smaller variance premium is due to the reduced volatility in conditional variances. As shown in Panel B in Table 1.5, a smaller variance premium results in the excessively-high regression coefficients and too small R^2s in the predictive regression of future equity returns on the variance premium.

The asset pricing implications of different preference specifications are further augmented by Table 1.6. The GDA specification produces a sizeable skew premium, which corresponds well to the historical value. The GDA model generates the positive skewness and excess kurtosis in the conditional distribution of the skew premium. Table 1.6 also demonstrates the first and second moments of the return skewness under risk-neutral and physical distributions. The conditional mean of the return skewness under both measures is significantly negative, although the model cannot fully capture the size observed in the data. The main drawback of the GDA model is lower volatility of the skew

Table 1.5
Variance premium and predictability.

	Data	GDA			DA			EZ		
		5%	50%	95%	5%	50%	95%	5%	50%	95%
Panel A: Variance premium										
$E(vp)$	10.27	8.13	12.32	17.14	1.34	2.11	3.39	3.15	4.92	7.23
$\sigma(vp)$	10.87	12.22	15.99	18.93	1.53	3.11	5.25	4.79	7.46	10.91
$skew(vp)$	2.33	0.91	1.49	2.25	0.49	2.76	4.17	−0.62	1.71	2.98
$kurt(vp)$	10.90	2.34	4.06	7.67	5.88	12.03	24.96	4.00	7.27	13.59
$\sigma(var_t^{\mathbb{P}}(r_e))$	29.32	17.34	25.44	32.68	5.02	14.25	36.48	13.00	25.50	40.47
$ac1(var_t^{\mathbb{P}}(r_e))$	0.79	0.70	0.81	0.88	0.61	0.79	0.92	0.66	0.82	0.91
$\sigma(var_t^{\mathbb{Q}}(r_e))$	33.76	29.58	40.60	49.11	6.55	17.24	38.91	17.79	31.46	45.00
$ac1(var_t^{\mathbb{Q}}(r_e))$	0.80	0.69	0.79	0.85	0.62	0.80	0.92	0.66	0.82	0.89
$skew(var_t^{\mathbb{Q}}(r_e))$	3.53	0.86	1.47	2.21	2.40	3.73	5.75	1.47	2.30	3.34
$kurt(var_t^{\mathbb{Q}}(r_e))$	21.47	2.30	4.13	7.78	8.46	19.47	45.54	3.87	8.24	16.18
Panel B: Predictability of excess returns										
$\beta(1m)$	0.76	0.19	0.75	1.38	−1.43	1.81	5.56	−0.95	0.93	2.77
$R^2(1m)$	2.70	0.15	2.39	6.99	0.02	1.09	4.19	0.01	1.31	6.01
$\beta(3m)$	0.83	0.18	0.63	1.09	−1.31	1.55	4.02	−0.73	0.87	2.11
$R^2(3m)$	8.61	0.47	5.57	15.24	0.02	2.39	8.67	0.03	3.28	12.87
$\beta(6m)$	0.57	0.15	0.50	0.82	−1.05	1.24	3.06	−0.63	0.74	1.60
$R^2(6m)$	7.55	0.68	7.78	20.79	0.08	3.32	13.14	0.04	4.67	16.95

Panel A reports moments of the conditional variance premium vp, market return variances $var_t^{\mathbb{P}}(r_e)$ and $var_t^{\mathbb{Q}}(r_e)$ under the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ probability measures, respectively. The Panel A entries are monthly statistics. Panel B reports results of the predictive regression of h-month future excess log equity returns constructed as $r^{ex}_{t+1\to t+h} = \sum_{i=1}^{h}\left(r_{e,t+i}-r_{f,t-1+i}\right)$ on the lagged variance premium vp_t. Specifically, the slope estimates $\beta(h)$ and $R^2(h)$ are based on the linear projection:

$$100 \times r^{ex}_{t+1\to t+h} = \text{Intercept} + \beta(h) \times vp_t + \varepsilon_{t+h}, \quad h = 1, 3, 6.$$

The moments and regression outputs are for the data and the three models: GDA, DA, and EZ. The empirical statistics are for the US data from January 1990 to December 2016. For each model I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report percentiles of sample statistics based on these artificial series. I use the common notations for the sample mean E, standard deviation σ, autocorrelation $ac1$, skewness $skew$, and kurtosis $kurt$.

premium, realized and implied skew. Since conditional dynamics of the model are driven by a single state, allowing the model to operate through multiple states (for example, via time-varying expected growth and volatility, jumps in consumption, etc.) would make the economy more flexible to jointly match all moments. In particular, it would be able to generate more time-variation in the higher moments of returns through different channels.

Table 1.6 also shows that disappointment aversion alone cannot reproduce the sign of the skew premium, which proves to be positive in the DA model. Furthermore, the bottom part of Table 1.6 shows that the DA model predicts the smallest first and second moments of return skewness across the three models. The last block in Table 1.6 displays the impact of relative risk aversion on the skew premium. The results for the

Table 1.6
Skew premium.

	Data	GDA			DA			EZ		
		5%	50%	95%	5%	50%	95%	5%	50%	95%
$E(sp)$	−42.20	−39.11	−34.58	−30.78	26.58	34.88	56.36	−22.79	−19.34	−12.84
$\sigma(sp)$	81.81	11.23	26.42	46.52	24.44	29.53	377.79	3.03	21.31	91.65
$skew(sp)$	3.57	−4.32	3.28	8.56	−3.37	1.24	13.98	−11.70	3.23	13.69
$kurt(sp)$	16.26	1.93	43.48	112.04	3.10	4.65	215.60	2.04	80.40	219.37
$ar1(sp)$	0.04	−0.12	0.15	0.62	−0.01	0.61	0.70	−0.27	0.11	0.58
$E(skew_t^{\mathbb{P}}(r_e))$	−87.52	−42.55	−39.99	−33.83	−20.49	−15.82	−12.34	−38.00	−33.98	−29.43
$\sigma(skew_t^{\mathbb{P}}(r_e))$	173.59	8.99	11.79	22.21	7.68	11.41	15.56	12.03	13.61	23.97
$E(skew_t^{\mathbb{Q}}(r_e))$	−177.73	−70.44	−64.13	−53.83	−17.39	−13.22	−9.82	−47.60	−42.42	−36.45
$\sigma(skew_t^{\mathbb{Q}}(r_e))$	92.33	23.20	28.27	41.96	7.91	11.14	15.54	16.44	18.68	28.90

This table reports moments of the conditional skew premium sp, market return skewness $skew_t^{\mathbb{P}}(r_e)$ and $skew_t^{\mathbb{Q}}(r_e)$ under the physical $\mathbb{P}$ and risk-neutral $\mathbb{Q}$ probability measures, respectively. The entries are monthly statistics. The moments are for the data and the three models: GDA, DA, and EZ. The empirical statistics are for the US data from January 1996 to January 2016. For each model I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report percentiles of sample statistics based on these artificial series. I use the common notations for the sample mean E, standard deviation σ, skewness $skew$, and kurtosis $kurt$.

EZ model show that the risk-neutral return density becomes more distorted towards the left tail; however, the model can generate less than half of the average skew premium observed in the data. Although the EZ framework predicts the correct size, it significantly understates the magnitude. Overall, these results indicate the important role of generalized disappointment aversion in generating a correct risk neutral distribution of equity returns.

1.5.5 The Term Structure of Implied Volatilities

I further examine the asset pricing implications of all models for equity index options. The top graph of Figure 1.5 compares the 1-month volatility curves for the data and the three models. The implied volatilities are expressed as a function of moneyness ranging from 0.90 to 1.05. The plot shows that the empirical implied volatilities decline in moneyness, a pattern also known in the literature as the implied volatility skew. The top panel of Figure 1.5 shows that the DA implied volatilities for the 1-month maturity are flat and approximately equal to the realized stock market volatility. One apparent candidate to generate a steep volatility skew is high risk aversion. Although raising risk aversion in Epstein-Zin preferences improves the model's performance, this cannot fully account for the level in implied volatilities. In contrast, the GDA framework can fit the option prices much better. The bottom plots of Figure 1.5 additionally present the term structure of

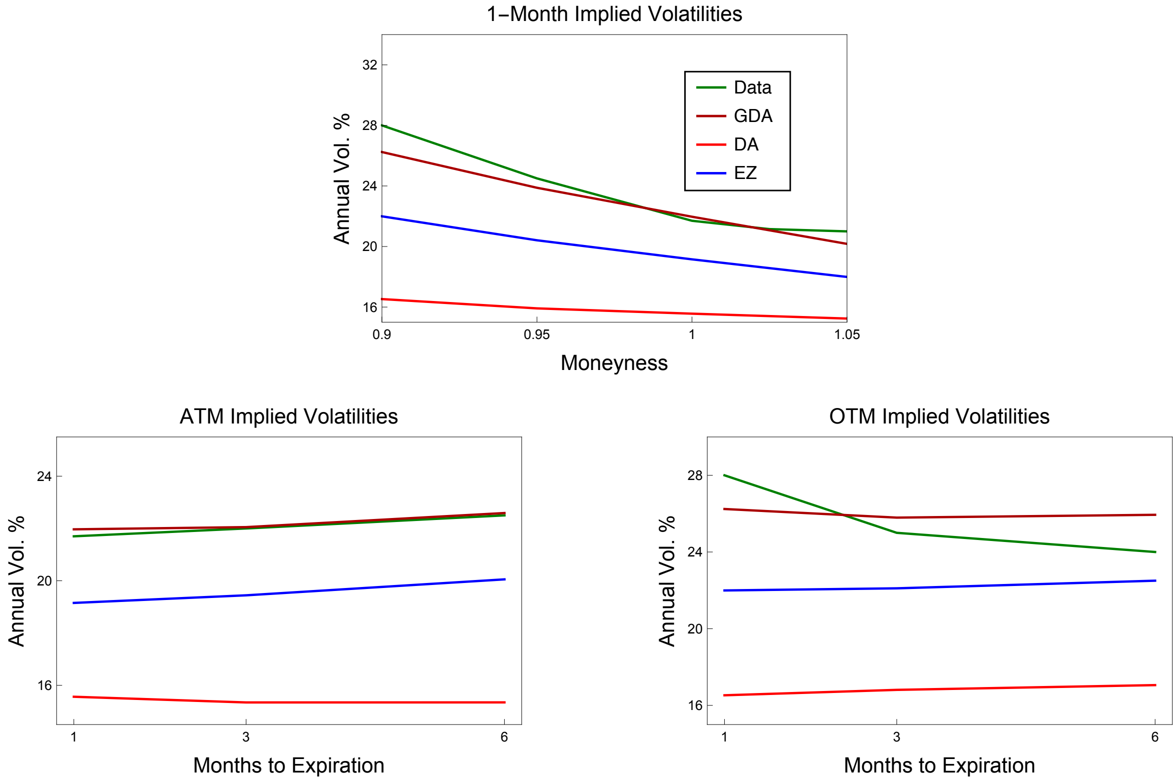

Figure 1.5: Implied volatilities. The top panel plots the 1-month implied volatility curve as a function of moneyness for the data and the three models: GDA, DA, and EZ. The bottom panels plot the empirical and model-based implied volatility curves for ATM (left) and OTM (right) options as functions of the time to maturity (in months). The empirical statistics are for the US data from January 1996 to December 2016. The model-based curves are calculated for option prices using the annualized model-implied interest rate $r_t(\pi_t)$ and dividend-yield $q_t(\pi_t)$ in each period. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report the medians fo sample statistics based on these artificial series.

implied volatilities for ATM and 0.90 OTM options. In the data, ATM volatilities slightly increase over the horizon, while a downward trend can be observed for OTM volatilities. The model-based results clearly indicate that neither DA nor EZ specification can match the level of the empirical curves. In contrast, the model with generalized disappointment aversion can explain overall patterns and magnitudes of the empirical implied volatilities for one-, three-, and six-month maturities well.

1.6 Sensitivity Analysis

In this section, I conduct an extensive sensitivity analysis to examine the robustness of the results to alternative calibrations of preference specifications. To address the concern that some of my results are driven by a particular choice of parameters, I consider the asset pricing implications of the three preference specifications for a large set of values of the key parameters.

1.6.1 The Variance Term Structure

Figure 1.6 depicts simulated Sharpe ratios and prices for variance forwards in various calibrations of the framework with GDA preferences. It shows how the term structures depend on the choices of preference parameters θ and δ. The shape of variance swap prices flattens and the term structure of Sharpe ratios becomes upward sloping with the higher disappointment threshold or disappointment aversion. The intuition for the result obtained is as follows. In very bad times, volatility risk is amplified more in the short-term than in the long-term in the GDA economy. As shown in Section 1.5.3, this generates the downward and upward sloping patterns in prices and Sharpe ratios, respectively. In normal and good times, variance swap prices are slightly increasing in the horizon; however, only short-term volatility shocks earn a significant premium (as measured by Sharpe ratios), generating relatively large and negative Sharpe ratios for one- and two-month periods, but insignificant ratios for longer horizons. Since higher disappointment risk reinforces the first effect, the models with the higher disappointment threshold and disappointment aversion generate on average negative Sharpe ratios for one- and two-month maturities but positive and upward sloping ones for maturities between three and 12 months.

Figure 1.7 examines the impact of disappointment aversion and risk aversion parameters on variance swaps in the models with Gul and Epstein-Zin preferences, respectively. In the DA specification, the slope of variance swap prices is increasing in θ. The reason for this result is that the disappointment-averse investor strongly dislikes both low and high variance. Hence, larger θ only increases already high insurance premia against shocks to realized and future volatility. In the EZ economy, the slope of variance swap prices is decreasing in $1 - \gamma$. As shown in Figure 1.7, in order to generate a close-to-zero slope at least after the ten-month maturity, the required risk aversion should be at least 7. For this value, however, the model would generate a Sharpe ratio of less than -2.0 for the one-month claim compared to -1.3 in the data. Furthermore, with this calibration of relative risk aversion, the mean equity premium has a median value of 8% in the EZ model, well above the empirical estimate of around 5%. Raising risk aversion even more would only worsen the fit of the model with the variance term structure at the one-month maturity and with the equity moments and higher-moment risk premiums (see Section

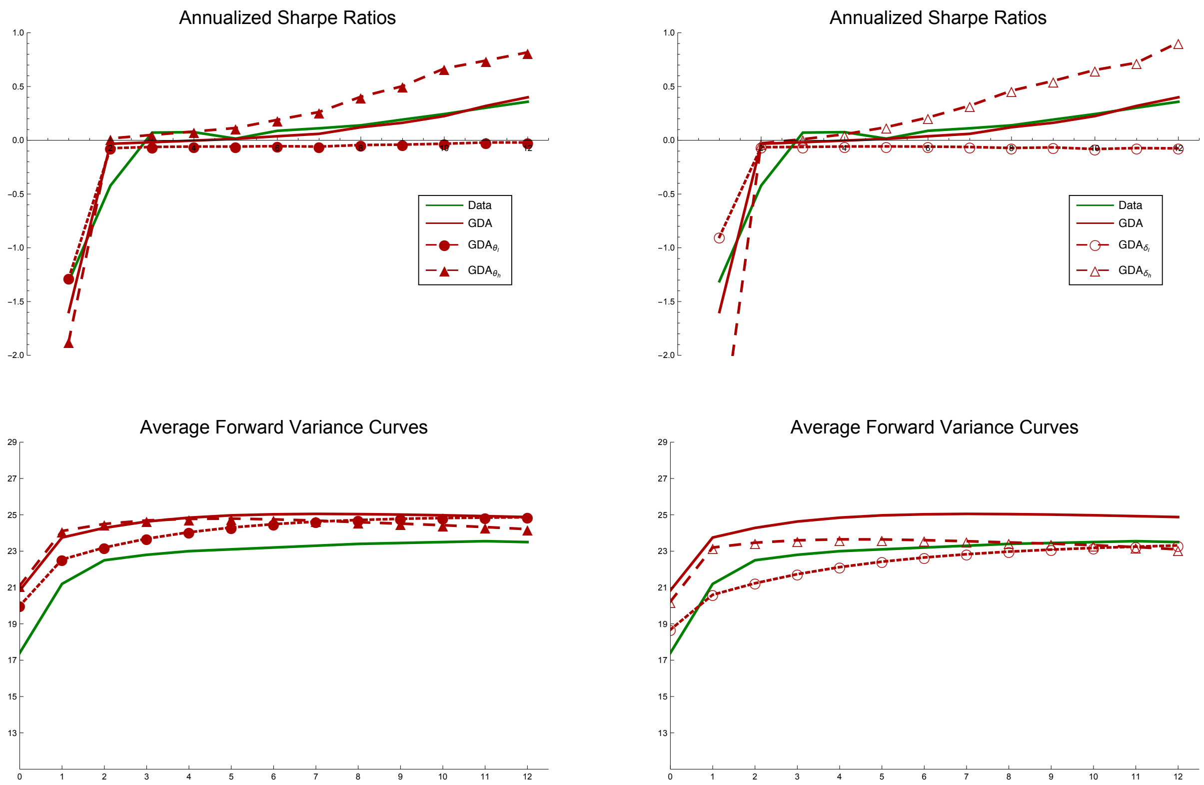

Figure 1.6: Sensitivity of Sharpe ratios and forward variance claim prices: GDA. The figure plots annualized Sharpe ratios (top) and average prices for forward variance claims (bottom) for different model calibrations with generalised disappointment aversion preferences. GDA corresponds to the original GDA model. In GDA_{θ_l} and GDA_{θ_h}, the disappointment aversion parameters are $\theta_l = 6.41$ and $\theta_h = 10.41$, respectively. In GDA_{δ_l} and GDA_{δ_h}, the disappointment threshold parameters are $\delta_l = 0.920$ and $\delta_h = 0.940$, respectively. If not stated otherwise, the remaining parameters in all specifications are set at the original values in the GDA model. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report medians of sample statistics based on these artificial series.

1.6.2 for more details). Hence, a high risk aversion in the EZ framework cannot explain the shape of the variance term structure.

1.6.2 Equity Returns and Moment Risk Premiums

Figure 1.8 provides sensitivity results for the risk-free rate, the equity premium, the price-dividend ratio and the moment risk premiums, for a broad range of parameter choices in the three models. In this sensitivity exercise, I consider three preference specifications and change a key parameter in each of them, while holding the remaining parameters at the values in the original calibration. In the GDA model, I vary the disappointment threshold between 0.915 and 0.945. In the DA model, I change the disappointment aversion parameter between 0.45 and 0.75. In the EZ model, the results are provided for the coefficient of relative risk aversion ranging from 4.5 to 7.5. The panels in Figure 1.8

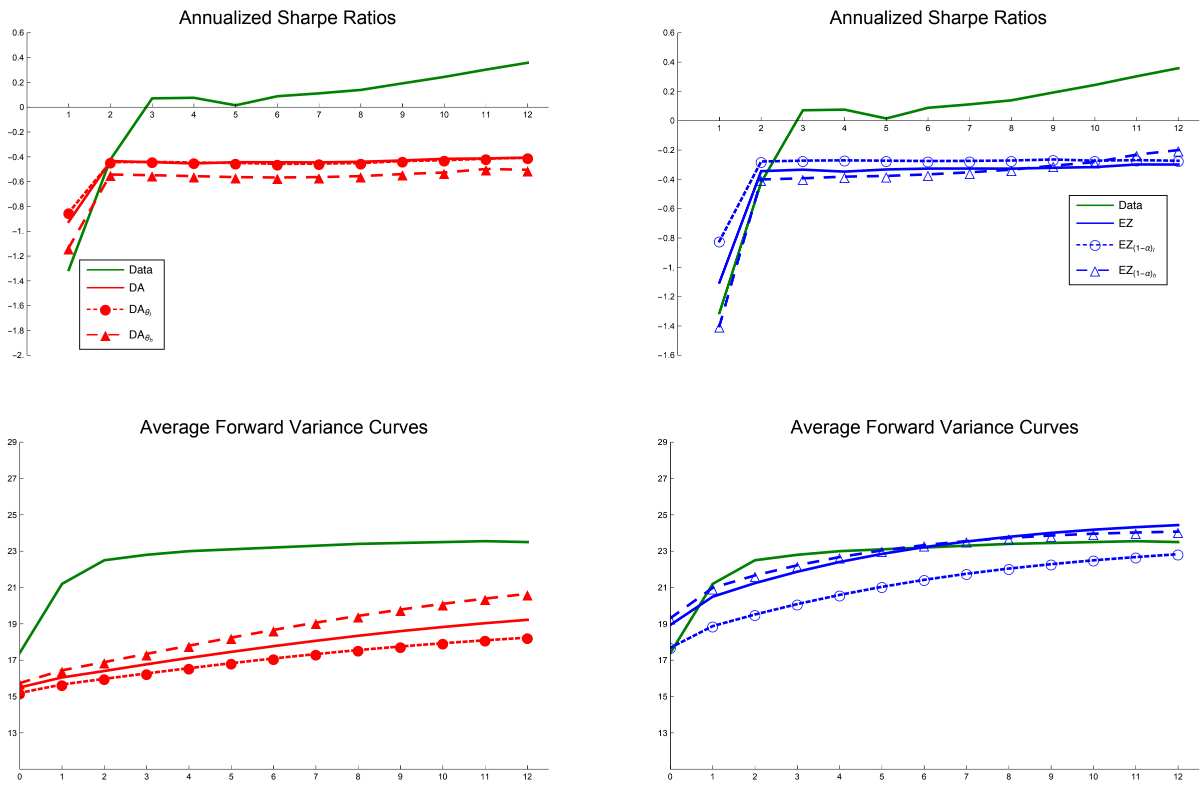

Figure 1.7: Sensitivity of Sharpe ratios and forward variance claim prices: DA and EZ. The figure plots annualized Sharpe ratios (top) and average prices for forward variance claims (bottom) for different model calibrations with disappointment aversion and Epstein-Zin preferences. DA and EZ correspond to the original DA and EZ models. In DA_{θ_l} and DA_{θ_h}, the disappointment aversion parameters are $\theta_l = 0.5$ and $\theta_h = 0.7$, respectively. In $\mathrm{EZ}_{(1-\alpha)_l}$ and $\mathrm{EZ}_{(1-\alpha)_h}$, the risk aversion parameters are $(1-\alpha)_l = 5$ and $(1-\alpha)_h = 7$, respectively. If not stated otherwise, the remaining parameters in all specifications are set at the original values in the DA and EZ models. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report medians of sample statistics based on these artificial series.

present the model-based average statistics implied by the GDA, DA, and EZ frameworks. The asset pricing moments are expressed as a function of a varying parameter, which is indicated on the corresponding axis.

Figure 1.8 shows that the risk-free rate decreases with the disappointment threshold, disappointment aversion and relative risk aversion in the GDA, DA, and EZ models, respectively. Further, the equity premium increases and equity prices decline in δ, θ, and $1-\alpha$. Intuitively, when the agent faces more disappointing outcomes or becomes more averse to low consumption growth rates, he demands larger premiums in expected returns for bearing the additional risk in consumption growth. The impact of δ and $1-\alpha$ on the volatility of asset prices is similar across the GDA and EZ models: a higher disappointment threshold or a higher risk aversion leads to a more volatile risk-free rate, while the volatility of equity returns and the price-dividend ratio exhibits a hump-shaped

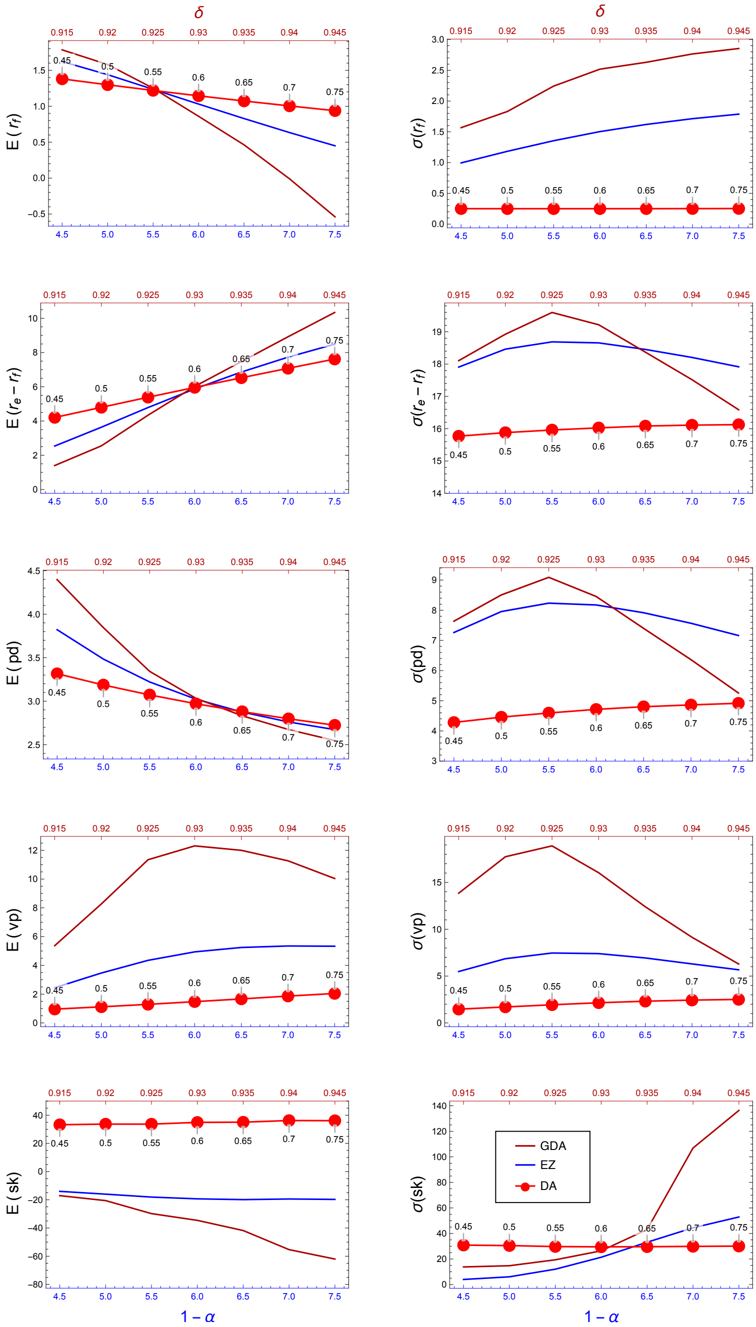

Figure 1.8: Sensitivity of asset prices: GDA, DA, and EZ. The figure plots asset pricing moments in the original GDA, DA, and EZ models, in which a single parameter is changed while others are fixed at the original values. Specifically, I change the disappointment threshold, the disappointment aversion parameter, and the coefficient of risk aversion in the original GDA, DA and EZ models, respectively, over a range of values. For each recalibration, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart. The entries of the figure are medians of sample statistics (annualized for the risk-free rate, the equity premium and the price-dividend ratio; monthly for the variance and skew risk premiums) based on these artificial series. I use the common notations for the sample mean E and standard deviation σ.

pattern with a maximum approximately in the middle of the parameter intervals considered. In the DA model, raising disappointment aversion slightly increases the volatility of the risk-free rate, equity returns, and prices. Overall, the magnitude of changes in the risk-free rate, equity returns, and the price-dividend ratio are quite comparable across the three models, especially when looking at the GDA and EZ frameworks. These findings suggest that all three preference specifications can reasonably explain the first and second moments of equity returns by adjusting a key preference parameter. In contrast, the four bottom panels in Figure 1.8 indicate the crucial importance of generalized disappointment aversion for generating significant risk premiums in higher moments of equity returns.

Figure 1.8 shows that, in the DA setting, changing the disappointment aversion for a wide range of values does not improve the model's performance, as the variance and skew risk premium moments are not very sensitive to changes in θ. Moreover, no value of the disappointment aversion parameter can support the negative skew premium. Figure 1.8 also shows that Epstein-Zin preferences provide a better fit of the model with the data. In particular, when the risk aversion increases from 4.5 to 6, the mean variance premium increases from less than 2 to around 5, while the skew premium declines from around -10% to -20%. However, the mean and volatility of the variance premium actually start to decline at some point, and thus the higher risk aversion will move the model away from the data. Finally, the comparative analysis with respect to the disappointment threshold in the GDA model generates patterns in the variance and skew risk premiums similar to those predicted by different risk aversion parameters in the EZ economy. However, with generalized disappointment aversion, the magnitude and time-variation of variance and skew risk premiums are significantly amplified. Overall, the sensitivity analysis in Figure 1.8 confirms that the distribution of the stochastic discount factor, necessary to reconcile the empirical asset pricing moments, is attributable to the agent's generalized disappointment aversion and cannot be supported by any parameter values in alternative preferences.

1.6.3 Implied Volatilities

Figures 1.9 and 1.10 further provide comparative statics of the implied volatility curves in the three preference specifications. Several observations are noteworthy. First, in all economies, the implied volatility curve for one-month options is not very sensitive to a further increase in effective risk aversion. In all cases, an incremental increase is less than 1% for any particular maturity and moneyness. Second, in the model with Epstein-Zin

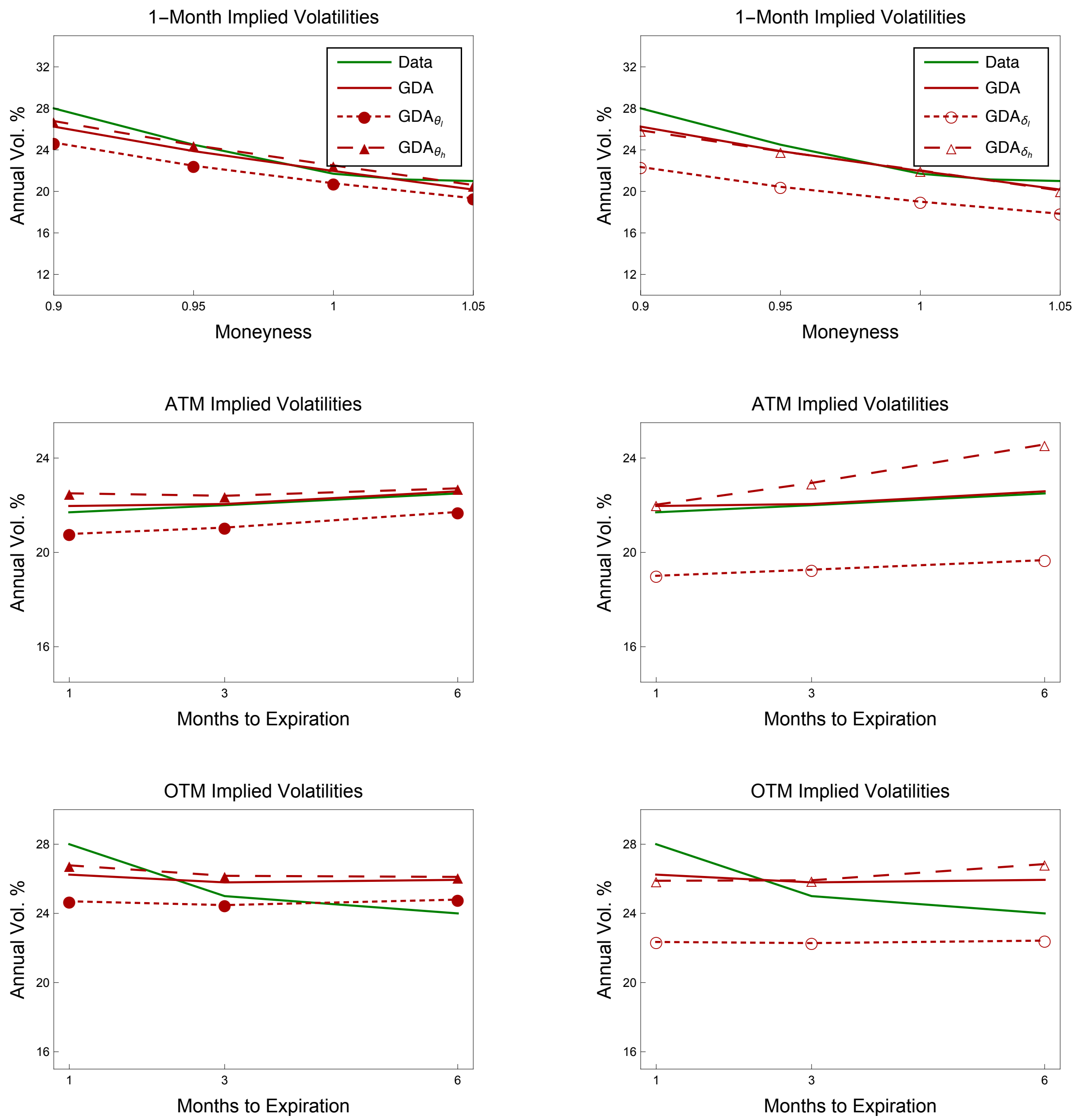

Figure 1.9: Sensitivity of implied volatilities: GDA. The figure plots the 1-month implied volatility curve (top) as a function of moneyness, implied volatility curves for ATM (middle) and OTM (bottom) options as functions of the time to maturity (in months) for different model calibrations with generalised disappointment aversion preferences. GDA corresponds to the original GDA model. In GDA$_{\theta_l}$ and GDA$_{\theta_h}$, the disappointment aversion parameters are $\theta_l = 6.41$ and $\theta_h = 10.41$, respectively. In GDA$_{\delta_l}$ and GDA$_{\delta_h}$, the disappointment threshold parameters are $\delta_l = 0.920$ and $\delta_h = 0.940$, respectively. If not stated otherwise, the remaining parameters in all specifications are set at the original values in the GDA model. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart and report medians of sample statistics based on these artificial series.

preferences, the slope of tre ATM and OTM volatilities stays the same for higher risk aversion. In the DA economy, even though ATM volatilities for longer maturities seem to increase more in response to raising disappointment aversion, the levels are significantly

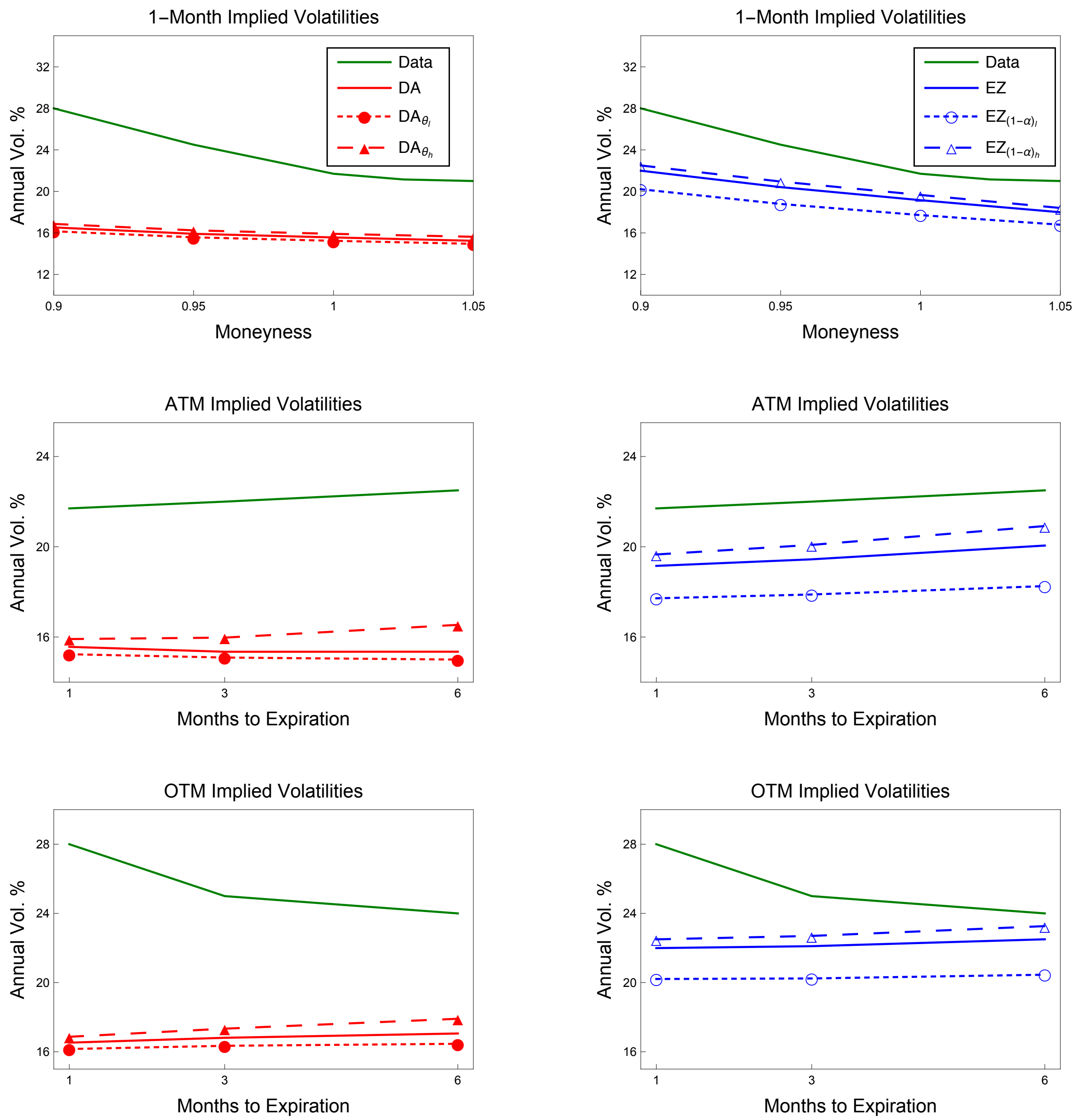

Figure 1.10: Sensitivity of implied volatilities: DA and EZ. The figure plots the 1-month implied volatility curve (top) as a function of moneyness, implied volatility curves for ATM (middle) and OTM (bottom) options as functions of the time to maturity (in months) for different model calibrations with disappointment aversion and Epstein-Zin preferences. DA and EZ correspond to the original DA and EZ models. In DA_{θ_l} and DA_{θ_h}, the disappointment aversion parameters are $\theta_l = 0.5$ and $\theta_h = 0.7$, respectively. In $EZ_{(1-\alpha)_l}$ and $EZ_{(1-\alpha)_h}$, the risk aversion parameters are $(1-\alpha)_l = 5$ and $(1-\alpha)_h = 7$, respectively. If not stated otherwise, the remaining parameters in all specifications are set at the original values in the DA and EZ models. For each model, I simulate 10,000 economies at a monthly frequency with a sample size equal to its empirical counterpart report medians of sample statistics based on these artificial series.

below the empirical curves. In the GDA economy, changes in θ and δ have a larger impact on the term structure of ATM and OTM volatilities. Specifically, Figure 1.9 suggests that a higher disappointment threshold increases prices of options with longer maturities, helping to explain a slightly upward sloping shape in ATM volatilities. Meanwhile, a higher disappointment aversion parameter seems to increase prices of short-term OTM options more than those with longer maturities, helping to explain a slightly downward sloping pattern in OTM volatilities. Therefore, in the setting of my model, simultaneously increasing θ and decreasing δ could allow one to keep the one-month implied volatilities close to the empirical curves while even better matching the salient statistics of ATM and OTM volatilities. Finally, a lower degree of effective risk aversion implies that the implied volatility curves become flatter and shift down in all models, especially in the economies with GDA and Epstein-Zin preferences.

1.7 Conclusion

This paper builds an equilibrium model with GDA preferences and rare events in consumption growth. I show that the combination of the investor's tail aversion and fluctuating economic uncertainty due to learning about a hidden depression state of the economy can explain a wide variety of asset pricing phenomena. Most notably, the model rationalizes the variance term structure, a new stylized fact of the variance swap data. In particular, the model predicts large and negative Sharpe ratios on one-month variance claims and produces a slightly positive term structure for maturities longer than two months, consistent with the empirical evidence. Furthermore, it accounts for the large variance and skew risk premiums in equity returns, and generates a realistic volatility surface implied by index options, while simultaneously matching the salient features of equity returns and the risk-free rate. I show that the success of the model is attributable to the generalized disappointment aversion channel by comparing the framework with GDA preferences to the models with nested utility functions: disappointment aversion and Epstein-Zin preferences. Although all three specifications can reasonably match moments of equity returns, only GDA preferences can explain the variance term structure, moment risk premiums, and option prices. These results suggest the important role of generalized disappointment aversion in asset pricing models, especially in the pricing of variance risk.

There are several interesting avenues for future research. First, the asset pricing results of my paper emphasize the importance of GDA preferences and the specific levels of the disappointment threshold and disappointment aversion. Although Delikouras (2017) provides the empirical estimate of a disappointment aversion parameter in Gul (1991), a joint estimation of the parameters in Routledge and Zin (2010) has not been addressed by the existing literature. Second, it seems to be a fruitful area to explore the implications of the richer model for the term structure of dividend strips and interest rates. For instance, an extension of the presented framework to include post-depression recoveries (Hasler and Márfe 2016) has the potential to provide a unified explanation of the term structures of interest rates, equity and variance risk premia. Third, generalized disappointment aversion is likely to have additional asset pricing implications for the size and time-variation of risk premia when combined with a multi-dimensional learning problem (Johannes, Lochstoer, and Mou 2016) or rational parameter learning (Collin-Dufresne, Johannes, and Lochstoer 2016). Finally, it would be interesting to investigate the interaction between GDA preferences and other behavioural biases, for instance, alternative learning rules (Brandt, Zeng, and Zhang 2004).

Appendix A

A.1 Data

A.1.1 Consumption, Dividends, and Market Returns

I follow Bansal and Yaron (2004) and construct real per capita consumption growth series (annual, due to the frequency restriction) for the longest sample available, 1930-2016. In the literature, consumption is defined as a sum of personal consumption expenditures on nondurable goods and services. I download the data from the US National Income and Product Accounts (NIPA) as provided by the Bureau of Economic Analysis. I apply the seasonally adjusted annual quantity indexes from Table 2.3.3. (Real Personal Consumption Expenditures by Major Type of Product, Quantity Indexes, A:1929-2016) to the corresponding series from Table 2.3.6. (Real Personal Consumption Expenditures by Major Type of Product, Chained Dollars, A:1995-2016) to obtain real personal consumption expenditures on nondurable goods and services for the sample period 1929-2016. I further retrieve mid-month population data from NIPA Table 7.1. to convert real consumption series to per capita terms.

I measure the total market return as the value-weighted return including dividends, and the dividends as the sum of total dividends, on all stocks traded on the NYSE, AMEX, and NASDAQ. The dividends and value-weighted market return data are monthly and are retrieved from the Center for Research in Security Prices (CRSP). To construct the monthly nominal dividend series, I use the CRSP value-weighted returns including and excluding dividends of CRSP common stock market indexes (NYSE/AMEX/NASDAQ/ARCA), denoted by RI_t and RE_t, respectively. Following Hodrick (1992), I construct the price series P_t by initializing $P_0 = 1$ and iterating recursively $P_t = (1 + RI_t)P_{t-1}$. Next, I compute normalized nominal monthly dividends $D_t = (RI_t - RE_t)P_t$. The proxy of the risk-free return $R_{f,t+1}$ is the 1-month nominal Treasury bill. The nominal annualized dividends are constructed by summing the corresponding monthly dividends within the year. Finally, I retrieve the inflation index from CRSP to deflate all quantities to real values.

A.1.2 The Variance Premium Data

For the variance risk premium, I closely follow Bollerslev, Tauchen, and Zhou (2009), Bollerslev, Gibson, and Zhou (2011), Drechsler and Yaron (2011), and Drechsler (2013). Under the no-arbitrage assumption, the risk-neutral conditional expectation of the return variance is equal to the price of a variance swap, which is a forward contract on the realized variance of the asset. Since the CBOE calculates the VIX index as a measure of the 30-days ahead risk-neutral expectation of the variance of the S&P 500 index, I use the VIX index as a proxy for the risk-neutral expectation of the market's return variation. The VIX is quoted in an annualized standard deviation. Hence, I first take it to a second power to transform to variance units and then divide by 12 to obtain monthly frequency. Thus, I obtain a new series defined as $[\mathrm{VIX}]_t^2 = \frac{\mathrm{VIX}_t^2}{12}$. I further use the last available observation of $[\mathrm{VIX}]_t^2$ in a particular month as a measure of the risk-neutral expectation of return variance in that month.

For the objective expectation of return variance, a second component in the variance premium, I calculate a one-step-ahead forecast from a simple regression similar to Drechsler and Yaron (2011) and Drechsler (2013). I first calculate the measure of the realized variance by summing the squared daily log returns on the S&P 500 futures and S&P 500 index obtained from the CBOE. The constructed series are denoted by FUT_t^2 and IND_t^2, respectively. Subsequently, I estimate the following regression:

$$\mathrm{FUT}_{t+1}^2 = \beta_0 + \beta_1 \cdot \mathrm{IND}_t^2 + \beta_2 \cdot [\mathrm{VIX}]_t^2 + \varepsilon_{t+1}. \tag{A.1}$$

The actual expectation is measured as the one-period ahead forecast given by (A.1). I refer to the resulting series as the realized variance and denote it by RV_t. Theoretically, the variance premium should be non-negative in each period. Thus, I truncate the difference between the implied series of $[\mathrm{VIX}]_t^2$ and RV_t from below by 0.

For the empirical strategy above, I obtain the daily data series of the VIX index, S&P 500 index futures, and the S&P 500 index from the CBOE. The main restriction on the length of the constructed monthly variance premium is the VIX index, reported by the CBOE from January 1990. Using high-frequency data would provide a finer estimation precision of the quantities in the variance premium, but my estimates remain largely consistent with the numbers reported by the existing literature.

A.1.3 Option Data for the Skew Premium and Implied Volatility Skew

The empirical strategy and key definitions of the skew risk premium are in line with Bakshi, Kapadia, and Madan (2003) and Kozhan, Neuberger, and Schneider (2013). For the empirical analysis of the skew risk premium and implied volatility surface, I use European options written on the S&P 500 index and traded on the CBOE. The option data set covers the period from January 1996 to December 2016 and is from OptionMetrics. Option data elements include the type of options (call/put) along with the contract's variables (strike price, time to expiration, Greeks, Black-Scholes implied volatilities, closing spot prices of the underlying) and trading statistics (volume, open interest, closing bid and ask quotes), among other details. The empirical estimates of the conditional skew risk premium are computed in line with Kozhan, Neuberger, and Schneider (2013). The empirical strategy consists of calculating fixed and floating legs for the skew swap, which correspond to the risk-neatral and physical expectations of the return skewness. For a detailed description of the methodology, see Kozhan, Neuberger, and Schneider (2013).

To construct the empirical implied volatility curves, I first compute the moneyness for each observed option using the daily S&P 500 index on a particular trading day. I filter out all data entries with non-standard settlements. I use the remaining observations to construct the implied volatility surface for a range of moneyness and maturities. In particular, I follow Christoffersen and Jacobs (2004) and perform polynomial extrapolation of volatilities in the maturity time and strike prices. This strategy makes use of all available options and not only those with a specific maturity time. The fitted values are further used to construct the implied volatility curves.

A.2 Representative Agent's Maximization Problem

A representative agent starts with an initial wealth denoted by W_0. Each period t, the agent consumes C_t consumption goods and invests in N assets traded on the competitive market. Denote the fraction of the total t-period wealth W_t invested in the i-th asset with gross real return $R_{i,t+1}$ by $\omega_{i,t}$. Then, the agent's budget constraint in period t takes the form:

$$W_{t+1} = (W_t - C_t)R^{\omega}_{t+1} \tag{A.2}$$

$$\sum_{i=1}^{N} \omega_{i,t} = 1 \quad \text{and} \quad R^{\omega}_{t+1} = \sum_{i=1}^{N} \omega_{i,t} R_{i,t+1}. \tag{A.3}$$

The agent chooses $\{C_t, \omega_{1,t}, ..., \omega_{N,t}\}$ in period t to maximize (1.1) subject to (A.2)-(A.3).

The Bellman equation becomes:

$$J_t = \max_{C_t, \omega_{1,t}, \ldots, \omega_{N,t}} \left\{ (1-\beta)C_t^{\rho} + \beta \left[\mu_t(J_{t+1})\right]^{\rho} \right\}^{1/\rho}$$

subject to (A.2) and (A.3). I guess optimal value function of the form $J_t = \phi_t W_t$. Using this conjecture of J_t and the form of μ_t from (1.2), I rewrite the Bellman equation as:

$$\phi_t W_t = \max_{C_t, \omega_{1,t}, \ldots, \omega_{N,t}} \left\{ (1-\beta)C_t^{\rho} + \beta \left[\mathbb{E}_t \Big[(\phi_{t+1}W_{t+1})^{\alpha} \mathcal{K}(\phi_{t+1}W_{t+1}) \Big]^{\rho/\alpha} \right\}^{1/\rho},$$
$$\mathcal{K}(x) = \frac{1 + \theta \mathbb{I}\{x \leqslant \delta \mu_t(x)\}}{1 + \theta \delta^{\alpha} \mathbb{E}_t \Big[\mathbb{I}\{x \leqslant \delta \mu_t(x)\} \Big]}.$$

Note that the function $\mathcal{K}$ defined above is homogeneous of degree zero.

The Return on the Aggregate Consumption Claim Asset. I further conjecture that the consumption C_t is homogeneous of degree one in wealth at the optimum, that is $C_t = b_t W_t$. Then, I obtain the Bellman equation:

$$\phi_t^{\rho} = \left\{ (1-\beta) \left(\frac{C_t}{W_t} \right)^{\rho} + \beta \left(1 - \frac{C_t}{W_t} \right)^{\rho} \left[\mathbb{E}_t \Big[(\phi_{t+1}R^{\omega}_{t+1})^{\alpha} \mathcal{K}(\phi_{t+1}R^{\omega}_{t+1}) \Big]^{\rho/\alpha} \right\} \tag{A.4}$$

or equivalently

$$\phi_t^{\rho} = \{(1-\beta)b_t^{\rho} + \beta (1-b_t)^{\rho} y_t^*\} \tag{A.5}$$
$$y_t^* = \left[\mathbb{E}_t \Big[(\phi_{t+1}R^{\omega}_{t+1})^{\alpha} \mathcal{K}(\phi_{t+1}R^{\omega}_{t+1}) \Big]^{\rho/\alpha}.$$

Taking the FOC of the right side of a simplified Bellman equation (A.4) with respect to

C_t, I find:

$$(1-\beta)\left(\frac{C_t}{W_t}\right)^{\rho-1} = \beta\left(1-\frac{C_t}{W_t}\right)^{\rho-1} y_t^*.$$

or using the notations:

$$(1-\beta)b_t^{\rho-1} = \beta(1-b_t)^{\rho-1}y_t^*. \tag{A.6}$$

Solving for y_t^* from the last equation and substituting it into (A.5), I deduce:

$$\phi_t = (1-\beta)^{\frac{1}{\rho}} b_t^{\frac{\rho-1}{\rho}} = (1-\beta)^{\frac{1}{\rho}}\left(\frac{C_t}{W_t}\right)^{\frac{\rho-1}{\rho}}$$

Shifting one period ahead the formula for ϕ_t and substituting ϕ_{t+1} into (A.6), I obtain:

$$(1-\beta)C_t^{\rho-1} = \beta(W_t-C_t)^{\rho-1}\left[\mathbb{E}_t\left[(1-\beta)^{\alpha/\rho}\left(\frac{C_{t+1}}{W_{t+1}}\right)^{\alpha\frac{\rho-1}{\rho}}\left(R_{t+1}^{\omega}\right)^{\alpha}\mathcal{K}\left(\phi_{t+1}R_{t+1}^{\omega}\right)\right]\right]^{\rho/\alpha}.$$

Then, I rewrite the equation above as:

$$C_t^{\rho-1} = \beta\mathbb{E}_t\left[\left(\frac{C_{t+1}}{\frac{W_{t+1}}{(W_t-C_t)}}\right)^{\alpha\frac{\rho-1}{\rho}}\left(R_{t+1}^{\omega}\right)^{\alpha}\mathcal{K}\left(\left(\frac{C_{t+1}}{\frac{W_{t+1}}{W_t-C_t}}\right)^{\frac{\rho-1}{\rho}}R_{t+1}^{\omega}\right)\right]^{\rho/\alpha}.$$

and derive the asset pricing restriction for the return on the total wealth R_{t+1}^{ω} :

$$\mathbb{E}_t\left[\left\{\underbrace{\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}R_{t+1}^{\omega}\right)^{1/\rho}}_{z_{t+1}}\right\}^{\alpha}\mathcal{K}\left(\underbrace{\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}R_{t+1}^{\omega}\right)^{1/\rho}}_{z_{t+1}}\right)\right]^{1/\alpha} = 1.$$

Define R_{t+1}^c the return on the consumption endowment. In equilibrium, $R_{t+1}^c = R_{t+1}^{\omega}$ and, as in (Routledge and Zin 2010), using the definition of the certainty equivalent (1.2) and the function $\mathcal{K}$, the return R_{t+1}^c should satisfy the equation:

$$\mu_t(z_{t+1}) = 1, \quad z_{t+1} = \left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho-1}R_{t+1}^c\right)^{1/\rho}. \tag{A.7}$$

Rewriting R_{t+1}^c in the form:

$$R_{t+1}^c = \frac{W_{t+1}}{W_t-C_t} = \frac{\frac{W_{t+1}}{C_{t+1}}}{\frac{W_t}{C_t}-1}\cdot\frac{C_{t+1}}{C_t} = \frac{\xi_{t+1}}{\xi_t-1}\cdot\frac{C_{t+1}}{C_t},$$

the wealth-consumption ratio $\xi_t = \frac{W_t}{C_t}$ can be found from the equation:

$$\mathbb{E}_t \left[\beta^{\frac{\alpha}{\rho}} \left(\frac{C_{t+1}}{C_t} \right)^{\alpha} \cdot \left(\frac{\xi_{t+1}}{\xi_t - 1} \right)^{\frac{\alpha}{\rho}} \cdot \mathcal{K}(z_{t+1}) \right] = 1.$$

The Return on the Aggregate Dividend Asset. Following Routledge and Zin (2010), the portfolio problem for the obtained values ϕ_{t+1} reads as follows:

$$\max_{\omega_{1,t},\dots,\omega_{N,t}} \mu_t(\phi_{t+1} R^{\omega}_{t+1}),$$

subject to the constraints $\sum_{i=1}^{N} \omega_{i,t} = 1$ and $R^{\omega}_{t+1} = \sum_{i=1}^{N} \omega_{i,t} R_{i,t+1}$. Taking the FOC with respect to the weight $\omega_{i,t}$, I derive:

$$\mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] R_{i,t+1} \right] = 0.$$

Taking the difference between the i-th and j-th FOCs, I thus obtain:

$$\mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] (R_{i,t+1} - R_{j,t+1}) \right] = 0.$$

Multiplying the last equation by $\omega_{j,t}$ and summing over j, I further obtain:

$$\mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] R_{i,t+1} \underbrace{\sum_{j=1}^{N} \omega_{j,t}}_{=1} \right] =$$

$$= \mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] \underbrace{\sum_{j=1}^{N} R_{j,t+1} \omega_{j,t}}_{=R^{\omega}_{t+1}} \right]$$

or

$$\begin{aligned} &\mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha-1} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] R_{i,t+1} \right] = \\ &= \mathbb{E}_t \left[\phi^{\alpha}_{t+1} (R^{\omega}_{t+1})^{\alpha} [1 + \theta \mathbb{I}(\phi_{t+1} R^{\omega}_{t+1} < \delta \mu_t)] \right]. \end{aligned} \tag{A.8}$$

Following Epstein and Zin (1989), it is straightforward to show that $\phi_{t+1} = \frac{z_{t+1}}{R^{\omega}_{t+1}}$ holds in equilibrium. Using these equilibrium conditions and the definition of μ_t, I have:

$$\mathbb{E}_t\left[\phi_{t+1}^{\alpha}(R_{t+1}^{\omega})^{\alpha}[1+\theta\mathbb{I}(\phi_{t+1}R_{t+1}^{\omega}<\delta\mu_t)]\right]=\mathbb{E}_t\left[z_{t+1}^{\alpha}[1+\theta\mathbb{I}(z_{t+1}<\delta\mu_t)]\right]=$$
$$\mathbb{E}_t\left[1+\theta\delta^{\alpha}\mathbb{I}(z_{t+1}<\delta\underbrace{\mu_t(z_{t+1})}_{=1})]\right]\underbrace{\mu_t(z_{t+1})^{\alpha}}_{=1}=\mathbb{E}_t\left[1+\theta\delta^{\alpha}\mathbb{I}(z_{t+1}<\delta]\right]. \tag{A.9}$$

Combining (A.8)-(A.9) and using the equilibrium condition $R_{t+1}^{c}=R_{t+1}^{\omega}$, I finally obtain the asset pricing restriction for the gross return $R_{i,t+1}$:

$$\mathbb{E}_t\left[\frac{z_{t+1}^{\alpha}(R_{t+1}^{c})^{-1}(1+\theta\mathbb{I}(z_{t+1}<\delta)R_{i,t+1}}{1+\theta\delta^{\alpha}\mathbb{E}_t\left[\mathbb{I}(z_{t+1}<\delta)\right]}\right]=1, \tag{A.10}$$

Moreover, the pricing kernel M_{t+1} is:

$$M_{t+1}=\frac{z_{t+1}^{\alpha}(R_{t+1}^{c})^{-1}(1+\theta\mathbb{I}(z_{t+1}<\delta))}{1+\theta\delta^{\alpha}\mathbb{E}\left[\mathbb{I}(z_{t+1}<\delta)\right]}.$$

Rewriting $R_{i,t+1}$ in the form:

$$R_{i,t+1}=\frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}=\frac{\frac{P_{i,t+1}}{D_{i,t+1}}+1}{\frac{P_{i,t}}{D_{i,t}}}\cdot\frac{D_{i,t+1}}{D_{i,t}}=\frac{\lambda_{t+1}+1}{\lambda_t}\cdot\frac{D_{i,t+1}}{D_{i,t}},$$

the price-dividend ratio of the i-th asset $\lambda_t=\frac{P_{i,t}}{D_{i,t}}$ can be found from the equation:

$$\mathbb{E}_t\left[\beta^{\frac{\alpha}{\rho}}\left(\frac{C_{t+1}}{C_t}\right)^{\alpha-1}\frac{D_{i,t+1}}{D_{i,t}}\cdot\left(\frac{\xi_{t+1}}{\xi_t-1}\right)^{\frac{\alpha}{\rho}-1}\cdot\mathcal{K}(z_{t+1})\cdot(\lambda_{t+1}+1)\right]=\lambda_t.$$

A.3 The Numerical Solution

Following the notation from the paper, aggregate consumption growth is

$$\Delta c_{t+1}=\mu_{s_{t+1}}+\sigma\varepsilon_{t+1},\quad \varepsilon_{t+1}\sim N(0,1).$$

The consumption volatility σ is constant, whereas the mean growth rate $\mu_{s_{t+1}}$ is driven by a two-state Markov-switching process s_{t+1} with a state space:

$$\mathcal{S}=\{1=\text{expansion},\ 2=\text{recession}\},$$

a transition matrix

$$\mathcal{P} = \begin{pmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{pmatrix}$$

and transition probabilities $\pi_{ii} \in (0,1),\ i = 1,2$. Let

$$\mathcal{X}(y_1, y_2, y_3) = \frac{1 + \theta \mathbb{I}\left\{\beta e^{\rho y_1}\left(\frac{y_2}{y_3 - 1}\right) \leqslant \delta^\rho\right\}}{1 + \theta\delta^\alpha \mathbb{E}_t\left[\mathbb{I}\left\{\beta e^{\rho y_1}\left(\frac{y_2}{y_3-1}\right) \leqslant \delta^\rho\right\}\right]},$$

then, the wealth-consumption ratio $\xi_t = \frac{W_t}{C_t}$ satisfies the equation:

$$\mathbb{E}_t\left[\beta^{\frac{\alpha}{\rho}} e^{\alpha\Delta c_{t+1}} \cdot \left(\frac{\xi_{t+1}}{\xi_t - 1}\right)^{\frac{\alpha}{\rho}} \cdot \mathcal{X}\Big(\Delta c_{t+1}, \xi_{t+1}, \xi_t\Big)\right] = 1, \tag{A.1}$$

and the price-dividend ratio $\lambda_t = \frac{P_t}{D_t}$ of the asset with gross return R_{t+1} (I skip the subscript i for convenience) is given by:

$$\mathbb{E}_t\left[\beta^{\frac{\alpha}{\rho}} e^{(\alpha-1)\Delta c_{t+1} + \Delta d_{t+1}} \cdot \left(\frac{\xi_{t+1}}{\xi_t - 1}\right)^{\frac{\alpha}{\rho}-1} \cdot \mathcal{X}\Big(\Delta c_{t+1}, \xi_{t+1}, \xi_t\Big) \cdot \frac{\lambda_{t+1}+1}{\lambda_t}\right] = 1. \tag{A.2}$$

A.3.1 The Projection Method

Following Pohl, Schmedders, and Wilms (2018), I apply a projection method of Judd (1992) to solve for the equilibrium pricing functions defined by (A.1) and (A.2). The model solution consists of two steps. First, I find the wealth-consumption ratio from the functional equation (A.1). Second, I use the wealth return from the first step and substitute it into (A.2) to find the price-dividend ratio for the equity claim.

The Return on the Aggregate Consumption Claim Asset. I conjecture the wealth-consumption ratio of the form $\xi_t = G(\pi_t)$, in which π_t is the posterior belief defined by (1.4). I seek to approximate the functional form of $G(\pi_t)$ by a basis of complete Chebyshev polynomials $\Psi = \{\Psi_k(\pi_t)\}_{k=0}^n$ of order n with coefficients $\psi = \{\psi_k\}_{k=0}^n$:

$$G(\pi_t) = \sum_{k=0}^{n} \psi_k \Psi_k(\pi_t) \quad \pi_t \in [1-p,\ q]. \tag{A.3}$$

I further define the function:

$$\Gamma(\pi_t; j) = \mathbb{E}_{t,j}\left[\beta^{\frac{\alpha}{\rho}} e^{\alpha \Delta c_{t+1}} \cdot \left(\frac{\xi_{t+1}}{\xi_t - 1}\right)^{\frac{\alpha}{\rho}} \cdot \mathcal{X}\Big(\Delta c_{t+1}, \xi_{t+1}, \xi_t\Big)\right] =$$

$$= \beta^{\frac{\alpha}{\rho}} \int e^{\alpha y} \left(\frac{G(B(y,\pi_t))}{G(\pi_t) - 1}\right)^{\frac{\alpha}{\rho}} \cdot \mathcal{X}\Big(y, G(B(y,\pi_t)), G(\pi_t)\Big) f(y,j) dy, \qquad \text{(A.4)}$$

$$B(y,\pi_t) = \frac{(1-q)f(y,1)(1-\pi_t) + pf(y,2)\pi_t}{f(y,1)(1-\pi_t) + f(y,2)\pi_t},$$

$f(y,j)$ is the probability density function of a normal distribution $N(\mu_{s_t}, \sigma^2)$ conditional on $s_t = 1, 2$. I further apply Gauss-Hermite quadrature to calculate expectations in (A.4). Substituting $G(\pi_t)$ from (A.3) and $\Gamma(\pi_t; j)$ from (A.4) into (A.1), I obtain:

$$R^c(\pi_t; \psi) = (1-\pi_t)\Gamma(\pi_t, 1) + \pi_t \Gamma(\pi_t, 2) - 1.$$

The objective is to choose the unknown coefficients ψ to make $R^c(\pi_t; \psi)$ close to zero $\forall \pi_t \in [1-p,\ q]$. I apply the orthogonal collocation method. Formally, I evaluate the residual function in the collocation points $\{r_k\}_{k=1}^{n+1}$ given by the roots of the $n+1$ order Chebyshev polynomial and then solve the system of $n+1$ equations:

$$R^c(r_k; \psi) = 0 \quad k = 1, ..., n+1$$

for $n+1$ unknowns $\psi = \{\psi_k\}_{k=0}^n$. Let $\tilde{\xi}_t = \tilde{G}(\pi_t) = \sum_{k=0}^{n} \tilde{\psi}_k \Psi_k(\pi_t)$ denote an approximation of the wealth-consumption ratio, which will be used in the second step.

The Return on the Aggregate Dividend Asset. I conjecture the price-dividend ratio of the form $\lambda_t = H(\pi_t)$. Now, I seek to approximate the functional form of $H(\pi_t)$, which solves the equation (A.2). I approximate $H(\pi_t)$ by a basis of complete Chebyshev polynomials $\Upsilon = \{\Upsilon_k(\pi_t)\}_{k=0}^n$ of order n with coefficients $\upsilon = \{\upsilon_k\}_{k=0}^n$:

$$H(\pi_t) = \sum_{k=0}^{n} \upsilon_k \Upsilon_k(\pi_t) \quad \pi_t \in [1-p,\ q]. \qquad \text{(A.5)}$$

I define the function:

$$\Lambda(\pi_t; j) = \mathbb{E}_{t,j}\left[\beta^{\frac{\alpha}{\rho}} e^{(\alpha-1)\Delta c_{t+1} + \Delta d_{t+1}} \left(\frac{\tilde{\xi}_{t+1}}{\tilde{\xi}_t - 1}\right)^{\frac{\alpha}{\rho}-1} \cdot \mathcal{X}\Big(\Delta c_{t+1}, \tilde{\xi}_{t+1}, \tilde{\xi}_t\Big) \cdot \frac{\lambda_{t+1}+1}{\lambda_t}\right] =$$

$$= \beta^{\frac{\alpha}{\rho}} \iint e^{(\alpha+\lambda-1)y+g_d+z} \left(\frac{\tilde{G}(B(y,\pi_t))}{\tilde{G}(\pi_t)-1}\right)^{\frac{\alpha}{\rho}-1} \cdot \mathcal{X}\Big(y, \tilde{G}(B(y,\pi_t), \tilde{G}(\pi_t)\Big) \cdot \tag{A.6}$$

$$\cdot \frac{H(B(y,\pi_t))}{H(\pi_t)-1} f(y,j) g(z,j) dy dz,$$

in which $f(y,j)$ and $g(z,j)$ are probability density functions of normal distributions $N(\mu_{s_{t+1}}, \sigma)$ and $N(g_d, \sigma_d)$, respectively, conditional on $s_{t+1} = 1, 2$. Substituting $H(\pi_t)$ from (A.5) and $\Lambda(\pi_t; j)$ from (A.6) into (A.2), I obtain:

$$R^d(\pi_t; \upsilon) = (1-\pi_t)\Lambda(\pi_t, 1) + \pi_t \Lambda(\pi_t, 2) - 1.$$

Again, I apply the orthogonal collocation method. Formally, I evaluate $R^d(\pi_t; \psi)$ in the collocation points $\{s_k\}_{k=1}^{n+1}$ given by the roots of the $n+1$ order Chebyshev polynomial and solve the system of $n+1$ equations

$$R^d(s_k; \upsilon) = 0 \quad \forall k = 1, ..., n+1$$

for $n+1$ unknowns $\upsilon = \{\upsilon_k\}_{k=0}^{n}$.

A.3.2 Implementation in Matlab

This paper implements a one-dimensional projection method for solving the functional equations. I approximate unknown functions using Chebyshev polynomials of the first kind and compute them recursively as:

$$T_0(z) = 1, \quad T_1(z) = z, \quad T_k(z) = 2zT_k(z) - T_{k-1}(z), \quad k = 2, ..., n \ \wedge \ z \in [-1, 1].$$

I adjust the domain of Chebyshev polynomials to the state space of pricing ratios and use modified polynomials in the approximation. Thus, the following equalities hold on the interval $[\pi_{\min}, \pi_{\max}] = [1-p, \ q]$:

$$\Psi_k(\pi_t) = \Upsilon_k(\pi_t) = T_k\left(2\left[\frac{\pi_t - \pi_{\min}}{\pi_{\max} - \pi_{\min}}\right] - 1\right), \quad k = 0, ..., n.$$

I present the results based on the collocation method. For this purpose, I evaluate residual functions in a set of nodes corresponding to $n+1$ zeros of the $(n+1)$-order Chebyshev

Table A.1
Accuracy of the projection method: Euler errors.

Model	$n = 200$ $N_{GH} = 100$	$n = 200$ $N_{GH} = 150$	$n = 400$ $N_{GH} = 100$	$n = 400$ $N_{GH} = 150$
GDA	4.71e-07	4.18e-07	1.83e-07	1.47e-07
GDA_{δ_l}	3.40e-07	2.83e-07	1.25e-07	1.16e-07
GDA_{δ_h}	5.01e-07	4.40e-07	1.83e-07	1.75e-07
GDA_{θ_l}	4.28e-07	4.12e-07	1.61e-07	1.53e-07
GDA_{θ_h}	5.42e-07	4.76e-07	1.86e-07	1.84e-07
DA	1.28e-08	9.48e-09	3.97e-09	3.47e-09
DA_{θ_l}	8.82e-09	7.95e-09	3.16e-09	2.21e-09
DA_{θ_h}	1.30e-08	1.17e-08	4.63e-09	4.39e-09
EZ	7.59e-14	7.37e-14	9.15e-14	9.32e-14
$\text{EZ}_{(1-\alpha)_l}$	5.57e-14	4.95e-14	6.85e-14	6.58e-14
$\text{EZ}_{(1-\alpha)_h}$	9.94e-14	9.86e-14	1.34e-13	1.26e-13

The table reports the RMSE for different models. For each specification, it shows the results for two different degrees of Chebyshev polynomials n and two different numbers of Gauss-Hermite quadrature points N_{GH}. The Euler errors are computed using the equation (A.7) with 10,000 points equally spaced on the interval $[\pi_{\min}, \pi_{\max}]$.

polynomial, which are formally defined as:

$$z_k = \cos\left(\frac{2k+1}{2n+2}\pi\right), \quad k = 0, ..., n.$$

I adjust the nodes $z_k \in [-1, 1]$ to the domain of the state variable π_t :

$$\pi_k = \pi_{\min} + \frac{\pi_{\max} - \pi_{\min}}{2}(1 + z_k), \quad k = 0, ..., n.$$

The numerical algorithm, which requires solving a system of nonlinear equations, is efficiently programmed in Matlab. I experiment with different nonlinear solvers to achieve better performance of the code. Initially, I use the simple solver "fsolve". Then I find the solution of the system of nonlinear equations through minimizing a constant subject to the system of nonlinear functions. I apply the nonlinear programming solver "fmincon" with the SQP algorithm for this purpose. Similar to Pohl, Schmedders, and Wilms (2018), I find that "fmincon" provides faster running of the code and a more accurate solution compared to "fsolve". Thus, I present the results of all models considered in my paper based on the "fmincon" approach.

Additional numerical details involve the choices of an order of Chebychev polynomials used in the approximation of unknown functions (n), a number of Gauss-Hermite quadrature points used in the numerical integration of expectations in the residual functions (N_{GH}), and a number of draws used in Monte-Carlo simulations to compute model-based European put prices (N_{MC}). I report the results of all models in the main text based on the numerical solution, in which $n = 400$, $N_{GH} = 150$, and $N_{MC} = 2,000,000$. The next section performs sensitivity analysis of the asset pricing results to alternative approximation and simulation choices.

A.3.3 Accuracy of the Numerical Methods

To better assess the numerical accuracy, I first calculate the root mean squared error (RMSE) in the residual function for the wealth-consumption ratio. I evaluate $R^c(\pi_t; \psi)$ on a dense grid of points $\{\pi_i\}_{i=1}^{N_{\text{RMSE}}}$ that are equally spaced on the interval $[\pi_{\min},\ \pi_{\max}]$. I choose $N_{\text{RMSE}} = 10,000$ of these points. The RMSE is calculated as:

$$\text{RMSE}^c = \sqrt{\frac{1}{N_{\text{RMSE}}} \sum_{k=1}^{N_{\text{RMSE}}} \Big[R^c(\pi_k; \psi)\Big]^2}, \tag{A.7}$$

$$\pi_k = \pi_{min} + \frac{\pi_{\max} - \pi_{\min}}{N_{\text{RMSE}} - 1}(k-1), \quad k = 1, ..., N_{\text{RMSE}}.$$

I consider four pairs of $(n, N_{GH}) : (200, 100), (200, 150), (400, 100), (400, 150)$. For each pair, I solve different model calibrations of this paper and compute the RMSE.

Table A.1 reports the Euler errors implied by various approximation and integration choices. Several observations are noteworthy. First, the numerical solution technique is highly accurate, producing errors consistently below 6e-7 for all cases. Second, the projection method generates smaller RMSE for the models with Epstein-Zin preferences relative to the calibrations with disappointment aversion and generalized disappointment aversion utility functions. This result is expected in the light of nonlinearities in the pricing kernel implied by disappointing outcomes in consumption growth. Third, increasing either the degree of Chebyshev polynomials or the number of quadrature points generally leads to a better approximation precision.

Figure A.1 conducts further robustness checks. It compares the results of the two solutions of the original GDA calibration. First, the "GDA" lines correspond to the variance term structures as presented in the main text. Second, the "GDA2" curves represent the results of the same calibration, which is solved with a twice larger order of Chebyshev polynomials and where variance swap prices are calculated with a twice larger number of Monte Carlo simulations. The panels in Figure A.1 show that the results across

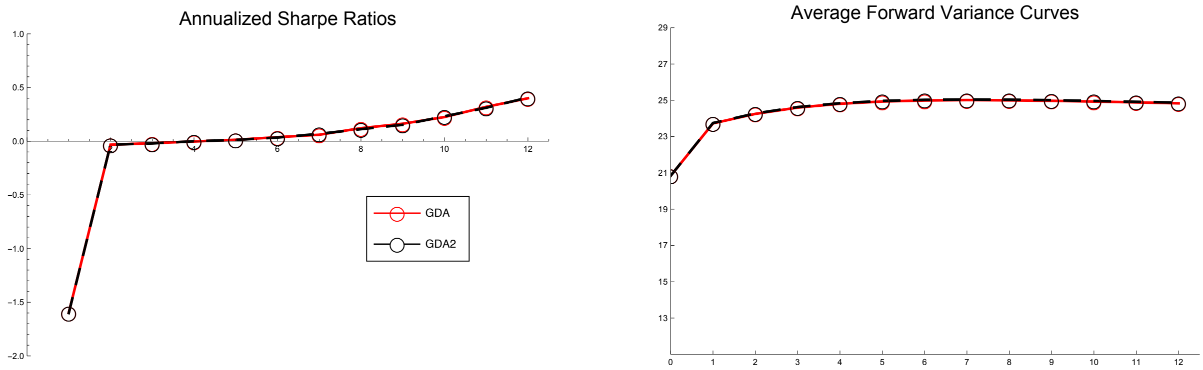

Figure A.1: Accuracy of the projection method: Sharpe ratios and forward variance claim prices. The figure plots annualized Sharpe ratios and average prices for variance forwards for the original GDA calibration, which is solved and simulated with different precisions. "GDA" denotes the results of the original solution. "GDA2" shows the results of the original calibration, which is solved with a twice larger order of Chebyshev polynomials and where variance swap prices are calculated with a twice larger number of Monte Carlo simulations.

the two solutions are very similar, confirming the high-precision solution obtained by the projection method.

Chapter 2

Parameter Learning in Production Economies

Mykola Babiak[1] and **Roman Kozhan**[2]

Abstract

We examine how parameter learning amplifies the impact of macroeconomic shocks on equity prices and quantities in a standard production economy where a representative agent has Epstein-Zin preferences. An investor observes technology shocks that follow a regime-switching process, but does not know the underlying model parameters governing the short-term and long-run perspectives of economic growth. We show that rational parameter learning endogenously generates long-run productivity and consumption risks that help explain a wide array of dynamic pricing phenomena. The asset pricing implications of subjective long-run risks crucially depend on the introduction of a procyclical dividend process consistent with the data.

[1]We would like to thank Frederico Belo (SAFE discussant), Andrea Gamba, Alessandro Graniero (EFA discussant), Michal Kejak, Ian Khrashchevskyi, Ctirad Slavik, Sergey Slobodyan and conference/seminar participants at the 2019 EFA Meeting, the 2019 SAFE Asset Pricing Workshop, the 2018 Lancaster-Warwick (LaWa) Workshop on Financial Econometrics and Asset Pricing, Warwick Business School, CERGE-EI and Università Ca' Foscari Venezia for their discussions and comments. The research support with the supercomputing clusters from the Centre for Scientific Computing at the University of Warwick is gratefully acknowledged.

[2]Warwick Business School, University of Warwick, Scarman Road, Coventry, CV4 7AL, UK.

2.1 Introduction

Parameter learning has recently been proposed as an amplification mechanism for the pricing of macroeconomic shocks used to explain standard asset pricing moments. In the endowment economy, parameter uncertainty helps explain the observed equity premium, the high volatility of equity returns, the market price-dividend ratio and the equity Sharpe ratio (Collin-Dufresne, Johannes, and Lochstoer 2016; Johannes, Lochstoer, and Mou 2016). In contrast to the consumption-based approach, a production dynamic stochastic general equilibrium (DSGE) model endogenously generates consumption and dividends and, as a result, it becomes more challenging to explain asset pricing puzzles in a production-based setting while simultaneously matching the moments of macroeconomic fundamentals. In this paper, we study how the macroeconomic risks arising from parameter uncertainty improve the performance of a standard DSGE model in jointly reproducing salient features of the macroeconomic quantities and equity returns.

Kaltenbrunner and Lochstoer (2010) and Croce (2014) have argued that the presence of a small but persistent long-run risk component in the productivity growth process can endogenously generate long-run risks in consumption growth that help boost up moments of financial variables. However, these long-run risk components are difficult to identify in the data.[3] In contrast, we demonstrate that rational pricing of parameter uncertainty is a source of these subjective long-run risks in productivity growth. This suggests the importance of accounting for parameter uncertainty in the productivity growth process. It is not clear, however, if macroeconomic risks associated with rational learning about productivity growth amplify the moments of financial variables. If so, what is the magnitude of the effect? In this paper, we document a considerable amplification mechanism of rational parameter learning on asset prices.

We introduce parameter uncertainty in the technology growth process of an otherwise standard production-based asset pricing model. We depart from the extant macro-finance literature by presuming that the representative investor does not know the parameters of the technology process and learns about true parameter values from the data. In each period, he updates his beliefs in a Bayesian fashion upon observing newly arrived data. Rational learning about unknown parameters together with recursive preferences gives

[3] Croce (2014) empirically demonstrates the existence of such a predictable component; however, the results are not robust to estimation method and sample choice. Moreover, low values for goodness-of-fit statistics lead to a conclusion that there is considerable uncertainty about the model specification for productivity growth.

rise to subjective long-lasting macroeconomic risks. Coupled with endogenous long-run consumption risks due to consumption smoothing (Kaltenbrunner and Lochstoer 2010) these risks are priced under the investor's preference for early resolution of uncertainty. The model generates higher equity Sharpe ratios, risk premia and volatility, as well as lower interest rates and price-dividend ratios relative to the standard framework. Additionally, the model with rational belief updating reproduces the excess return predictability pattern observed in the data. We further show that under certain calibrations of the elasticity of intertemporal substitution and a capital adjustment cost, parameter learning significantly magnifies propagation of shocks and hence helps to match the second moments and comovements of macroeconomic variables.

In our analysis, we restrict our attention to uncertainty about parameters governing the magnitude and persistence of productivity growth over the various phases of the business cycle. In particular, we examine the implications of learning about the transition probabilities and mean growth rates in a two-state Markov-switching process for productivity growth, where volatility of productivity growth is homoskedastic and known.[4] We consider two approaches to dealing with parameter uncertainty in the equilibrium models: anticipated utility (AU) and priced parameter uncertainty (PPU). The AU approach is common for most existing models, and assumes that economic agents learn about unknown parameters over time, but treat their current beliefs as true and fixed parameter values in the decision-making. For the PPU case, the representative investor calculates his utility and prices in the current period, assuming that posterior beliefs can be changed in the future. We quantify the impact of each type of parameter uncertainty pricing by comparing the results of AU and PPU with the full information (FI) model.

We begin our investigation by illustrating the economic importance of parameter uncertainty in the standard production economy with convex capital adjustment costs. The increased uncertainty due to unknown parameters in the productivity growth process creates a stronger precautionary saving motive, which leads to a lower risk-free rate. Fully rational learning about unknown parameters generates endogenous long-run risks in the economy, which in turn increase the mean and volatility of levered returns to the firm's payouts (Jermann 1998). In contrast, fluctuations in parameter beliefs are not priced in

[4] There is a large strand of the literature emphasizing the importance of time-varying macroeconomic uncertainty (see, for example, Justiniano and Primiceri (2008, Bloom (2009, Fernandez-Villaverde et al. (2011, Born and Pfeifer (2014, Christiano, Motto, and Rostagno (2014, Gilchrist, Sim, and Zakrajsek (2014, Liu and Miao (2014) and more recent studies by Leduc and Liu (2016, Basu and Bundick (2017, Bloom et al. (2018)). We leave the investigation of learning about volatility risks for future research.

the AU case. Thus, the PPU approach leads to around a two-fold increase in the risk premium (in addition to higher return volatility) on a levered firm's dividends, relative to the FI and AU cases. The combination of time-varying posterior beliefs and rational parameter learning is crucial for generating long-term predictability of excess returns by Tobin's Q, investment-capital, price-dividend and consumption-wealth ratios, as found in the empirical literature. The time-variation in beliefs leads to fluctuations in the equity risk premium and hence generates more predictability in the models with parameter uncertainty relative to the known parameter frameworks. Fully rational learning further magnifies the impact of belief revisions on the conditional equity premium and therefore there is more significant return predictability with PPU compared to AU. Specifically, the model with PPU closely replicates the increasing patterns (in absolute terms) of the regression coefficients and R^2's. In contrast, both the FI and AU models generate less predictability power and cannot match the magnitude of slope coefficients.

In terms of the macroeconomic variables, the benchmark model with parameter learning has a small effect on the unconditional second moments and a large impact on comovements of consumption, investment and output. In the sensitivity analysis, we further investigate how the impact of parameter learning on quantities changes for alternative calibrations of the inter-temporal elasticity of substitution and a capital adjustment cost. We find that a lower value of the inter-temporal elasticity of substitution, or a smaller capital adjustment cost, magnifies the effect of rational parameter learning on comovements between macroeconomic quantities. In particular, decreasing IES or adjustment costs for capital leads to significantly lower correlations between consumption, investment and output in the PPU model, while aggregate variables still remain highly correlated in the FI and AU economies. Thus, our evidence indicates that fully rational parameter learning generates additional macroeconomic risks, which interact with adjustment costs and elasticity of inter-temporal substitution, allowing us to better match the macro dynamics. These findings complement the results of Tallarini (2000), Campanale, Castro, and Clementi (2010) and Liu and Miao (2015), who find no effect of increasing agents' sensitivity to risk on the macroeconomic variables.

There are however several issues that this version of the model with parameter learning does not resolve. Although parameter learning increases the equity premium, the magnitudes are still too small compared to the historically observed statistics. The main reason for this underperformance of the model is found in countercyclical dynamics of a firm's endogenous dividends in the production economy. Therefore, we further consider pricing a claim to exogenous market dividends that are directly calibrated to reconcile dividend dynamics. In this way, we are able to verify that the low equity premium arises not because of an insufficient amplification effect of parameter uncertainty, but due to the inability of the production economy to generate procyclical dividends.

When pricing a claim to calibrated dividends, we find that the PPU model with a century-long prior learning period and unbiased prior beliefs generates an average equity premium, equity volatility, equity Sharpe ratio, risk-free rate, and a level and autocorrelation of the price-dividend ratio close to the values observed in the data. Learning provides a significant improvement in the performance of the production model relative to the FI and AU cases, which cannot match these standard asset pricing moments. Furthermore, learning generates long-lasting effects on asset prices as the size of the risk premium and its volatility remain high even after 200 years of a prior learning period. To better understand the source of the model's improvement, we look at the conditional dynamics of the key asset prices and conditional moments. We find that parameter learning generates a much stronger amplification mechanism in bad times than in good, generating countercyclical fluctuations in the conditional risk premium, volatility and Sharpe ratios that are consistent with the data.

In sum, fully rational pricing of parameter uncertainty improves the fit of the standard production economy to a large array of empirical regularities, though parameter learning alone cannot fix a common problem of countercyclical firm dividends in the frictionless economy. In order to maintain a desired feature of endogenous dividends, we consider an extension of the benchmark model that would generate a more procyclical firms' levered payouts consistent with the data. In particular, the extension incorporates the idea of costly reversibility (see, for example, Abel and Eberly (1994, Abel and Eberly (1996), Hall (2001) and Zhang (2005) among others), which means that firms face higher costs in cutting than in expanding capital stocks. Intuitively, the mechanism of investment frictions works as follows. In bad times, it is more difficult for a representative firm to reduce investment, due to higher costs that would lead to a smaller drop in investment compared to the symmetric capital adjustment cost. Thus, net profits after deducting investment appear less countercyclical. With the financial leverage, a firm's dividends are the sum of a firm's profits and the net balance of the long-term debt. The latter is proportional to capital and therefore declines in the recession. The overall sum of the profits and net issuance of the long-term debt results in procyclical dividends.

We find that the unconditional statistics of the levered returns to endogenous firm dividends are now much closer to the data. In particular, the extended model accounts for a large equity premium, around two thirds of equity volatility, and furthermore, it matches well to the mean, volatility and autocorrelation of the price-dividend ratio. The quadratic asymmetric adjustment cost function further lowers the correlations between macroeconomic variables. The results of the benchmark calibration and the extended model confirm our findings that for all relevant moments parameter learning provides a substantial improvement relative to the FI and AU cases.

The main mechanism of this paper is closely related to the work of Collin-Dufresne, Johannes, and Lochstoer (2016) who study a similar learning problem in the endowment economy. Our analysis differs from theirs in the following ways. First, we extend their methodology to a production economy setting and explore joint implications of parameter uncertainty for macroeconomic quantities and asset prices. Second, relative to the endowment model, one needs to generate procyclical dividends in the production economy to obtain a significant amplification of equity moments by parameter learning. We document this result by pricing a claim to exogenous calibrated dividends. We further confirm this finding in the extension of the model with costly reversibility, which generates endogenous procyclical firm's payoffs. Third, rather than exploring the impact of learning in a rare events model (Rietz 1988; Barro 2006), we instead estimate the production parameters by the expectation maximization algorithm from the postwar U.S. data. Even though the estimated process for productivity growth does not reflect rare states that are naturally difficult to be learned about due to their rareness, fully rational parameter learning still matches well financial moments in our setting with more frequent states. The main reason for this is that long-run consumption risks generated by consumption smoothing (Kaltenbrunner and Lochstoer 2010) magnify the impact of endogenous long-run productivity risks originating from belief revisions on asset prices; therefore, less is needed in terms of the speed of parameter learning.

Our paper also speaks to macro-finance research in the production-based economies. Cagetti et al. (2002) is one of the first examples of a business cycle model with parameter learning. In their paper, Cagetti et al. (2002) consider a signal extraction problem about the unobservable mean growth rate of technology shocks. However, they do not study the implications of incomplete information for quantities and asset prices, a key focus of our analysis. In a recent paper, Jahan-Parvar and Liu (2014) examine a production economy with learning about a latent state in a productivity growth process following a two-state hidden Markov chain. Their paper is an adaption of the endowment economy with ambiguity preferences (Ju and Miao 2012) to a production setting. The key differentiators of our study from Jahan-Parvar and Liu (2014), as well as the extant literature on learning in a business cycle model, is a multidimensional learning problem and rational pricing of parameter beliefs.

Our paper is also related to the long-run risks models introduced by Bansal and Yaron (2004). Kaltenbrunner and Lochstoer (2010) and Croce (2014) investigate the original source and implications of long-run productivity and consumption risks. In relation to these studies, we do not explicitly incorporate long-run dynamics in productivity growth by adopting the model of Bansal and Yaron (2004). In our paper, the subjective long-run macroeconomic risks appear as a result of Bayesian learning about true parameter values. Our approach is complementary to the existing long-run risks literature and in fact provides the empirical investigation of possible origins of long-run productivity risks.

The paper proceeds as follows. Section 2.2 presents the formal model, Section 2.3 investigates the quantitative implications of parameter learning for quantities and asset prices. Section 2.4 performs sensitivity analysis. Section 2.5 concludes.

2.2 The Model

In this section, we present a production-based asset pricing model (Jermann 1998; Campanale, Castro, and Clementi 2010; Croce 2014; Kaltenbrunner and Lochstoer 2010). The model is a standard business cycle framework (Kydland and Prescott 1982; Long and Plosser 1983) populated by a representative firm with Cobb-Douglas production technology and capital adjustment costs, and a representative household with Epstein-Zin preferences. The firm produces a single consumption-investment good using labor and capital as inputs subject to productivity shocks. The household participates in the production process by working for the firm and investing in capital. Additionally, the representative investor trades firm shares and risk-free bonds to maximize lifetime utility of a consumption stream subject to a sequential budget constraint. Ultimately, the representative firm maximizes its value by choosing labor and investment demand. Our objective is to investigate the impact of learning about parameters in the productivity process on the moments of macroeconomic quantities and equity returns.

2.2.1 The Representative Household

We assume that a representative household has the recursive utility of Epstein and Zin (1989):

$$U_t = \left\{ (1-\beta) C_t^{1-1/\psi} + \beta \left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}} \tag{2.1}$$

where U_t denotes the household's continuation utility, C_t denotes aggregate consumption, E_t denotes the expectation operator, $\beta \in (0,1)$ is the discount factor, $\psi > 0$ represents the elasticity of inter-temporal elasticity of substitution (EIS), $\gamma > 0$ represents the risk aversion parameter. For simplicity, we will assume that the household inelastically supplies one unit of labor; thus, the household's intra-period utility depends only on consumption.

It is straightforward to derive the stochastic discount factor:

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi} \left(\frac{U_{t+1}}{\left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}} \right)^{1/\psi - \gamma} \tag{2.2}$$

The key feature is a separation of agent's relative risk aversion from the elasticity of inter-temporal substitution. If $\gamma \neq \frac{1}{\psi}$, the utility function is not time-additive, and the stochastic discount factor has two components. The first term represents the kernel of the power utility, while the second term is the adjustment of the Epstein-Zin utility. In this paper, we set $\gamma > \frac{1}{\psi}$ and, thus, the household prefers earlier resolution of uncertainty. When the household's continuation utility U_{t+1} is below the certainty equivalent of this continuation utility, the second ingredient in the pricing kernel increases, raising a premium for long-run risks.

We also aim to investigate the impact of learning about unknown parameters governing the technology process. Although we do not introduce long-run productivity risks directly by assuming a persistent component in productivity growth (Croce 2014; Kaltenbrunner and Lochstoer 2010), Bayesian belief updating will generate subjective long-run productivity risks. Since the representative household has a preference for early resolution of uncertainty and is particularly averse to such long-run risks, parameter learning will generate quantitatively significant macroeconomic risks, improving the performance of the model in explaining salient features of the data.

2.2.2 The Representative Firm

The representative firm produces the consumption good using a constant returns to scale Cobb-Douglas production function:

$$Y_t = K_t^{\alpha}(A_t N_t)^{1-\alpha}, \tag{2.3}$$

where Y_t is the output, K_t is the capital stock, N_t is labor hours, and A_t is an exogenous, labor-enhancing technology level (which we also refer to as productivity). For simplicity, we assume that the representative household supplies the fixed amount of labor hours, which are exogenously set $N_t = 1$.

The firm's capital accumulation equation incorporates capital adjustment costs and is formally defined by:

$$K_{t+1} = (1-\delta)K_t + \varphi(I_t/K_t)K_t,$$

where $\delta \in (0,1)$ is the capital depreciation rate, $I_t = Y_t - C_t$ denotes gross investment, and $\varphi(\cdot)$ is the capital adjustment cost function given by:

$$\varphi(x) = a_1 + \frac{a_2}{1-1/\xi} x^{1-1/\xi}, \tag{2.4}$$

where ξ is the elasticity of the investment rate to Tobin's q. We follow Boldrin, Christiano, and Fisher (2001) and choose the constants a_1 and a_2 such that there are no adjustment costs in the non-stochastic steady state.[5]

2.2.3 Technology

We consider a parsimonious two-state Markov switching model for the productivity growth rate $\Delta a_t = \ln\left(\frac{A_t}{A_{t-1}}\right)$:

$$\Delta a_t = \mu_{s_t} + \sigma\varepsilon_t,$$

where $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, s_t is a two state Markov chain with transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix},$$

[5] Specifically, $a_1 = \frac{1}{\xi-1}\left(1-\delta-\exp(\bar{\mu})\right)$, $a_2 = \left(\exp(\bar{\mu})-1+\delta\right)$, where $\bar{\mu}$ is the unconditional mean of μ_{s_t}. We find steady state values of the remaining quantities from the conditions $\varphi\left(\frac{I}{K}\right) = 1$, $\varphi^{'}\left(\frac{I}{K}\right) = 1$. In particular, the steady state investment-capital ratio is $\frac{I}{K} = \exp(\bar{\mu}) - 1 + \delta$.

where $\pi_{ii} \in (0,1)$. We label $s_t = 1$ the "good" regime with high productivity growth and $s_t = 2$ the "bad" regime with low productivity growth.

2.2.4 Equilibrium Asset Prices

In the competitive equilibrium of the economy, the representative household works for the firm and trades its shares to maximize the lifetime utility over a consumption stream. The representative firm chooses labor and capital inputs (through investment) to maximizes the firm's value, the present value of its future cash flows. The firm's maximization problem implies the following equilibrium conditions for gross return $R_{j,t+1}$ of the asset j between period t and $t+1$:

$$E_t\left[M_{t+1}R_{j,t+1}\right] = 1. \tag{2.5}$$

In particular, the equation above is satisfied by the investment return, $R_{I,t+1}$, defined by:

$$R_{I,t+1} = \frac{1}{Q_t}\left[Q_{t+1}\left(1-\delta+\varphi\left(\frac{I_{t+1}}{K_{t+1}}\right)\right) + \frac{\alpha Y_{t+1} - I_{t+1}}{K_{t+1}}\right], \tag{2.6}$$

where Q_t is Tobin's marginal Q :

$$Q_t = \frac{1}{\varphi'\left(\frac{I_t}{K_t}\right)} = \frac{1}{a_2}\left(\frac{I_t}{K_t}\right)^{1/\xi}.$$

The return on investment can be interpreted as the return of an equity claim to the unlevered firm's payouts (Restoy and Rockinger 1994). As the firm behaves competitively, the labor input is chosen at a level equal to its marginal product: $w_t = \partial Y_t/\partial N_t = (1-\alpha)A_t^{1-\alpha}K_t^{\alpha}N_t^{-\alpha} = (1-\alpha)Y_t/N_t$. The unlevered firm value, FV_t, is given by $FV_t = Q_tK_{t+1}$, and the firm's unlevered dividends, D_t, are defined by:

$$D_t = Y_t - w_tN_t - I_t = \alpha Y_t - I_t. \tag{2.7}$$

Since the observed aggregate stock market dividends are not directly comparable to the endogenous payouts defined above,[6] we consider pricing levered equity claims. We introduce financial leverage in the spirit of Jermann (1998) by presuming that in

[6] As noted by other studies, unlevered cash flows and investment returns are not directly observed in reality. Additionally, the equity prices observed on the market are for leveraged corporations, in contrast to unlevered dividend payments of production companies in the model.

each period the firm issues long-term bonds for a fixed fraction of capital and pays the outstanding debt from previous periods. Note that Modigliani and Miller conditions hold in our model and, thus, introducing the financial leverage does not change the equilibrium allocations. It only influences the dynamics of a firm's payouts and the way we report the returns on a claim to the firm's dividends. In particular, the financial leverage increases volatility of dividends and makes equity returns more risky.

Following Jermann (1998), we assume that the firm issues n period discount bonds and pays back its outstanding debt of n period maturity in each period. The fraction ω of the firm's capital K_t at time t is invested in long-term bonds. Denoting the price of the n period discount bonds at time t by $B_{t,n}$ the dividends stream is given by:

$$D_t^l = Y_t - w_t N_t - I_t + \omega K_t - \omega K_{t-n}/B_{t-n,n}, \tag{2.8}$$

where the first part, $Y_t - w_t N_t - I_t$, represents the operating cash flow of an unlevered claim, whereas the second part, $\omega K_t - \omega K_{t-n}/B_{t-n,n}$, is the difference between proceeds from newly issued bonds in period t at the price $B_{t,n}$ and repayments of the bonds purchased in period $t-n$ at the price $B_{t-n,n}$. We assume that the Modigliani and Miller Theorem holds in this setting. This implies that the financial policy above does not affect a firm's value and investment decision.

The price of the n-period bonds is defined recursively by:

$$B_{t,n} = E_t\left[M_{t+1} B_{t+1,n-1}\right], \tag{2.9}$$

with the boundary condition $B_{t,0} = 1$ for any t. We denote the price of the (levered) equity claim by P_t^l, and the (levered) equity return by $R_{t+1}^l = (P_{t+1}^l + D_{t+1}^l)/P_t^l$. By (2.5) and (2.8), the equity price satisfies $P_t^l = E_t(M_{t+1}(D_{t+1}^l + P_{t+1}^l))$ and can be readily computed by the formula $P_t^l = FV_t - DV_t$, where FV_t represents a firm's value, DV_t denotes debt value of all outstanding bonds from period $t-n+1$ to period t. Specifically:

$$DV_t = \sum_{j=1}^{n} \frac{B_{t,j} \omega K_{t-n+j}}{B_{t-n+j,n}}.$$

2.3 Results

We start with calibrating a benchmark model and analyzing the implications of parameter uncertainty in the productivity growth process for macroeconomic quantities and asset returns. We focus our attention on the stylized facts observed in the U.S. post-World War II data. Specifically, we compare the model-generated statistics with the historical data for 1952:Q1-2016:Q4. Macroeconomic data on consumption, investment, capital, and output are taken from the U.S. National Income and Product Accounts (NIPA) as provided by the Bureau of Economic Analysis (BEA). The asset returns data and dividends are from the Center for Research in Security Prices (CRSP). The model is calibrated at a quarterly frequency.

Since the model does not admit an analytical solution, we solve for equilibrium allocations numerically through value function iteration. We extend the Collin-Dufresne, Johannes, and Lochstoer (2016) solution methodology to the production-based setting. The detailed description of the numerical methods is presented in the Appendix. Having solved the model, we generate 1,000 simulations of the economy with the sample length of 260 periods and report statistics of asset returns and macroeconomic quantities corresponding to their empirical counterparts.

2.3.1 Parameter Values

Panel A in Table 2.1 reports the parameter values of an investor's preferences, production and capital adjustment cost functions. We choose these parameter values similarly to the existing real-business cycle models. In particular, the constant capital share in a Cobb-Douglas production function (α) is 0.36, and the quarterly depreciation rate (δ) is 0.02. We set the capital adjustment cost parameter (ξ) equal to 4, which yields volatility of investment growth and consumption growth relatively close to the data. We further choose the constants (a_1, a_2) in the capital adjustment cost function such that there are no adjustment costs in the non-stochastic steady state.

The preference parameters are also consistent with the macroeconomic literature. The coefficient of relative risk aversion (γ) is equal to 10, the upper bound of an interval considered plausible by Mehra and Prescott (1985). The subjective discount factor (β) is set to 0.9945. This value allows the benchmark calibration to generate the low unconditional risk-free rate. There is no consensus in the literature about the value of the elasticity of inter-temporal substitution. We follow the disaster risk literature (Gourio 2012) and

Table 2.1
Benchmark Calibration

Parameter	Description	Value
Panel A: Preferences, Production and Capital Adjustment Costs Functions, and Financial Leverage		
β	Discount factor	0.9945
γ	Risk aversion	10
ψ	EIS	2
α	Capital share	0.36
δ	Depreciation rate	0.02
ξ	Adjustment costs parameter	4
a_1	Normalization	-0.0075
a_2	Normalization	0.3877
Panel B: Markov-switching Model of Productivity Growth		
π_{11}	Transition probability from expansion to expansion	0.947
π_{22}	Transition probability from recession to recession	0.662
μ_1	Productivity growth in expansion	0.54
μ_2	Productivity growth in recession	-1.53
σ	Productivity volatility	1.36

This table reports the parameter values in the benchmark calibration. Panel A presents preferences parameters, values in the production and adjustment costs functions. Panel B shows the maximum likelihood estimates of parameters in a two-state Markov-switching model for productivity growth. We obtain these estimates by applying the expectation maximization algorithm (Hamilton 1990) to quarterly total factor productivity growth rates from 1952:Q1 to 2016:Q4.

long-run risks models (Bansal and Yaron 2004; Ai, Croce, and Li 2013; Bansal et al. 2014) by setting EIS (ψ) to 2.

Following the methodology of Stock and Watson (1999), we use the macroeconomic data to construct the cumulative Solow residuals. We further scale these residuals by the labor share $(1-\alpha)$ in order to interpret them as labor-augmenting technology. We estimate a two-state Markov switching process of quarterly productivity growth rates by applying the expectation maximization algorithm developed by Hamilton (1990). Panel B in Table 2.1 reports the maximum likelihood estimates for the transition probabilities (π_{ii}), productivity growth rates (μ_i) as well as the constant volatility (σ). Productivity is estimated to grow at the quarterly rate of about 0.54 percent in expansions and about -1.53 percent in recessions. The productivity volatility comes out around 1.36 percent. The transition probability to the expansion (recession) conditional on being in the expansion (recession) is estimated around 0.947 (0.662). These numbers imply the average duration of the high-growth expansion state of about 18.87 quarters and the average duration of the low-growth recession of about 2.96 quarters. Our maximum likelihood estimates are broadly consistent with the values reported by Hamilton (1989) and Cagetti et al. (2002).

Once we solve the model according to the calibration above, we introduce financial leverage by assuming the representative firm issues long-term bonds with a maturity of fifteen years. Because the Modigliani-Miller theorem holds, this only changes the levered returns and dividends but does not influence the equilibrium allocation of the economy. With financial leverage, equity value depends on the market value of the firm and the total debt oustanding bonds. All valuations endogenously depend on the equilibrium investment decision. For each model, we calibrate (ω) in order to match the average debt-to-equity ratio of around 1:1. Therefore, the leverage parameter across different models is in the interval $[1\%, 1.1\%]$.

2.3.2 Parameter Uncertainty

In this paper, we consider five parameters in the productivity growth process. We employ conjugate priors for each unknown parameter in order to obtain conjugate posteriors via Bayesian updating. If all parameters are assumed to be unknown for the agent, we obtain a 10-dimensional vector of state variables including the current regime of the Markov chain, capital stock, time and hyperparameters of prior distributions. In addition to the curse of dimensionality, the numerical solution methodology in the production-based setting requires the solution of the agent's maximization problem for every combination of state variables in each period. This makes the model solution especially slow. To mitigate complexity in the model solution, we investigate the impact of uncertainty about the transition probabilities and mean growth rates, whereas a volatility parameter is assumed to be known.[7] Furthermore, our analysis assumes homoskedastic volatility of productivity growth, though a large strand of the macroeconomic literature documents the importance of time-varying uncertainty on macroeconomic variables and asset returns. We leave the important investigation of the implications of learning about volatility risk and regime switches in volatility of productivity growth for future research.

Having decided which parameters are unknown for the agent in the production economy, we consider the two approaches to dealing with parameter uncertainty: priced parameter uncertainty and anticipated utility. PPU implies the economic agents learn about unknown parameters from the data and rationally take into account the changing beliefs while making their decisions. AU assumes the decision-makers learn about un-

[7]We motivate our choice of unknown parameters by the results in the consumption-based asset pricing model. Specifically, Collin-Dufresne, Johannes, and Lochstoer (2016) conclude that uncertainty about variance has a negligible effect on asset prices. Meanwhile, learning about transition probabilities has long-lasting asset pricing implications. Mean growth rates are harder to learn than volatilities, though the implications are less pronounced compared to learning about unknown transition probabilities.

known parameters over time but in each period of time they treat their current beliefs as "true" values. Thus, AU agents ignore the possibility that parameters might actually change in the future. Since the posterior beliefs are martingales, Bayesian learning generates subjective long-run risks in the economy, which would be priced under rational beliefs pricing, unlike the AU case. To evaluate the impact of these additional macroeconomic risks, we consider different specifications of the production economy for the comparison analysis. We start by solving three frameworks: our preferred benchmark with full information about parameters in the productivity growth process and an identical calibration with unknown parameters incorporating either PPU or AU. This comparison allows us to isolate the impact of rational pricing of parameter uncertainty. Furthermore, we study the role of an investor's prior knowledge by injecting training samples with different lengths into the model. By experimenting with the prior samples, we can evaluate how persistent the impact of rational beliefs updating is.

In sum, we use standard, conjugate priors distributions for the unknown parameters: beta and normal distributions for the transition probabilities and mean growth rates, respectively. We choose the hyperparameters of the distributions such that initial beliefs are centered at the true values of uncertain parameters estimated from the postwar sample. Furthermore, we solve the models with parameter uncertainty based on the prior training samples of 100, 150 or 200 years of initial learning. Thus, although calibrating initial beliefs based on the historical data may be a realistic feature and will certainly improve the model's performance due to pessimism induced by the Great Depression and both World Wars, our results do not require pessimistic prior specifications and are based on the information contained in the postwar data. For each specification, we numerically solve the production economy using the methodology outlined in the Appendix.

2.3.3 Pricing a Claim to Firm's Levered Dividends

In this section, we quantitatively analyze the impact of uncertainty about the transition probabilities (π_{11}, π_{22}) and mean growth rates (μ_1, μ_2) in the production economy of this paper.[8] The macroeconomic variables of our interest are consumption, investment and output. The financial variables include short-term (one quarter) and long-term (15

[8]The Appendix provides extensive details of the numerical solution methodology in different settings. The model-generated statistics with uncertainty about the transition probabilities are similar to the results presented in the main text with both unknown probabilities and mean growth rates. For brevity, these results are not reported but are available upon request.

years) risk-less bonds and equity claim on the firm's leveraged dividends. First, we assess the implications of parameter uncertainty in the production economy by comparing model-implied unconditional moments of quantities and asset returns to their sample counterparts. Second, we study the impulse responses of quantities to a regime switch in the productivity growth. Finally, we check the ability of the economy to reproduce the long-horizon predictability of excess equity returns.

Unconditional Moments

Panel A in Table 2.2 presents business cycle moments of macroeconomic variables from simulations of models as well as the U.S. post WWII statistics. The data column shows that output is more volatile than consumption but less volatile than investment. Also, there is a significant correlation between the three series, especially between investment and output growth. Comparing the empirical moments with the model-generated statistics, all three models with PPU, AU and fixed parameters explain the empirical moments reasonably well. Relative to the case of known parameters and AU, rational pricing of beliefs with 100 years of prior learning slightly increases investment growth volatility, lowers consumption growth volatility and brings the correlations between the macro quantities closer to the data. However, parameter learning has quantitatively marginal effects on the macro dynamics.

In contrast, Panel B in Table 2.2 shows that priced parameter uncertainty improves more significantly the performance of the real business cycle model in terms of financial moments. The last two columns in Table 2.2 show that the production economy with known parameters, or with unknown parameters but AU pricing (for the AU case, we report the results only with a prior period of 100 years), generates a too high average risk-free rate and price-dividend ratio as well as a too low mean and volatility of excess equity returns compared to the data. Columns 3 to 6 shows that rationally taking into account parameter uncertainty in the productivity growth process leads to a lower risk-free rate and price-dividend ratio. The risk premium is almost two times higher with parameter learning, equity volatility and the price of risk also increase in this case.

Although the financial moments are amplified in the model with priced parameter uncertainty, they are still too small compared to the data. The reason for a very small equity premium and equity volatility in the production economy is the countercyclical dynamics of dividends growth as documented by Kaltenbrunner and Lochstoer (2010)

Table 2.2
Sample Moments

	Data	PPU				AU	FI
		100 yrs	150 yrs	200 yrs	∞ yrs		
Panel A: Macroeconomic Quantities							
$\sigma(\Delta c)$	1.26	1.30	1.32	1.32	1.33	1.31	1.33
$\sigma(\Delta i)$	4.51	3.51	3.46	3.45	3.47	3.47	3.44
$\sigma(\Delta y)$	2.41	1.97	1.96	1.96	1.96	1.94	1.94
$ar1(\Delta c)$	0.32	0.14	0.16	0.17	0.19	0.19	0.18
$\rho(\Delta i, \Delta y)$	0.72	0.97	0.98	0.99	0.99	0.99	0.99
$\rho(\Delta c, \Delta y)$	0.52	0.96	0.97	0.98	0.99	0.98	0.98
$\rho(\Delta c, \Delta i)$	0.36	0.90	0.92	0.94	0.96	0.95	0.95
Panel B: Financial Variables							
$E(R_f) - 1$	1.44	1.64	1.78	1.87	2.01	2.13	2.14
$\sigma(R_f)$	1.07	0.41	0.38	0.37	0.34	0.31	0.31
$\sigma(M)/E(M)$		0.29	0.27	0.25	0.23	0.19	0.19
$E(\Delta d^l)$	2.06	0.65	0.77	0.84	0.95	1.07	1.08
$\sigma(\Delta d^l)$	10.38	13.35	13.74	14.20	15.77	17.09	16.19
$ar1(\Delta d^l)$	0.25	−0.01	−0.01	−0.01	−0.01	−0.01	−0.02
$\rho(\Delta c, \Delta d^l)$	0.44	−0.53	−0.58	−0.61	−0.63	−0.62	−0.60
$E(R^l - R_f)$	5.51	2.92	2.60	2.25	1.84	1.56	1.52
$\sigma(R^l - R_f)$	16.55	5.41	5.25	4.83	4.49	4.41	4.38
$E(p^l - d^l)$	3.19	3.38	3.45	3.55	3.65	3.67	3.67
$\sigma(p^l - d^l)$	0.33	0.31	0.31	0.32	0.35	0.36	0.36
$ar1(p^l - d^l)$	0.97	0.95	0.95	0.95	0.95	0.95	0.95

This table reports the average moments from 1,000 simulations of 260 quarters of the data from the production economy considered in this paper, where the transition probabilities and mean growth rates are assumed to be unknown. The historical data moments are reported in the data column and correspond to the U.S. data from 1952:Q1 to 2016:Q4. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. The FI column presents the results of the full information case where the parameters are known. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All statistics are expressed in annualized terms, except for market price of risk given in percent, whereas correlations and autocorrelations are expressed in quarterly terms.

among others. This is in contrast to the observed procyclical dividends in the data. Panel B in Table 2.2 indicates that a firm's payouts are strongly negatively correlated with consumption growth in all models, while we document that the corresponding correlation between consumption and dividend growth is about 0.44 in the data. A number of studies (Uhlig 2007; Belo, Lin, and Bazdresch 2014; Favilukis and Lin 2016) introduce wage rigidity in the standard production model in order to generate more volatile and procyclical dividends. This extension of the model can further improve our results and possibly magnify the effect of parameter learning. However, we leave the investigation of the interplay between sticky wages and parameter uncertainty for future research. In this paper, we will directly calibrate the firm's dividends process to the empirical counterpart.

Impulse Response Functions

Figure 2.1 illustrates the response of the economy with unknown transition probabilities to a typical recession lasting for 1 quarter, 3 quarters and 2 years. The economy is assumed to grow at the mean growth μ_1 and μ_2 in each state. Before the economy enters the recession, the representative investor holds unbiased beliefs about the uncertain parameters (the transition probabilities π_{11} and π_{22}) assuming a 100-year prior period. We feed these simulated paths of beliefs and productivity growth series into the model and calculate the equilibrium quantities as described in Appendix B.3.

The top panels in Figure 2.1 show the mean beliefs about the transition probabilities. Upon the onset of the recession, the mean belief about staying in the good regime falls sharply and stays at the same level during the recession. Once the economy returns back to the high growth state, the investor gradually updates his beliefs about π_{11} upward. In contrast, learning about π_{22} happens only in the recession. The random durations of 1 quarter, 3 quarters and 2 years correspond to the realization of a short economic decline, an average recession and a long downturn, respectively. When the agent experiences an average duration of the recession, his belief about π_{22} increases but then returns back to the initial value. The mean belief about π_{22} remains permanently lower (higher) relative to the initial belief in the case of the recession that is shorter (longer) than the average downturn.

The middle panels of Figure 2.1 present the impulse responses of macroeconomic quantities and equity prices. Given that productivity growth declines and the investor's probability beliefs about π_{11} drop, the capital stock declines upon the bad news in the economy and consequently leads to a reduction in investment and consumption. As productivity stays low and probability beliefs become more pessimistic, macroeconomic variables continue to fall and start to recover only after the economy exits the recession. The stock prices fall in response to switching to the low productivity growth regime. Also, the realized equity returns are smaller in the recession due to low productivity growth and bounce back to the original rate in the expansion. Since consumption dynamics predicts the high marginal utility when productivity growth is low, equity returns are positively

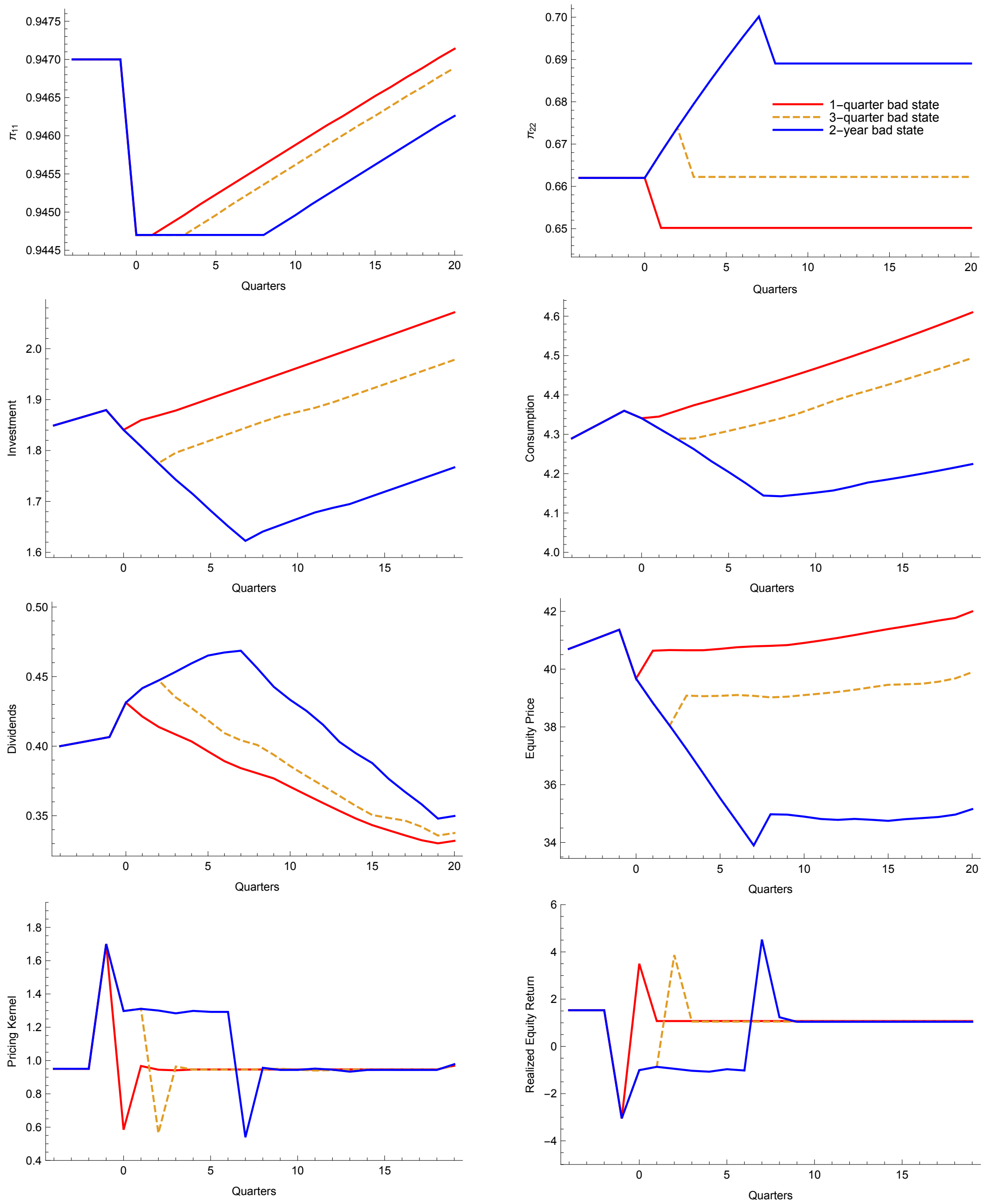

Figure 2.1: Beliefs, Impulse Responses and Recession Realizations. The figure shows the dynamics of investor's beliefs and the impulse responses of macroeconomic and financial variables to a typical recession in the model with unknown transition probabilities. Initially, the economy stays in the expansion for a long period, and the investor holds unbiased mean beliefs about the transition probabilities (π_{11} and π_{22}) based on a 100-year prior period. The panels show three simulated paths with the duration of the recession state equal to 1 quarter, 3 quarters and 2 years.

exposed to the regime switching in mean productivity growth. Overall, the model predicts the dynamics of consumption, investment, equity prices and equity returns consistent with the data.

Turning to endogenous firm's payouts, the dividends increase following the recession realization, in contrast to a procyclical movement observed in the data. The unlevered dividends, which are not reported in Figure 2.1, are approximately equal to profits minus investment. Since profits in the model are smooth relative to investment and investment is procyclical, the endogenous unlevered firm's dividends are strongly countercyclical and would initially increase on impact and then grow at the original rate. Similarly, the levered firm's dividends reported in Figure 2.1 increase upon entering the recession but start to decline when the economy returns to the growth state. The reason is that a sharp decline in the mean beliefs about π_{11} has a negative and long-lasting impact on the capital stock. Since the agent invests a constant proportion of the firm's capital in the long-term bonds, the current profits from trading the long-term bonds remain negative until capital recovers to the initial level.

Return Predictability

A large strand of the empirical literature documents that excess returns at an aggregate level can be predicted by variables like the investment-capital ratio (Cochrane 1991; Bansal and Yaron 2004), Tobin's Q (Pontiff and Schall 1998; Lewellen 2004), the dividend-price ratio (Campbell and Shiller 1988; Fama and French 1989) and the consumption-wealth ratio (Lettau and Ludvigson 2001). In this section, we compare the long-term predictability patterns generated by the production economy with parameter uncertainty (both the PPU and AU cases) and fixed parameters to the predictability observed in the post-war data. The conclusion of the extensive empirical literature is that high dividend yields, high book-to-market and consumption-wealth ratios predict high future excess returns, whereas high investment rates forecast low future excess returns. Furthermore, the predictive regressions suggest that the slope coefficients (in absolute terms) and R^2's are relatively large and tend to increase over the forecast horizon. These regularities pose a significant challenge for the standard real business cycle model.

Tobin's Q, the investment-capital and consumption-wealth ratios are endogenously specified in our production economy. Furthermore, we follow Epstein and Zin (1989) and calculate the wealth-consumption ratio as:

$$\frac{W_t}{C_t} = \frac{1}{1-\beta}\left(\frac{U_t}{C_t}\right)^{1-1/\psi},$$

where the equilibrium allocations of the agent's utility and consumption are endogenously determined. Using these model-generated quantities, we run the abovementioned predictive regressions and report results in Table 2.3. We find that all three models can generate monotonicity in the slope coefficients and R^2's over the forecast horizon. Furthermore, the model with priced parameter uncertainty produces dramatically larger (in absolute terms) slopes and R^2's relative to the AU approach and especially to the model with known parameters.

2.3.4 Pricing a Claim to Calibrated Dividends

In the previous section, we studied the implications of parameter uncertainty for the dynamics of macroeconomic quantities and equity returns. We introduced financial leverage in the spirit of Jermann (1998) in order to make the equity return more risky. The main drawback of financial leverage considered in the previous section is that the leveraged firm's dividends still remained significantly procyclical in the model and, thus, the equity premium and its volatility were too small compared to the data. Furthermore, following the discussion of Kaltenbrunner and Lochstoer (2010), one can argue that the aggregate stock market dividends are only a small part of the payouts of the productive sector and, thus, cannot be directly interpreted as the firm's dividends in our model. Therefore, we follow a consumption-based asset pricing literature by directly calibrating an exogenous dividend process to replicate the stock market dividends.

Following Bansal and Yaron (2004), we price a levered consumption claim with a leverage factor λ. We formally define quarterly log dividend growth as follows:

$$\Delta d_t^M = g_d + \lambda \Delta c_t + \sigma_d \varepsilon_t^d, \tag{2.10}$$

where $\varepsilon_t^d \overset{\text{iid}}{\sim} N(0,1)$, g_d and σ_d are the dividend growth rate and volatility, respectively. We calibrate the parameters g_d, σ_d, and λ to make model implied statistics of dividend growth consistent with the historical data. Panel B in Table 2.1 reports the parameter values in an exogenous dividend stream. We set the mean adjustment (g_d) and the idiosyncratic dividend volatility (σ_d) to match the observed annual mean growth (2.06 percent) and volatility (10.38 percent) of dividends for the considered period. The leverage parameter (λ) is equal to 3.5, a midpoint of the range from 2.5 to 4.5 used in other studies.

Let R_{t+1}^M denote the return on a claim delivering stochastic dividends given by (2.10). Then:

$$R_{t+1}^M = \frac{P_{t+1}^M + D_{t+1}^M}{P_t^M} = \frac{P_{t+1}^M / D_{t+1}^M + 1}{P_t^M / D_t^M} \cdot \frac{D_{t+1}^M}{D_t^M}.$$

Table 2.3
Return Predictability

h	Data Slope	Data R^2	PPU Slope	PPU R^2	AU Slope	AU R^2	FI Slope	FI R^2
Panel A: Investment-capital ratio $(i-k)$								
1Y	−0.394	0.071	−0.504	0.037	−0.108	0.043	−0.090	0.035
2Y	−0.603	0.123	−0.910	0.071	−0.198	0.079	−0.165	0.065
3Y	−1.293	0.245	−1.287	0.106	−0.280	0.114	−0.236	0.095
4Y	−1.831	0.333	−1.639	0.138	−0.362	0.150	−0.307	0.126
5Y	−2.453	0.372	−1.970	0.163	−0.439	0.182	−0.376	0.155
Panel B: Tobin's Q								
1Y	−0.335	0.078	−0.493	0.052	−0.434	0.042	−0.362	0.035
2Y	−1.573	0.133	−0.875	0.096	−0.792	0.079	−0.661	0.065
3Y	−1.952	0.165	−1.228	0.137	−1.123	0.114	−0.945	0.095
4Y	−2.443	0.192	−1.578	0.179	−1.452	0.150	−1.232	0.126
5Y	−2.815	0.231	−1.907	0.214	−1.758	0.181	−1.507	0.155
Panel C: Dividend-price ratio (d^l-p^l)								
1Y	0.083	0.041	0.024	0.030	0.013	0.018	0.011	0.016
2Y	0.122	0.055	0.040	0.051	0.023	0.033	0.018	0.029
3Y	0.175	0.074	0.055	0.070	0.030	0.045	0.024	0.040
4Y	0.212	0.093	0.069	0.089	0.037	0.055	0.029	0.050
5Y	0.227	0.105	0.080	0.103	0.042	0.063	0.034	0.058
Panel D: Consumption-wealth ratio $(c-w)$								
1Y	3.173	0.086	2.050	0.085	1.788	0.071	2.110	0.039
2Y	5.944	0.182	3.429	0.138	3.195	0.124	3.717	0.067
3Y	7.845	0.274	4.620	0.184	4.412	0.171	5.159	0.093
4Y	9.352	0.297	5.821	0.229	5.627	0.217	6.659	0.122
5Y	11.134	0.327	6.933	0.269	6.747	0.259	8.110	0.148

This table reports univariate regressions of cumulative excess log equity returns on several valuation and macroeconomic variables over various forecasting horizons (h years; 1 to 5). We use investment-capital ratio, Tobin's Q, dividend-price and consumption-wealth ratios as the right-hand side variable (x_t) in the linear projection:

$$r^{ex}_{t+1\to t+h} = \text{Intercept} + \beta(h) \times x_t + \varepsilon_{t+h},$$

where $r^{ex}_{t+1\to t+h}$ are h-year future excess log equity returns. The empirical statistics are for the U.S. data from 1952:Q1 to 2016:Q4. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. The FI column presents the results of the full information case where the parameters are known. For each model, we simulate 1,000 economies at a quarterly frequency with a sample size equal to the empirical counterpart. We obtain the slope coefficients and R^2's for each simulation and report average sample statistics over all 1,000 artificial series.

Substituting this expression into the equilibrium condition (2.5), the price-dividend ratio of a claim on the aggregate stock market dividends satisfies the equation:

$$\frac{P_t^M}{D_t^M} = E_t \left[M_{t+1} \left(1 + \frac{P_{t+1}^M}{D_{t+1}^M} \right) \frac{D_{t+1}^M}{D_t^M} \right]. \tag{2.11}$$

Unconditional Moments

Now we take a closer look at the equity claim paying stochastic dividends as a leverage on consumption similarly to Bansal and Yaron (2004). The numerical methods used to solve for the equilibrium price-dividend ratio are presented in the Appendix. Table 2.4 shows the model-implied statistics of dividend growth, excess equity returns, the Sharpe ratio and the price-dividend ratio.

The calibrated dividends closely replicate the empirical first and second moments as well as a positive correlation between dividends and consumption observed in the data. Our conservative choice of a leverage parameter produces a slightly lower correlation between dividend and consumption growth rates, but it is crucial that the correlation remains positive in all models. Turning to equity moments, parameter uncertainty with AU pricing produces similar results to the production model with known parameters. Relative to the FI and AU cases, a priced parameter uncertainty approach significantly improves the fit of the model with the data. The model with parameter uncertainty and a prior sample of learning of 100 years match the sample equity premium, its volatility, the equity Sharpe ratio and the level of the price-dividend ratio well. Furthermore, the volatility of the price-dividend ratio comes out two to three times its value with fixed parameters, though it still remains lower than in the data. In the data, the log price-dividend ratio is highly persistent and the model with parameter learning reconciles this feature. Furthermore, looking at the results based on different training samples, one can see that Bayesian learning and rational pricing of an investor's subjective beliefs generates permanent shocks in the production economy.

It is important to stress that the implications of parameter learning in the production-based setting are based on a productivity growth process that is estimated over the post-war data. Even though the parameter estimates in our model reflect the business cycle fluctuations rather than rare and bad macroeconomic events, learning about the true productivity growth process has significant quantitative effects. This is mainly due to the fact that the impact of endogenous long-run risks originating from belief revisions is magnified by long-run risks in consumption growth through consumption smoothing, as

Table 2.4
Calibrated Stock Market Dividend Claim

	Data	PPU				AU	FI
		100 yrs	150 yrs	200 yrs	∞ yrs		
$E(\Delta d^M)$	2.06	1.89	1.80	1.75	1.68	1.60	1.60
$\sigma(\Delta d^M)$	10.38	11.48	11.51	11.52	11.54	11.54	11.56
$ar1(\Delta d^M)$	0.25	0.01	0.01	0.01	0.01	0.01	0.01
$\rho(\Delta c, \Delta d^M)$	0.44	0.32	0.32	0.32	0.32	0.32	0.32
$E(R^M - R_f)$	5.51	5.92	5.18	4.64	3.75	2.88	2.80
$\sigma(R^M - R_f)$	16.55	15.90	15.43	15.18	14.78	14.81	14.42
$SR(R^M - R_f)$	0.33	0.33	0.27	0.24	0.20	0.15	0.15
$E(p^M - d^M)$	3.19	3.15	3.28	3.37	3.54	3.76	3.78
$\sigma(p^M - d^M)$	0.33	0.07	0.05	0.04	0.03	0.05	0.02
$ar1(p^M - d^M)$	0.97	0.90	0.87	0.84	0.79	0.90	0.76

This table reports the average moments from 1,000 simulations of 260 quarters of the data from the production economy considered in this paper, where the transition probabilities and mean growth rates are assumed to be unknown. As in Bansal and Yaron (2004), equity is a claim to an exogenous dividend stream. The historical data moments are reported in the data column and correspond to the U.S. data from 1952:Q1 to 2016:Q4. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. The FI column presents the results of the full information case where the parameters are known. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All statistics are expressed in annualized terms, except for correlations and autocorrelations expressed in quarterly terms.

documented by Kaltenbrunner and Lochstoer (2010). In the consumption-based setting, in order to match the financial moments one needs to add either learning about rare events observed in the pre-war data (Collin-Dufresne, Johannes, and Lochstoer 2016) or a more complex learning process (Johannes, Lochstoer, and Mou 2016). Furthermore, given *confounding* effects[9] documented by Johannes, Lochstoer, and Mou (2016), we expect that learning additionally about volatility risks and especially introducing a multidimensional learning problem with model, state and parameter uncertainty are expected to slow down the speed of learning and, thus, will improve our results. Finally, adding a more rare state into the productivity growth is expected to amplify the impact of parameter uncertainty, following the results of Collin-Dufresne, Johannes, and Lochstoer (2016). We view the investigation of these Bayesian approaches as an interesting avenue for future research.

[9] *Confounding* effectively means that uncertainty about one variable makes learning about another variable more difficult.

Conditional Dynamics

Figure 2.2 plots the responses of several key variables to a bad state realization, that lasts for 1 quarter, 3 quarters and 2 years. The sharp decline in beliefs about the probability of staying in the good state leads to a reduction in the interest rate, a decline in the price-dividend ratio as well as an increase in the risk premium and equity volatility. As long as the economy stays in the low productivity growth regime, the agent learns about the persistence of the bad state by revising his beliefs upward. During this period, the interest rates are low, the price-dividend ratio keeps declining, while the equity Sharpe ratio, the conditional equity premium and volatility remain elevated. Although both AU and PPU pricing predict similar paths of financial variables in response to a negative long-run risk shock to the expected productivity growth, the magnitude of their responses is substantially different.

For the anticipated utility case, one can observe very moderate responses in the returns, prices and conditional moments upon the onset of the bad state. Before the regime switch, both the conditional equity premium and the conditional Sharpe ratio are too low relative to the data and then they approximately double in response to the negative shock. Meanwhile, the conditional equity volatility increases only marginally in this case. In contrast, rationally priced parameter uncertainty predicts around 6-fold and 3-fold increases in the conditional risk premium and the equity Sharpe ratio, respectively. The equity volatility turns out to be highly countercyclical as it increases by a factor of about 2.5. The interest rate drops more in bad times with parameter uncertainty, while the realized equity returns are more volatile. The price-dividend ratio experiences about the same percentage decline upon the realization of the low productivity growth state in both cases. However, the level of the price-dividend ratio is substantially higher with AU, while parameter learning generates reasonable levels of the price-dividend ratio.

2.3.5 Adding Costly Reversibility

In the previous section we demonstrated that parameter learning and rational pricing are able to reproduce salient features of the macroeconomic quantities and equity returns, as long as the dividends exhibit a positive correlation with business cycle. However, it is important to maintain the endogeneity of the dividend process within the model. To fix this issue, we present an extension of the model where we include investment frictions in the form of costly reversibility. This will endogenously generate more procyclical

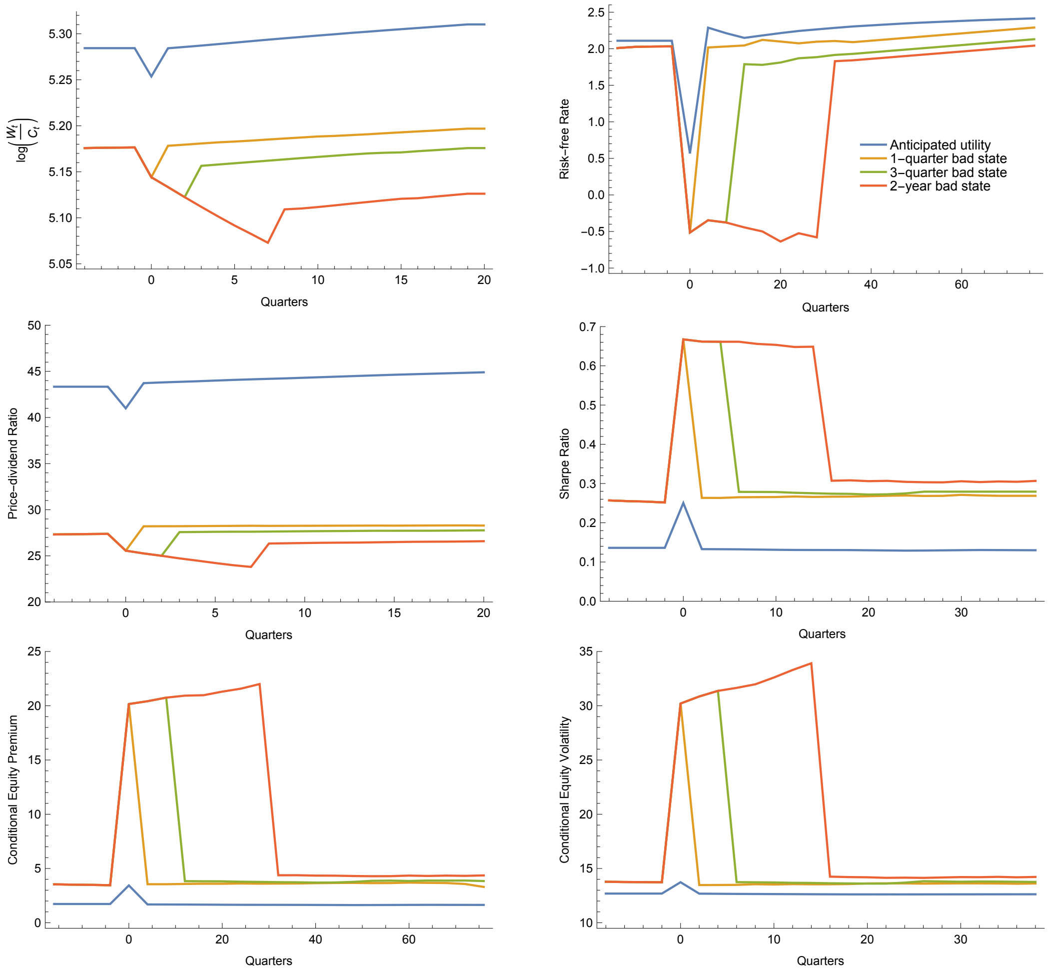

Figure 2.2: Conditional Prices and Moments. This figure shows the conditional risk-free rate, the price-dividend and equity Sharpe ratios, as well as the conditional equity premium and its volatility. The simulated variables are impulse response functions to the realization of a bad state of 1 quarter, 3 quarters and 2 years in the production economy, considered in this paper for the case of a 100-year prior. The economy is assumed to stay in the high productivity growth steady-state for a long period, and the representative agent holds unbiased initial mean beliefs. We report the conditional dynamics of the variables for the AU and PPU cases. For the sake of a convenient exposition, the former one includes only the responses to a 1-quarter bad state realization. The Appendix describes the numerical approach used.

dividends consistent with the data.

Formally, we model costly reversibility by adopting the asymmetric capital adjustment cost function, which takes a quadratic form:

$$\varphi(x_t) = x_t - \frac{\theta_t}{2} \cdot (x_t - x_0)^2 ,$$

where

$$\theta_t = \theta^+ \cdot \mathbb{I}(x_t \geq x_0) + \theta^- \cdot \mathbb{I}(x_t < x_0)$$

and $\mathbb{I}(\cdot)$ denotes the indicator operator that equals 1 if the condition is satisfied and 0 otherwise. We choose the constant x_0 such that there are no adjustment costs in the non-stochastic steady state, which implies $x_0 = \exp(\bar{\mu}) - 1 + \delta$. The remaining two parameters θ^+ and θ^- satisfy the condition $0 < \theta^+ < \theta^-$ to capture the idea of costly reversibility: the representative firm faces higher capital adjustment costs for the investment decisions leading to the capital stock being below a non-stochastic steady state value. In the quantitative exercise here, we calibrate the parameters θ^+ and θ^- consistent with the literature. The empirical estimates of θ^+ vary from 2 to 8. We choose a middle point of this range and set $\theta^+ = 5$ as in Zhang (2005). We further follow Zhang (2005) by assuming the degree of asymmetry equal $\theta^-/\theta^+ = 10$ that would imply $\theta^- = 50$.

Table 2.5 shows the results of the model with the asymmetric adjustment cost calibrated above and other parameters fixed at the values in Table 2.1. As shown in Panel A of Table 2.5, the model generates volatility of macroeconomic quantities relatively close to the data, though consumption is slightly more volatile and investment is smoother compared to the case with a convex adjustment cost function. Also, the quadratic adjustment cost better matches the comovements between macroeconomic variables than the convex adjustment cost.

Panel B summarizes the model-generated statistics of financial variables. Our calibration with investment reversibility predicts dividend dynamics quite similar to the data. Most importantly, the correlation between consumption and dividends becomes slightly positive, which in turn has a large impact on equity returns. In particular, the unconditional risk premium compares quite well with the sample estimate. Even though the excess volatility puzzle remains unresolved, the rationally priced parameter uncertainty magnifies the unconditional second moment of excess equity returns compared to the full information case and, in turn, explains around two thirds of the equity volatility in the data. Further, the mean, volatility and autocorrelation of the log price-dividend ratio compare surprisingly well to the observed point estimates. The introduction of additional channels such as, for example, a combination of wage rigidity and a constant elasticity of substitution (CES) production function (Favilukis and Lin 2016) or learning about time-varying volatility risks can further improve the model performance; however, we leave a rigorous investigation of a more complex model for future research.

Table 2.5
Sample Moments: The Extended Model with Costly Reversibility

	Data	PPU				AU	FI
		100 yrs	150 yrs	200 yrs	∞ yrs		
Panel A: Macroeconomic Quantities							
$\sigma(\Delta c)$	1.26	1.59	1.60	1.61	1.62	1.61	1.65
$\sigma(\Delta i)$	4.51	3.31	3.29	3.27	3.25	3.30	3.23
$\sigma(\Delta y)$	2.41	1.97	1.97	1.97	1.97	1.97	1.97
$ar1(\Delta c)$	0.32	0.18	0.18	0.18	0.18	0.18	0.18
$\rho(\Delta i, \Delta y)$	0.72	0.90	0.90	0.90	0.90	0.89	0.88
$\rho(\Delta c, \Delta y)$	0.52	0.85	0.85	0.85	0.85	0.85	0.85
$\rho(\Delta c, \Delta i)$	0.36	0.55	0.55	0.55	0.55	0.52	0.49
Panel B: Financial Variables							
$E(R_f) - 1$	1.44	1.64	1.77	1.85	1.98	2.02	2.03
$\sigma(R_f)$	1.07	0.59	0.55	0.53	0.49	0.44	0.44
$\sigma(M)/E(M)$		0.28	0.26	0.25	0.21	0.19	0.19
$E(\Delta d^l)$	2.06	0.77	0.88	0.95	1.05	1.14	1.13
$\sigma(\Delta d^l)$	10.38	9.98	10.77	11.34	12.55	15.12	14.40
$ar1(\Delta d^l)$	0.25	−0.06	−0.06	−0.06	−0.06	−0.06	−0.06
$\rho(\Delta c, \Delta d^l)$	0.44	0.01	0.01	0.02	0.03	0.03	0.07
$E(R^l - R_f)$	5.51	5.71	5.14	4.53	3.59	3.50	3.29
$\sigma(R^l - R_f)$	16.55	11.04	10.73	9.88	8.56	9.80	8.93
$E(p^l - d^l)$	3.19	2.89	2.97	3.06	3.23	3.29	3.32
$\sigma(p^l - d^l)$	0.33	0.26	0.26	0.26	0.26	0.30	0.29
$ar1(p^l - d^l)$	0.97	0.92	0.92	0.92	0.92	0.92	0.92

This table reports the average moments from 1,000 simulations of 260 quarters of the data from the production economy considered in this paper, where the transition probabilities and mean growth rates are assumed to be unknown. The historical data moments are reported in the data column and correspond to the U.S. data from 1952:Q1 to 2016:Q4. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. The FI column presents the results of the full information case where the parameters are known. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All statistics are expressed in annualized terms, except for market price of risk given in percent, whereas correlations and autocorrelations are expressed in quarterly terms.

2.4 Sensitivity Analysis

We examine the sensitivity of our results to the choice of the two parameters (ψ, ξ), which determine how the agent is willing to substitute consumption intertemporally and how the capital stock can be adjusted over time. These channels are the two natural candidates to influence the propagation of and the interplay between the productivity shocks and subjective long-run risks due to Bayesian learning in our model. Table 2.6

presents a two-part sensitivity analysis by decreasing the EIS and considering different capital adjustment cost parameters, while keeping other values as in the benchmark calibration. For convenience, we report the results of the simulations for the setting with PPU and AU pricing based on a 100-year prior period.

The elasticity of intertemporal substitution is an important parameter for matching the moments of macroeconomic variables as shown in Panel A of Table 2.6. For both PPU and AU, a lower EIS reduces the volatility of investment growth and makes consumption growth more volatile relative to the benchmark calibration. This is due to the fact that the investor is less willing to substitute consumption intertemporally. The AU columns of Panel A also suggest that the EIS does not affect the correlations between macroeconomic quantities. Interestingly, the PPU case predicts significantly different moments of macroeconomic variables for the smaller EIS. Indeed, consumption, investment and output become less correlated mainly due to larger short-run risks in consumption growth.

The bottom panel of Table 2.6 shows the impact of the EIS on financial moments. As expected, the risk-free rate is inversely related to the EIS parameter. The prices of equity on the endogenous levered firm's payouts are not markedly affected by the EIS. Turning to the equity claim on aggregate market dividends, there are large differences in the average equity premium and equity volatility. In this case, more volatile consumption growth predicts riskier dividends, which are modeled as a leverage on consumption. Therefore, the equity as a leveraged consumption claim implies the lower price-dividend ratios, the higher equity premium and equity volatility for a smaller value of the EIS. Notably, this impact on the financial moment is magnified in the PPU case.

As an additional exercise, we change the degree of capital adjustment costs in the production economy. The lower values of ξ introduce higher costs for capital adjustment. Decreasing the value of ξ to 2.5 leads to more volatile consumption growth and smoother investment growth. This additionally generates more volatile Tobin's Q and increases the mean and volatility of the investment return in the model. In general, the higher capital adjustment costs generate stronger short-run risks in the model and reduce the impact of the long-run risks generated by Bayesian learning and rational belief pricing. Since the latter shocks are the dominant drivers of high and volatile equity returns in the model, a claim to aggregate dividends becomes less risky as reflected in the higher price-dividend ratio, lower equity premium and equity volatility. In the light of this observation, a lower capital adjustment cost helps jointly match salient moments of macroeconomic quantities

Table 2.6
Sensitivity Analysis

	$\psi = 1.2$		$\psi = 1.5$		$\xi = 2.5$		$\xi = 5.5$	
	PPU	AU	PPU	AU	PPU	AU	PPU	AU
Panel A: Macroeconomic Quantities								
$\sigma(\Delta c)$	1.54	1.51	1.44	1.43	1.46	1.49	1.21	1.22
$\sigma(\Delta i)$	2.99	3.00	3.20	3.18	3.13	3.12	3.74	3.68
$\sigma(\Delta y)$	1.95	1.94	1.96	1.94	1.96	1.96	1.95	1.93
$\rho(\Delta i, \Delta y)$	0.91	0.98	0.96	0.99	0.99	0.98	0.97	0.99
$\rho(\Delta c, \Delta y)$	0.93	0.98	0.96	0.99	0.99	0.98	0.92	0.98
$\rho(\Delta c, \Delta i)$	0.69	0.92	0.83	0.97	0.97	0.92	0.79	0.96
Panel B: Financial Variables								
$E(R_f) - 1$	1.96	2.52	1.84	2.34	1.64	2.10	1.64	2.17
$\sigma(R_f)$	0.43	0.37	0.42	0.35	0.48	0.37	0.36	0.28
$E(R^l - R_f)$	2.86	1.59	2.89	1.67	3.69	2.23	2.55	1.50
$\sigma(R^l - R_f)$	5.23	4.72	5.35	4.72	6.98	6.12	4.64	4.04
$E(R^M - R_f)$	9.90	3.55	8.03	3.31	5.60	3.06	6.56	2.82
$\sigma(R^M - R_f)$	21.41	15.98	18.50	15.41	15.39	14.98	16.57	14.75
$SR(R^M - R_f)$	0.40	0.19	0.36	0.17	0.30	0.16	0.35	0.15
$E(p^M - d^M)$	2.64	3.45	2.84	3.56	3.22	3.71	3.10	3.78
$\sigma(p^M - d^M)$	0.14	0.05	0.11	0.05	0.06	0.05	0.08	0.05
$ar1(p^M - d^M)$	0.90	0.84	0.90	0.86	0.90	0.90	0.90	0.90

This table reports the average moments from 1,000 simulations of 260 quarters of the data from the production economy considered in this paper, where the transition probabilities and mean growth rates are assumed to be unknown. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All statistics are expressed in annualized terms, except for correlations and autocorrelations expressed in quarterly terms.

and financial returns. In particular, increasing the value of ξ to 5.5 moves the model-implied volatilities of consumption and investment closer to the data. Most importantly, parameter uncertainty generates stronger propagation of productivity shocks in this case by lowering the correlations between macroeconomic variables. In addition, stronger long-run risks originating from rational belief pricing further lead to a substantial increase in risk premia compared to the AU case, as evidenced in the last two columns of Table 2.6.

To assess the impact of asymmetric adjustment costs, we conduct sensitivity analysis to alternative parameter choices of θ^+ and θ^-. Table 2.7 shows that the model with symmetric quadratic adjustment costs $\theta^+ = \theta^- = 5$ displays a small equity premium and levered equity volatility originating from a wrong business cycle movement of dividends.

Table 2.7
Sensitivity Analysis: The Extended Model with Costly Reversibility

	$\theta^+ = 5$ $\theta^- = 5$		$\theta^+ = 4$ $\theta^- = 40$		$\theta^+ = 4$ $\theta^- = 50$		$\theta^+ = 4$ $\theta^- = 60$	
	PPU	AU	PPU	AU	PPU	AU	PPU	AU
Panel A: Macroeconomic Quantities								
$\sigma(\Delta c)$	1.14	1.12	1.51	1.54	1.58	1.60	1.63	1.64
$\sigma(\Delta i)$	3.98	3.91	3.42	3.41	3.36	3.34	3.27	3.22
$\sigma(\Delta y)$	1.93	1.92	1.97	1.97	1.97	1.96	1.97	1.96
$\rho(\Delta i, \Delta y)$	0.91	0.99	0.93	0.92	0.91	0.89	0.90	0.89
$\rho(\Delta c, \Delta y)$	0.78	0.96	0.86	0.84	0.83	0.83	0.84	0.84
$\rho(\Delta c, \Delta i)$	0.45	0.91	0.62	0.57	0.53	0.50	0.53	0.52
Panel B: Financial Variables								
$E(R_f) - 1$	1.62	2.23	1.66	2.03	1.65	2.00	1.62	2.00
$\sigma(R_f)$	0.28	0.24	0.55	0.43	0.59	0.45	0.59	0.45
$E(\Delta d^l)$	0.58	1.24	0.72	1.13	0.76	1.14	0.77	1.12
$\sigma(\Delta d^l)$	24.45	27.45	11.29	16.72	10.39	15.65	9.86	15.11
$ar1(\Delta d^l)$	−0.06	−0.03	−0.04	−0.05	−0.06	−0.06	−0.06	−0.05
$\rho(\Delta c, \Delta d^l)$	−0.11	−0.61	−0.10	−0.05	0.01	0.05	0.06	0.08
$E(R^l - R_f)$	2.03	1.15	4.88	3.00	5.68	3.53	6.49	3.82
$\sigma(R^l - R_f)$	3.65	3.21	9.32	8.09	11.02	9.93	13.25	10.90

This table reports the average moments from 1,000 simulations of 260 quarters of the data from the production economy considered in this paper, where the transition probabilities and mean growth rates are assumed to be unknown. The PPU column refers to the production economy with rational pricing of parameter uncertainty, whereas the AU column refers to the production economy with AU pricing. In both cases, parameter uncertainty includes unknown transition probabilities and mean growth rates. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All statistics are expressed in annualized terms, except for correlations and autocorrelations expressed in quarterly terms.

This result mimics our findings in the benchmark model with convex adjustment costs. Introducing costly reversibility improves the asset pricing implications of the model. We quantify this improvement by varying the degree of asymmetry. In particular, we fix $\theta^+ = 4$ and consider three cases for θ^-/θ^+ : 10, 12.5 and 15. Table 2.7 reports that a higer degree of asymmetry increases the average mean and volatility of levered equity returns. This comes as a result of a more procyclical firm's dividends consistent with our main result: the importance of parameter learning in the production economy is conditional on the introduction of procyclical dividends in the economy.

2.5 Conclusion

In this paper, we show that introducing rational parameter learning into an otherwise standard real business cycle model improves its ability to match asset return data. The model with priced parameter uncertainty has a small effect on the second moments of macroeconomic variables and more significant impact on the comovements between quantities. Parameter learning generates a substantial amplification of the risk premium on a levered firm's payouts and reproduces the long-horizon predictability of excess returns by macroeconomic and valuation variables. Furthermore, we show that rational belief pricing considered in this paper has the largest impact on equity returns when introducing and pricing a procyclical dividend growth process. In this case, the production economy can closely replicate the first and second moments of risk-free rates and excess equity returns, the equity Sharpe ratio and the level of the price-dividend ratio, while generating smooth consumption and volatile investment. Finally, we show that introducing investment friction in the form of costly reversibility helps endogenously generate a pro-cyclical dividend process and, at the same time, to maintain all the desired pricing effects that come from multidimensional learning and rational pricing.

Future research may consider extending our mechanism to a richer model with sticky prices and financial frictions. In particular, modeling wage rigidity in the spirit of Favilukis and Lin (2016) can help endogenously generate procyclical dividend growth in the model. The interaction between sticky prices and learning effects may have additional interesting implications for the labor market. Motivated by a large strand of the literature on time-varying macroeconomic uncertainty, it is interesting and straightforward to extend our methodology to learning about volatility risks. This might have additional asset pricing implications, especially for volatility sensitive assets, as well as interesting effects for the real economy.

Appendix B

B.1 Numerical Algorithm: Anticipated Utility

In the AU case, the representative household learns about the unknown parameters by updating his beliefs upon the realization of new data, but ignores parameter uncertainty when making decisions. Thus, although the beliefs vary over time, the household centers the "true" parameters at the current posterior means and keeps these subjective estimates constant while solving for the continuation utility (and a levered equity claim) in each period.

In this paper, we focus on two learning about parameters economies with unknown transition probabilities, and unknown transition probabilities and mean growth rates.[1] The numerical solution for both models under AU pricing simplifies to solving for the equilibrium pricing ratios when all parameters are actually known by the household. We find the solution of these simplest economies on a dense grid for unknown parameters (that is, unknown transition probabilities in the former model; unknown transition probabilities and mean growth rates in the latter model). Then the household uses these equilibrium pricing functions for the decision making and asset pricing based on the current beliefs.

[1]The methodology for the AU case (as well as the priced parameter uncertainty case in Appendix C.3) can be further extended for learning about the volatility of productivity growth. However, we leave this investigation for the future research.

B.1.1 All Known Parameters

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma\varepsilon_t,$$

where $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, s_t is a two state Markov chain with transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix},$$

where $\pi_{ii} \in (0,1)$. The regimes switches in s_t are independent of the Gaussian shocks ε_t.

Here, we give details on how the continuation utility (and a levered equity claim) is computed for the economy with all parameters known. We define the following stationary variables:

$$\left\{\tilde{C}_t, \tilde{I}_t, \tilde{Y}_t, \tilde{K}_t, \tilde{U}_t\right\} = \left\{\frac{C_t}{A_t}, \frac{I_t}{A_t}, \frac{Y_t}{A_t}, \frac{K_t}{A_t}, \frac{U_t}{A_t}\right\}$$

The household's problem is:

$$\tilde{U}_t = \max_{\tilde{C}_t, \tilde{I}_t} \left\{(1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta\left(E_t\left[\tilde{U}_{t+1}^{1-\gamma}\cdot\left(\frac{A_{t+1}}{A_t}\right)^{1-\gamma}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right\}^{\frac{1}{1-\psi}} \tag{B.1}$$

subject to the constraints:

$$\tilde{C}_t + \tilde{I}_t = \tilde{K}_t^{\alpha}\bar{N}^{1-\alpha} \tag{B.2}$$

$$e^{\Delta a_{t+1}}\tilde{K}_{t+1} = (1-\delta)\tilde{K}_t + \varphi\left(\frac{\tilde{I}_t}{\tilde{K}_t}\right)\tilde{K}_t \tag{B.3}$$

$$\Delta a_t = \mu_{s_t} + \sigma\varepsilon_t,\ \ \varepsilon_t \sim N(0,1) \tag{B.4}$$

$$\tilde{C}_t \geq 0,\ \ \tilde{K}_{t+1} \geq 0 \tag{B.5}$$

where the subscript t indicates the time, $E_t(\cdot)$ denotes the expectation conditional on the information available at time t. Because the parameters are assumed known, s_t and $\tilde{K}_t$ are the only state variables in the economy. Ultimately, the recursive equation (C.1) can be rewritten as:

$$\tilde{U}_t(s_t, \tilde{K}_t) \tag{B.6}$$

$$= \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1} \left(s_{t+1}, \tilde{K}_{t+1} \right)^{1-\gamma} \cdot e^{(1-\gamma)\Delta a_{t+1}} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

To solve the recursion (C.6), we use the the value function iteration algorithm. In particular, the numerical algorithm proceeds as follows:

1. We find the de-trended steady state capital $\tilde{K}_{ss}$, assuming the productivity growth equals the steady state level predicted by a Markov-switching model. The state space for capital normalized by technology is set at $[0.2\tilde{K}_{ss}, 2.2\tilde{K}_{ss}]$. We further use $n_k = 100$ points on a grid for capital in the numerical computation. A denser grid does not lead to significantly different results.

2. For any level of capital $\tilde{K}_t$ at time t, we construct a grid for $\tilde{I}_t$ with uniformly distributed points between 0 and $\tilde{K}_t^{\alpha}\bar{N}^{1-\alpha}$. Specifically, we use $n_i = 400$ points.

3. For the expectation, we use the Gauss-Hermite quadrature with $n_{gh} = 8$ points. Using the quadrature weights and nodes, we can calculate the expression on the right hand side.

4. We solve the optimization problem in the Bellman equation (C.6) subject to (C.2)-(C.5) and update a new value function $\tilde{U}_t = \tilde{U}_t(s_t, \tilde{K}_t)$ given an old one $\tilde{U}_{t+1} = \tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1})$.

5. We iterate Steps 2-4 by updating the continuation utility on each iteration until a suitable convergence is achieved. Specifically, the stopping rule is that the distance between the new value function and the old value function satisfies $|\tilde{U}_{t+1} - \tilde{U}_t|/|\tilde{U}_t| < 10^{-12}$.

B.2 Numerical Algorithm: Priced Parameter Uncertainty

The numerical solution for the case of priced parameter uncertainty consists of two main steps[2]. First, we solve for the equilibrium pricing ratios when true parameters are actu-

[2](Johnson 2007) uses this solution methodology in a case with parameter learning and power utility. Johannes, Lochstoer, and Mou (2016) and Collin-Dufresne, Johannes, and Lochstoer (2016) extend this approach to the case of Epstein-Zin utility in the endowment economy. We further extend the numerical solution to the case of Epstein-Zin utility in the production economy.

ally known by the household (by assumption, these are learned at $T = \infty$). We find the solution of this simplest limiting economy on a dense grid of state variables. Second, we use the known parameters boundary economies as terminal values in the backward recursion to obtain the equilibrium function at time t. For the first step, Appendix C.2 outlines details of the numerical algorithm for all known parameters. Therefore, we present the solution methodology employed at the second step for two models with unknown transition probabilities, and unknown transition probabilities and mean growth rates.

B.2.1 Unknown Transition Probabilities

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma \varepsilon_t,$$

where $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, s_t is a two state Markov chain with a transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1 - \pi_{11} \\ 1 - \pi_{22} & \pi_{22} \end{bmatrix},$$

where $\pi_{ii} \in (0,1)$. The regimes switches in s_t are independent of the Gaussian shocks ε_t.

In the case of unknown transition probabilities, the representative household knows true values of the parameters within each state (μ_1, μ_2, σ) and observes states (s_t) but does not know the transition probabilities (π_{11}, π_{22}). At time $t = 0$, the household holds priors about uncertain probabilities in the transition matrix and updates beliefs each period upon realization of new series and regimes. We assume a Beta distributed prior and, thus, posterior beliefs are also Beta distributed.

The Beta distribution has the probability density function of the form:

$$p(\pi|a,b) = \frac{\pi^{a-1}(1-\pi)^{b-1}}{B(a,b)},$$

where $B(a,b)$ is the Beta function (a normalization constant), a and b are two positive shape parameters. We are particularly interested in the expected value of the Beta distribution defined by:

$$E[\pi|a,b] = \frac{a}{a+b}.$$

Furthermore, we use two pairs of hyperparameters parameters (a_1, b_1) and (a_2, b_2) for unknown transition probabilities in the states π_{11} and π_{22}, respectively. At time t, the household uses Bayes' rule and the fact that states are observable to update hyperparameters for each state i as follows:

$$a_{i,t} = a_{i,0} + \#(\text{state } i \text{ has been followed by state } i), \tag{B.7}$$

$$b_{i,t} = b_{i,0} + \#(\text{state } i \text{ has been followed by state } j), \tag{B.8}$$

given the initial prior beliefs $a_{i,0}$ and $b_{i,0}$.

Once we find the limiting boundary economies on the first step, we perform a backward recursion using the following state variables:

$$\tau_{1,t} = a_{1,t} - a_{1,0} + b_{1,t} - b_{1,0} \tag{B.9}$$

$$\lambda_{1,t} = E_t[\pi_{11}] = \frac{a_{1,t}}{a_{1,t} + b_{1,t}} \tag{B.10}$$

$$\tau_{2,t} = a_{2,t} - a_{2,0} + b_{2,t} - b_{2,0} \tag{B.11}$$

$$\lambda_{2,t} = E_t[\pi_{22}] = \frac{a_{2,t}}{a_{2,t} + b_{2,t}} \tag{B.12}$$

Note that $X_t = \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}\}$ are sufficient statistics for the agent's priors. Also, we can update X_{t+1} using the equations (C.7)-(C.12), the next period regime, and sufficient statistics:

$$X_{t+1} = f(s_{t+1}, s_t, X_t).$$

For notational purposes, it might be useful to denote $X_t^s \equiv \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}\}$ and $X_t^{\Delta a} \equiv \{\tilde{K}_t\}$, where the superscripts s and Δa indicate that variables in the vectors X_t^s and $X_t^{\Delta a}$ are a function only of the observed state realization s_t and a function of (also) the realized productivity growth, respectively. Thus, $X_t = \left[X_t^s, X_t^{\Delta a}\right]$. Using these notations, we can rewrite

$$\tilde{U}_{t+1}(s_{t+1}, X_{t+1}) = \tilde{U}_{t+1}(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a})$$

to better indicate the dependence of state variables on specific shocks. Ultimately, the recursive equation (C.1) can be rewritten as:

$$\begin{aligned} & \tilde{U}_t(s_t, X_t) \\ = \quad & \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, X_t \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}, \end{aligned} \tag{B.13}$$

where the expectation on the right hand side is equivalent to:

$$\begin{aligned}
& E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}}\middle| s_t, X_t\right] \\
= \quad & E_t\left[E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}}\middle| s_{t+1}, s_t, X_t\right]\middle| s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2} \mathbb{P}(s_{t+1}|s_t, X_t^s) \quad \ldots \\
\times \quad & E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}}\middle| s_{t+1}, s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2} E_t(\pi_{s_{t+1},s_t}|s_t, X_t^s) \quad \ldots \\
\times \quad & E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}}\middle| s_{t+1}, s_t, X_t\right], \qquad \text{(B.14)}
\end{aligned}$$

where the first and second equalities follow from the independency of the regime changes and the Gaussian shocks to productivity growth (s_{t+1} and ε_{t+1}). Let the conditional density of π_{s_{t+1},s_t} be $g(\pi_{s_{t+1},s_t}|s_t, X_t^s)$, then the third equality follows from:

$$\mathbb{P}(s_{t+1}|s_t, X_t^s) = \int_0^1 \pi_{s_{t+1},s_t} g(\pi_{s_{t+1},s_t}|s_t, X_t) d\pi_{s_{t+1},s_t} = E_t(\pi_{s_{t+1},s_t}|s_t, X_t)$$

Furthermore, using the definition of our state variables, this last conditional expectation equals $\lambda_{s_t,t}$ or $1 - \lambda_{s_t,t}$.

Note that before choosing the optimal consumption and investment in (B.13), we need to solve numerically first the inner expectation, which is equivalently represented by (B.14). Hopefully, we have an analytical expression for the conditional expectation of transition probabilities in (B.14), which is either $\lambda_{s_t,t}$ or $1 - \lambda_{s_t,t}$. For the second conditional expectation in (B.14), we do not have a closed form since the continuation utility depends on the realized productivity growth through $\tilde{K}_{t+1}$. Therefore, we use quadrature-type numerical methods to evaluate this expectation as follows:

$$\begin{aligned}
& E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}}\middle| s_{t+1}, s_t, X_t\right] \\
\approx \quad & \sum_{j=1}^{J} \omega_\varepsilon(j)\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a(j), X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a(j)}\middle| s_{t+1}, s_t, X_t\right], \qquad \text{(B.15)}
\end{aligned}$$

where $\omega_\varepsilon(j)$ is the quadrature weight corresponding to the quadrature node $n_\varepsilon(j)$ used for the integration of a standard normal shock ε_{t+1} in productivity growth. The observed realized productivity growth, $\Delta a(j)$, and a state variable, $X^{\Delta a}_{t+1}(j) = \tilde{K}_{t+1}(j)$, are updated as follows:

$$\Delta a(j) = \mu_{s_{t+1}} + \sigma \cdot n_{\varepsilon}(j) \tag{B.16}$$

$$e^{\Delta a(j)} \tilde{K}_{t+1}(j) = (1-\delta)\tilde{K}_t + \varphi\left(\frac{\tilde{I}_t}{\tilde{K}_t}\right)\tilde{K}_t, \tag{B.17}$$

where

$$\tilde{I}_t = \tilde{K}_t^{\alpha} \bar{N}^{1-\alpha} - \tilde{C}_t. \tag{B.18}$$

Finally, the numerical backward recursion can be performed by using (B.13)-(B.18). The boundary conditions are defined by the limiting economies $\tau_{1,\infty}$ and $\tau_{2,\infty}$, where the transition probabilities π_{11} and π_{22} are known.

Solving for a Dividend Claim

We also solve for the price-dividend ratio of the equity claim written on aggregate dividends, which are defined as a leverage to aggregate consumption. Let exogenous aggregate dividends be given by:

$$\Delta d_{t+1} = g_d + \lambda \Delta c_{t+1} + \sigma_d \varepsilon_{d,t+1},$$

where $g_d = \Big(1-\lambda\Big)\Big(E(\mathbb{P}(s_\infty = 1|\pi_{11},\pi_{22}))\mu_1 + E(\mathbb{P}(s_\infty = 2|\pi_{11},\pi_{22}))\mu_2\Big)$ and $\mathbb{P}(s_\infty = i|\pi_{11},\pi_{22})$ is the ergodic probability of being in state i conditional on the transition probabilities π_{11} and π_{22}. Note that the long run mean of dividends growth, g_d, is changing under the household's filtration, though the true long run growth is constant. The subjective beliefs about the true parameter values induce fluctuations in g_d, which can be expressed as $g_d = g_d(s_{t+1}, s_t, X_t)$.

The equilibrium condition for the price-dividend ratio is standard in the Epstein-Zin economy and is given by:

$$PD_t = E_t\left[\beta\left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t}\right)^{-\frac{1}{\psi}}\left(\frac{A_{t+1}}{A_t}\right)^{-\frac{1}{\psi}}\left(\frac{\tilde{U}_{t+1}\cdot\left(\frac{A_{t+1}}{A_t}\right)}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot\left(\frac{A_{t+1}}{A_t}\right)\right)}\right)^{\frac{1}{\psi}-\gamma}\left(\frac{D_{t+1}}{D_t}\right)(PD_{t+1}+1)\right] \tag{B.19}$$

Similarly to the solution for the value function, we rewrite all variables in the recursion (C.23) as a function of the state variables and further use quadrature-type numerical methods to evaluate expectations on the right hand side of (C.23). Additionally, we update the long run dividends growth, $g_d(s_{t+1}, s_t, X_t)$, which is in fact random. Consequently, the equilibrium recursion used to solve the model is then:

$$
\begin{aligned}
& PD_t(s_t, X_t,) \\
= \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma}\dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_t, X_t\right] \\
= \quad & E_t\left[E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma}\dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right]\middle| s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2}\mathbb{P}(s_{t+1}|s_t, X_t^s) \quad \dots \\
\times \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma}\dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2}E_t(\pi_{s_{t+1},s_t}|s_t, X_t^s) \quad \dots \\
\times \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma}\dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right]
\end{aligned}
$$

Again, the conditional expectation of transition probabilities under the household's filtration permits an analytical formula, while the inner expectation in the expression above can be evaluated using the quadrature-type integration methods.

Limiting Economies - Boundary Values for General Case

The key assumption of the numerical solution is that the household eventually learns the true values of all uncertain parameters in the productivity growth. Thus, the simplest limiting economy is the one where all parameters are known, including both transition probabilities π_{11} and π_{22}. In this case, s_t and K_t are the only state variables in the economy. We employ the numerical solution methodology outlined for AU pricing for this limiting economy. Specifically, we find the continuation utility (and the price-dividend ratio of the equity claim) for a grid on π_{11} and π_{22}.

B.2.2 Unknown Transition Probabilities and Unknown Mean Growth Rates

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma\varepsilon_t,$$

where $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, s_t is a two state Markov chain with the transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix},$$

where $\pi_{ii} \in (0,1)$. The regimes switches in s_t are independent of the Gaussian shocks ε_t.

As before, we assume that the representative household does not know the transition probabilities (π_{11}, π_{22}). Additionally, the mean growth rates within each state (μ_1, μ_2) are assumed to be unknown, while the realization of states (s_t) and productivity volatility (σ_t) remain observable. Due to the limitations of the numerical solution algorithm under the prices parameter uncertainty case, we are unable to extend the economy to unobservable regimes, while it is still possible to assume that the household does not know a volatility parameter. Nevertheless, the extension to the case with all parameters unknown, including volatility except for states, is quite straightforward, and we leave the investigation of learning about volatility parameters for future research.

Regarding priors, we assume a conjugate prior for transition probabilities and mean growth rates within each state i : the Beta distributed prior and the truncated normal distributed prior, respectively. The updating equations for two pairs of hyperparameters (a_1, b_1) and (a_2, b_2) remain as before. Additionally, we denote hyperparameters of the truncated normal distributed prior for mean growth in state i by $\mu_{i,t}$ and $\sigma_{i,t}$, which are updated by the Bayes' rule as follows:

$$\mu_{i,t+1} = \mu_{i,t} + \mathbf{1}_{s_{t+1}=i} \frac{\sigma_{i,t}^2}{\sigma_i^2 + \sigma_{i,t}^2} \left(\Delta a_{t+1} - \mu_{i,t}\right) \tag{B.20}$$

$$\sigma_{i,t+1}^{-2} = \mathbf{1}_{s_{t+1}=i} \cdot \sigma_i^{-2} + \sigma_{i,t}^{-2}, \tag{B.21}$$

where $\mathbf{1}$ is an indicator function that equals 1 if the condition in subscript is true and 0 otherwise.

Note that since the variance hyperparameters $\sigma_{1,t}^2$ and $\sigma_{2,t}^2$ are a function of the time, the following 6-dimensional vector $X_t \equiv \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}, \mu_{1,t}, \mu_{2,t}\}$ is sufficient statistics for the priors. Thus, we can define X_{t+1} using the equations (C.7)-(C.12), (C.13)-(C.16), the next period regime, and sufficient statistics at time t :

$$X_{t+1} = f(s_{t+1}, s_t, X_t).$$

Following the notations of a previous section, we define $X_t^s \equiv \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}\}$ and $X_t^{\Delta a} \equiv \{\tilde{K}_t, \mu_{1,t}, \mu_{2,t}\}$, where the superscripts s and Δa indicate that variables in the vectors X_t^s and $X_t^{\Delta a}$ are a function only of the observed state realization s_t and a function of (also) the realized productivity growth, respectively. Thus, $X_t = \left[X_t^s, X_t^{\Delta a}\right]$. Using these notations, we can rewrite

$$\tilde{U}_{t+1}(s_{t+1}, X_{t+1}) = \tilde{U}_{t+1}(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a})$$

to better indicate the dependence of state variables on specific shocks. Ultimately, the recursive equation (C.1) is of the same form:

$$\begin{aligned} & \tilde{U}_t(s_t, X_t) \qquad \qquad \text{(B.22)} \\ = \quad & \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, X_t \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}, \end{aligned}$$

where the expectation on the right hand side is equivalent to:

$$\begin{aligned} & E_t \left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, X_t \right] \\ = \quad & \sum_{s_{t+1}=1}^{2} E_t(\pi_{s_{t+1}, s_t} | s_t, X_t^s) \quad \ldots \\ \times \quad & E_t \left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, X_t \right]. \qquad \text{(B.23)} \end{aligned}$$

In this case, we compute the conditional expectation in (C.18) by integrating over conditional distribution of mean growth rates as well as Gaussian distribution of the error term in productivity growth. In particular:

$$
\begin{aligned}
& E_t\left[\tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, X_t\right] \\
\approx \quad & \sum_{j=1}^{J} \omega_{\varepsilon}(j)\left[\sum_{k=1}^{K} \omega_{\mu_{s_{t+1}}}(k) \cdot \tilde{U}_{t+1}^{1-\gamma}\left(s_{t+1}, s_t, X_t^s, \Delta a(j,k), X_t^{\Delta a}\right) \cdot e^{(1-\gamma)\Delta a(j,k)} \middle| s_{t+1}, s_t, X_t\right],
\end{aligned} \tag{B.24}
$$

where $\omega_\varepsilon(j)$ is the quadrature weight corresponding to the quadrature node $n_\varepsilon(j)$ used for the integration of a standard normal shock ε_{t+1} in productivity growth, and $\omega_{\mu_{s_{t+1}}}(k)$ is the quadrature weight corresponding to the quadrature node $n_{\mu_{s_{t+1}}}(k)$ used for the integration of a truncated standard normal variable $\mu_{s_{t+1}}$. The observed realized productivity growth, $\Delta a(j,k)$, and a state variable, $X_{t+1}^{\Delta a}(j,k) = \tilde{K}_{t+1}(j,k)$, are updated as follows:

$$
\Delta a(j,k) = n_{\mu_{s_{t+1}}}(k) + \sigma \cdot n_\varepsilon(j) \tag{B.25}
$$

$$
e^{\Delta a(j,k)}\tilde{K}_{t+1}(j,k) = (1-\delta)\tilde{K}_t + \varphi\left(\frac{\tilde{I}_t}{\tilde{K}_t}\right)\tilde{K}_t, \tag{B.26}
$$

where

$$
\tilde{I}_t = \tilde{K}_t^{\alpha}\bar{N}^{1-\alpha} - \tilde{C}_t. \tag{B.27}
$$

Finally, the numerical backward recursion can be performed by using (C.17)-(C.22). The boundary conditions are defined by the limiting economies $\tau_{1,\infty}$ and $\tau_{2,\infty}$, where the transition probabilities π_{11} and π_{22}, and mean growth rates μ_1 and μ_2, are known.

Solving for a Dividend Claim

We also solve for the price-dividend ratio of the equity claim written on aggregate dividends, which are defined as a leverage to aggregate consumption. Let exogenous aggregate dividends be given by:

$$
\Delta d_{t+1} = g_d + \lambda\Delta c_{t+1} + \sigma_d\varepsilon_{d,t+1},
$$

where $g_d = \Big(1-\lambda\Big)\Big(E(\mathbb{P}(s_\infty = 1|\pi_{11},\pi_{22}))\mu_1 + E(\mathbb{P}(s_\infty = 2|\pi_{11},\pi_{22}))\mu_2\Big)$ and $\mathbb{P}(s_\infty = i|\pi_{11},\pi_{22})$ is the ergodic probability of being in state i conditional on the transition probabilities π_{11} and π_{22}.

Note that the long run mean of dividends growth, g_d, is changing under the household's filtration, though the true long run growth is constant. The subjective beliefs about the true parameter values induce fluctuations in g_d, which can be expressed as $g_d = g_d(s_{t+1}, s_t, X_t)$. The equilibrium condition for the price-dividend ratio and the equilibrium recursion remain the same as in the "unknown transition probabilities" model. The only difference between the two models lie in the way we calculate the conditional expectations. With unknown transition probabilities and mean growth rates in the productivity growth process, we employ quadrature-type integration methods analogous to solving for the continuation utility in this economy.

Limiting Economies - Boundary Values for General Case

The key assumption of the numerical solution is that the household eventually learns the true values of all uncertain parameters in the productivity growth. Thus, the simplest limiting economy is the one where all parameters are known, including both transition probabilities π_{11} and π_{22}, mean growth rates μ_1 and μ_2. In this case, s_t and K_t are the only state variables in the economy. We employ the numerical solution methodology outlined for AU pricing for this limiting economy. Specifically, we find the continuation utility (and the price-dividend ratio of the equity claim) for a grid on $\pi_{11}, \pi_{22}, \mu_1$ and μ_2.

B.2.3 Existence of Equilibrium

Similarly to Collin-Dufresne, Johannes, and Lochstoer (2016) and Johannes, Lochstoer, and Mou (2016), the existence of the equilibrium in our production-based economy relies on the fact that the value function is concave and finite for all parameters known economies. Therefore, we verify that these conditions are satisfied for all limiting boundary economies.

B.3 Impulse Responses

In this section, we consider the numerical procedure used to obtain impulse responses of key macroeconomic and financial variables to a regime switch in the mean growth rate of productivity. In particular, we assume that the economy stays in the high growth state for a long period and then moves to a low growth regime at time 0. We further consider three possible scenarios where the economy remains in the bad regime for one quarter, three quarters, or two years before returning to the good state. The details of the numerical algorithm look as follows.

First, we find the steady state of capital, $\tilde{K}$, in the high growth regime, $s_t = 1$, assuming unbiased parameter beliefs, X_t, which are centered at the true values. Formally, $\tilde{K}$ solves the equation:

$$\tilde{K} = f^k(s_{-1} = 1, X_{-1}, \tilde{K}),$$

where $f^K(\cdot)$ is the policy function for capital assuming the productivity growth is high forever.

Second, suppose that the economy starts in the high growth steady state before time 0 and the investor holds unbiased parameter beliefs. Then unexpectedly the economy shifts to to the bad state at time 0 and stays there for τ periods. Using the policy function, capital is computed recursively as:

$$\begin{aligned}
\tilde{K}_{-1} &= \tilde{K}, \\
\tilde{K}_0 &= f^k(s_0 = 2, X_0, \tilde{K}_{-1}), ... \\
\tilde{K}_\tau &= f^k(s_\tau = 2, X_\tau, \tilde{K}_{\tau-1}), \\
\tilde{K}_{\tau+1} &= f^k(s_{\tau+1} = 1, X_{\tau+1}, \tilde{K}_\tau), ... \\
\tilde{K}_t &= f^k(s_t = 1, X_t, \tilde{K}_{t-1}), \quad \forall t,
\end{aligned}$$

where investor's parameter beliefs are updated in each period.

Third, we use policy functions for investment and consumption to obtain equilibrium values of $\tilde{I}_t$ and $\tilde{C}_t$. Finally, we calculate the remaining macroeconomic and financial variables using the updated state variables.

Chapter 3

Option Prices and Learning about Productivity Dynamics

Mykola Babiak[1] and **Roman Kozhan**[2]

Abstract

We demonstrate that incorporating time-varying productivity volatility and priced parameter uncertainty in a production economy can explain index option prices, equity returns, the risk-free rate, and macroeconomic quantities. A Bayesian investor learns about the true parameters governing mean, persistence, and volatility of productivity growth. Rational parameter learning amplifies the conditional risk premium and volatility especially at the onset of recessions. We estimate the model based on post-war U.S. data and find that it can capture the implied volatility surface and the variance premium. Intuitively, the agent pays a large premium for index options because they hedge future belief revisions.

[1]We would like to thank Andrea Gamba, Michal Kejak, Ian Khrashchevskyi, Ctirad Slavik, Sergey Slobodyan and conference/seminar participants at the 2018 Lancaster-Warwick (LaWa) Workshop on Financial Econometrics and Asset Pricing, Warwick Business School, CERGE-EI and Universitȧ Ca' Foscari Venezia for their discussions and comments. We also appreciate the research support with the supercomputing clusters from the Centre for Scientific Computing at the University of Warwick. The financial support from the Charles University Grant Agency (GAUK No. 744218) is gratefully acknowledged.

[2]Warwick Business School, University of Warwick, Scarman Road, Coventry, CV4 7AL, UK.

3.1 Introduction

Economic uncertainty, broadly defined, is an important determinant of asset valuations and business cycles. Recent work shows that rational pricing of uncertainty about the true structure of the economy can help capture prominent features of equity returns and macroeconomic fundamentals.[3] Derivative prices, in turn, directly speak to investor perceptions of macroeconomic uncertainty and, therefore, offer an attractive opportunity to explore further links between asset prices and incomplete structural knowledge of agents. In this paper, we show that incorporating rational parameter learning into an otherwise standard real business cycle framework can explain two puzzling features of index options: (i) a large variance risk premium, and (ii) a steep implied volatility surface. Simultaneously, the model with priced parameter uncertainty is able to match the large equity premium and excess return volatility, the low risk-free rate, and the level of the price-dividend ratio. At the same time, it remains consistent with second moments of consumption, investment and output, and comovements between the series, while capturing the relation between risk-neutral variance and macroeconomic quantities. The framework is the first, to our knowledge, to provide a pure learning-based explanation of such a wide array of pricing phenomena without introducing long-run risks, tail events, and non-standard assumptions about preferences or beliefs. The model instead reflects information in post-World War II data and uses a modest risk aversion of seven.

In this paper, a representative investor has non-time-separable preferences and learns about a true model structure for productivity growth, which follows a two-state Markov switching process with regimes in drift and volatility. An inference problem entails learning about unknown transition probabilities, and the mean and volatility of productivity growth in the two regimes, while the state of the economy is assumed to be observable. In each period, the agent uses Bayes' rule to update his beliefs about unknown parameters upon observing new data. Furthermore, he rationally incorporates parameter uncertainty in the decision-making process by acknowledging future changes in his beliefs when choosing optimal investment and when pricing assets. In terms of equity prices, prior literature documents the failure of the standard production model to explain equity returns due to a countercyclical firm's dividends, which make the equity work like a hedge.[4] In this paper,

[3] Collin-Dufresne, Johannes, and Lochstoer (2016) develop the mechanism of priced parameter uncertainty in the consumption-based model and explore the implications for equities, whereas Babiak and Kozhan (2018) extend the methodology to the production economy.

[4] Favilukis and Lin (2016) employ wage rigidity in order to produce procyclical dividends endogenously, while Babiak and Kozhan (2018) use a combination of asymmetric capital adjustment costs and a financial

we abstract from introducing additional frictions in the economy that would complicate the solution and we instead price exogenous stock market dividends. In doing so, we are able to preserve the procyclical dynamics of dividends in the model consistent with the behaviour of aggregate stock market dividends.

The key mechanism of the model is as follows. First, in the presence of parameter uncertainty, learning generates time-variation in posterior estimates of unknown parameters creating an additional channel by which shocks to productivity growth introduce extra fluctuations in an investor's marginal utility. Second, rational pricing of subjective beliefs amplifies the impact of parameter uncertainty on the stochastic discount factor, conditional moments of returns, and asset prices. The agent is concerned about future revisions, especially those in response to negative news to technology growth, and hence he is willing to pay a large premium for insurance against pessimistic updates. The deep out-of-the-money put options on the aggregate stock market index provide such insurance and therefore bear high parameter uncertainty premiums. We show that this mechanism generates a steep implied volatility skew, which closely replicates the shape observed in the data. Furthermore, the conditional volatility of equity return variance is amplified, thus raising the investor's concerns about high realized variance in the future. In order to hedge his concerns, the agent is willing to pay large prices for variance swaps, which would provide a high payoff in states of high return volatility. This substantially increases the mean and volatility statistics of the variance premium that become closer to empirical estimates.

The model with priced parameter uncertainty also helps capture stylized facts of equity returns and macroeconomic quantities apart from index option prices. In particular, the mechanism of this paper generates low interest rates, high equity premiums and equity Sharpe ratios, and excess volatility of equity returns. While matching these standard asset pricing moments, the model generates low volatility of consumption and output, and high volatility of investment in line with the data. Furthermore, we show that rational parameter learning resolves the problem of perfectly correlated quantities in the real business cycle models. In our model, macroeconomic risks associated with rational pricing of parameter uncertainty substantially reduce correlations between macroeconomic variables, consistent with the empirical estimates. Finally, the framework reasonably reconciles the correlations between the squared VIX index and investment growth, consumption growth, and equity returns at different leads and lags. We further show that the

leverage in the production economy with priced parameter uncertainty.

model is consistent with the empirical lead-lag relations between the risk-neutral variance and variance of aforementioned variables.

Related literature. This paper belongs to a new and growing literature on parameter learning models in macroeconomics and finance as advocated by Hansen (2007), Weitzman (2007) and Cogley and Sargent (2008). The early studies in this vein focus on learning about a latent state or a single parameter in endowment economies (see Pastor and Veronesi (2009) for a survey of the early literature on learning in financial markets). These studies use Bayesian updating as a driving process to explain salient moments of equity returns. Their findings suggest that learning may in fact improve the model performance with respect to some features of the data. However, several long-standing puzzles, including the equity premium and excess return volatility, remain unresolved in endowment economies.

One potential reason is the lack of persistent effects in a setting with learning about a single variable. Cogley and Sargent (2008) find that injecting a pessimistic initial prior into the model with learning about the mean duration of recessions can yield long-lasting effects on asset prices and, in particular, generate the large equity premium. Pakos (2013) and Gillman, Kejak, and Pakos (2015) document that Bayesian learning about a protracted recession in the hidden three-state Markov process can generate a large number of pricing phenomena in the equity and bond markets. Andrei, Carlin, and Hasler (2017) construct an equilibrium model with a disagreement mechanism about the length of business cycles to explain stock return volatility and equity risk premium. Andrei, Hasler, and Jeanneret (2018) further emphasize the importance of learning about persistence risk for reproducing observed dynamics of expected returns, return volatility, and the price of risk, as well as excess return predictability patterns. The key features of the literature above are learning about a built-in persistence in the economy and learning about a single state variable. In contrast, we study the asset pricing implications of a high-dimensional learning problem, which in turn gives rise to endogenous slow learning due to confounding (Johannes, Lochstoer, and Mou 2016). Further, our paper generates persistent macroeconomic risks via fully rational pricing of posterior parameter beliefs.

This paper is also related to several consumption-based models that specifically target index options with certain types of learning. In particular, learning about a persistent component of consumption growth (Benzoni, Collin-Dufresne, and Goldstein 2011; Shaliastovich 2015) or about model uncertainty with rare disasters (Liu, Pan, and Wang 2005) can explain the skew in index option implied volatilities. Further, Drechsler (2013) constructs the generalized long-run risks framework with ambiguity over the true model governing endowments to simultaneously explain equity returns and option-related puzzles. Recently, Babiak (2019) shows that the combination of learning about a latent state in consumption growth and an investor's asymmetric preferences can reproduce both the variance premium and the implied volatility surface observed in the data. While learning in these models generates time-variation in asset prices, tail risks induced by disastrous events or asymmetric preferences are the key ingredient generating the large variance premium and the steep implied volatility curves. In contrast, our paper does not rely on any of these channels. Our model is estimated based on post-war US data and hence does not allow tail outcomes in productivity growth. Further, the representative investor has standard Epstein-Zin preferences and does not exhibit tail risk attitude. Our framework instead considers rational pricing of parameter uncertainty as a key amplification mechanism of productivity shocks on risk premiums embedded in option prices.

In the production-based setting, the learning literature is rather scarce since it is more challenging to explain asset prices with endogenous consumption and dividends.[5] Recently, Jahan-Parvar and Liu (2014) adopt an endowment economy of Ju and Miao (2012) with ambiguity preferences and learning to the real business cycle model. In their framework, however, the equity prices are mainly driven by the investor's ambiguity aversion. To our knowledge, Liu and Zhang (2018) is the only study that targets the variance premium and the implied volatility skew in the production economy similar to our analysis. Liu and Zhang (2018) find that large risk premiums embedded in option prices are the manifestation of an investor's pessimism due to ambiguity aversion. In contrast, our study employs Epstein-Zin preferences that would not generate the same results in the model of Liu and Zhang (2018). Further, our paper considers rational parameter learning as a key driver of option prices, which is again different from the central ingredient considered by Liu and Zhang (2018). Finally, our framework can capture the volatility term structure, whereas Liu and Zhang (2018) focus only on the three-month implied volatility curve.

Methodologically, this article is related to Collin-Dufresne, Johannes, and Lochstoer (2016), who develop the idea of priced parameter uncertainty in the consumption-based models, and to Babiak and Kozhan (2018), who extend the methodology to the production-

[5]In fact, there are several learning papers with the production sector including Andrei, Mann, and Moyen (2018), Kozlowski, Veldkamp, and Venkateswaran (2018a) and Kozlowski, Veldkamp, and Venkateswaran (2018b), but their key mechanism and focus are different from those in our article.

based setting. In terms of the asset pricing implications, both studies limit their attention to equity returns and do not explore the role of parameter learning for the variance risk premium and the implied volatility surface, the key focus of this paper. Unlike Collin-Dufresne, Johannes, and Lochstoer's (2016) model with learning about rare events in consumption growth, our framework incorporates learning about parameters governing business cycle fluctuations in productivity growth. Unlike Babiak and Kozhan's (2018) model with homoskedastic and known volatility of productivity growth, our analysis additionally introduces learning about regime-switching volatility of technology growth. We show that this extended specification with learning about volatility parameters helps quantitatively capture a large variance risk premium and a steep implied volatility surface. Turning off time-varying productivity volatility yields risk premiums in option prices that are excessively low compared to the data.

The paper proceeds as follows. Section 3.2 presents the formal model. Section 3.3 investigates the quantitative implications of parameter learning for quantities and asset prices. Section 3.4 performs a sensitivity analysis. Section 3.5 concludes. Sections C.1, C.2, and C.3 of Appendix C outline the numerical solution methodology for the models with known parameters, and parameter uncertainty under anticipated utility and fully rational learning, respectively.

3.2 The Model

3.2.1 The Representative Household

We consider a standard production-based asset pricing framework with a representative household that has the utility function of Epstein and Zin (1989) defined recursively as:

$$U_t = \left\{ (1-\beta) C_t^{1-1/\psi} + \beta \left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \tag{3.1}$$

in which U_t denotes the household's continuation utility, C_t denotes aggregate consumption, E_t denotes the expectation operator, $\beta \in (0,1)$ is the discount factor, $\psi > 0$ represents the elasticity of inter-temporal substitution (EIS), and $\gamma > 0$ represents the risk aversion parameter. For simplicity, we will assume that the household inelastically supplies one unit of labor, and thus the household's intra-period utility depends only on consumption.

These recursive preferences allow the separation between the agent's relative risk aversion and the elasticity of inter-temporal substitution. In this paper, we consider a representative household with a preference for early resolution of uncertainty by setting $\gamma > \frac{1}{\psi}$. This calibration is crucial for our results since subjective long-run beliefs will be priced in equilibrium. The stochastic discount factor is:

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi} \left(\frac{U_{t+1}}{\left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}} \right)^{1/\psi - \gamma}. \tag{3.2}$$

3.2.2 The Representative Firm

We assume a representative firm produces the consumption good using a constant returns to scale Cobb-Douglas production function:

$$Y_t = K_t^{\alpha} (A_t N_t)^{1-\alpha}, \tag{3.3}$$

in which Y_t is the output, K_t is the capital stock, N_t is labor hours, and A_t is an exogenous, labor-enhancing technology level. For simplicity, we assume that the representative household supplies a fixed amount of labor hours, which are exogenously set $N_t = 1$.

The firm chooses investment according to the resource constraint $I_t = Y_t - C_t$ and faces capital adjustment costs while accumulating capital stock. Formally, the law of motion for capital is defined by:

$$K_{t+1} = (1 - \delta) K_t + \varphi(I_t / K_t) K_t,$$

in which $\delta \in (0, 1)$ is the capital depreciation rate, and $\varphi(\cdot)$ is the adjustment cost function given by:

$$\varphi(x) = a_1 + \frac{a_2}{1 - 1/\xi} x^{1-1/\xi}, \tag{3.4}$$

in which ξ is the elasticity of the investment rate to Tobin's q. The lower value of ξ implies higher capital adjustment costs, while the extreme case of $\xi = \infty$ means that capital adjustment costs are zero. We follow Boldrin, Christiano, and Fisher (2001) and choose the constants a_1 and a_2 such that there are no adjustment costs in the non-stochastic steady state.[1]

3.2.3 Technology

We assume a two-state Markov switching model for productivity growth,

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t,$$

in which Δa_t is log-technology growth, $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, and s_t is a two state Markov chain with transition matrix Π defined by:

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix}, \quad \pi_{11}, \pi_{22} \in (0,1).$$

The mean μ_{s_t} and volatility σ_{s_t} of productivity growth depend on the state variable s_t. We label $s_t = 1$ the "good" regime with the high mean and low volatility of productivity growth and $s_t = 2$ the "bad" regime with the low mean and high volatility.

3.2.4 Equilibrium Asset Prices

In the competitive equilibrium of the economy, the representative household works for the firm and maximizes the lifetime utility over a consumption stream. The representative firm chooses labor and capital inputs (through investment) to maximize the firm's value, the present value of its future cash flows. The firm's maximization problem implies the following equilibrium conditions for asset's j gross return $R_{j,t+1}$:

$$E_t \left[M_{t+1} R_{j,t+1} \right] = 1. \tag{3.5}$$

In particular, the equation above is satisfied by the investment return, $R_{I,t+1}$, defined by:

$$R_{I,t+1} = \frac{1}{Q_t} \left[Q_{t+1} \left(1 - \delta + \varphi \left(\frac{I_{t+1}}{K_{t+1}} \right) \right) + \frac{\alpha Y_{t+1} - I_{t+1}}{K_{t+1}} \right], \quad Q_t = \frac{1}{\varphi' \left(\frac{I_t}{K_t} \right)}. \tag{3.6}$$

[1]Specifically, $a_1 = \frac{1}{\xi - 1} \left(1 - \delta - \exp(\bar{\mu}) \right)$, $a_2 = \left(\exp(\bar{\mu}) - 1 + \delta \right)$, in which $\bar{\mu}$ is the unconditional mean μ_{s_t}. We find state values of remaining quantities from the conditions $\varphi\left(\frac{I}{K}\right) = 1$, $\varphi'\left(\frac{I}{K}\right) = 1$. In particular, the steady state investment-capital ratio is $\frac{I}{K} = \exp(\bar{\mu}) - 1 + \delta$.

Aggregate Dividends

The return on investment can be interpreted as the return of an equity claim to the unlevered firm's payouts (Restoy and Rockinger 1994):

$$D_t = Y_t - w_t N_t - I_t = \alpha Y_t - I_t. \tag{3.7}$$

However, the aggregate stock market dividends observed in reality are not directly comparable to the endogenous firm's payouts in the model mainly because the equity prices in the data are for leveraged corporations, in contrast to unlevered payments of a production firm considered in our setting. Further, the standard frictionless model generates the countercyclical firm's payouts and hence this would substantially decrease the equity risk premium and equity volatility. The main aim of this paper is to explore the link between rational parameter learning and risk premiums embedded in option prices. For the sake of a convenient interpretation of our results, we do not consider extensions of the model that would resolve the problem with endogenous dividends. We model aggregate dividends as a leverage to aggregate consumption and price an equity claim to these calibrated dividends. Specifically, log-dividend growth is defined as:

$$\Delta d_t = g_d + \lambda \Delta c_t + \sigma_d \varepsilon_t^d, \tag{3.8}$$

in which $\varepsilon_t^d \overset{\text{iid}}{\sim} N(0,1)$, λ is a leverage factor, and g_d and σ_d are the dividend growth rate and volatility, respectively. We choose g_d and σ_d to match the first and second moments of dividend growth in the data. Our choice of λ allows us to closely match the observed correlation between dividends and consumption in the data.

Macroeconomic Quantities and Equity Returns

The model does not admit a closed-form solution for the equilibrium quantities and, therefore, it is solved numerically through value function iteration. The Appendix contains the details of the solution algorithm. Having solved the model numerically, we look at the asset pricing implications of rational parameter learning in the production-based setting. We start our quantitative investigation by looking at the standard moments of macroeconomic quantities and equity returns. The model solution provides equilibrium investment and consumption decisions as functions of state variables, in addition to the price-dividend ratio of an equity claim to the calibrated aggregate dividends. Therefore, we can readily simulate equity returns as follows:

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{P_{t+1}/D_{t+1} + 1}{P_t/D_t} \cdot e^{\Delta d_{t+1}}.$$

The Variance Risk Premium

The main contribution of our paper is to rationalize the salient features of index options, while explaining the salient moments of fundamentals and equity prices in the production-based setting. The first puzzling feature associated with option prices is the variance risk premium defined as the difference between the risk-neutral and the physical expectations of the aggregate stock market return variance for a given horizon. Following Bollerslev, Tauchen, and Zhou (2009) and Carr and Wu (2009), we define the variance premium between the periods t and $t+1$ as:

$$vp_t = VIX_t^2 - VOL_t^2,$$

in which VIX_t^2 and VOL_t^2 denote expectations of return variance under the risk-neutral $\mathbb{Q}$ and the physical $\mathbb{P}$ probability measures, respectively. The Radon-Nykodim density ratio

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t+1}}{\mathbb{E}_t(M_{t+1})}$$

associated with the pricing kernel allows us to compute the risk-neutral expectations and to evaluate vp_t in each period t. Formally, variance measures are calculated as:

$$\begin{aligned} VIX_t^2 &= \mathbb{E}_t^{\mathbb{Q}}\left[var_{t+1}(r_{t+2})\right] = \mathbb{E}_t\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \cdot var_{t+1}(r_{t+2})\right], \\ VOL_t^2 &= \mathbb{E}_t\left[var_{t+1}(r_{t+2})\right]. \end{aligned}$$

in which

$$var_{t+1}(r_{t+2}) = \mathbb{E}_{t+1}\left[r_{t+2}^2\right] - \left[\mathbb{E}_{t+1}\left[r_{t+2}\right]\right]^2.$$

Since the model is calibrated at the quarterly frequency, the quantity vp_t effectively measures the variance premium over a quarterly horizon. For convenience of the comparison with empirical estimates in the existing literature, we report the descriptive statistics of the variance premium at the monthly frequency.

Implied Volatilities

We further compute the model-based prices of European put options P_t^o and solve for their Black-Scholes implied volatilities σ_t^{imp}. Consider a European put option written on the ex-dividend price of the equity P_t. Denote the relative price of the τ-period European put option as $\mathcal{O}_t(\tau, K) = \frac{P_t^o(\tau,K)}{P_t}$, where the strike price K is expressed as a ratio to the price of the equity. Substituting the return on the put option into the equilibrium condition (3.5), the relative price should satisfy:

$$\mathcal{O}_t(\tau, K) = \mathbb{E}_t\left[\prod_{k=1}^{\tau} M_{t+k} \cdot \max\left(K - \frac{P_{t+\tau}}{P_t}, 0\right)\right]. \tag{3.9}$$

We express $\frac{P_{t+\tau}^e}{P_t^e}$ in terms of the dividend growth rates and price-dividend ratios on the equity and compute model-based European put prices $\mathcal{O}_t = \mathcal{O}_t(\tau, K)$ via Monte Carlo simulations. We convert them into Black-Scholes implied volatilities with properly annualized continuous interest rate r_t and dividend yield q_t. Thus, given the time to maturity τ, the strike price K, the risk-free rate r_t and dividend yield q_t, the implied volatility $\sigma_t^{\text{imp}} = \sigma_t^{\text{imp}}(\tau, K)$ solves the equation:

$$\mathcal{O}_t = e^{-r_t\tau} \cdot K \cdot N(-d_2) - e^{-q_t\tau} \cdot N(-d_1), \tag{3.10}$$
$$d_{1,2} = \frac{\ln\left(\frac{1}{K}\right) + \tau\left(r_t - q_t \pm \frac{\left(\sigma_t^{\text{imp}}\right)^2}{2}\right)}{\sigma_t^{\text{imp}}\sqrt{\tau}}.$$

It is worth noting that the option prices are calculated conditional on other state variables in the economy. For convenience, we do not write the extra arguments, which would include the capital stock and the regime of the economy as well as the subjective investor's beliefs in the case of the model with unknown parameters.

3.3 Quantitative Analysis

We now calibrate the production economy to illustrate quantitatively the role of parameter uncertainty for explaining the salient features of macroeconomic quantities, equity returns and option prices. We use the U.S. National Income and Product Accounts (NIPA) tables to construct the historical U.S. time series of consumption, investment, capital and output for the period 1947:Q1 to 2016:Q4. We further retrieve the data from the Center for Research in Security Prices (CRSP) to obtain aggregate equity market dividends and asset returns for the corresponding time horizon. The data related to the variance premium measures cover the period 1990:Q1 to 2016:Q4 and is obtained from the Chicago Board of Options Exchange (CBOE). Finally, we calculate implied volatility curves using the prices of European options written on the S&P 500 index and traded on the CBOE as provided by OptionMetrics. The option data set spans the period 1996:Q1 to 2016:Q4. A detailed description of our data construction of U.S. time series is provided in the Appendix. We then calibrate the economy at a quarterly frequency. Analytical solutions of equilibrum conditions are not available either for the full information case or for the incomplete information setting. Thus, we solve the model numerically for each case using the methodology provided in the Appendix. Having found the numerical solution, we compare the historical moments with model-implied statistics of quantities and asset prices based on the 1,000 simulations of the economy.

3.3.1 Calibration

Table 3.1 summarizes the choice of parameters in the production economy of this paper. Consistent with the real business cycle literature, we set the capital share in a Cobb-Douglas production function at $\alpha = 0.36$ and the quarterly capital depreciation rate at $\delta = 0.02$. The constants a_1 and a_2 are chosen such that there are no adjustment costs in the non-stochastic steady state. The value of the EIS is under a long-standing debate in the literature. Following Bansal and Yaron (2004), Gourio (2012), Ai, Croce, and Li (2013) and Bansal et al. (2014), we choose $\psi = 2$. Further, the subjective discount factor is set at $\beta = 0.995$ to produce a low mean of the risk-free rate. The relative risk aversion is equal to $\gamma = 7$. This is a conservative value within a range of plausible values considered by Mehra and Prescott (1985). The costs for adjusting capital are set to $\xi = 7$. These choices of risk aversion and capital adjustment costs jointly generate the large equity premium and volatility of equity returns, smooth consumption and volatile investment, while reasonably matching the correlations between consumption, investment and output. Even though the salient features of the variance premium and the volatility curves implied by option prices are not directly targeted during calibration, we show that the model with fully rational parameter learning can reasonably replicate these hard-to-match features of the data by generating a sizable variance premium and a realistic implied volatility surface.

Table 3.1
Benchmark Calibration

Parameter	Description	Value
Panel A: Preferences, Production and Capital Adjustment Costs Functions, and Financial Leverage		
β	Discount factor	0.995
γ	Risk aversion	7
ψ	EIS	2
α	Capital share	0.36
δ	Depreciation rate	0.02
ξ	Adjustment costs parameter	7
a_1	Normalization	-0.0075
a_2	Normalization	0.3877
Panel B: Markov-switching Model of Productivity Growth		
π_{11}	Transition probability from expansion to expansion	0.966
π_{22}	Transition probability from recession to recession	0.712
μ_1	Productivity growth in expansion	0.48
μ_2	Productivity growth in recession	-1.25
σ_1	Productivity volatility in expansion	1.35
σ_2	Productivity volatility in recession	2.36
Panel C: Dividends Growth Process		
λ	Leverage ratio	4.5
g_d	Mean adjustment of dividend growth	-1.05
σ_d	Std. deviation of dividend growth shock	6.2

This table reports the parameter values in the benchmark calibration. Panel A presents preferences parameters, values in the production and adjustment costs functions. Panel B shows the maximum likelihood estimates of parameters in a two-state Markov-switching model for productivity growth. We obtain these estimates by applying the expectation maximization algorithm (Hamilton 1990) to quarterly total factor productivity growth rates from 1947:Q1 to 2016:Q4.

Panel B in Table 3.1 also shows the maximum likelihood estimates for the transition probabilities π_{ii}, productivity growth rates μ_i as well as the volatilities σ_i for each state $i = 1, 2$ in a parsimonious two-state Markov switching model.[2] The results indicate two separate states of the economy: an expansion with high mean and low volatility, and a recession with low mean and high volatility of productivity growth. The expansion is persistent with the mean duration of around 7.5 years and, according to our estimates, the recession is a brief economic slowdown with the mean duration of less than 1 year. Productivity is estimated to grow at the quarterly rate of about 0.48% in expansions and about -1.25% in recessions, while the volatility almost doubles when switching from a high

[2]Note that it is possible to estimate a model with a larger number of states or to assume independent regime changes in mean and volatility of productivity growth. Even though a more complex specification should better capture time-variation in mean and volatility of productivity growth, this would result in an increased number of parameters, making the model solution very costly numerically due to the curse of dimensionality. Since a simple two-state model of this paper can reasonably explain option prices, a richer specification would certainly improve our predictions.

growth state to a low growth regime. Our specification extends the models of Cagetti et al. (2002) and Babiak and Kozhan (2018) by introducing regime switches in the mean and volatility of productivity growth. Our estimates remain generally consistent with those reported in Cagetti et al. (2002) and Babiak and Kozhan (2018).

In their production-based economy, Babiak and Kozhan (2018) find that the impact of fully rational pricing of parameter uncertainty crucially depends on the introduction of procyclical dividends. They further introduce investment frictions in the form of costly reversibility that would generate the more procyclical endogenous firm's payouts. In this paper, we do not consider any extensions of the standard business cycle model and focus on pricing a claim to exogenous calibrated dividends defined as a leverage to aggregate consumption (Bansal and Yaron 2004). Relative to Babiak and Kozhan (2018), we study a learning problem in a setting in which productivity volatility switches between the states. Furthermore, we show that the lregime switches in productivity volatility are critical for quantitatively explaining the large and volatile variance premium as well as the slope of volatility curves implied by index option prices.

Panel C in Table 3.1 reports the parameter values of the calibrated dividend process. We set a leverage factor at $\lambda = 4.5$. The annual consumption volatility in our simulations turns out to be around 1.2% thus the systematic annual dividend volatility is around 5.4%. Further, we fix the remaining two parameters at $g_d = -1.05$ and $\sigma_d = 6.2$ to approximately capture the observed mean and volatility of aggregate stock market dividends. The choice of λ and σ_d also implies a positive sample correlation between consumption and dividends, which ranges between 0.35 and 0.45 in the simulations of different models. These sample correlations correspond well to the empirical point estimate of 0.45.

3.3.2 Unconditional Moments

We start our quantitative investigation by comparing the observed features of the data with the model-based statistics in the two versions of the benchmark calibration. First, we report the results for the case in which the agent knows true parameter values in the productivity growth process. Second, we consider the framework with parameter uncertainty in which the investor learns about π_{ii} and μ_i for each state $i = 1, 2$, while σ_i's are assumed to be known. As it is common in the Bayesian literature, we employ conjugate beta and normal distributions for the transition probabilities and mean growth rates. We further calibrate the hyper-parameters to embed the realistic prior information of agents in the model. First, we consider various lengths of a prior learning period incorporating the information based on 100, 150 and 200 years of prior learning. Since we start our asset pricing exercise after World War II, training samples of 100 and 150 years effectively means that the representative investor started learning about the unknown structure of the economy in the middle of the nineteenth century or at the end of the eighteenth century, respectively. These dates approximately correspond to the beginning of the historical U.S. consumption and GDP growth series in the Barro-Ursua Macroeconomic Database. Second, we center the mean prior beliefs at the true MLE parameter estimates obtained from the post-war data. Thus, our results are not driven by the pessimistic experience of the Great Depression and the two World Wars, but are rather the manifestation of rational parameter learning and the information contained in the post-war data.

Quantities, Cash Flows and Equity Returns

Panel A in Table 3.2 shows that rational parameter learning improves the model's fit with the empirical moments of macroeconomic quantities compared to the full information case. In particular, the model with known parameters predicts strongly correlated consumption, investment and output, whereas priced parameter uncertainty significantly reduces the correlation between quantities. In particular, the unknown parameter model with a century long training sample and unbiased prior beliefs generates the correlation between investment growth and output growth of about $\rho(\Delta i, \Delta y) = 0.91$, whereas the correlations of consumption growth with output growth and investment growth come out about $\rho(\Delta c, \Delta y) = 0.77$. and $\rho(\Delta c, \Delta i) = 0.45$. Thus, rational parameter learning generates additional risks that help better capture the macro dynamics of the economy compared to the complete information model and broadly to other existing frameworks (Kaltenbrunner and Lochstoer 2010; Liu and Miao 2015). Notice that the impact of parameter learning on correlations descends very slowly over time. The quantities gradually become more correlated as the length of a training period becomes longer. This result is in line with the observed pattern of empirical macroeconomic series that appear to be more correlated in the post-war period compared to a longer sample.

Panel A in Table 3.2 also shows that both models with parameter uncertainty and complete investor knowledge can reasonably capture the volatility of macroeconomic quantities, though parameter learning produces lower autocorrelation in consumption growth compared to the data. The latter feature echoes the empirical results of the growing literature on alternative measures of consumption. In particular, Savov (2011) documents that a measure of consumption called "garbage" has several times lower autocorrelation compared to the reported NIPA consumption. Recently, Kroencke (2017) suggest that one possible explanation is the filtering process used to generate the series of NIPA consumption. Our paper shows that rational parameter learning can endogenously account for the low consumption autocorrelation in line with this new evidence. The bottom of Panel A shows that the calibrated models for dividends match well the empirical statistics of aggregate stock market dividends.

Panel B in Table 3.2 summarizes the average annualized financial moments observed in the data and generated by various models. The "Data" column reports the low mean and volatility of the risk-free rate, the large equity premium and excess return volatility, the large equity Sharpe ratio and the low mean of the log price-dividend ratio. These stylized features pose a challenge for standard asset pricing models. Indeed, the framework with known parameters cannot match any of the salient moments. In contrast, rational pricing of parameter uncertainty has a large impact on equity valuations. Parameter learning produces more than a three-fold increase in the equity premium compared to the known-parameter model. Rational belief revisions amplify the impact of macroeconomic shocks that help generate substantial excess return volatility and increase the mean Sharpe ratio. The table further shows that the increased economic uncertainty in the model with unknown parameters lowers the interest rates and equity valuations, hence allowing to better match the average risk-free rate and price-dividend ratio. Even though the volatility of the log price-dividend ratio remains below the sample estimate, it is around three times higher than in the complete information setting.

Variance Premium and Option Prices

The top of Panel C in Table 3.2 provides monthly statistics for the variance risk premium in the data and various models. As shown in the "Data" column, there is indeed a large and volatile variance premium in the data. The variance premium has a positive skewness and an excess kurtosis that indicate a fat-tailed distribution of the quantity. The bottom of Panel C in Table 3.2 further reports unconditional volatility of return variance under the physical measure, VOL^2, and summary statistics of return variance under the risk-neutral measure as captured by the squared VIX index, VIX^2. The empirical quantities

Table 3.2
Sample Moments

	Data	PPU				FI
		100 yrs	150 yrs	200 yrs	∞ yrs	
Panel A: Macroeconomic Quantities and Cash Flows						
$\sigma(\Delta c)$	1.34	1.20	1.19	1.18	1.17	1.17
$\sigma(\Delta i)$	4.79	3.78	3.74	3.72	3.72	3.77
$\sigma(\Delta y)$	2.56	1.92	1.92	1.92	1.92	1.92
$ar1(\Delta c)$	0.29	0.01	0.05	0.08	0.14	0.18
$\rho(\Delta i, \Delta y)$	0.72	0.91	0.94	0.96	0.98	0.99
$\rho(\Delta c, \Delta y)$	0.45	0.77	0.84	0.89	0.95	0.97
$\rho(\Delta c, \Delta i)$	0.35	0.45	0.62	0.72	0.87	0.94
$E(\Delta d)$	1.95	1.40	1.24	1.16	1.03	0.90
$\sigma(\Delta d)$	10.58	11.37	11.34	11.33	11.31	11.32
$ar1(\Delta d)$	0.20	0.01	0.01	0.01	0.01	0.01
$\rho(\Delta c, \Delta d)$	0.45	0.45	0.42	0.40	0.37	0.36
Panel B: Returns						
$E(R_f) - 1$	1.01	1.72	1.84	1.92	2.06	2.22
$\sigma(R_f)$	1.09	0.35	0.33	0.31	0.29	0.27
$E(R - R_f)$	6.34	6.87	5.51	4.70	3.32	2.04
$\sigma(R - R_f)$	18.65	20.50	19.11	18.18	16.44	14.74
$SR(R - R_f)$	0.31	0.30	0.26	0.22	0.18	0.13
$E(p - d)$	3.01	3.03	3.19	3.30	3.54	3.92
$\sigma(p - d)$	0.34	0.13	0.11	0.09	0.06	0.04
$ar1(p - d)$	0.95	0.82	0.81	0.80	0.78	0.79
Panel C: Variance Premium						
$E(VP)$	10.24	8.53	5.12	3.47	1.39	0.31
$\sigma(VP)$	10.49	7.09	4.35	2.94	1.17	0.26
$skew(VP)$	2.62	3.28	3.23	3.21	3.16	3.08
$kurt(VP)$	14.15	16.55	16.30	16.06	15.59	14.67
$\sigma(VOL^2)$	26.14	40.50	26.30	18.27	8.03	2.29
$E(VIX^2)$	40.10	54.33	40.91	33.41	23.63	17.66
$\sigma(VIX^2)$	34.34	47.55	30.39	20.87	8.97	2.48
$skew(VIX^2)$	3.45	3.32	3.27	3.24	3.23	3.20
$kurt(VIX^2)$	20.72	16.37	16.02	15.86	15.75	15.55

This table reports the average moments from 1,000 simulations of the production economy considered in this paper. The historical data moments are reported in the data column. The macroeconomic quantities, dividends, and return variables correspond to the U.S. data from 1947:Q1 to 2016:Q4. The variance premium statistics are based on the U.S. data from 1990:Q1 to 2016:Q4. The PPU column refers to the production economy with rational pricing of parameter uncertainty that entails unknown transition probabilities, mean growth rates and volatilities. The FI column presents the results of the full information case in which the parameters are known. $E(x)$ and $\sigma(x)$ denote the average sample mean and standard deviations of x, respectively. $ar1(x)$ and $\rho(x, y)$ denote the average sample autocorrelation of x and correlation between x and y, respectively. All macro and return statistics are expressed in annualized terms, except for correlations and autocorrelations expressed in quarterly terms. The variance premium statistics are expressed in monthly terms.

of VOL^2 and VIX^2 exhibit a large time-variation, with the latter one being more volatile in the data. The squared VIX time series has been historically large, especially during periods of high stock market volatility, which leads to the sizable mean, positive skewness and excess kurtosis of the squared VIX. These salient moments of the data are difficult to reproduce in the standard asset pricing models.

The last column in Panel C of Table 3.2 documents the failure of a rational expectations economy to reconcile the empirical moments. Indeed, in that case, the average variance premium and its volatility are almost zero. A poor performance of the model with complete investor knowledge originates from the low volatility of return variance under both probability measures. The table shows that time-variation in VOL^2 and VIX^2 is smaller than the empirical numbers by an order of magnitude. In contrast, the "PPU" columns in Panel C of Table 3.2 show that priced parameter uncertainty increases the relevant moments of the variance premium and return variance measures that become comparable to the observed statistics. In particular, the model with fully rational parameter learning increases the first and second moments of the variance premium by almost a factor of 30 relative to the known parameter case. Further, rational pricing of belief revisions has a long-lasting impact on the variance premium, which remains significant even after 200 years of prior learning. Quantitatively, the mean and volatility of the variance premium in the economy with rational parameter learning is more than ten times larger compared to the full information setting even after a 200-year prior period. The model better captures the size and volatility of the variance premium since rational parameter learning produces significantly different magnitudes for the return variance under the physical and risk-neutral measures. The table shows that there is more than a 15-fold increase in the unconditional volatility of the VOL^2 and VIX^2 time series compared to the case of no parameter uncertainty. Overall, rationally accounting for parameter uncertainty helps to reconcile the salient moments of the variance premium and conditional return variances.

Figure 3.1 shows the implied volatility curves for the data and various models. The three panels plot the implied volatilities as a function of moneyness defined as a ratio of a spot price to a strike price for 3, 6, and 12-month maturities. The empirical curves are obtained from the polynomial extrapolation of the historical implied volatilities as described in the Appendix. The model-based results are the sample averages of the implied volatilities predicted by the corresponding model and are calculated as outlined in Section 3.2.4. Several features in the empirical curves are noteworthy. First, the implied volatil-

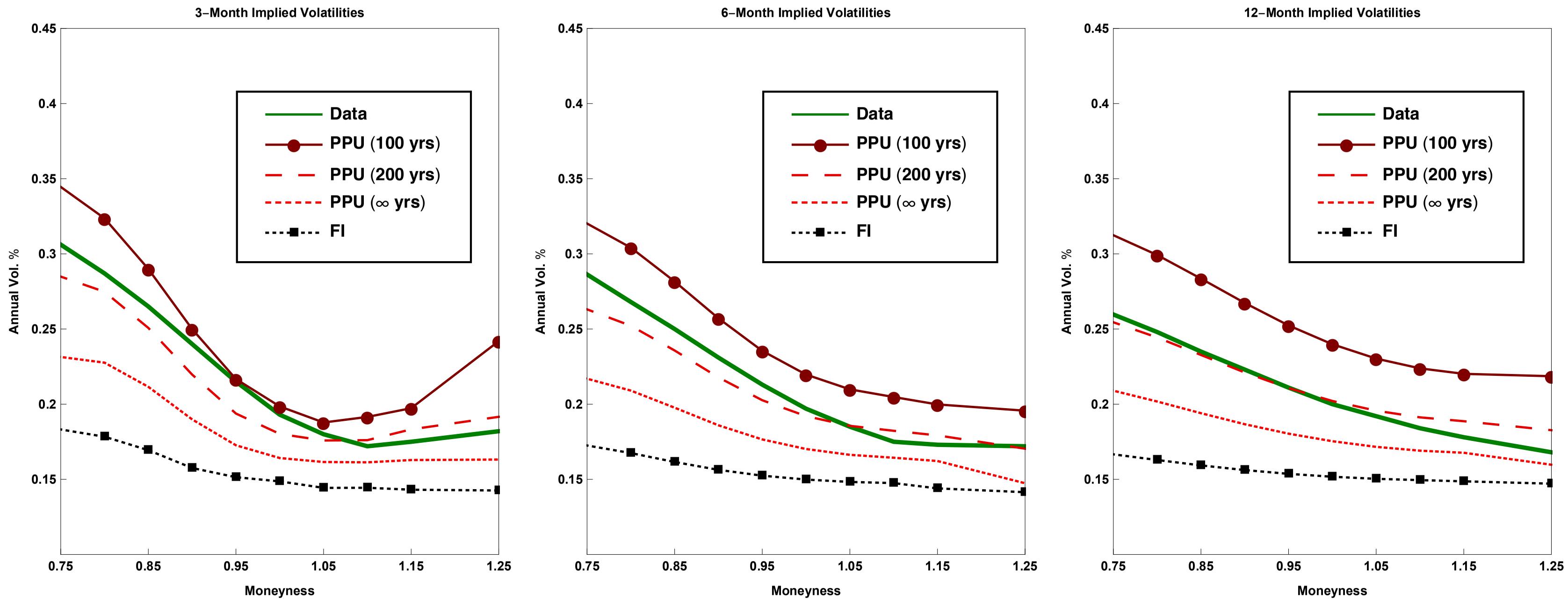

Figure 3.1: 3-, 6-, and 12-month Implied Volatilities: Data, PPU, and FI. The three panels plot the 3-, 6-, and 12-month implied volatility curve as a function of moneyness (Strike/Spot Price) for the data and the models with fully rational parameter learning (PPU) and full information (FI). The empirical statistics are for the U.S. data from January 1996 to December 2016. For each model, I simulate 1,000 economies at a quarterly frequency with a sample size equal to its empirical counterpart. The model-based curves are calculated for option prices using the annualized model-implied interest rate r_t and dividend-yield q_t in each period. The simulation results are averages of implied volatilities based on these 1,000 artificial series.

ities for out-of-the-money put options exhibit a pronounced downward slopping pattern called the skew. Second, the implied volatilities for the 3-month maturity options slightly increase at high moneyness, a feature called the smirk. Third, the implied volatilities for all moneyness values and maturities appear to be higher than the annualized stock market volatility. These level and slope patterns of the implied volatility surface constitute a challenge for the equilibrium asset pricing models.

Turning to the model-implied results, the panels in Figure 3.1 show the failure of the model with complete investor knowledge in replicating the empirical regularities. The rational expectations framework generates an almost flat term structure of the implied volatilities (a dotted black line). In the absence of parameter uncertainty, the three volatility curves for 3, 6, and 12-month maturities are approximately equal to the equity return volatility. Further, the plots in Figure 3.1 illustrate the success of the model with rational parameter learning in capturing the main properties of the empirical data. Several insights are noteworthy. First, priced parameter uncertainty inflates the level of the implied volatility curves. The higher degree of parameter uncertainty as measured by a shorter prior sample leads to an upward shift in the implied volatility surface. Second, parameter learning helps to capture the steep skew at all three maturities and to reproduce the smirk pattern in the 3-month implied volatilities consistent with the empirical findings. Similarly to a level shift, the implied volatility curves become steeper at both ends of the moneyness range in response to higher structural uncertainty about the unknown parameters. Third, the impact of parameter learning on the level and the overall shape of the implied volatility curves is persistent as one can see by comparing the results of the models with a 100-year training sample (a solid red line with dots) and a 200-year prior period (a dashed red line). In the latter case, the curves flatten and shift downward but remain largely consistent with the empirical lines. Interestingly, the influence of belief revisions with priced parameter uncertainty does not disappear even after a very long period of learning as the infinite-horizon model (a dotted red line) dominates the results of the full information framework.

3.3.3 Conditional Moments

To better understand the source of the model improvement, we examine the conditional dynamics of excess equity returns and variance measures in the models with known parameters and parameter uncertainty. Due to the multidimensional nature of learning, we

Table 3.3
Conditional Moments

	PPU				FI
	100 yrs	150 yrs	200 yrs	∞ yrs	
Panel A: Expansion					
$E_t(R_{t+1} - R_{f,t+1})$	6.58	4.92	4.03	2.77	1.64
$\sigma_t(R_{t+1} - R_{f,t+1})$	19.34	17.84	16.82	15.45	14.21
$SR_t(R_{t+1} - R_{f,t+1})$	0.30	0.25	0.22	0.17	0.11
$E_t(VP_{t+1})$	8.10	4.44	2.68	0.95	0.21
$\sigma_t(VP_{t+1})$	4.85	2.99	1.83	0.69	0.16
$\sigma_t(VIX^2_{t+1})$	33.13	22.00	14.81	6.45	2.15
$\sigma_t(VOL^2_{t+1})$	28.28	19.01	12.99	5.77	1.99
Panel B: Recession					
$E_t(R_{t+1} - R_{f,t+1})$	30.30	21.19	16.12	9.32	4.53
$\sigma_t(R_{t+1} - R_{f,t+1})$	46.45	37.54	32.01	23.37	17.58
$SR_t(R_{t+1} - R_{f,t+1})$	0.58	0.50	0.45	0.37	0.25
$E_t(VP_{t+1})$	28.75	16.14	9.69	3.36	0.74
$\sigma_t(VP_{t+1})$	13.85	8.02	4.82	1.66	0.38
$\sigma_t(VIX^2_{t+1})$	89.79	56.20	37.54	15.24	4.89
$\sigma_t(VOL^2_{t+1})$	75.94	48.19	32.72	13.58	4.52

This table reports the conditional asset-pricing moments in the benchmark calibration with rational parameter learning (PPU) and full information (FI). These conditional moments are dependent on the observable state of the economy, while other state variables (beliefs about unknown parameters) are centered at the true parameter values. Symbols are defined in Table 3.2.

do not focus on a particular trajectory of the productivity growth series that would lead to belief revisions of all unknown parameters. For convenience, we illustrate the moment statistics of selected quantities at the onset of each regime conditional on the unbiased parameter beliefs.

Table 3.3 presents the conditional distributions of excess equity returns, equity Sharpe ratios, the variance premium and return variances under both probability measures over the various phases of the business cycle. Qualitatively, the two models with complete investor knowledge and rationally priced parameter uncertainty predict similar behaviour of asset prices across the two state. Namely, the mean and volatility of the equity and variance risk premiums in short recessions come out above the corresponding values during expansions. The same observation is true for the equity Sharpe ratio and volatility of return variances. Although these responses are in line with countercyclical dynamics of risk premiums and volatility observed in the data, the two models predict significantly different magnitudes. Quantitatively, in expansions priced parameter uncertainty increases the average equity and variance risk premiums as well as the volatility of the variance premium, the equity return variance and the squared VIX index by the factors of 4, 38, and 30, 14, 15, respectively, compared to the full information case. Further, the annualized Sharpe ratio of excess equity returns almost triples from 0.11 to 0.30, while the return volatility increase from 14.21% to 19.34%. The amplification mechanism turns out to be stronger during recessions. For instance, the first and second moments of the excess equity returns in the parameter learning model are now around 6.5 and 2.5 times higher than with known parameters, whereas the variance premium volatility and time-variation of VOL^2 and VIX^2 are up by a factor of more than 36, 16, and 18, respectively.

Overall, the above analysis strongly suggests that fully rational pricing of parameter uncertainty in the productivity growth process is an important amplification mechanism of the conditional moments of asset prices, especially equity return variances under the physical and risk-neutral probability measures. In the presence of priced parameter uncertainty, the representative investor is strongly concerned about the high realized equity variance and the aggregate stock market declines in response to pessimistic belief revisions during recessions. Consequently, he is willing to pay a high premium for the variance swaps and European put options that would provide a high payoff in the states of low productivity growth. Interestingly, although very parsimonious, the model with a realistic learning problem and rational pricing of belief changes is able to capture a large number of features in the data without relying on the peso-type events or exotic preferences.

3.4 The Impact of Priced Parameter Uncertainty and Volatility Risks

This section performs a two-part comparative statics exercise to examine the impact of fully rational parameter learning and regime-switching volatility in the productivity process. First, we study the role of priced parameter uncertainty by comparing the benchmark framework and the model with anticipated utility, a commonly assumed approach for dealing with parameter uncertainty. Under anticipated utility pricing, the representative agent updates his beliefs about unknown parameters upon arrival of the new data, but he treats his current beliefs as true parameter values in the decision-making process. The numerical solution methodology for this case is presented in the Appendix. Second, we shut down regime shifts in productivity growth volatility by setting the volatility param-

Table 3.4
Sensitivity Analysis

	Data	PPU	AU	PPU_σ	AU_σ
		$\sigma_1 \neq \sigma_2$		$\sigma_1 = \sigma_2$	
Panel A: Macroeconomic Quantities					
$\sigma(\Delta c)$	1.34	1.20	1.18	1.18	1.20
$\sigma(\Delta i)$	4.79	3.78	3.75	3.90	3.93
$\sigma(\Delta y)$	2.56	1.91	1.92	1.98	1.99
$ar1(\Delta c)$	0.29	0.01	0.18	0.11	0.24
$\rho(\Delta i, \Delta y)$	0.72	0.91	0.99	0.96	0.99
$\rho(\Delta c, \Delta y)$	0.45	0.77	0.97	0.89	0.98
$\rho(\Delta c, \Delta i)$	0.35	0.45	0.93	0.73	0.94
$E(\Delta d)$	1.95	1.40	0.90	0.57	0.24
$\sigma(\Delta d)$	10.58	11.37	11.34	11.32	11.37
$ar1(\Delta d)$	0.20	0.01	0.01	0.01	0.02
$\rho(\Delta c, \Delta d)$	0.45	0.45	0.36	0.39	0.35
Panel B: Returns					
$E(R_f) - 1$	1.01	1.72	2.22	1.79	2.16
$\sigma(R_f)$	1.09	0.35	0.27	0.29	0.26
$E(R - R_f)$	6.34	6.87	2.04	3.89	1.78
$\sigma(R - R_f)$	18.65	20.50	14.77	16.14	14.32
$SR(R - R_f)$	0.31	0.30	0.13	0.22	0.12
$E(p - d)$	3.01	3.03	3.92	3.46	4.00
$\sigma(p - d)$	0.34	0.13	0.08	0.07	0.03
$ar1(p - d)$	0.95	0.82	0.90	0.86	0.76
Panel C: Variance Premium					
$E(VP)$	10.24	8.53	0.31	1.69	0.16
$\sigma(VP)$	10.49	7.09	0.27	1.17	0.10
$skew(VP)$	2.62	3.28	3.10	2.80	2.61
$kurt(VP)$	14.15	16.55	15.03	11.43	10.03
$\sigma(VOL^2)$	26.14	40.50	2.29	21.84	1.30
$E(VIX^2)$	40.10	54.33	17.66	24.34	13.36
$\sigma(VIX^2)$	34.34	47.55	2.48	25.38	1.93
$skew(VIX^2)$	3.45	3.32	3.20	3.32	3.20
$kurt(VIX^2)$	20.72	16.37	15.55	11.24	10.34

This table presents the results of the comparative statics exercise for the benchmark production economy. The PPU and AU columns refer to the benchmark calibration with regime-switching volatilities under rational pricing of parameter uncertainty and an anticipated utility case. The PPU_σ and AU_σ columns refer to the economy with no regime shifts in the productivity growth volatility under the two approaches of dealing with parameter uncertainty. Symbols are defined in Table 3.2.

eter to a constant steady-state value implied by the estimated Markov switching process. We solve this model with homoskeadastic volatility for the two cases: priced parameter

Table 3.5
Conditional Moments

	PPU	AU		PPU_σ	AU_σ
Moments	$\sigma_1 \neq \sigma_2$			$\sigma_1 = \sigma_2$	
Panel A: Expansion					
$E_t(R_{t+1} - R_{f,t+1})$	6.58	1.64		3.41	1.67
$\sigma_t(R_{t+1} - R_{f,t+1})$	19.34	14.21		15.26	14.09
$SR_t(R_{t+1} - R_{f,t+1})$	0.30	0.12		0.22	0.12
$E_t(VP_{t+1})$	8.10	0.21		1.28	0.09
$\sigma_t(VP_{t+1})$	4.85	0.16		0.96	0.08
$\sigma_t(VIX^2_{t+1})$	33.13	2.15		17.05	1.15
$\sigma_t(VOL^2_{t+1})$	28.28	1.99		14.79	1.07
Panel B: Recession					
$E_t(R_{t+1} - R_{f,t+1})$	30.30	4.66		15.16	3.38
$\sigma_t(R_{t+1} - R_{f,t+1})$	46.45	17.83		34.60	16.64
$SR_t(R_{t+1} - R_{f,t+1})$	0.58	0.25		0.50	0.20
$E_t(VP_{t+1})$	28.75	0.77		12.88	0.45
$\sigma_t(VP_{t+1})$	13.85	0.40		6.08	0.23
$\sigma_t(VIX^2_{t+1})$	89.79	5.27		43.74	3.19
$\sigma_t(VOL^2_{t+1})$	75.94	4.86		37.66	2.97

This table reports the conditional asset-pricing moments in the benchmark calibration and the economy with no regime shifts in the productivity growth volatility. For each model, the table presents the results with rational parameter learning and anticipated. These conditional moments are dependent on the observable state of the economy, while other state variables (beliefs about unknown parameters) are centered at the true parameter values. Symbols are defined in Table 3.2.

uncertainty and anticipated utility. The robustness exercise conducts the comparison between different models to evaluate the contribution of alternative assumptions about the productivity process and learning.

Table 3.4 presents the moments of quantities and asset prices. The "PPU" and "AU" columns show that using anticipated utility pricing of parameter uncertainty greatly reduces the size and volatility of risk premiums in the benchmark specification. The average excess equity returns are more than three times smaller compared to the benchmark model with rational parameter learning, while the first and second moments of the variance premium are reduced by an order of magnitude and are almost equal to zero. Generally, when looking at all results for macroeconomic quantities and asset prices, one can observe that the anticipated utility model performs almost identical to the full information case. The "PPU_σ" and "AU_σ" columns in Table 3.4 show the performance of the model with no regime shifts in the productivity growth volatility under the two learning types. Similarly to the benchmark results, anticipated utility generates the moments far from the

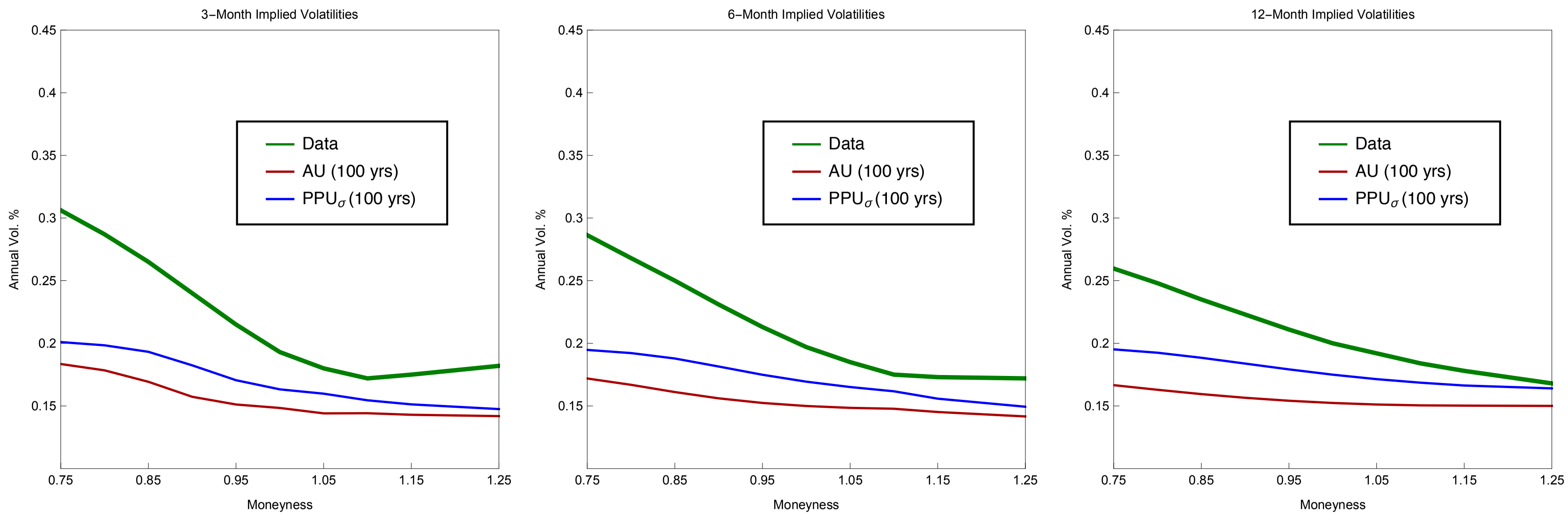

Figure 3.2: 3-, 6-, and 12-month Implied Volatilities: Data, AU and No Switches in Productivity Volatility. The three panels plot the 3-, 6-, and 12-month implied volatility curve as a function of moneyness (Strike/Spot Price) for the data and the two models: a benchmark anticipated utility framework (AU) and a calibration with rational parameter learning and no regime-switches in productivity volatility (PPU$_\sigma$). The empirical statistics are for the U.S. data from January 1996 to December 2016. For each model, I simulate 1,000 economies at a quarterly frequency with a sample size equal to its empirical counterpart. The model-based curves are calculated for option prices using the annualized model-implied interest rate r_t and dividend-yield q_t in each period. The simulation results are averages of implied volatilities based on these 1,000 artificial series.

data estimates, while rational parameter learning somehow improves the results. However, shutting down the regime switches in volatility of productivity growth significantly reduces the amplification mechanism of fully rational parameter learning on asset prices. Thus, although the risk premiums remain significant in this case, the magnitudes become much lower compared to the data. For instance, the mean and volatility of the variance premium are, respectively, six and eight times smaller than the empirical estimates.

Figure 3.2 further augments the sensitivity results by plotting the implied volatility curves for the benchmark calibration with anticipated utility (a red line) and the constant volatility framework with priced parameter uncertainty (a blue line). The plots show that the model-generated implied volatility skew at the 3, 6, and 12-month maturities significantly flattens and becomes lower than in the data, though the latter model performs slightly better.

Finally, Table 3.5 looks at the conditional moments of asset prices. It confirms that introducing time-varying productivity volatility further reinforces the impact of fully rational parameter learning on the risk premiums and volatility of return variances. The presence of regime-dependent volatility raises the concerns of the representative investor about the high volatility of innovations shocks to productivity growth in bad times. These concerns are priced under investor rational parameter learning that become instrumental in capturing the large and volatile variance premium as well as high prices of index options.

3.5 Conclusion

This paper studies a production economy with regime switches in the conditional mean and volatility of productivity growth. The representative investor faces uncertainty about the true parameters of the productivity process and rationally learns about the unknown parameters from the data. We estimate the model using the expectation maximization algorithm on U.S. post-war productivity data and show that it can generate a large variance risk premium and a realistic implied volatility surface, while simultaneously capturing the salient properties of cash-flows, equity returns, the risk-free rate and macroeconomic variables. Further, we show that the volatility risk in productivity growth, combined with priced parameter uncertainty, carries a quantitatively significant risk premium in the prices of equity index options. Shutting down time-varying volatility risk in productivity growth reduces the impact of fully rational parameter learning on the model's asset prices, especially the variance premium and the volatilities implied by index option prices. This suggests the important role of rational investor learning and fluctuating economic volatility in macro-finance models.

Appendix C

C.1 Numerical Algorithm: All Known Parameters

This section reviews the numerical solution methodology for the model with full information and parameter uncertainty. In this paper, we focus on an economy with learning about unknown transition probabilities, mean growth rates, and volatility of productivity growth.[1] For the unknown parameter case, we further provide numerical solutions for anticipated utility pricing and priced parameter uncertainty. Here, we give details on how the continuation utility and the levered equity claim are computed for the economy with all parameters known. This case simplifies solving a standard rational expectations model in which the agent knows true parameters of the economy.

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t,$$

in which $\varepsilon_t \overset{\text{iid}}{\sim} N(0,1)$, s_t is a two state Markov chain with transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1-\pi_{11} \\ 1-\pi_{22} & \pi_{22} \end{bmatrix},$$

in which $\pi_{ii} \in (0,1)$. The regimes switches in s_t are independent of the Gaussian shocks

[1] It might be instructive to consider simpler models with learning about the transition probabilities only or learning about the transition probabilities and mean growth rates. Please refer to Babiak and Kozhan (2018) for details about the numerical solution methodology in these two cases.

ε_t. We define the following stationary variables:

$$\left\{\tilde{C}_t, \tilde{I}_t, \tilde{Y}_t, \tilde{K}_t, \tilde{U}_t\right\} = \left\{\frac{C_t}{A_t}, \frac{I_t}{A_t}, \frac{Y_t}{A_t}, \frac{K_t}{A_t}, \frac{U_t}{A_t}\right\}$$

The household's problem is:

$$\tilde{U}_t = \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma} \cdot \left(\frac{A_{t+1}}{A_t} \right)^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \tag{C.1}$$

subject to the constraints:

$$\tilde{C}_t + \tilde{I}_t = \tilde{K}_t^{\alpha} \bar{N}^{1-\alpha} \tag{C.2}$$

$$e^{\Delta a_{t+1}} \tilde{K}_{t+1} = (1-\delta)\tilde{K}_t + \varphi \left(\frac{\tilde{I}_t}{\tilde{K}_t} \right) \tilde{K}_t \tag{C.3}$$

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t, \;\; \varepsilon_t \sim N(0,1) \tag{C.4}$$

$$\tilde{C}_t \geq 0, \;\; \tilde{K}_{t+1} \geq 0 \tag{C.5}$$

in which the subscript t indicates the time, $E_t(\cdot)$ denotes the expectation conditional on the information available at time t. Because the parameters are assumed known, s_t and $\tilde{K}_t$ are the only state variables in the economy. Ultimately, the recursive equation (C.1) can be rewritten as:

$$\begin{aligned} & \tilde{U}_t(s_t, \tilde{K}_t) \\ = \;\; & \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1} \left(s_{t+1}, \tilde{K}_{t+1} \right)^{1-\gamma} \cdot e^{(1-\gamma)\Delta a_{t+1}} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}} \end{aligned} \tag{C.6}$$

To solve the recursion (C.6), we use the the value function iteration algorithm. In particular, the numerical algorithm proceeds as follows:

1. We find the de-trended steady state capital $\tilde{K}_{ss}$, assuming the productivity growth equals the steady state level predicted by a Markov-switching model. The state space for capital normalized by technology is set at $[0.1\tilde{K}_{ss}, 2.6\tilde{K}_{ss}]$. We further use $n_k = 100$ points on a grid for capital in the numerical computation. A denser grid does not lead to significantly different results.

2. For any level of capital $\tilde{K}_t$ at time t, we construct a grid for $\tilde{I}_t$ with uniformly distributed points between 0 and $\tilde{K}_t^{\alpha}\bar{N}^{1-\alpha}$. Specifically, we use $n_i = 200$ points.

3. For the expectation, we use Gauss-Hermite quadrature with $n_{gh} = 8$ points. Using the quadrature weights and nodes, we can calculate the expression on the right hand side.

4. We solve the optimization problem in the Bellman equation (C.6) subject to (C.2)-(C.5) and update a new value function $\tilde{U}_t = \tilde{U}_t(s_t, \tilde{K}_t)$ given an old one $\tilde{U}_{t+1} = \tilde{U}_{t+1}(s_{t+1}, \tilde{K}_{t+1})$.

5. We iterate Steps 2-4 by updating the continuation utility on each iteration until a suitable convergence is achieved. Specifically, the stopping rule is that the distance between the new value function and the old value function satisfies $|\tilde{U}_{t+1} - \tilde{U}_t|/|\tilde{U}_t| < 10^{-12}$.

C.2 Numerical Algorithm: Anticipated Utility

In the anticipated utility case, the representative household learns about unknown parameters but ignores parameter uncertainty when making decisions. The numerical solution proceeds as follows. At each time t, the household holds his current beliefs and solves for the continuation utility and the levered equity claim in the rational expectations model in which the true parameter values in the productivity growth process are centered at the time t posterior means. In the next period $t + 1$, the household updates his beliefs upon observing new data and resolves the rational expectations economy in which the true parameters are centered at the time $t + 1$ posterior means. In sum, the numerical algorithm reduces to applying the methodology for the full information case with a set of model parameters, which are equal to the mean beliefs at each point in time t.

C.3 Numerical Algorithm: Priced Parameter Uncertainty

The numerical solution for the case of priced parameter uncertainty consists of two main steps.[2] First, we solve for the equilibrium pricing ratios when true parameters are actually known by the household (by assumption, these are learned at $T = \infty$). We find the solution for this simplest limiting economy on a dense set of state variables by applying the methods outlined in Appendix C.1. Second, we use the known parameters boundary economy as a terminal value in the backward recursion to obtain the equilibrium model solution at each time t.

C.3.1 Unknown Transition Probabilities, Unknown Mean Growth Rates, and Unknown Volatilities

Here, we outline the details of the numerical solution for the model with the unknown transition probabilities, mean growth rates and the volatility of productivity growth.

Solving for the Continuation Utility

Productivity growth is given by:

$$\Delta a_t = \mu_{s_t} + \sigma_{s_t} \cdot \varepsilon_t,$$

in which $\varepsilon_t \overset{\text{iid}}{\sim} N(0, 1)$, s_t is a two state Markov chain with transition matrix:

$$\Pi = \begin{bmatrix} \pi_{11} & 1 - \pi_{11} \\ 1 - \pi_{22} & \pi_{22} \end{bmatrix},$$

in which $\pi_{ii} \in (0, 1)$. The regimes switches in s_t are independent of the Gaussian shocks ε_t.

The representative household does not know the true values of the transition probabilities (π_{11}, π_{22}), the mean growth rates (μ_1, μ_2) and the volatilities (σ_1, σ_2), but observes states (s_t) of the economy. At time $t = 0$, the household holds priors about unknown

[2] Johnson (2007) uses this solution methodology in a case with parameter learning and power utility. Johannes, Lochstoer, and Mou (2016) and Collin-Dufresne, Johannes, and Lochstoer (2016) extend this approach to the case of Epstein-Zin utility in the endowment economy. We further extend the numerical solution to the case of Epstein-Zin utility in the production economy.

parameters and updates beliefs each period upon realization of new series and regimes. We assume a conjugate prior for all parameters: the Beta distributed prior and the truncated normal-inverse-gamma prior for the transition probabilities, the mean growth rates and volatilities, respectively.

The Beta distribution has the probability density function of the form:

$$p(\pi|a,b) = \frac{\pi^{a-1}(1-\pi)^{b-1}}{B(a,b)},$$

in which $B(a,b)$ is the Beta function (a normalization constant), a and b are two positive shape parameters. We are particularly interested in the expected value of the Beta distribution defined by:

$$E[\pi|a,b] = \frac{a}{a+b}.$$

We use two pairs of hyper-parameters (a_1, b_1) and (a_2, b_2) for unknown transition probabilities π_{11} and π_{22}, respectively. At time t, the household uses Bayes' rule and the fact that states are observable to update hyper-parameters for each state i as follows:

$$a_{i,t} = a_{i,0} + \#(\text{state } i \text{ has been followed by state } i), \tag{C.7}$$

$$b_{i,t} = b_{i,0} + \#(\text{state } i \text{ has been followed by state } j), \tag{C.8}$$

given the initial prior beliefs $a_{i,0}$ and $b_{i,0}$. Once we find the limiting boundary economies on the first step, we perform a backward recursion using the following state variables

$$\tau_{1,t} = a_{1,t} - a_{1,0} + b_{1,t} - b_{1,0} \tag{C.9}$$

$$\lambda_{1,t} = E_t[\pi_{11}] = \frac{a_{1,t}}{a_{1,t} + b_{1,t}} \tag{C.10}$$

$$\tau_{2,t} = a_{2,t} - a_{2,0} + b_{2,t} - b_{2,0} \tag{C.11}$$

$$\lambda_{2,t} = E_t[\pi_{22}] = \frac{a_{2,t}}{a_{2,t} + b_{2,t}} \tag{C.12}$$

Furthermore, we denote hyper-parameters of the truncated normal-inverse-gamma distributed prior for the mean and variance of productivity growth in each state i by $(\mu_{i,t}, A_{i,t})$ and $(b_{i,t}, B_{i,t})$. Formally, at time t the joint prior over the mean μ_i and variance σ_i^2 conditional on the data Δa^t and the states of the economy s^t is:

$$p(\mu_i, \sigma_i^2|\Delta a^t, s^t) = p(\mu_i|\sigma_i^2, \Delta a^t, s^t)p(\sigma_i^2|\Delta a^t, s^t),$$

in which

$$p(\sigma_i|\Delta a^t, s^t) = IG\left(\frac{b_{i,t}}{2}, \frac{B_{i,t}}{2}\right),$$
$$p(\mu_i|\sigma_i^2, \Delta a^t, s^t) = N(\mu_{i,t}, A_{i,t}\sigma_i^2).$$

We update these hyper-parameters by the Bayes' rule as follows:

$$\mu_{i,t+1} = \mu_{i,t} + \mathbf{1}_{s_{t+1}=i}\frac{A_{i,t}}{A_i+1}\left(\Delta a_{t+1} - \mu_{i,t}\right) \tag{C.13}$$
$$A_{i,t+1}^{-1} = A_{i,t}^{-1} + \mathbf{1}_{s_{t+1}=i}, \tag{C.14}$$
$$b_{i,t+1} = b_{i,t} + \mathbf{1}_{s_{t+1}=i}, \tag{C.15}$$
$$B_{i,t+1} = B_{i,t} + \mathbf{1}_{s_{t+1}=i}\frac{\left(\Delta a_{t+1} - \mu_{i,t}\right)^2}{1 + A_{i,t}} \tag{C.16}$$

in which $i \in \{1, 2\}$, $\mathbf{1}$ is an indicator function that equals 1 if the condition in subscript is true and 0 otherwise.

Note that since the hyper-parameters $A_{i,t}$'s and $b_{i,t}$'s are a function of the time spent in each period, the following 8-dimensional vector

$$X_t \equiv \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}, \mu_{1,t}, \mu_{2,t}, B_{1,t}, B_{2,t}\}$$

is sufficient statistics for the priors. Thus, we can define X_{t+1} using the equations (C.7)-(C.12), (C.13)-(C.16), the next period regime, and sufficient statistics at time t :

$$X_{t+1} = f(s_{t+1}, s_t, X_t).$$

We further define $X_t^s \equiv \{\tau_{1,t}, \lambda_{1,t}, \tau_{2,t}, \lambda_{2,t}\}$ and $X_t^{\Delta a} \equiv \{\tilde{K}_t, \mu_{1,t}, \mu_{2,t}, B_{1,t}, B_{2,t}\}$, in which the superscripts s and Δa indicate that variables in the vectors X_t^s and $X_t^{\Delta a}$ are a function only of the observed state realization s_t and a function of the realized productivity growth as well. Thus, $X_t = \left[X_t^s, X_t^{\Delta a}\right]$. Using these notations, we can rewrite

$$\tilde{U}_{t+1}(s_{t+1}, X_{t+1}) = \tilde{U}_{t+1}(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a})$$

to better indicate the dependence of state variables on specific shocks. Ultimately, the recursive equation (C.1) is of the same form:

$$
\begin{aligned}
& \tilde{U}_t(s_t, X_t) && \text{(C.17)} \\
= \quad & \max_{\tilde{C}_t, \tilde{I}_t} \left\{ (1-\beta)\tilde{C}_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[\tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a} \right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, X_t \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\psi}},
\end{aligned}
$$

in which the expectation on the right hand side is equivalent to:

$$
\begin{aligned}
& E_t \left[\tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a} \right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_t, X_t \right] \\
= \quad & \sum_{s_{t+1}=1}^{2} E_t(\pi_{s_{t+1}, s_t} | s_t, X_t^s) \quad \ldots \\
\times \quad & E_t \left[\tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a} \right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, X_t \right]. \qquad \text{(C.18)}
\end{aligned}
$$

In this case, we compute the conditional expectation in (C.18) by integrating over the conditional distribution of mean growth rates and volatilities as well as Gaussian distribution of the error term in productivity growth. In particular:

$$
\begin{aligned}
& E_t \left[\tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a_{t+1}, X_t^{\Delta a} \right) \cdot e^{(1-\gamma)\Delta a_{t+1}} \middle| s_{t+1}, s_t, X_t \right] \\
\approx \quad & \sum_{j=1}^{J} \omega_\varepsilon(j) \left[\sum_{k=1}^{K} \omega_{\sigma^2_{s_{t+1}}}(k) \sum_{l=1}^{L} \omega_{\mu_{s_{t+1}}}(l) \cdot \tilde{U}_{t+1}^{1-\gamma} \left(s_{t+1}, s_t, X_t^s, \Delta a(j,k,l), X_t^{\Delta a} \right) \cdot e^{(1-\gamma)\Delta a(j,k,l)} \middle| s_{t+1}, s_t, X_t \right] \\
& \text{(C.19)}
\end{aligned}
$$

in which $\omega_\varepsilon(j)$ is the quadrature weight corresponding to the quadrature node $n_\varepsilon(j)$ used for the integration of a standard normal shock ε_{t+1} in productivity growth, $\omega_{\sigma^2_{s_{t+1}}}(k)$ and $\omega_{\mu_{s_{t+1}}}(l)$ are the quadrature weights corresponding to the quadrature nodes $n_{\sigma^2_{s_{t+1}}}(k)$ and $n_{\mu_{s_{t+1}}}(l)$ used for the integration of a truncated inverse gamma variable $\sigma^2_{s_{t+1}}$ and a truncated standard normal variable $\mu_{s_{t+1}}$, respectively. The observed realized productivity growth, $\Delta a(j,k,l)$, and a state variable, $\tilde{K}_{t+1}(j,k,l)$, are updated as follows:

$$
\Delta a(j,k,l) = n_{\mu_{s_{t+1}}}(l) + \sqrt{n_{\sigma^2_{s_{t+1}}}(k)} \cdot n_\varepsilon(j) \tag{C.20}
$$

$$
e^{\Delta a(j,k,l)} \tilde{K}_{t+1}(j,k,l) = (1-\delta)\tilde{K}_t + \varphi\left(\frac{\tilde{I}_t}{\tilde{K}_t}\right)\tilde{K}_t, \tag{C.21}
$$

in which

$$\tilde{I}_t = \tilde{K}_t^{\alpha} \bar{N}^{1-\alpha} - \tilde{C}_t. \tag{C.22}$$

Finally, the numerical backward recursion can be perform by using (C.17)-(C.22). The boundary conditions are defined by the limiting economies $\tau_{1,\infty}$ and $\tau_{2,\infty}$, in which the transition probabilities π_{11} and π_{22}, mean growth rates μ_1 and μ_2, volatilities σ_1 and σ_2, are known.

Solving for a Dividend Claim

We also solve for the price-dividend ratio of the equity claim written on aggregate dividends, which are defined as a leverage to aggregate consumption. Let exogenous aggregate dividends be given by:

$$\Delta d_{t+1} = g_d + \lambda \Delta c_{t+1} + \sigma_d \varepsilon_{d,t+1},$$

in which $g_d = \Big(1-\lambda\Big)\Big(E(\mathbb{P}(s_\infty = 1|\pi_{11},\pi_{22}))\mu_1 + E(\mathbb{P}(s_\infty = 2|\pi_{11},\pi_{22}))\mu_2\Big)$ and $\mathbb{P}(s_\infty = i|\pi_{11},\pi_{22})$ is the ergodic probability of being in state i conditional on the transition probabilities π_{11} and π_{22}. Note that the long run mean of dividends growth, g_d, changes under the household's filtration, though the true long run growth is constant. The subjective beliefs about the true parameter values induce fluctuations in g_d, which can be expressed as $g_d = g_d(s_{t+1}, s_t, X_t)$.

The equilibrium condition for the price-dividend ratio is standard in the Epstein-Zin economy and is given by:

$$PD_t = E_t\left[\beta\left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t}\right)^{-\frac{1}{\psi}}\left(\frac{A_{t+1}}{A_t}\right)^{-\frac{1}{\psi}}\left(\frac{\tilde{U}_{t+1}\cdot\left(\frac{A_{t+1}}{A_t}\right)}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot\left(\frac{A_{t+1}}{A_t}\right)\right)}\right)^{\frac{1}{\psi}-\gamma}\left(\frac{D_{t+1}}{D_t}\right)(PD_{t+1}+1)\right] \tag{C.23}$$

Similarly to the solution for the value function, we rewrite all variables in the recursion (C.23) as a function of the state variables and further use quadrature-type numerical methods to evaluate expectations on the right hand side of (C.23). Additionally, we update the long run dividends growth, $g_d(s_{t+1}, s_t, X_t)$, which is in fact random. Consequently, the equilibrium recursion used to solve the model is then:

$$
\begin{aligned}
& PD_t(s_t, X_t,) \\
= \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma} \dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1},s_t,X_t^s,\Delta a_{t+1},X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_t, X_t\right] \\
= \quad & E_t\left[E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma} \dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1},s_t,X_t^s,\Delta a_{t+1},X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right]\middle| s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2}\mathbb{P}(s_{t+1}|s_t,X_t^s) \quad \dots \\
\times \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma} \dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1},s_t,X_t^s,\Delta a_{t+1},X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right] \\
= \quad & \sum_{s_{t+1}=1}^{2}E_t(\pi_{s_{t+1},s_t}|s_t,X_t^s) \quad \dots \\
\times \quad & E_t\left[\begin{array}{c} \beta e^{\left(\lambda-\frac{1}{\psi}\right)(\Delta\tilde{c}_{t+1}+\Delta a_{t+1})}\left(\frac{\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}}{\mathcal{R}_t\left(\tilde{U}_{t+1}\cdot e^{\Delta a_{t+1}}\right)}\right)^{\frac{1}{\psi}-\gamma} \dots \\ \times \quad e^{g_d(s_{t+1},s_t,X_t)+0.5\sigma_d^2}\cdot\left(PD_{t+1}\left(s_{t+1},s_t,X_t^s,\Delta a_{t+1},X_t^{\Delta a}\right)+1\right) \end{array}\middle| s_{t+1}, s_t, X_t\right]
\end{aligned}
$$

Again, the conditional expectation of transition probabilities under the household's filtration permits an analytical formula, while the inner expectation in the expression above can be evaluated using the quadrature-type integration methods.

Limiting Economies - Boundary Values for General Case

The key assumption of the numerical solution is that the household eventually learns the true values of all uncertain parameters in productivity growth. Thus, the simplest limiting economy is the one in which all parameters are known, including both transition probabilities π_{11} and π_{22}, mean growth rates μ_1 and μ_2, volatilities σ_1 and σ_2. In this case, s_t and K_t are the only state variables in the economy. We employ the numerical solution methodology outlined for all known parameters. Specifically, we find the continuation utility and the price-dividend ratio of the equity claim for a set of parameter values

$\pi_{11}, \pi_{22}, \mu_1, \mu_2, \sigma_1$ and σ_2.

C.3.2 Existence of Equilibrium

Similarly to Collin-Dufresne, Johannes, and Lochstoer (2016) and Johannes, Lochstoer, and Mou (2016), the existence of the equilibrium in our production-based economy relies on the fact that the value function is concave and finite for all parameters-known economies. Therefore, we verify that these conditions are satisfied for all limiting boundary economies.

Bibliography

Abel, A. B., and J. C. Eberly. 1994. "A Unified Model of Investment under Uncertainty." *American Economic Review* 84:1369–1384.

———. 1996. "Optimal Investment with Costly Reversibility." *Review of Economic Studies* 63:581–593.

Ai, H., M. Croce, and K. Li. 2013. "Toward a quantitative general equilibrium asset pricing model with intangible capital." *Review of Financial Studies* 26:491–530.

Ai, Hengjie, Mariano Max Croce, Anthony M Diercks, and Kai Li. 2018. "News shocks and the production-based term structure of equity returns." *The Review of Financial Studies* 31 (7): 2423–2467.

Ait-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini. 2019. "The term structure of variance swaps and risk premia." *Journal of Econometrics*, p. Forthcoming.

Andrei, D., B. Carlin, and M. Hasler. 2017. "Asset Pricing with Disagreement and Uncertainty about the Length of Business Cycles." *Management Science*, vol. forthcoming.

Andrei, D., M. Hasler, and A. Jeanneret. 2018. "Asset Pricing with Persistence Risk." *Working Paper.*

Andrei, D., W. Mann, and N. Moyen. 2018. "Learning and the improving relationship between investment and q." *Working Paper.*

Augustin, P., and R. Tedongap. 2016. "Real economic shocks and sovereign credit risk." *Journal of Financial and Quantitative Analysi* 51:541–87.

Babiak, M. 2019. "Generalized Disappointment Aversion and the Variance Term Structure." *CERGE-EI Working Paper.*

Babiak, M., and R. Kozhan. 2018. "Parameter Learning in Production Economies." *Working Paper.*

Backus, D., M. Chernov, and I. Martin. 2011. "Disasters Implied by Equity Index Options." *The Journal of Finance* 66:1969–2012.

Bakshi, G., N. Kapadia, and D. Madan. 2003. "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options." *The Review of Financial Studies* 16(1):101–143.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2014. "Volatility, the macroeconomy and asset prices." *Journal of Finance* 69:2471–2511.

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

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

Basu, S., and B. Bundick. 2017. "Uncertainty Shocks in a Model of Effective Demand." *Econometrica* 85 (3):937–958.

Belo, F., P. Collin-Dufresne, and Goldstein. R. 2015. "Dividend dynamics and the term structure of dividend strips." *Journal of Finance* 70:1115–160.

Belo, F., X. Lin, and S. Bazdresch. 2014. "Labor hiring, investment and stock return predictability in the cross section." *Journal of Political Economy* 122:129–77.

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

Berger, David, Ian Dew-Becker, and Stefano Giglio. 2019. "Contractionary volatility or volatile contractions?" *Review of Economic Studies*, p. Forthcoming.

Berrada, Tony, Jérôme Detemple, and Marcel Rindisbacher. 2018. "Asset pricing with beliefs-dependent risk aversion and learning." *Journal of Financial Economics* 128 (3): 504–534.

Bloom, N. 2009. "The impact of uncertainty shocks." *Econometrica* 77:623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S.J. Terry. 2018. "Really Uncertain Business Cycles." *Forthcoming Econometrica.*

Boldrin, M., L. J. Christiano, and J. D. M. Fisher. 2001. "Habit Persistence, Asset Returns, and the Business Cycle." *American Economic Review* 91:149–66.

Bollerslev, T., M. Gibson, and H. Zhou. 2011. "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-implied and Realized Volatilities." *Journal of Econometrics* 160:102–118.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies* 22:4463–4492.

Bonomo, M., R. Garcia, N. Meddahi, and R. Tedongap. 2011. "Generalized disappointment aversion, long-run volatility risk, and asset prices." *Review of Financial Studies* 24:82–122.

———. 2015. "The long and the short of the risk-return trade-off." *Journal of Econometrics, Elsevier* 187(2):580–592.

Born, B., and J. Pfeifer. 2014. "Policy Risk and the Business Cycle." *Journal of Monetary Economics* 68:68–85.

Brandt, M. W., Q. Zeng, and L. Zhang. 2004. "Equilibrium Stock Return Dynamics Under Alternative Rules of Learning About Hidden States." *Journal of Economic Dynamics and Control* 28:1925–1954.

Britten-Jones, M., and A. Neuberger. 2000. "Option prices, implied price processes, and stochastic volatility." *Journal of Finance* 55:839–866.

Cagetti, M., L. P. Hansen, T. Sargent, and N. Williams. 2002. "Robustness and Pricing with Uncertain Growth." *The Review of Financial Studies* 15 (2): 363–404.

Campanale, C., R. Castro, and G.L. Clementi. 2010. "Asset pricing in a production economy with Chew-Dekel preferences." *Review of Economic Dynamics* 13:379–402.

Campbell, J., and J. Cochrane. 1999. "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy* 107:205–251.

Campbell, J., and R. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies* 1:195–228.

Carr, P., and L. Wu. 2009. "Variance risk premiums." *Review of Financial Studies* 22:1311–1341.

Cecchetti, S. G., P. S. Lam, and N. C. Mark. 1990. "Mean Reversion in Equilibrium Asset Prices." *American Economic Review* 80(3):398–418.

Choi, Hoyong, Philippe Mueller, and Andrea Vedolin. 2017. "Bond Variance Risk Premiums*." *Review of Finance* 21 (3): 987–1022 (01).

Christiano, L., R. Motto, and M. Rostagno. 2014. "Risk Shocks." *The American Economic Review* 104 (1): 27–65.

Christoffersen, P., and K. Jacobs. 2004. "The importance of the loss function in option valuation." *Journal of Financial Economics* 72:291–318.

Cochrane, J. 1991. "Production-based asset pricing and the link between stock returns and economic fluctuations." *Journal of Finance* 46:209–237.

Cogley, T., and T. J. Sargent. 2008. "The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?" *Journal of Monetary Economics* 55(3):454–476.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer. 2016. "Parameter Learning in General Equilibrium: The Asset Pricing Implications." *American Economic Review* 106(3):664–98.

Croce, M. M. 2014. "Long-run productivity risk: A new hope for production-based asset pricing?" *Journal of Monetary Economics* 66:13–31.

Croce, Mariano M, Martin Lettau, and Sydney C Ludvigson. 2014. "Investor information, long-run risk, and the term structure of equity." *The Review of Financial Studies* 28 (3): 706–742.

Dahlquist, Magnus, Adam Farago, and Roméo Tédongap. 2016. "Asymmetries and portfolio choice." *The Review of Financial Studies* 30 (2): 667–702.

Delikouras, S. 2017. "Where's the kink? Disappointment events in consumption growth and equilibrium asset prices." *Review of Financial Studies*, vol. forthcoming.

Delikouras, Stefanos, and Alexandros Kostakis. 2019. "A Single-Factor Consumption-Based Asset Pricing Model." *Journal of Financial and Quantitative Analysis* 54 (2): 789–827.

Dew-Becker, Ian, and Stefano Giglio. 2019. "Cross-sectional uncertainty and the business cycle: evidence from 40 years of options data."

Dew-Becker, Ian, Stefano Giglio, and Bryan T Kelly. 2019. "Hedging macroeconomic and financial uncertainty and volatility." *Available at SSRN.*

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez. 2017. "The Price of Variance Risk." *Journal of Financial Economics* 123 (2): 225–250.

Drechsler, I. 2013. "Uncertainty, Time-Varying Fear, and Asset Prices." *Journal of Finance* 68(5):1837–1883.

Drechsler, I., and A. Yaron. 2011. "What's Vol Got to Do With It." *Review of Financial Studies* 24:1–45.

Du, D. 2011. "General equilibrium pricing of options with habit formation and event risks." *Journal of Financial Economics* 99:400–426.

Epstein, L. G., and S. E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica: Journal of the Econometric Society* 57(4):937–969.

Eraker, B., and I. Shaliastovich. 2008. "An equilibrium guide to designing affine pricing models." *Mathematical Finance* 18:519–543.

Fama, E. F., and K. R. French. 1989. "Business conditions and expected returns on stocks and bonds." *Journal of Financial Economics* 25:23–49.

Farago, Adam, and Roméo Tédongap. 2018. "Downside risks and the cross-section of asset returns." *Journal of Financial Economics* 129 (1): 69–86.

Favilukis, J., and X. Lin. 2016. "Wage Rigidity: A Quantitative Solution to Several Asset Pricing Puzzles." *Review of Financial Studies* 29 (1):148–192.

Favilukis, Jack, and Xiaoji Lin. 2015. "Wage rigidity: A quantitative solution to several asset pricing puzzles." *The Review of Financial Studies* 29 (1): 148–192.

Fernandez-Villaverde, J., P. Guerron-Quintana, J. Rubio-Ramirez, and M. Uribe. 2011. "Risk matters: the real effects of volatility shocks." *American Economic Review* 101:2530–2561.

Gabaix, X. 2012. "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." *The Quarterly Journal of Economics* 127(2):645–700.

Gilchrist, S., J.W. Sim, and E. Zakrajsek. 2014. "Uncertainty, Financial Frictions, and Irreversible Investment." *Working Paper.*

Gillman, M., M. Kejak, and M. Pakos. 2015. "Learning about rare disasters: Implications for consumption and asset prices." *Review of Finance* 19(3):1053–1104.

Gourio, F. 2012. "Disaster Risk and Business Cycles." *American Economic Review* 102 (6):2734–66.

Gul, F. 1991. "A theory of disappointment aversion." *Econometrica* 59:667–686.

Hall, R. E. 2001. "The Stock Market and Capital Accumulation." *American Economic Review* 91:1185–1202.

Hamilton, J. 1989. “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle.” *Econometrica* 57(2):357–384.

Hamilton, J.D. 1990. “Analysis of time series subject to changes in regimes.” *Journal of Econometrics* 45:39–70.

Hansen, L. P. 2007. “Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk.” *American Economic Review* 97:1–30.

Hasler, M., M. Khapko, and R. Márfe. 2019. “Rational learning and term structures.” *Working Paper, University of Toronto and Collegio Carlo Alberto*.

Hasler, Michael, and Roberto Márfe. 2016. “Disaster recovery and the term structure of dividend strips.” *Journal of Financial Economics* 122 (1): 116–134.

Hodrick, R. 1992. “Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement.” *Review of Financial Studies* 5:357–386.

Jahan-Parvar, M., and H. Liu. 2014. “Ambiguity aversion and asset prices in production economies.” *Review of Financial Studies* 27:3060–3097.

Jermann, U. J. 1998. “Asset Pricing in Production Economics.” *Journal of Monetary Economics* 41:257–75.

Johannes, M., L. A. Lochstoer, and Y. Mou. 2016. “Learning about Consumption Dynamics.” *Journal of Finance* 71:551–600.

Johnson, T. 2007. “Optimal Learning and New Technology Bubbles.” *Journal of Monetary Economics* 87:2486–2511.

Ju, N., and J. Miao. 2012. “Ambiguity, learning, and asset returns.” *Econometrica* 80:559–591.

Judd, K. L. 1992. “Projection Methods for Solving Aggregate Growth Models.” *Journal of Economic Theory* 58(2):410–452.

Justiniano, A., and G. Primiceri. 2008. “The time varying volatility of macroeconomic fluctuations.” *American Economic Review* 98:604–641.

Kaltenbrunner, G., and L. Lochstoer. 2010. “Long-run risk through consumption smoothing.” *Review of Financial Studies* 23:3190–224.

Kozhan, R., A. Neuberger, and P. Schneider. 2013. “The Skew Risk Premium in the Equity Index Market.” *The Review of Financial Studies* 26(9):2174–2203.

Kozlowski, J., L. Veldkamp, and V. Venkateswaran. 2018a. “The Tail that Keeps the Riskless Rate Low.” *NBER Macroeconomics Annual*, vol. forthcoming.

———. 2018b. “The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation.” *Working Paper*.

Kroencke, T. A. 2017. “Asset pricing without garbage.” *Journal of Finance* 72 (1):47–98.

Kydland, F. E., and E. C. Prescott. 1982. “Time-to-build and aggregate fluctuations.” *Econometrica* 50:1345–1370.

Leduc, S., and S. Liu. 2016. “Uncertainty shocks are aggregate demand shocks.” *Journal of Monetary Economics* 82:20–35.

Lettau, M., and S. Ludvigson. 2001. "Consumption, aggregate wealth, and expected stock returns." *Journal of Finance* LVI(3):815–849.

Lewellen, J. 2004. "Predicting returns with financial ratios." *Journal of Financial Economics* 74:209–35.

Liu, H., and J. Miao. 2014. "Growth Uncertainty, Generalized Disappointment Aversion and Production-based Asset Pricing." *Journal of Monetary Economics* 69:70–89.

———. 2015. "Growth uncertainty, generalized disappointment aversion and production-based asset pricing." *Journal of Monetary Economics* 69:70–89.

Liu, H., and Y. Zhang. 2018. "Financial Uncertainty with Ambiguity and Learning." *Working Paper.*

Liu, J., J. Pan, and T. Wang. 2005. "An equilibrium model of rare-event premia and its implication for option smirks." *Review of Financial Studies* 18:131–164.

Long, J. B., and C. I. Plosser. 1983. "Real Business Cycles." *Journal of Political Economy* 91:39–69.

Márfe, Roberto. 2017. "Income insurance and the equilibrium term structure of equity." *The Journal of Finance* 72 (5): 2073–2130.

Mehra, R., and E. C. Prescott. 1985. "The equity premium: a puzzle." *Journal of Monetary Economics* 15:145–161.

Nakamura, Emi, Jon Steinsson, Robert Barro, and Jose Ursua. 2013. "Crises and Recoveries in an Empirical Model of Consumption Disasters." *American Economic Journal: Macroeconomics*, no. 5:135–74.

Pakos, M. 2013. "Long-Run Risk and Hidden Growth Persistence." *Journal of Economic Dynamics and Control* 37 (9):1911–1928.

Pastor, L., and P. Veronesi. 2009. "Learning in Financial Markets." *Annual Review of Financial Economics* 1:361–381.

Pohl, Walter, Karl Schmedders, and Ole Wilms. 2018. "Higher Order Effects in Asset Pricing Models with Long-Run Risks." *The Journal of Finance* 73 (3): 1061–1111.

Pontiff, J., and L. D. Schall. 1998. "Book-to-market ratios as predictors of market returns." *Journal of Financial Economics* 49:141–60.

Restoy, F., and M. Rockinger. 1994. "On Stock Market Returns and Returns on Investment." *Journal of Finance* 49:543–56.

Rietz, T. A. 1988. "The Equity Risk Premium: A Solution." *Journal of Monetary Economics* 22(1):117–131.

Routledge, B. R., and S. E. Zin. 2010. "Generalized Disappointment Aversion and Asset Prices." *Journal of Finance* 65(4):1303–1332.

Savov, A. 2011. "Asset Pricing with Garbage." *Journal of Finance* 66:177–201.

Schreindorfer, D. 2018. "Macroeconomic Tail Risks and Asset Prices." *Working Paper.*

Seo, S., and J. Wachter. 2018. "Option Prices in a Model with Stochastic Disaster Risk." *Management Science.*

Shaliastovich, I. 2015. “Learning, confidence, and option prices.” *Journal of Econometrics* 187(1):18–42.

Stock, J., and M. Watson. 1999. “Business cycle fluctuations in us macroeconomic time series.” *J. Taylor and M. Woodford (eds.), Handbook of Macroeconomics, Amsterdam, Elsevier.*

Tallarini, T.D. 2000. “Risk-sensitive real business cycles.” *Journal of Monetary Economics* 45:507–532.

Uhlig, H. 2007. “Explaining asset prices with external habits and wage rigidity in a DSGE model.” *American Economic Review* 97:239–43.

van Binsbergen, J., M. Brandt, and R. Koijen. 2012. “On the timing and pricing of dividends.” *American Economic Review*, no. 102:1596–618.

van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt. 2013. “Equity yields.” *Journal of Financial Economics*, no. 110:503–519.

van Binsbergen, Jules H, and Ralph SJ Koijen. 2017. “The term structure of returns: Facts and theory.” *Journal of Financial Economics* 124 (1): 1–21.

Veronesi, P. 1999. “Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model.” *Review of Financial Studies* 12(5):975–1007.

Wachter, J. 2013. “Can time-varying risk of rare disasters explain aggre- gate stock market volatility?” *Journal of Finance* 68 (3:987–1035.

Weitzman, M. 2007. “Subjective expectations and asset-return puzzles.” *American Economic Review* 97:1102–1130.

Zhang, L. 2005. “The Value Premium.” *Journal of Finance* 60 (1):67–103.

Zhou, G., and Y. Zhu. 2014. “Macroeconomic volatilities and long-run risks of asset prices.” *Management Science* 61(2):413–430.