Essays on Asset Pricing in Production Economies

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Andrew Yeh-Chi Chen, M.A.

Graduate Program in Business Administration

The Ohio State University
2014

Dissertation Committee
Lu Zhang, Advisor
Xiaoji Lin
René Stulz
Julia Thomas
Abstract

This dissertation examines the modeling of asset prices in production economies. Chapter 1 presents a model which endogenizes a key mechanism of many theories of aggregate asset prices. In order to generate time-varying risk premia, many theories assume time-varying volatility. Chapter 1 shows that this channel can be endogenized with precautionary saving motives. Precautionary motives prescribe that, in bad times, next period’s consumption should be very sensitive to economic news. High sensitivity in bad times results in time-varying consumption volatility, even in the presence of homoskedastic shocks. This channel is made visible by modeling production, and is amplified with external habit preferences. An estimated model featuring this channel quantitatively accounts for excess return and dividend predictability regressions. It also matches the first two moments of excess equity returns, the risk-free rate, and the second moments of consumption, output, and investment.

Chapter 2 shows that the model of Chapter 1 not only addresses aggregate asset prices, but can also be extended to address key facts about the cross section of stock returns. This result is important because a solution to the equity premium puzzle should be informative about risk in general. I add idiosyncratic productivity to the model from Chapter 1. I find that the model’s expected returns are log-linear in book-to-market equity, consistent with the data. Moreover, the slope of the relationship is similar. In both the model and the data, a 20% higher book-to-market implies a 100 b.p. increase in expected returns. The result is robust. It requires neither operating leverage nor asymmetric adjustment costs. Rather, value firms are low productivity firms, and mean reversion causes them to have high cash flow growth. This prediction is inconsistent with conventional wisdom, but consistent with recent empirical evidence. I present additional empirical evidence showing that value firms have high cash flow growth according to a number of definitions of cash flow. High cash flow growth means that value firms’ cash flows are distributed toward the future, and, as a result, their prices are more exposed to discount rate shocks that drive the external habit model. The value premium is compensation for this high exposure.
Abstract cont.

Chapter 3 examines general restrictions on production technologies implied by asset prices. It shows that representative firm models which are consistent with asset price data require either large capital adjustment costs, or volatile investment-specific technology shocks. These restrictions hold regardless of preferences, beliefs, operating leverage, or the completeness of asset markets. The restrictions summarize the sense in which asset prices are anomalous with respect to the theory of optimal investment.
Acknowledgements

I am indebted to Aubhik Khan, Xiaoji Lin, René Stulz, Julia Thomas, and especially Lu Zhang for their training, advice, and inspiration.
Vita

2004 ..........  B.A. Economics, University of Maryland
2004 ..........  B.S. Physics, with High Honors, University of Maryland
2009 ..........  M.B.A., Georgetown University
2010 ..........  M.A. Economics, The Ohio State University

Fields of Study

Major Field: Business Administration
Area of Specialization: Finance
# Contents

**Front Matter**

Abstract .............................................. ii
Acknowledgements .................................. iv
Vita ...................................................... v
List of Figures ....................................... viii
List of Tables ......................................... xiii

1 **Precautionary Volatility and Asset Prices** 1

1.1 Introduction ..................................... 1
1.2 A Two-period Model ............................. 7
1.3 The General Equilibrium Production Model .......... 14
1.4 Simulated Method of Moments Estimation .......... 22
1.5 Matching Asset Price and Business Cycle Moments .... 27
1.6 Inspecting the Mechanism ....................... 33
1.7 Conclusion ....................................... 59
1.8 Proofs, Data, and Solution Method ............ 60

2 **Habit, Production, and the Cross Section of Stock Returns** 82

2.1 Introduction ..................................... 82
2.2 Empirical Evidence on Value and Cash Flow Growth .... 86
2.3 A General Equilibrium Model with Heterogeneous Firms .... 90
2.4 Calibration to Post-War U.S. Data ................ 98
2.5 Quantitative Results ............................ 99
2.6 Inspecting the Mechanism ..................... 102
## List of Figures

1.1 **Precautionary Savings and Time-Varying Volatility.** ..... 3
1.2 **Two-period Model: Date 1 Consumption and MPC.** $C_1(W_1)$ is date 1 consumption as a function of date 1 wealth. $C'_1(W_1)$ is its derivative (the MPC). The different lines show consumption for various levels of habit $H_1$. ..... 12
1.3 **Two-period Model: Date 0 Consumption Volatility.** $\sigma_0[C_1(W_1)]$ is the volatility of date 1 consumption, given that date 0 wealth is $W_0$. The different lines show consumption volatility for various levels of habit $H_1$. ..... 12
1.4 **Time-Varying Consumption Volatility and Asset Prices.** The figures show scatterplots from model simulations. Consumption volatility and the equity premium are calculated using the model’s laws of motion. All values are annualized. ..... 34
1.5 **Consumption Volatility and the Campbell-Cochrane $\lambda(s_t)$.** Dots represent consumption volatility from model simulations. The dashed line represents Campbell-Cochrane $\lambda(s_t)$ implied consumption volatility (equation (1.24)). Figures are annualized. ..... 38
1.6 **Decomposition of the Risk-Free Rate.** ‘Intertemporal Substitution’ represents the 2nd and 3rd terms of equation (28). ‘Precautionary Saving’ represents the last term. Computed from model’s laws of motion. Risk-free rate is in annualized %. Capital in all panels is fixed at the mean capital stock. ..... 48
1.7 **Predictability of Consumption Growth: Effect of the Conditioning Set.** Figures are annualized and computed from the model's laws of motion. Price-dividend ratio is for a claim on the firm's dividends. The left panel shows scatterplots from model simulations. The middle and right panels show consumption growth as a function of state variables. ................................................................. 54

1.8 **Decomposition of the Risk-Free Rate: low persistence of habit.** \( \rho_s = 0.80 \). All other parameter values are from the estimation (Table 1). 'Intertemporal Substitution' represents the 2nd and 3rd terms of equation (1.28). 'Precautionary Saving' represents the last term. Computed from model's laws of motion. Risk-free rate is in annualized %. Capital in all panels is fixed at the mean capital stock. . . . 56

1.9 **Distribution of State Variables.** Panel A shows simple histograms from simulated data. Panel B shows 2D histograms as contour plots. 78

1.10 **Law of Motion for Consumption.** ........................................ 79

1.11 **Law of Motion for Consumption Growth Volatility.** Consumption growth volatility is the volatility of log first-differenced consumption. 80

1.12 **Law of Motion for the Hansen Jaganathan Bound.** ............... 81

2.1 **Expected Returns as a Function of B/M.** Figures are annual . . . . . 110

2.2 **B/M and Expected Returns as a Function of Firm States.** Figures are annual ................................................................. 110

2.3 **Cash Flow Growth of Book-to-Market Sorted Portfolios.** Figures are annual ................................................................. 111

2.4 **Cash Flow Growth of Book-to-Market Sorted Portfolios: No Habit.** Figures are annual ................................................................. 111

2.5 **The Cyclicality of Value and Growth Cash Flows.** Value firm plots are calculated using the median capital and productivity of firms in the 10th decile of B/M-sorted portfolios. Growth firm plots are from the 1st decile. Net investment is investment net of depreciation. Aggregate capital is fixed at its mean. ................................. 112
2.6 Partial Equilibrium Experiment: Cash Flow Growth of Book-to-Market Sorted Portfolios. ‘GE’ uses calibrated parameter values from Table 2.3 and equilibrium laws of motion. ‘PE’ uses equilibrium aggregate laws of motion, but firm-level decision rules consistent with adjustment costs which are 1/20th of the value from Table 2.3. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113

3.1 Admissible Region The shaded region shows admissible regions for the volatility of the investment-specific technology shock $\sigma_\mu$ and adjustment cost $\phi$ restricted by the mean and volatility of stock returns. The adjustment costs are constant elasticity with $1/\phi = $ the elasticity of the investment rate with respect to marginal Q. x’s and literature references represent parameter values used in the respective papers. Models in green are able to fit the equity premium. KL ’10 represents the benchmark permanent shock model from Kaltenbrunner and Lochstoer (2010). Gourio ’13 represents the unlevered equity return from his model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 145

3.2 Expected Investment Return: No Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows mean investment returns as a function of capital share $\alpha$ and adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = $ the elasticity of the investment rate with respect to marginal Q. The shaded area shows the admissible region implied by expected stock returns. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 146

3.3 Volatility of Investment Return: No Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows the standard deviation of investment returns as a function of capital share $\alpha$ and adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = $ the elasticity of the investment rate with respect to marginal Q. The shaded area shows the admissible region implied by stock return volatility. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 147
3.4 **Admissible Region: No Investment-Specific Technology Shocks.**
The shaded region shows admissible regions for the capital share $\alpha$ and adjustment cost $\phi$ restricted by the mean and volatility of stock returns. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal } Q$. $x$’s and literature references represent parameter values used in the respective papers. Models in green are able to fit the equity premium. KL ’10 represents the benchmark permanent shock model from Kaltenbrunner and Lochstoer (2010). Gourio ’13 represents the unlevered equity return from his model.  

3.5 **Expected Investment Return: Investment-Specific Technology Shocks.**
Figures are annualized. This contour plot shows mean investment returns as a function of the volatility of the investment-specific technology shock $\sigma_\mu$ and the adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal } Q$. The shaded area shows the admissible region implied by expected stock returns.  

3.6 **Volatility of the Investment Return: Investment-Specific Technology Shocks.**
Figures are annualized. This contour plot shows the volatility of investment returns as a function of the volatility of the investment-specific technology shock $\sigma_\mu$ and the adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal } Q$. The shaded area shows the admissible region implied by the volatility of stock returns.  

3.7 **Additional Restrictions on the Elasticity of Intertemporal Substitution.**
The contour plot shows the relative volatility of consumption growth as a function of the EIS and adjustment cost. Results are from a standard rbc model with Epstein-Zin preferences (Section 3.4.3). The shaded region represents admissible parameter values implied by the measured volatility of consumption growth.
3.8 **Investment Return and Stock Return Volatility Under Decreasing Returns.** All figures are annualized. The plot shows stock return and investment return volatility under various returns to scale and adjustment cost parameters. .......................... 151

3.9 **Mean Investment Return and Stock Return Under Decreasing Returns.** All figures are annualized. The plot shows stock return and investment return volatility under various returns to scale and adjustment cost parameters. .......................... 152
List of Tables

1.1 Parameter Estimates and Moment Targets ........................................ 26
1.2 Unconditional Asset Price Moments ............................................. 28
1.3 Predicting Dividends and Excess Returns with the Price-Dividend Ratio .......................................................... 30
1.4 Basic Business Cycle Moments ................................................... 32
1.5 GARCH Estimates of Time-Varying Consumption Volatility .......... 41
1.6 Regressions of Consumption Volatility on the Price-Dividend Ratio 43
1.7 Alternative Dividend Processes: Basic Statistics ......................... 50
1.8 Alternative Dividend Processes: Predictive Regressions ............... 52
1.9 Comparative Statics ................................................................. 55

2.1 Mean Cash Flow of Value and Growth ........................................ 114
2.2 Growth Rates of Mean Cash Flow for Value and Growth ............ 115
2.3 Calibration ................................................................................. 116
2.4 Unconditional Aggregate Asset Price Moments ......................... 117
2.5 Business Cycle Moments ............................................................ 117
2.6 Predicting Dividend Growth and Excess Returns with the Price- Dividend Ratio .......................................................... 118
2.7 Regressions of Future Returns on Book-to-Market ...................... 119
2.8 Summary Statistics from 10 Book-to-Market Sorted Portfolios .... 119
2.9 Partial Equilibrium Experiment: Regressions of Future Returns on Book-to-Market .......................................................... 120

3.1 Summary Statistics of Stock Returns and Components of the In- vestment Return ................................................................. 153
Chapter 1

Precautionary Volatility and Asset Prices

1.1 Introduction

Time-varying uncertainty plays a central role in theories of aggregate asset prices. This feature is frequently required in order to capture the nature of stock market volatility: stock market fluctuations have little relationship with movements in expected cash flows but are closely linked with movements in expected excess returns (LeRoy and Porter (1981), Shiller (1981), Cochrane (2011)). Time-varying uncertainty addresses these facts by generating fluctuations in risk which move asset prices, while leaving average cash flows unchanged. This mechanism exists in a broad array of asset pricing models, though the details of the modeling differ. In long-run risk models, time-varying uncertainty comes from the assumption of time-varying consumption volatility (Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2009)). In ambiguity aversion, it is the assumption of a time-varying magnitude of ambiguity (Sbuelz and Trojani (2008), Bianchi, Ilut,
and Schneider (2012)). In disasters, it is the assumed time-varying severity or probability of disaster (Gabaix (2012), Wachter (2012)). In external habit models, it is the assumption of time-varying habit sensitivity (Campbell and Cochrane (1999)). These modeling approaches provide coherent and tractable descriptions of stock market volatility, but by assuming time-varying uncertainty, they leave unanswered the deeper question: why does uncertainty vary over time?¹

This paper provides an answer to the question. I show that time-varying uncertainty is the natural result of a fundamental economic phenomenon: the desire for precautionary savings. The mechanism lies in the dynamics of this motive. Precautionary motives intensify in bad times, making investors hunker down and cut consumption. But in the following month, they may need to hunker down again. In fact, how much they will need to hunker down next month depends sharply on economic news. News that the bad times will get even worse would put them in austerity, leading to a sharp drop in consumption. News that the economy will recover relieves them of their concerns, and they can relax and sharply increase consumption. While they wait for this news, next month's consumption is very uncertain, risk is high, and asset prices are low. This volatility is absent in good times, when investors have a lot of wealth. High wealth shields investors from precautionary concerns and allows them to continue with their plans, regardless of economic news.

I formalize this “precautionary volatility” story in two ways. The first is a simple consumption-savings model which allows for proofs and propositions. The model features a precautionary savings motive, that is, the agent desires savings as protection against uncertainty. This motive leads to a concave consump-

¹Early versions of these models often do not feature time-varying volatility (e.g. Bansal and Yaron (2004) model I, Rietz (1988), Barro (2006), Constantinides (1990)), though later versions add this feature in order to address time-varying risk premia.
tion function (Carroll and Kimball (1996)), as depicted in Figure 1.1. Concave consumption means that in bad times, consumption is very sensitive to shocks (dashed lines), while in good times, there is little action (dotted lines). Countercyclical sensitivity then leads to time-varying consumption volatility, even in the presence of homoskedastic shocks. I also examine an amplification mechanism: external habit preferences à la Campbell and Cochrane (1999). Investors with external habit judge their consumption relative to an aggregate reference level. This perspective makes them feel poor, amplifying precautionary motives and the concavity of the consumption function (dash-dot line). These results are presented in a specific context for clarity, though the mechanism applies in a wide variety of settings. The model features a shock to wealth, though the underlying cause of the shock may be a shift in technology, a shift in monetary policy, or even an uncertainty shock in the sense of Bloom (2009). And though the model’s precautionary motive comes from prudent preferences (Kimball (1990)), financial constraints would produce similar dynamics (Carroll (2001)). Additionally, analogous effects exist for firm policies under financial constraints (i.e. Almeida, Campello, and Weisbach (2004)).

**Figure 1.1: Precautionary Savings and Time-Varying Volatility.**
The bulk of the paper is devoted to the second formalization: an estimated general equilibrium model. The general equilibrium model is a real business cycle model with capital adjustment costs and external habit preferences. Fluctuations originate from a standard AR1 productivity shock. Despite the lack of a heteroskedastic driving process, the model produces time-varying consumption volatility and captures the nature of stock market volatility: the price-dividend ratio has little forecasting power for cash flow growth but is a good predictor of excess returns. This result is due to two key features: real investment and external habit. Real investment means that precautionary savings motives are reflected in aggregates, and thus the economy exhibits the precautionary volatility mechanism described above. External habit plays two roles. Habit amplifies precautionary volatility, and also gives investors strong concerns about risk which allow the model to address aggregate asset prices.

The model also provides a unified description of aggregate asset prices and business cycles. The model captures a long list of asset price facts: it matches the mean, volatility, and persistence of excess returns, the mean, volatility, and persistence of the risk-free rate, the volatility and persistence of the price-dividend ratio, as well as the previously mentioned excess return and dividend predictability regressions using the price-dividend ratio. And these data-like asset price dynamics originate from a data-like business cycle: the model matches the volatilities of output, consumption, and investment, as well as their autocorrelations and cross-correlations.

The preferences build off of the external habit preferences of Campbell and Cochrane (1999) but feature a critical deviation. While Campbell and Cochrane assume a time-varying “habit sensitivity,” I make this preference parameter a constant. This changes the key economics. Constant sensitivity means that as-
set prices in my model are driven by time-varying consumption volatility, rather than a time-varying sensitivity of preferences. Indeed, the model’s time-varying consumption volatility quantitatively mimics the Campbell-Cochrane habit sensitivity function. Thus, the model can be considered a way of endogenizing the Campbell-Cochrane mechanism. This description conflicts with the common wisdom that time-varying risk aversion, not time-varying volatility of habit, drives the external habit model. While the common wisdom provides an intuitive description of asset prices, it is not an entirely accurate description of the Campbell-Cochrane model. Indeed, Campbell and Cochrane (1999) explicitly state that time-varying sensitivity is required for time-varying risk premia (p. 211).

I provide empirical support for the precautionary volatility mechanism. GARCH and GJR-GARCH estimates of U.S. data show only mild evidence of time-varying consumption volatility, and the model is consistent with this fact. The mechanism predicts that low asset prices are linked to high consumption volatility, and regressions of consumption volatility on the price-dividend ratio show that this prediction is borne out by the data, consistent with previous work (Kandel and Stambaugh (1990), Lettau, Ludvigson, and Wachter (2008)). Additionally, regressions on simulated data show that the model does a good job of describing the magnitude of this relationship. The model is able to generate significant time-varying risk premia out of mild time-variation in consumption volatility for two reasons. The first is that habit preferences make investors very sensitive to changes in consumption volatility. As a result, a small change in consumption volatility leads to a large change in risk premia. The second is that consumption volatility is persistent in both the model and data. This persistence makes it difficult to detect time-varying volatility in only 50 years of quarterly data.

This paper belongs to the literature which studies the macroeconomic ef-
fects of uncertainty fluctuations. Uncertainty has been linked to a broad set of aggregate economic phenomena, including business cycles (Bloom (2009), Bloom et al. (2012), Gourio (2012)), aggregate stock prices (Kandel and Stambaugh (1990), Bansal and Yaron (2004), Lettau, Ludvigson, and Wachter (2008), Gabaix (2012)), and exchange rates (Lustig et al. (2011), Farhi and Gabaix (2013)). The bulk of these papers take uncertainty fluctuations as given and then investigate its consequences. This paper differs by proposing a theory of the origins of uncertainty fluctuations. A handful of papers do provide an origin theory. Uncertainty may rise in bad times because low production makes it hard to learn about economic fundamentals (Van Nieuwerburgh and Veldkamp (2006), Fajgelbaum, Schaal, and Taschereau-Dumouchel (2012)), because bad times are a good time for experimentation (Pastor and Veronesi (2011), Bachmann and Moscarini (2011)), or because bad shocks lead to doubts about the true model of the economy (Orlik and Veldkamp (2013)). A distinct channel arising from disaggregated volatility is proposed in Carvalho and Gabaix (2013) and Koren and Tenreyro (2013) who show that changes in the ability to diversify lead to changes in volatility. The precautionary volatility channel complements these theories. Indeed, since the vast majority of models contain precautionary savings motives (Carroll and Kimball (1996)), this channel is active in many of the above papers.

The quantitative model belongs to the literature on asset pricing models with habit formation. This literature shows that that habit formation can account for a wide variety of asset price facts, including aggregate stock markets (Campbell and Cochrane (1999)), term structure (Wachter (2006), Bekaert, Engstrom, and Grenadier (2010)), foreign exchange (Verdelhan (2010)), and option markets (Bekaert and Engstrom (2010)). The aforementioned papers are all endowment economies and so cannot address output and investment. In contrast, this paper
shows that habit formation can be successfully extended into a production economy and account for asset prices as well as quantity fluctuations. Previous habit-production papers consider a short-lived, internal habit (Jermann (1998) and Boldrin, Christiano, and Fisher (2001)). In contrast, I consider a slow-moving, external habit, in the style of Campbell and Cochrane (1999). Without this modification, the model would produce a counterfactually volatile risk-free rate. Lettau and Uhlig (2000) do consider a slow-moving, external habit in a production economy, but they assume costless adjustment of the capital stock. In contrast, I model convex capital adjustment costs. Without such costs, RBC models imply a counterfactually smooth Tobin’s Q (Boldrin, Christiano, and Fisher (1999)).

The paper proceeds as follows. Section 2 examines a simple consumption-savings model which illustrates the mechanism. Sections 3-6 estimates the general equilibrium model and demonstrates that the mechanism can quantitatively account for a long list of asset price and business cycle facts.

1.2 A Two-period Model

This section illustrates the precautionary volatility mechanism with a two-period model. The model favors simplicity over generality. It is a consumption-savings problem for a prudent investor with external habit who faces uncertainty in the return on savings. I prove two propositions and two corollaries. Proposition 1 and Corollary 1 show that consumption volatility is countercyclical. Proposition 2 and Corollary 2 show that habit amplifies this countercyclicality. Proofs are found in the Appendix. The consumption-savings problem can be formalized as a general equilibrium model with a linear technology (à la Constantinides (1990)), but the notation is simpler as a consumption-savings problem.
Consider a two-period-lived investor with an external habit stock \( H > 0 \), period utility

\[
u(C - H) = \frac{(C - H)^{1-\gamma} - 1}{(1-\gamma)}
\]

and \( \gamma > 0 \). These preferences display precautionary savings motives because they are “prudent” in the sense of Kimball (1990). An alternative approach of modeling precautionary savings motives through financial constraints would produce similar effects. Habit is there to illustrate an amplification mechanism. For simplicity the investor has no time preference.

At date 0, the investor has wealth \( W_0 \). Nothing occurs at date 0, but the date helps to serve as a reference point. Throughout this section, I describe a variable as ‘countercyclical’ if it is negatively related to \( W_0 \). At date 1, the investor has a habit level of \( H_1 > 0 \), and he receives a wealth shock \( \Delta W_1 \), making his wealth \( W_1 = W_0 + \Delta W_1 \). He consumes \( C_1 \), and saves the rest of his wealth \( W_1 - C_1 \) in a risky asset. At date 2, the investor has a habit level of \( H_2 > 0 \). The return on the risky asset \( R_2 \) is realized, and he consumes his remaining wealth \( R_2(W_1 - C_1) \).

The model comes down to finding the investor’s optimal consumption at date 1. His optimal consumption policy solves

\[
C_1(W_1) = \arg\max_{C_1} u(C_1 - H_1) + \mathbb{E}_1[u(C_2 - H_2)]
\]

\[s.t. \quad C_2 = R_2(W_1 - C_1)\]  

Taking a Taylor of expansion of \( C_1(W_1) \) around \( W_0 \), the volatility of \( C_1 \) is ap-
proximately

\[ \sigma_0[C_1(W_1)] \approx \sigma_0[C_1(W_0) + C_1'(W_0)\Delta W_1] \]

\[ = |C_1'(W_0)|\sigma_0(\Delta W_1) \] (1.2)

That is, consumption volatility is proportional to the marginal propensity to consume (MPC). Intuitively, the MPC captures the responsiveness of consumption to shocks. So the higher the MPC, the more responsive, and thus the more volatile, is consumption. For the remainder of this section, I assume that \( \Delta W_1 \) is small enough so that equation (1.2) is a good approximation.

Unfortunately, even in a two-period consumption savings problem, the presence of precautionary savings motives typically precludes closed-form solutions (Carroll (2001)). However, using the methods from Carroll and Kimball (1996), I can still prove that the solution exhibits countercyclical volatility and that habit amplifies this cyclicality.\(^2\)

**Proposition 1.** The date 1 MPC is decreasing in wealth, that is, \( C_1''(W_1) < 0 \).

The essence of the proof is that, as long as the date 2 habit is positive, the model falls into the broad class of models for which consumption is strictly concave (Carroll and Kimball (1996)). Carroll and Kimball do not show the strict concavity result for this model, but I show in Appendix 1.8.1 that their proofs can be extended to include this setting. The proof is rather technical, but, intuitively, the investor saves for precautionary reasons in the presence of uncertainty. As the agent becomes wealthier, the uncertainty becomes less relevant, and this motive declines, creating convex savings and concave consumption.

\(^2\)I am grateful to Pok-Sang Lam for teaching me his version of the Carroll and Kimball (1996) proof.
Corollary 1. The volatility of $C_1$ is countercyclical, that is, $\frac{\partial}{\partial W_0} \sigma_0[C_1(W_1)] < 0$.

The proof is short and illustrates the role of the MPC, so I state it here.

Proof. Since the MPC is positive and decreasing in wealth, the absolute value of the MPC is decreasing in wealth. Since the volatility of $C_1$ is proportional to the absolute value of the MPC (equation (1.2)), the volatility of $C_1$ is also decreasing in wealth.  

Proposition 1 and its corollary capture “precautionary volatility,” that is, the mechanism by which precautionary motives generate countercyclical volatility. Countercyclical consumption volatility stems from the effect of the date 1 wealth shock on the investor’s desire for precautionary savings at date 1. A positive shock weakens this desire, decreasing savings and boosting consumption, while a negative shock encourages him to hunker down, with opposite effects. From the perspective of date 0, this uncertainty in the need for precautionary savings at date 1 causes additional consumption volatility. Since the need for precautionary savings intensifies at low levels of wealth, the effect on consumption volatility is countercyclical.

Mathematically, the precautionary volatility mechanism is manifested as a strictly concave consumption function. Strictly concave consumption is an implication of a broad class of models (Carroll and Kimball (1996)). Indeed, Posch (2011) shows that a standard real business cycle model generates time-varying risk premia. However, the magnitude in a standard model is too small to account for the data. The following proposition and corollary show that habit amplifies precautionary volatility.

Proposition 2. The date 1 habit increases the sensitivity of the MPC to wealth, that
is,
\[
\frac{\partial}{\partial H_1} C''_1(W_1) < 0
\]

To understand this result, note that an investor with high habit has become accustomed to a high standard of living. He judges consumption not by its absolute level, but by how much it exceeds this standard. As a result, he has a much stronger precautionary savings motive than an investor who is accustomed to living in poverty. The mechanism driving Proposition 1 becomes stronger, and the MPC becomes even more sensitive to wealth. Mathematically, an external and additive habit acts as a reduction in income, which reduces wealth and amplifies precautionary savings effects.

**Corollary 2.** *As the date 1 habit increases, the volatility of* \(C_1\) *becomes more countercyclical, that is, for any set of initial wealth* \(W_0\),

\[
\frac{\partial}{\partial H_1} \left\{ \max_{W_0 \in W_0} \sigma_0[C_1(W_1)] - \min_{W_0 \in W_0} \sigma_0[C_1(W_1)] \right\} > 0
\]

One can think of the set of initial wealth \(W_0\) as the range of the business cycle states in the economy. Corollary 2 states that, across this business cycle, the range of consumption volatilities is increasing in habit. The proof simply applies the link between the MPC and consumption volatility (equation (1.2)) to Proposition 2. Because habit makes the MPC more countercyclical, habit also makes consumption volatility more countercyclical.

Figures 1.2 and 1.3 illustrate the propositions of this section. The figures show results from numerical solutions of the two-period model for different levels of habit. In all solutions, I set \(H_1 = H_2\). This specification focuses on how habit “concavifies” the consumption function by eliminating intertemporal substitu-
tion effects which complicate the picture.

**Figure 1.2: Two-period Model: Date 1 Consumption and MPC.** $C_1(W_1)$ is date 1 consumption as a function of date 1 wealth. $C'_1(W_1)$ is its derivative (the MPC). The different lines show consumption for various levels of habit $H_1$.

![Figure 1.2: Two-period Model: Date 1 Consumption and MPC.](image)

**Figure 1.3: Two-period Model: Date 0 Consumption Volatility.** $\sigma_0[C_1(W_1)]$ is the volatility of date 1 consumption, given that date 0 wealth is $W_0$. The different lines show consumption volatility for various levels of habit $H_1$.

![Figure 1.3: Two-period Model: Date 0 Consumption Volatility.](image)

Figure 1.2 plots date 1 consumption and MPC as a function of wealth. The left panel shows that, as long as habit is positive, consumption is strictly con-
The right panel shows that this concavity is reflected in an MPC which decreases in wealth (Proposition 1). Figure 1.2 also shows that habit intensifies this relationship (Proposition 2). As the line gets lighter, habit increases, and the slope of the MPC gets steeper. Intuitively, habit makes the investor feel poorer and intensifies his precautionary savings motives.

Figure 1.3 shows how these consumption policies are reflected in consumption volatility. As long as habit is positive, consumption volatility decreases in wealth (Corollary 1). This is the essence of precautionary volatility. Because precautionary savings motives prescribe a countercyclical MPC, and because consumption volatility is proportional to the MPC (equation (1.2)), consumption volatility is countercyclical. The figure also shows that habit amplifies this countercyclicality (Corollary 2). As the line gets lighter, habit increases, and the range that is spanned by consumption volatility increases.

The results of this section may come as a surprise since many models in the finance literature produce consumption policies which are linear in wealth. These linear consumption policies are the result of the fact that much of this literature is set in continuous time (Sundaresan (1989), Constantinides (1990)) or relies on log-linear approximations (Campbell (1994), Lettau (2003), Kaltenbrunner and Lochstoer (2010)). Continuous time allows the investor to make an infinite number of trades within any trading period. This instantaneous trading allows for consumption policies which, if applied in a discrete time setting, would imply a positive probability that marginal utility becomes infinite (Brandt (2009)). It is exactly fear of hitting this condition of infinite marginal utility which seems to drive concave consumption behavior (Attanasio (1999)). Log-linear approximations, on the other hand, implicitly assume that consumption policies are (log) linear in state variables. The non-linearities introduced by combining power util-
ity functions with linear budget constraints are important for generating non-linear consumption (Posch (2011)).

The results may also appear to conflict with the common intuition that habit encourages smooth consumption. This intuition is straightforward in internal habit models, where an increase in consumption has a direct effect of lowering utility in later periods (Sundaresan (1989)). In an external habit model, however, the investor by assumption does not take into account this indirect effect. Prices may encourage smooth consumption via general equilibrium effects, but the habit itself acts very similarly to a reduction in income which increases the marginal propensity to consume (Carroll and Kimball (1996)).

1.3 The General Equilibrium Production Model

The previous section illustrates the intuition behind the precautionary volatility mechanism, but a quantitative evaluation requires a richer model. This section presents a richer model. The remainder of the paper examines the predictions of the quantitative model.

The model sticks as close to the standard real business cycle model as possible. The only features are external habit preferences and quadratic capital adjustment costs. External habit serves as an amplification mechanism. Capital adjustment costs are required for a volatile Tobin's Q. There is a representative household and representative firm. Time is discrete, the horizon is infinite, and markets are complete.

For the remainder of the paper, I denote log variables with lowercase, i.e. \( z_t \equiv \log Z_t \).
1.3.1 Representative Household

There is a continuum of identical households with external habit preferences. Each household chooses its asset holdings to maximize

\[
E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{(C_t - H_t)^{1-\gamma} - 1}{1 - \gamma} \right\}
\]

Where \( H_t \), the level of habit, is taken as external to the household. Habit is important for both amplifying the precautionary volatility mechanism and for providing strong concerns about risk which have the chance at being consistent with asset price facts. For simplicity, the household does not value leisure and is endowed with a unit of labor.

I specify the evolution of habit using surplus consumption, rather than the level of habit itself. That is, define surplus consumption as

\[
\hat{S}_t = \frac{\hat{C}_t - H_t}{\hat{C}_t}
\]

where the hats denote aggregates. This approach leads to a simple and symmetric stochastic discount factor and allows for some clean asset pricing analysis. It also eases comparison with the existing literature on external habit (Campbell and Cochrane (1999), Wachter (2006), among others).

Surplus consumption evolves according to an autoregressive process

\[
\hat{s}_{t+1} = (1 - \rho_s) \hat{s} + \rho_s \hat{s}_t + \lambda (\hat{c}_{t+1} - \hat{c}_t)
\]

where \( \lambda \) is a constant. Endowment economy external habit models specify \( \lambda \) as a decreasing function of surplus consumption. This assumption builds in a countercyclical volatility of marginal utility which is essential for addressing excess
market volatility. The model does not require this assumed countercyclicality because, as we will see, production and precautionary motives endogenously generate countercyclical consumption volatility. For comparability with the literature, I fix $\lambda$ at the Campbell and Cochrane (1999) steady state value

$$\lambda = \frac{1}{\bar{S}} - 1 \quad (1.6)$$

This modification causes the issue that habit may move negatively with consumption. However, that habit is still a geometric average of the history of consumption (Appendix 1.8.4).

The non-linear habit specification is not critical to the economic mechanism, but the persistent AR1 specification is. Persistent habit is important for capturing the persistence in aggregate asset prices such as the price-dividend ratio. This strong persistence contrasts with previous habit models in production economies (Jermann (1998) and Boldrin, Christiano, and Fisher (2001)) which assume that habit depends only on last quarter’s consumption.

The external nature of habit and the surplus consumption specification produce a simple stochastic discount factor

$$M_{t,t+1} = \beta \left( \frac{\hat{C}_{t+1}}{\hat{C}_t} \frac{\hat{S}_{t+1}}{\hat{S}_t} \right)^{-\gamma} \quad (1.7)$$

The stochastic discount factor has the traditional consumption growth term $\left( \frac{\hat{C}_{t+1}}{\hat{C}_t} \right)^{-\gamma}$, but habit adds a surplus consumption term $\left( \frac{\hat{S}_{t+1}}{\hat{S}_t} \right)^{-\gamma}$. This SDF is the same as that in Campbell and Cochrane (1999) but is very different from the Epstein-Zin-habit model of Dew-Becker (2011). Dew-Becker’s SDF also has two components, but the two components are those generated by Epstein and Zin (1989) preferences: a consumption growth component and a term related to the return on wealth. In
Dew-Becker’s model, habit enters the SDF by affecting the curvature (power parameter) of the return on wealth component. As a result, the mechanisms driving Dew-Becker’s model are very different from those in this paper.

1.3.2 Representative firm

The production side of the economy is standard. The only feature is quadratic capital adjustment costs. There is a unit measure of identical firms which produce consumption using capital $K_t$ and labor $N_t$. Production is given by

$$\Pi(K_t, Z_t, N_t) \equiv Z_t K_t^\alpha N_t^{1-\alpha}$$  \hspace{1cm} (1.8)

where $\alpha$ is capital’s share of output and $Z_t$ is productivity.

The Cobb-Douglas specification of production, combined with the fact that the household does not value leisure implies that wages are equal to the marginal product of labor

$$W_t = (1 - \alpha) Z_t \hat{K}_t^\alpha$$  \hspace{1cm} (1.9)

Productivity $Z_t$ follows the standard AR(1) process

$$z_{t+1} = (1 - \rho) \tilde{z} + \rho_z z_t + \sigma_z \epsilon_{z,t+1}$$  \hspace{1cm} (1.10)

Where $\epsilon_{z,t+1}$ is a standard normal i.i.d. shock. This is the only source of uncertainty in the model, and by assumption it is homoskedastic. $\tilde{z}$ is chosen so that the non-stochastic steady state capital stock is approximately one. The choice of $\tilde{z}$ does not have a material effect on the results, but keeping the steady state capital stock near one helps with the computation.
Capital accumulates according to the usual capital accumulation rule

$$K_{t+1} = I_t + (1 - \delta)K_t$$  \hspace{1cm} (1.11)

and firms face a convex capital adjustment cost

$$\Phi(I_t, K_t) = \frac{\phi}{2} \left( \frac{I_t}{K_t} - \delta \right)^2 K_t$$  \hspace{1cm} (1.12)

This formulation punishes the firm for deviating from the non-stochastic steady state investment rate of $\delta$. I assume that the adjustment costs are a pure loss. They do not represent payments to labor. Adjustment costs are included because production economies produce a counterfactually smooth Tobin's Q unless one includes an investment friction.

The firm's objective is standard

$$\max_{\{I_t, K_{t+1}, N_t\}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} M_{0,t} \left[ \Pi(K_t, Z_t, N_t) - W_t N_t - \Phi(I_t, K_t) - I_t \right] \right\}$$  \hspace{1cm} (1.13)

It chooses investment, capital, and labor to maximize future dividends, discounted with the household's stochastic discount factor.

### 1.3.3 Recursive Competitive Equilibrium

Market clearing is standard:

$$\hat{C}_t + \hat{I}_t = Z_t \hat{K}_t^\alpha - \Phi(\hat{I}_t, \hat{K}_t)$$  \hspace{1cm} (1.14)

Due to the consumption externality, the welfare theorems do not hold and the equilibrium cannot be easily described by a social planner's problem. Thus,
I define equilibrium competitively. The aggregate state variables are aggregate capital \( \hat{K} \), surplus consumption \( \hat{S} \), and productivity \( Z \).

**Definition.** Equilibrium is a firm decision rule for investment \( I(K; \hat{K}, \hat{S}, Z) \), cum-dividend value function \( V(K; \hat{K}, \hat{S}, Z) \), law of motion for aggregate consumption \( \hat{C}(\hat{K}, \hat{S}, Z) \), and a law of motion of aggregate capital \( \Gamma(\hat{K}, \hat{S}, Z) \) such that

(i) Firms optimize: \( I(K; \hat{K}, \hat{S}, Z) \) and \( V(K; \hat{K}, \hat{S}, Z) \) solve

\[
V(K; \hat{K}, \hat{S}, Z) = \max_{\{N,I,K'\}} \left\{ \Pi(K, Z, N) - W(\hat{K}, \hat{S}, Z)N - \Phi(I, K) - I \\
+ \int_{-\infty}^{\infty} dF(e')M(\hat{K}, \hat{S}, Z; Z')V(K'; \hat{K}', \hat{S}', Z') \right\}
\] (1.15)

subject to capital accumulation (1.11), competitive wages (1.9), the productivity process (1.10), the habit process (1.5), the SDF is equal to the household's IMRS (1.7), \( \hat{K}' = \Gamma(\hat{K}, \hat{S}, Z) \), and where \( F(e') \) is the standard normal CDF.

(ii) Markets clear and aggregates are consistent with individual behavior:

\[
\hat{C}(\hat{K}', \hat{S}', Z') = \Pi(\hat{K}, Z, 1) - \Phi(I(\hat{K}; \hat{K}, \hat{S}, Z), \hat{K}) - I(\hat{K}; \hat{K}, \hat{S}, Z)
\] (1.16)

\[
\Gamma(\hat{K}, \hat{S}, Z) = (1 - \delta)\hat{K} + I(\hat{K}; \hat{K}, \hat{S}, Z)
\] (1.17)

Since households and firms are identical, in equilibrium, \( \hat{K} = K, \hat{S} = S \) and \( \hat{C} = C \). Thus, in what follows, I drop the hats.

**1.3.4 Solution Method**

Capturing the economics of this model requires a global and non-linear solution method. This is important for both external habit (Campbell and Cochrane
(1999)) and precautionary savings (Carroll (1997)). The result is a challenging numerical problem which I discuss in this section.

**Projection Method**

I solve the model using a projection method (Judd (1992)). Specifically, I represent the law of motion for capital using cubic splines and then use Broyden's method (a quasi-Newton algorithm) to find cubic spline coefficients which satisfy the firm's Euler equation. The solution program makes extensive use of the Miranda and Fackler (2001) CompEcon toolbox.

The projection method is important for two reasons. The first is that, due to the consumption externality, the welfare theorems do not hold and the model cannot be easily solved by value function iteration. Projection avoids this problem by working with equilibrium conditions from the recursive competitive equilibrium. The second is that projection produces a global and non-linear solution. These properties are important for asset pricing models in general (Cochrane (2008b)), but are particularly important for capturing the mechanisms in this model. The precautionary savings channel which drives countercyclical risk premia are related to the third derivative of the utility function and are not captured by traditional linearizations. Moreover, precautionary effects are particularly pronounced when the investor is threatened with infinite marginal utility (Attanasio (1999)), and fully capturing this effect requires a global solution.

**Homotopy Method**

Projection methods require a good initial guess of the spline coefficients. There is no guarantee that a general non-linear equation solver will converge and the high dimensionality of the problem tends to make the solvers unstable. The stan-
The standard approach is to use the real business cycle model as an initial guess. Unfortunately, with external habit preferences, the real business cycle model is a poor initial guess.

To overcome this issue, I use a homotopy method. Specifically, I modify the firm’s problem so that it discounts future profits using the SDF

\[ M' = \beta \left( \frac{C'}{C} \left( \frac{S'}{S} \right)^\chi \right)^Y \]

Note that \( \chi = 0 \) corresponds to a model with no habit, and \( \chi = 1 \) corresponds to the full model. I begin by solving the model for \( \chi = 0 \), and then slowly increase \( \chi \), using the coefficients from the previous \( \chi \) as the initial guess for the current \( \chi \).

This homotopy algorithm is very computationally intensive. To aid in the speed of computation, I discretize the productivity process using the Rouwenhorst method.

**Solving for the Evolution of Surplus Consumption**

Another difficult issue which arises with external habit preferences is that they result in a state variable which is not predetermined and is endogenous. Surplus consumption tomorrow is not known today and it depends on consumption tomorrow, which is endogenous. This makes solving for the surplus consumption process rather difficult.

To see this clearly, it helps to write the evolution of surplus consumption (1.5) as functions of state variables

\[ \log S' = (1 - \rho_s) \bar{s} + \rho_s \log S + \lambda [c(K', S', Z_j) - c(K, S, Z_j)] \]  

(1.18)

Where \( Z_j \) is tomorrow’s discrete productivity state. Since \( K' \) is predetermined,
is a function of four variables $s'(K, S, Z_i, Z_j)$. Note that surplus consumption tomorrow $S'$ appears on both sides of this equation. Thus determining surplus consumption tomorrow requires solving this non-linear equation. I once again use Broyden's method to solve this equation. This calculation is very computationally intensive since it must be done at every collocation node for every potential productivity shock within every iteration of the big Broyden's method which is solving for the coefficients of the law of motion of capital.

For further details about the solution method, see Appendix 1.8.5.

### 1.4 Simulated Method of Moments Estimation

In order to capture economic magnitudes, I need to give the parameters numerical values. Most papers in this literature use an informal calibration procedure. I choose a more rigorous estimation method. I estimate the model using simulated method of moments (SMM), the simulated version of Hansen (1982)'s generalized method of moments (GMM). SMM differs from traditional GMM in that it uses simulation to compute model moments rather than closed-form expressions, which may be unavailable for moments of interest. Duffie and Singleton (1993) derive adjustments to GMM formulas and additional regularity conditions required by the use of simulation.

#### 1.4.1 Data

The estimation uses post-war (1948-2011) data from the CRSP, BEA, and BLS. Some authors argue for using the longest sample available when evaluating consumption-based asset pricing models (e.g. Bansal, Kiku, and Yaron (2009)). However, in a production economy I must also address the data on aggregate investment and
output. The nature of this data is significantly affected by using pre-war data. For example, the correlation between investment growth and output growth have a high correlation of 0.73 in the post-war sample, but have a mild correlation of 0.23 for the sample 1929-2011. This consideration leads me to target only post-war data.

All variables are real and per capita. Consumption is real per capita non-durable goods and services consumption. This measure excludes volatile consumer durables such as automobiles, which produce a smooth consumption flow over the life of the durable. Aggregate equity is represented by the CRSP value-weighted index, adjusted for inflation with the consumer price index. Dividends are calculated with the assumption that all dividends are reinvested in the stock market. This method of aggregation preserves the Campbell and Shiller (1988) present value identity. The risk-free rate is a forecast of the inflation-adjusted 90-day T-bill return using the previous year’s inflation rate and the nominal 90-day yield (following Beeler and Campbell (2009)). No adjustments are made for financial leverage. Further details regarding the data are found in the Appendix 1.8.2.

1.4.2 Estimation Method and Predefined Parameters

Because SMM is computationally intensive, I set the more traditional parameters outside of the estimation. I set the depreciation rate $\delta = 0.02$ to match the mean, growth-adjusted, investment rate. I set the capital share parameter $\alpha = 0.35$ to match the capital share of output implied by constant returns to scale. The persistence of productivity shocks $\rho_z = 0.979$ matches the persistence of the Solow residual with a fixed labor input. In an external habit model, the utility curvature $\gamma$ and the steady state surplus consumption ratio $\bar{S}$ jointly control risk.
aversion. As a result, it is difficult to identify these parameters separately. For ease of comparison with the literature, $\gamma = 2$ is set at Campbell and Cochrane (1999)’s value.\(^3\)

The remaining five parameters are estimated by SMM. Explicitly, let $\theta$ represent the five parameters as a vector and $\hat{M}^*$ represent a vector of target data moments. SMM estimates the parameters by minimizing the distance between data moments and simulated moments:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left[ \hat{M}^* - M(\theta) \right]^T W \left[ \hat{M}^* - M(\theta) \right]$$

(1.19)

where $M(\theta)$ is the model-simulated counterpart to $\hat{M}^*$ and $W$ is a weighting matrix chosen by the econometrician. I use one-stage, exactly-identified GMM, that is, I target five data moments which are economically informative about the five parameter values. This approach has the advantage of transparency. Seeing that the empirical targets are equal to the model moments immediately verifies that the minimization algorithm is successful. Under exact identification, a consistent estimator for the asymptotic variance of the estimated parameter values is

$$\hat{\text{Var}}[\sqrt{T}(\hat{\theta} - \theta_0)] = \left(1 + \frac{1}{S}\right)[DM(\hat{\theta})']^{-1}\hat{\text{Var}}(\sqrt{T}\hat{M}^*)[DM(\hat{\theta})]^{-1}$$

(1.20)

where $S$ is the number of simulations used to calculate model moments, $DM(\theta)$ is the derivative of the simulated moments with respect to $\theta$ and $\hat{\text{Var}}(\sqrt{T}\hat{M}^*)$ is a consistent estimate of the asymptotic variance of the estimated data moments. Intuitively, we have a good estimate if the moments are informative about the parameter values ($DM(\hat{\theta})$ is large), or if we have a precise estimate of the moments

---

\(^3\)Campbell and Cochrane (1999) choose $\gamma$ to fit the Sharpe ratio and $\bar{S}$ to produce a constant risk-free rate, but that cannot be done here because I set $\lambda$ as a constant in order eliminate exogenous time-varying volatility.
(\text{Var}(\sqrt{T \hat{M}^*}) \text{ is small}).^4

I optimize using Levenberg-Marquardt, a variant of Newton’s method. I choose this method rather than the more commonly-used simulated annealing algorithm for two reasons. The first is that the moment function does not display an extreme number of local minima, which is where simulated annealing has an advantage. With a relatively smooth objective function, a method which uses derivative information is much more efficient. Another advantage of LM is robustness. Simulated annealing tends to be sensitive to the choice of the optimization parameters (Press, Teukolsky, Vetterling, and Flannery (1992)). Additional details regarding the estimation method can be found in Appendix 1.8.3.

1.4.3 Parameter Estimates and Moment Targets

Table 1.1 summarizes the estimation. It shows estimated parameter values, standard errors, and targeted moments. For convenience, the predefined parameters are also shown, with standard errors omitted. The model is quarterly and all parameter values are quarterly. To eliminate seasonality in dividends, both the U.S. and simulated data are aggregated to the annual level. All moments are annual.

Preference parameters are identified with asset prices. Because time-preference \( \beta \) is reflected in the risk-free rate, I choose the mean 90-Day T-bill return as a target moment. The resulting \( \beta = 0.970 \) is rather low because the model features a non-trivial precautionary savings motive, and a low level of patience helps counteract that motive. Steady state surplus consumption \( \bar{S} \) controls the magnitude

---

^4The moments I use are not moments in the strict sense (i.e. the Sharpe ratio), and I estimate transformations of the parameters in order to avoid corner solutions. This leads to slightly more complicated formulas than those presented here. These details can be found in the Appendix 1.8.3.
Table 1.1: Parameter Estimates and Moment Targets

The model is quarterly, and all parameter values are quarterly. Empirical figures are annual. Standard errors are Newey-West with 10 lags and shown only for estimated parameters. The sample period is 1948-2011. Consumption is non-durable goods and services consumption. GDP and consumption are logged and HP-filtered with a smoothing parameter of 6.25. Further details are found in Section 1.8.2 and 1.8.3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Empirical Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β  Time Preference</td>
<td>0.971</td>
<td>(0.009)</td>
<td>Mean 90-Day T-bill Return (%) 0.98</td>
</tr>
<tr>
<td>ρs Persistence of Surplus</td>
<td>0.963</td>
<td>(0.006)</td>
<td>Volatility of Excess Return (%) 16.07</td>
</tr>
<tr>
<td>Consumption</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ś  Steady State Surplus</td>
<td>0.063</td>
<td>(0.010)</td>
<td>Mean Sharpe Ratio of CRSP Index 0.48</td>
</tr>
<tr>
<td>Consumption</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ  Utility Curvature</td>
<td>2</td>
<td></td>
<td>(Chosen to Match Campbell-Cochrane (1999))</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>φ  Adjustment Cost Parameter (Quadratic)</td>
<td>75.00</td>
<td>(10.65)</td>
<td>Relative Volatility of Consumption to GDP 0.47</td>
</tr>
<tr>
<td>σz Volatility of Productivity</td>
<td>0.014</td>
<td>(0.001)</td>
<td>Volatility of GDP 0.015</td>
</tr>
<tr>
<td>ρz Persistence of Productivity</td>
<td>0.979</td>
<td></td>
<td>Persistence of Solow Residual 0.979</td>
</tr>
<tr>
<td>α  Capital Share</td>
<td>0.35</td>
<td></td>
<td>Capital’s Share of GDP 0.35</td>
</tr>
<tr>
<td>δ  Depreciation Rate</td>
<td>0.02</td>
<td></td>
<td>Mean Investment Rate (Growth-Adjusted) 0.02</td>
</tr>
</tbody>
</table>

of habit, and, in effect, the degree of risk aversion in the model. I thus choose the mean Sharpe ratio of the CRSP index as a target moment. The resulting value Ś = 0.063 is close to the values used in the external habit literature (Campbell and Cochrane (1999), Wachter (2006), Santos and Veronesi (2010)). The persistence of surplus consumption ρs has a strong effect on the volatility of the market return, and so I choose the volatility of the excess return on the CRSP index as a target. The resulting value of ρs = 0.963 indicates a very persistent habit process, which is also consistent with values used in the external habit literature.

Technological parameters are identified with moments of the real economy. The volatility of productivity σz = 0.014 targets the volatility of HP-filtered log GDP of 0.014. Since the data is annual, I use the annual smoothing parameter
of 6.25 advocated by Ravn and Uhlig (2002). The quadratic adjustment cost parameter $\phi = 75.00$ targets the relative volatility of consumption to GDP (also HP-filtered). This estimated value results in mean adjustment costs as a percentage of output of less that 1%.

1.5 Matching Asset Price and Business Cycle

Moments

This section contains the main quantitative results. It shows that precautionary volatility channel, when amplified by external habit, provides a quantitative description of time-varying risk premia. The model also matches numerous other moments, and provides a unified description of asset prices and business cycles. I make no adjustments to account for un-modeled leverage or payout policy. The price of equity is simply the present value of dividends from the representative firm, discounted with the household’s SDF.

For quick navigation, this section does not discuss mechanisms. Readers focused on mechanisms may want to skip to Section 1.6.

1.5.1 Unconditional Asset Price Moments

Table 1.2 shows that the model produces a nice fit for all of the basic moments of asset prices. As intended by the estimation, it hits three of these moments exactly: the mean risk-free rate, the Sharpe Ratio, and the volatility of excess returns. Producing this large equity volatility is a difficult task in production economies (Gourio (2010), Kaltenbrunner and Lochstoer (2010), Croce (2010)).

The model also captures many asset market features beyond those used in
Table 1.2: Unconditional Asset Price Moments

Figures are annual. No adjustments are made to account for financial leverage. The model columns show means and percentiles across simulations of the same length as the empirical sample. \( r, p, \) and \( d \) are the logs of returns, prices, and dividends from the CRSP value-weighted index. Capital letters show levels rather than logs. \( r_f \) is a forecast of the ex-post real return on 90-day T-bills following Beeler and Campbell (2009). \( E, \sigma, \) and AC1 represent the sample mean, standard deviation and first-order autocorrelation. Further details are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th></th>
<th>Data 1948-2011</th>
<th>Model mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Identifying Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(r_f) ) (%)</td>
<td>0.98</td>
<td>0.98</td>
<td>-2.72</td>
<td>0.28</td>
<td>6.83</td>
</tr>
<tr>
<td>( E(R - R_f)/\sigma(R) )</td>
<td>0.48</td>
<td>0.48</td>
<td>0.29</td>
<td>0.48</td>
<td>0.68</td>
</tr>
<tr>
<td>( \sigma(r - r_f) ) (%)</td>
<td>16.07</td>
<td>16.07</td>
<td>10.77</td>
<td>15.71</td>
<td>22.53</td>
</tr>
<tr>
<td><strong>Untargeted Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(r - r_f) ) (%)</td>
<td>6.47</td>
<td>6.77</td>
<td>5.03</td>
<td>6.96</td>
<td>7.88</td>
</tr>
<tr>
<td>AC1( r - r_f )</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.28</td>
<td>-0.07</td>
<td>0.16</td>
</tr>
<tr>
<td>( \sigma(r_f) ) (%)</td>
<td>2.24</td>
<td>2.96</td>
<td>0.55</td>
<td>1.88</td>
<td>8.60</td>
</tr>
<tr>
<td>AC1( r_f )</td>
<td>0.56</td>
<td>0.88</td>
<td>0.69</td>
<td>0.91</td>
<td>0.99</td>
</tr>
<tr>
<td>( E(p - d) )</td>
<td>3.42</td>
<td>2.67</td>
<td>2.09</td>
<td>2.66</td>
<td>3.27</td>
</tr>
<tr>
<td>( \sigma(p - d) )</td>
<td>0.43</td>
<td>0.46</td>
<td>0.26</td>
<td>0.44</td>
<td>0.76</td>
</tr>
<tr>
<td>AC1( p - d )</td>
<td>0.95</td>
<td>0.90</td>
<td>0.79</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>( E(r_{10yr} - r_f) ) (%)</td>
<td>2.03</td>
<td>0.93</td>
<td>2.10</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>( \sigma(r_{10yr} - r_f) ) (%)</td>
<td>1.42</td>
<td>0.62</td>
<td>1.37</td>
<td>2.44</td>
<td></td>
</tr>
<tr>
<td>AC1( r_{10yr} - r_f )</td>
<td>0.85</td>
<td>0.70</td>
<td>0.86</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

the identification. The model produces a risk-free rate volatility of 2.96% which is close to the data estimate of 2.24%. This moment is difficult to match in habit models, which tend to produce an excessively volatile risk-free rate (i.e. 11% in Jermann (1998) and 25% in Boldrin, Christiano, and Fisher (1999)). This low risk-free rate volatility is reflected in a reasonable term premium. The model produces a mean excess return on 10-year bonds of 2%, indicating that the model's
equity premium is distinct from the term premium, as in the data. The volatility of the log price-dividend ratio is 0.46, which is close to the data value of 0.43. This moment has been difficult to capture even in endowment economies (Bansal, Gallant, and Tauchen (2007), Bansal, Kiku, and Yaron (2009)).

Additionally, the model generates data-like persistence. As in the data, the excess market return is mildly negatively autocorrelated, and the risk-free rate and price-dividend ratio are highly positively autocorrelated. The model somewhat overstates the persistence of the risk-free rate, but this deviation contains some uncertainty because risk-free rate is not directly observable.

1.5.2 Excess Return and Dividend Predictability

The previous section shows that the model generates a volatile, data-like stock market. This large volatility does not mean that the model captures the nature of stock market fluctuations, however. The data show that these fluctuations are driven by time-varying risk premia: stock price movements have little relationship with movements in expected cash flows, but are closely linked to movements in expected excess returns (LeRoy and Porter (1981), Shiller (1981), Campbell and Shiller (1988), Cochrane (1992)). In this section, I consider dividends as cash flows, but other notions of cash flow (i.e. consumption or profits) produce similar results.

Table 1.3 shows that the model captures key elements of time-varying risk premia. Panel A shows regressions of future dividend growth on the log price-dividend ratio at various horizons. The model predicts no relationship between the price-dividend ratio and future dividend growth at the one-year horizon, as is seen in the data. As in the data, this predictability increases with the horizon, but
Table 1.3: Predicting Dividends and Excess Returns with the Price-Dividend Ratio

Figures are annual. The model columns show means and percentiles across simulations of the same length as the empirical sample. \( r_t, p_t, \) and \( d_t \) are the log-returns, prices, and dividends from the CRSP value-weighted index. \( r_{f,t} \) is a forecast of the ex-post real return on 90-day T-bills following Beeler and Campbell (2009). Standard errors are Newey-West with \( 2(L-1) \) lags. Further details are found in Appendix 1.8.2.

### Panel A: Predicting dividend growth
\[ \sum^{L}_{j=1} \Delta d_{t+j} = \alpha + \beta (p_t - d_t) + \epsilon_{t+L} \]

<table>
<thead>
<tr>
<th>L</th>
<th>1948-2011 Data</th>
<th>Model Mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.03</td>
<td>-0.00</td>
<td>-0.04</td>
<td>-0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>( \hat{\beta} )</td>
<td>3</td>
<td>0.01</td>
<td>0.07</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.03</td>
<td>0.13</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>SE(( \hat{\beta} ))</td>
<td>3</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.09</td>
<td>0.07</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.00</td>
<td>0.10</td>
<td>0.01</td>
<td>0.09</td>
</tr>
</tbody>
</table>

### Panel B: Predicting excess returns
\[ \sum^{L}_{j=1} r_{t+j} - r_{f,t+j} = \alpha + \beta (p_t - d_t) + \epsilon_{t+L} \]

<table>
<thead>
<tr>
<th>L</th>
<th>1948-2010 Data</th>
<th>Model Mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.23</td>
<td>-0.11</td>
<td>-0.04</td>
</tr>
<tr>
<td>( \hat{\beta} )</td>
<td>3</td>
<td>-0.27</td>
<td>-0.30</td>
<td>-0.55</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-0.40</td>
<td>-0.44</td>
<td>-0.78</td>
<td>-0.44</td>
</tr>
<tr>
<td>SE(( \hat{\beta} ))</td>
<td>3</td>
<td>0.05</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.12</td>
<td>0.10</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>3</td>
<td>0.09</td>
<td>0.10</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.26</td>
<td>0.37</td>
<td>0.11</td>
<td>0.38</td>
</tr>
</tbody>
</table>

The coefficients remain small and statistically insignificant. The predictability is somewhat overstated at the 5-year horizon, but Section 1.6.6 shows that this issue can be fixed with a richer dividend process which accounts for deviations.
between neoclassical firm cash flows and dividends.

If the price-dividend ratio does not forecast dividends, then it must forecast returns (Campbell and Shiller (1988), Cochrane (2008a)). Panel B shows the ability of the model to capture this flip side of dividend predictability. It shows regressions of future excess returns on the log price-dividend ratio. Regression coefficients, standard errors, and \( R^2 \)'s, are close to the data for all forecasting horizons. The coefficients on the price-dividend ratio are negative, and they are both economically and statistically significant. At the one-year horizon, the model exactly matches the data coefficient of -0.12. To understand this economically, recall that the volatility of the log price-dividend ratio is roughly 0.40 in both the model and the data. This means that a one standard deviation rise in the price-dividend ratio predicts a huge 4% reduction in the equity premium over the next year. This forecasting power increases with the forecasting horizon, reaching \( R^2 \)'s of roughly 30% at the 5-year horizon in both model and data.

Overall, the model captures both sides of predictability. Asset price fluctuations have little relationship with fluctuations in future dividends, but are tightly linked to fluctuations in future excess returns.

1.5.3 Unconditional Business Cycle Moments

The previous sections show that the model produces a good description of numerous asset price moments. Table 1.4 shows that these asset price moments come with data-like fluctuations in the real economy. The model hits the volatility of output and relative volatility of consumption to output, as intended by the estimation. Investment is more volatile than output, as in the data. Like the data, the model moments display strong co-movement between output, consumption, and investment. Output, consumption, and investment are highly
persistent, and are nearly as persistent as the data. First-differenced log consumption has a low volatility. In both the model and data this volatility is about 1% per year. Lastly, the table also shows that average adjustment costs are small at less than 1% of output.

Table 1.4: Basic Business Cycle Moments

Figures are annual. The model columns show means and percentiles across simulations of the same length as the empirical sample. $y$ is log GDP, $c$ is log nondurable and services consumption, and $i$ is the log of fixed investment plus durable goods plus government investment. The subscript $hp$ indicates that the moment is calculated from HP-filtered data with a smoothing parameter of 6.25. $\Delta$ indicates first-differences. Further details are found in the Appendix 1.8.2.

<table>
<thead>
<tr>
<th>Identifying Moments</th>
<th>Data 1948-2011</th>
<th>Model mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma(y_{hp})$ (%)</td>
<td>1.50</td>
<td>1.50</td>
<td>1.11</td>
<td>1.49</td>
<td>1.93</td>
</tr>
<tr>
<td>$\sigma(c_{hp})/\sigma(y_{hp})$</td>
<td>0.47</td>
<td>0.47</td>
<td>0.40</td>
<td>0.47</td>
<td>0.55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Untargeted Moments</th>
<th>Data 1948-2011</th>
<th>Model mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho(c_{hp}, y_{hp})$</td>
<td>0.83</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho(i_{hp}, y_{hp})$</td>
<td>0.85</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>AC1($y_{hp}$)</td>
<td>0.12</td>
<td>0.21</td>
<td>0.01</td>
<td>0.21</td>
<td>0.38</td>
</tr>
<tr>
<td>AC1($c_{hp}$)</td>
<td>0.32</td>
<td>0.21</td>
<td>0.01</td>
<td>0.22</td>
<td>0.39</td>
</tr>
<tr>
<td>AC1($i_{hp}$)</td>
<td>0.27</td>
<td>0.20</td>
<td>0.01</td>
<td>0.21</td>
<td>0.38</td>
</tr>
<tr>
<td>$\sigma(\Delta c)$ (%)</td>
<td>1.11</td>
<td>1.41</td>
<td>0.97</td>
<td>1.38</td>
<td>1.90</td>
</tr>
<tr>
<td>$\mathbb{E}(\text{Adj Cost}/Y)$ (%)</td>
<td>0.54</td>
<td>0.21</td>
<td>0.49</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>$\mathbb{E}(\text{Adj Cost}/I)$ (%)</td>
<td>3.21</td>
<td>1.13</td>
<td>2.70</td>
<td>6.88</td>
<td></td>
</tr>
</tbody>
</table>
1.6 Inspecting the Mechanism

I have shown that the model provides a unified description of numerous asset price and business cycle moments, including time-varying risk premia. The remainder of the paper is devoted to explaining how.

Most of the new economics can be seen by examining time-varying risk premia. The model contains the key features of the simple model of Section 1.2, suggesting that precautionary volatility is the underlying mechanism. However, the model is closely related to Campbell and Cochrane (1999), and so a closer look is informative. Section 1.6.1 shows that time-varying risk premia come from time-varying consumption volatility. Section 1.6.2 compares this mechanism with Campbell and Cochrane (1999). Section 1.6.3 shows that the model’s time-varying consumption volatility is consistent with the data.

The other asset pricing results are examined later in this section. The large and volatile equity premium is analyzed in Section 1.6.4. The low and smooth risk-free rate is examined in Section 1.6.5. Section 1.6.6 investigates cash flow dynamics and alternative dividend processes. Section 1.6.7 ends the inspection by examining comparative statics.

1.6.1 Time-varying Risk Premia and Precautionary Volatility

The model addresses time-varying risk premia through the precautionary volatility mechanism described in Section 1.2. To demonstrate this, I’ll work backward, starting with time-varying risk premia and ending with precautionary motives.

The model produces fluctuations in risk by producing time-varying consumption volatility. This can be seen in Figure 1.4, which shows scatterplots of as-
set prices against consumption volatility produced by model simulations. Consumption volatility varies over time, and both the price-dividend ratio and equity premium are nicely linked to consumption volatility. The price-dividend ratio declines in consumption volatility, while the equity premium increases.

**Figure 1.4: Time-Varying Consumption Volatility and Asset Prices.** The figures show scatterplots from model simulations. Consumption volatility and the equity premium are calculated using the model’s laws of motion. All values are annualized.

![Figure 1.4](image)

This result has a very simple intuition. Times of high consumption volatility are risky times with large risk premia and low asset prices. But a little formalism provides additional insight and builds confidence in the mechanism. A log-normal approximation of the SDF shows that the conditional maximum Sharpe ratio is

$$\max_{\text{all assets}} \left\{ \frac{\mathbb{E}_t(R_{t+1} - R_{f,t+1})}{\sigma_t(R_{t+1})} \right\} \approx \gamma(\lambda + 1)\sigma_t(\Delta c_{t+1})$$

(1.21)

---

5The log-SDF is $m_{t+1} = \log \beta - \gamma \Delta s_{t+1} - \gamma \Delta c_{t+1}$ and the habit process is $\Delta s_{t+1} = -(1 - \rho_s)(s_t - \bar{s}) + \lambda \Delta c_{t+1}$. Plug the habit process into the log-SDF, then assume that the SDF is log-normal, and we have $q_t(M_{t+1}) = \sqrt{\text{var}(m_{t+1}) - 1} = \sigma_t(m_{t+1}) = \gamma(\lambda + 1)\sigma_t(\Delta c_{t+1})$. 

34
The conditional maximum Sharpe ratio is the conditional volatility of consumption growth, multiplied by preference parameters: $\gamma$ is the utility curvature (1.3), and $\lambda$ is the conditional volatility of the habit process (1.5). This expression shows that consumption volatility is critical. Since consumption volatility is the sole term on the RHS which can vary over time, it must be the driver of time-varying Sharpe ratios.

Time-varying consumption volatility in turn is the result of the precautionary volatility channel described in the simple model (Section 1.2). To see this, note that the model is driven by a single, homoskedastic productivity shock (equation (1.10)). Moreover, the conditional volatility of the habit process is assumed to be constant (equation (1.5)). Thus, there are no exogenous drivers of time-varying volatility. On the other hand, the model features the same preferences as the simple model. These preferences produce precautionary savings motives which lead to strictly concave consumption. This shape means that consumption has time-varying volatility, even though the underlying shock is homoskedastic. External habit amplifies this channel and produces the quantitative result seen in Figure 1.4. Unlike previous external habit models, this model features real investment, which allows the precautionary volatility channel to be visible in equilibrium quantities.

1.6.2 Endogenizing the Campbell-Cochrane Mechanism

Though the model features external habit preferences, the mechanism is distinct from the existing literature. Existing external habit models drive time-varying risk premia through a time-varying preference parameter, i.e. the Campbell and Cochrane (1999) $\lambda(s_t)$ function. My model drives time-varying risk premia through time-varying consumption volatility. In this section, I show that these two chan-
nels are economically distinct, but have similar quantitative properties. Thus, my model can be considered a way of endogenizing the Campbell and Cochrane (1999) mechanism.

To compare the channels, it helps to examine the maximum Sharpe ratio from Campbell and Cochrane (1999)

\[
\max_{\text{all assets}} \left\{ \frac{E_t(R_{t+1} - R_{t+1}^f)}{\sigma_t(R_{t+1})} \right\} \approx \gamma [\lambda(s_t) + 1] \sigma(\Delta c_{t+1})
\]

(1.22)

where

\[
\lambda(s_t) = \begin{cases} 
\frac{1}{3} \sqrt{1 - 2(s_t - \bar{s}) - 1} & \text{if } s_t \leq \bar{s} + \frac{1}{2}(1 - \bar{S}^2) \\
0 & \text{if } s_t \geq \bar{s} + \frac{1}{2}(1 - \bar{S}^2)
\end{cases}
\]

(1.23)

The volatility of consumption growth \( \sigma(\Delta c_{t+1}) \) is assumed to be constant. So the only way the model can generate time-varying Sharpe ratios is through the preference parameter \( \lambda(s_t) \). This is a direct contrast to my model, where time-varying Sharpe ratios must come from consumption volatility (equation (1.21)). \( \lambda(s_t) \) is the “habit sensitivity function,” that is, it controls how sensitive habit is to shocks. Existing models assume that this sensitivity declines in surplus consumption (equation (1.23)). This assumption builds in a time-varying volatility of marginal utility and thus a time-varying maximum Sharpe ratio.

The above analysis also shows that the common intuition regarding the external habit mechanism is misplaced. The common intuition is that time-varying risk premia are driven by time-varying risk aversion. But time-varying risk aversion is not sufficient for generating time-varying Sharpe ratios. To see this, note that a constant \( \lambda(s_t) \) still results in time-varying surplus consumption \( S_t \) (see equation (1.5)). Since risk aversion is approximately \( \gamma S_t^{-1} \), a constant \( \lambda(s_t) \) still
produces time-varying risk aversion. But equation (1.22) shows that, even though risk aversion is time-varying, a constant $\lambda(s_t)$ implies a constant maximum Sharpe ratio. It is time-varying habit sensitivity $\lambda(s_t)$, not time-varying risk aversion, which drives the traditional external habit model. The critical role of the sensitivity function is acknowledged in Campbell and Cochrane (1999). This is not to say that the common intuition cannot be modeled. Dew-Becker (2011) uses Epstein-Zin preferences to create a habit model which captures the common intuition.

The fact that, in traditional external habit models, time-varying habit sensitivity drives time-varying risk leaves open many questions. How do we interpret a time-varying habit sensitivity? Where does this time-varying sensitivity come from? This model can help address these questions since it replaces the time-varying habit sensitivity function with time-varying consumption volatility, which has a very simple interpretation. And time-varying consumption volatility originates from the precautionary savings motives, that is, bad times make investors unsure of how much they can consume next period.

Indeed, the model's time-varying consumption volatility does a good job of mimicking the habit sensitivity function of Campbell and Cochrane. To demonstrate this I construct an ‘implied consumption volatility’ from the Campbell-Cochrane sensitivity function by equating the two expressions for the maximum Sharpe ratio (equations (1.21) and (1.22)). I call this implied consumption volatility because one can replace the Campbell-Cochrane sensitivity function with this implied consumption volatility and get similar quantitative results. Explicitly,

$$\text{Implied Consumption Volatility} = \left( \frac{\lambda(s_t) + 1}{\lambda + 1} \right) \sigma(\Delta c_{t+1}) \quad (1.24)$$

Implied consumption volatility is simply the habit sensitivity $\lambda(s_t)$, rescaled by
its steady state value $\lambda$, and multiplied by the unconditional volatility $\sigma(\Delta c_{t+1})$.

Figure 1.5 compares the Campbell-Cochrane implied consumption volatility to the consumption volatility of my model. It shows scatterplots of the volatilities against surplus consumption. Both are countercyclical, that is, they decline in surplus consumption. Moreover, for much of the plot, implied consumption volatility runs right through the middle of the cloud of dots representing consumption volatility. The two channels are both qualitatively and quantitatively similar. Thus, this production economy could be considered a method for endogenizing the external habit mechanism.

Figure 1.5: Consumption Volatility and the Campbell-Cochrane $\lambda(s_t)$. Dots represent consumption volatility from model simulations. The dashed line represents Campbell-Cochrane $\lambda(s_t)$ implied consumption volatility (equation (1.24)). Figures are annualized.
1.6.3 Empirical Support for the Precautionary Volatility Mechanism

I have shown that the precautionary volatility channel generates time-varying consumption volatility, and that this time-varying consumption volatility produces time-varying risk premia. This section provides direct empirical support for this mechanism. I show that the mechanism generates mild time-variation in consumption volatility, as is seen in the data. I also show that, in the data, consumption volatility has a negative relationship with asset prices. The model captures both the sign and the magnitude of this relationship.

I use a number of measures of conditional consumption volatility. All measures are constructed by first fitting an AR(1) model to log consumption growth to remove an expected growth component:

$$\Delta c_{t+1} = b_0 + b_1 \Delta c_t + \epsilon_{c,t}$$  \hspace{1cm} (1.25)

I then either estimate GARCH-type models on the residual $\epsilon_{c,t+1}$ or use the mean absolute residual as a non-parametric measure of conditional volatility. I use quarterly data because it is difficult to detect time-varying volatility with the post-war annual sample of only 50 observations. This procedure follows Bansal, Khatchatrian, and Yaron (2005a), Bansal, Kiku, and Yaron (2009) and Beeler and Campbell (2009).

Table 1.5 compares GARCH estimates of consumption volatility from the data and model. Panel A shows results from a GARCH(1,1) process

$$\sigma_{c,t+1}^2 = \omega_0 + \omega_1 \epsilon_{c,t}^2 + \omega_2 \sigma_{c,t}^2$$  \hspace{1cm} (1.26)
The data columns show modest time-variation in consumption volatility. Consistent with Bansal, Khatchatrian, and Yaron (2005a), the ARCH parameter $\omega_1$ just makes it to the 95% significance level if one uses non-robust standard errors. However, using standard errors which are robust to the assumption of non-normal shocks (Bollerslev and Wooldridge (1992)), one cannot reject the hypothesis of no time-varying volatility. This modest time-variation in volatility is consistent with the model columns. The model's mean estimate of the ARCH parameter is close to zero, and 5% of simulations produce a value of 0. Supposing that ARCH effects do exist, the data produce GARCH parameter estimates of $\hat{\omega}_2 = 0.79$ indicating that consumption volatility is persistent. The model captures this persistence, producing a GARCH parameter of 0.76. Indeed, this high persistence of consumption volatility may be the reason why time-varying volatility is hard to detect in the data.

Panel B compares estimates of a GJR-GARCH(1,1,1) model (Glosten, Jagannathan, and Runkle (1993))

$$\sigma_{c,t+1}^2 = \omega_0 + \omega_1 \epsilon_{c,t}^2 + \omega_2 \sigma_{c,t}^2 + \omega_3 I(\epsilon_{c,t} > 0)\epsilon_{c,t}^2$$

GJR-GARCH introduces an additional term which allows negative shocks to have a larger effect on volatility, as predicted by the analysis of Section 1.2. Though the data show significant sampling uncertainty, the point estimates are consistent with the intuition. The asymmetric GJR-GARCH parameter $\hat{\omega}_3$ is much larger than the symmetric ARCH $\hat{\omega}_1$ in both the model and the data. In terms of magnitudes, the model's parameter estimates are all smaller than the data's. Consistent with Panel A, the model reproduces the modest time-variation in consumption volatility, as seen in the data.
Table 1.5: GARCH Estimates of Time-Varying Consumption Volatility

Data and figures are quarterly. This table shows measures GARCH estimates of the residual from an AR(1) model of consumption growth

$$\Delta c_{t+1} = b_0 + b_1 \Delta c_t + \epsilon_{c,t+1}$$

GARCH estimation is done by quasi maximum likelihood. Robust standard errors use the Bollerslev-Wooldridge method. The model columns show means and percentiles across simulations of the same length as the empirical sample. Further details are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th>Panel A: GARCH</th>
<th>Data: 1948Q1-2011Q4</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>1.39E-06</td>
<td>8.51E-07</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>0.79</td>
<td>0.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: GJR-GARCH</th>
<th>Data: 1948Q1-2011Q4</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>1.97E-06</td>
<td>1.25E-06</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>0.76</td>
<td>0.10</td>
</tr>
<tr>
<td>$\omega_3$</td>
<td>0.14</td>
<td>0.11</td>
</tr>
</tbody>
</table>

There are two reasons why the model can generate substantial fluctuations in risk premia out of modest GARCH effects. The first is that habit preferences amplify fluctuations in consumption volatility. This can be seen in the surplus consumption process (1.5). Surplus consumption is basically an AR1 process, where the shocks are consumption growth and the conditional volatility $\lambda$ is a free parameter. The calibration chooses a large $\lambda = 1/0.06 - 1 \approx 15$ which provides substantial amplification. Thus small movements in the volatility of consumption growth produce large fluctuations in the volatility of marginal utility. The second is that consumption volatility generated by the model is very persistent. This
persistent consumption volatility makes it hard to measure time-varying volatility using only 50 years of quarterly data.

A key prediction of the precautionary volatility channel is that consumption volatility is high when asset prices are low. Table 1.6 shows that this prediction is supported by the data. The table shows regressions of various proxies for consumption volatility on the price-dividend ratio. The data column shows that, using all proxies, consumption volatility and the price-dividend ratio are negatively related. A high price-dividend ratio indicates a safe time of low volatility. The model replicates this pattern. The large standard errors in Table 1.5 suggest that comparing magnitudes should be done with caution, but with that in mind, the model coefficients are of similar magnitude to those from the data.

1.6.4 The Large and Volatile Equity Premium

The high volatility of the equity premium comes from high capital adjustment costs and the low elasticity of intertemporal substitution (EIS) of habit preferences. High capital adjustment costs mean that productivity shocks are absorbed by asset prices rather than investment. This channel is not new (Jermann (1998), Kogan (2004), Jermann (2010), Kogan and Papanikolaou (2012)) and so I provide only a brief discussion. The role of the low EIS is that a low EIS pins down high adjustment costs via estimation and general equilibrium effects. The importance of the EIS is less well understood and so I focus on it.

The link between equity volatility and adjustment costs can be seen in the investment return - stock return identity (Cochrane (1991), Restoy and Rockinger (1994)). Since the model has a homogenous production technology, this identity
Table 1.6: Regressions of Consumption Volatility on the Price-Dividend Ratio

This table shows regressions of the form

\[ cvol_t = \alpha + \beta (p_t - d_t) + \epsilon_t \]

where \( cvol_t \) is a measure of conditional consumption volatility, \( p_t \) is the log equity price, and \( d_t \) is the log dividend. To generate \( cvol_t \), first an AR(1) model (1.25) is run on log consumption growth. Panel A estimates either a GARCH(1,1) or GJR-GARCH(1,1,1) model on the residuals. Panel B uses a non-parametric measure: \( cvol_t(L) \equiv \log \left( \sum_{j=1}^{L} |\epsilon_{c,t+j}| \right) \), where \( \epsilon_{c,t+j} \) is the residual from the AR(1) model. Consumption data is quarterly and price-dividend ratio data is annual, which results in some abuse of notation. The model columns show means and percentiles across simulations of the same length as the empirical sample. Further details are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th>Panel A: GARCH estimates of consumption volatility</th>
<th>Data Model</th>
<th>1948Q1-2011Q4</th>
<th>mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model: GARCH</td>
<td>( \hat{\beta} )</td>
<td>-0.43</td>
<td>-0.19</td>
<td>-0.50</td>
<td>-0.15</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>SE(( \hat{\beta} ))</td>
<td>0.14</td>
<td>0.05</td>
<td>0.00</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.37</td>
<td>0.23</td>
<td>-0.01</td>
<td>0.18</td>
<td>0.61</td>
</tr>
<tr>
<td>Model: GJR-GARCH</td>
<td>( \hat{\beta} )</td>
<td>-0.43</td>
<td>-0.32</td>
<td>-0.75</td>
<td>-0.32</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>SE(( \hat{\beta} ))</td>
<td>0.16</td>
<td>0.07</td>
<td>0.00</td>
<td>0.06</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.31</td>
<td>0.33</td>
<td>-0.01</td>
<td>0.34</td>
<td>0.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Non-parametric estimates of consumption volatility</th>
<th>Data Model</th>
<th>1948Q1-2011Q4</th>
<th>mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model: cvol(<em>t)( (L) \equiv \log \left( \sum</em>{j=1}^{L}</td>
<td>\epsilon_{c,t+j}</td>
<td>\right) )</td>
<td>( \hat{\beta} )</td>
<td>-0.66</td>
<td>-0.21</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>SE(( \hat{\beta} ))</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.23</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>( R^2 )</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
<td>0.17</td>
<td>0.36</td>
</tr>
</tbody>
</table>

means we can express the stock return as

\[ R_{t,t+1} = \frac{\alpha(Y_{t+1}/K_{t+1}) + (1 + \phi(I_{t+1}/K_{t+1}))(1 - \delta) + \frac{\phi}{2} (I_{t+1}/K_{t+1})^2}{1 + \phi(I_t/K_t)} \] (1.27)
The stock return is related to the marginal product of capital \( \alpha(Y_{t+1}/K_{t+1}) \), the investment rate \( I_{t+1}/K_{t+1} \), and the adjustment cost parameter \( \phi \). This equality holds state-by-state, which means that the volatilities of the two sides are equal.

This identity shows the difficulty of matching equity volatility in a production economy. The marginal product of capital and the investment rate have a very low volatility, less than 2% per year. On the other hand, the equity return has a huge volatility of about 20% per year. The only way to reconcile these two sides, then, is with capital adjustment costs. Adjustments costs increase the curvature of the RHS. Jensen’s inequality then implies that adjustment costs increase the volatility of the investment return. The large discrepancy between the equity volatility and the volatility of the investment rate imply a very large adjustment cost parameter.

The role of the EIS comes through general equilibrium effects and standard empirical restrictions. General equilibrium means that both the EIS and capital adjustment costs affect consumption volatility. The EIS channel works through the SDF. Since the firm uses the investor’s IMRS as an SDF, the firm is rewarded for providing consumption which matches the investor’s preferences. A low EIS results in a strong incentive to produce smooth dividends, and through market clearing, smooth consumption. Capital adjustment costs also affect consumption volatility through market clearing. Large adjustment costs encourage the firm to keep investment smooth in the face of shocks. Since shocks must be absorbed by either investment or dividends, this encourages volatile dividends, and, through market clearing, volatile consumption.

Since both the EIS and capital adjustment cost affect consumption volatility, the ubiquitous requirement that asset pricing models should fit consumption data means that these two elements must match each other. A low EIS requires
high capital adjustment costs, and vice versa. In this model, the large Sharpe ratio pins down a low EIS. Given the low EIS, the volatility of consumption pins down a high adjustment cost. Note that the adjustment cost is not pinned down by equity volatility, and so equity volatility forms an overidentified restriction which is satisfied by the model.

This EIS-adjustment cost link exists in long-run risk and disaster models, but it tends to hurt rather than help their asset pricing results. The vast majority of long-run risk and disaster papers do not calibrate the EIS. However, the EIS must be larger than one in both models in order to match certain qualitative facts, and the common wisdom is that the larger the EIS, the better the asset pricing results. For long-run risk models, this is required to qualitatively match the fact that consumption volatility and the price-dividend ratio are negatively related (Bansal and Yaron (2004)). In disaster models, this is required to match the intuitive notion that a rise in disaster risk lowers investment and increases excess returns (Gourio (2010)). This large EIS then implies low adjustment costs and counterfactually low equity volatility (Kaltenbrunner and Lochstoer (2010), Gourio (2010)).

Given that equity volatility is high, the large equity premium is the consequence of a large Sharpe ratio. The large Sharpe ratio is simply due to the fact that habit preferences offer an additional degree of freedom for the modeler. Habit results in an additional term in the SDF (1.7). The modeler can choose the volatility of this term to be high by adjusting the preference parameter $\lambda$ in equation (1.5).

Economically, $\lambda$ can be interpreted as the ‘moodiness’ of the economy. The surplus consumption process says that if consumption growth goes up by 1%, surplus consumption gets boosted by $\lambda\%$. The calibration chooses $\lambda = 1/0.065 - 1 \sim 15$, meaning that changes in ‘mood’ are responsible for the vast majority
of changes in marginal utility. Checking this magnitude by introspection is, of course, a dangerous activity. But Section 1.5 shows that it is consistent with numerous overidentifying restrictions regarding asset prices.

1.6.5 The Low and Smooth Risk-Free Rate

The low risk-free rate is simply the result of time preference. Time preference is effectively a free parameter which allows me to hit the low risk-free rate in the data. Economically speaking, this results in an intuitive time preference of \( \beta^4 \approx 0.90 \) annually, that is, consumption one year from today is worth 90% of consumption today.

The smooth risk-free rate comes from an interplay of intertemporal substitution and precautionary savings effects. This channel has a simple intuition. In bad times, people want to borrow in order to consume today. But in bad times, the economy is particularly volatile, and the desire for precautionary savings prevents them from borrowing. This channel is similar to that of Campbell and Cochrane (1999), but with endogenous consumption volatility driving precautionary savings effects rather than the exogenous habit sensitivity function. Indeed, Section 1.6.1 shows that the model does a good job of mimicking the Campbell-Cochrane sensitivity function and so the smoothness of the risk-free rate should not be surprising.

These intuitions are fleshed out by examining the log-normal approximation of the risk-free rate\(^6\)

\[ r^f_{t+1} = -\log \left[ \mathbb{E}_t (e^{m_{t+1}}) \right] = -\mathbb{E}_t (m_{t+1}) - \frac{1}{2} \text{Var}_t (m_{t+1}) \]

Then just plug in the log SDF \( m_{t+1} = \log \beta - \gamma \Delta s_{t+1} - \gamma c_{t+1} \) and habit process \( \Delta s_{t+1} = -(1 - \rho_s)(s_t - \bar{s}) + \lambda \Delta c_{t+1} \).

---

\(^6\)If \( m_{t+1} \) is normal,
The 1st term reflects time preference. Time preference has a small impact on other unconditional moments, and so it is essentially a free parameter which one can use to fit the low mean risk-free rate in the data.

The 2nd and 3rd terms are due to intertemporal substitution and tend to create excessive volatility in habit models (Jermann (1998), Boldrin, Christiano, and Fisher (2001)). They reflect the Friedman (1957) permanent income hypothesis. In bad times, investors want to borrow from the future in order to consume today. This motive pushes down the price on the risk-free bond and pushes up the risk-free rate, leading to a countercyclical effect on a risk-free rate. Habit models imply a very strong consumption smoothing motive which makes this channel very volatile.

This model has the volatile intertemporal substitution effect typical of habit models, but it counters this effect with a precautionary savings effect. The precautionary savings effect runs through the 4th term, which is decreasing in the conditional volatility of consumption. Intuitively, in bad times, high consumption volatility creates a desire for savings. Investors buy bonds, pushing up the price and down the risk-free rate. This channel creates a procyclical effect on the risk-free rate, which helps counteract the countercyclical effect of the intertemporal substitution channel.

The previous discussion shows that, qualitatively, precautionary savings effects help counteract intertemporal substitution effects. Whether the quantitative effect is enough generate a smooth risk-free rate is another question. Figure
1.6 examines the quantitative effect. It plots the risk-free rate decomposition (1.28) against surplus consumption for various levels of productivity. The solid red lines represent intertemporal substitution effects. The dashed blue lines represent the precautionary savings effects. The two effects are near mirror-images of each other, showing that, quantitatively, the channels balance each other quite nicely.

Figure 1.6: Decomposition of the Risk-Free Rate. ‘Intertemporal Substitution’ represents the 2nd and 3rd terms of equation (28). ‘Precautionary Saving’ represents the last term. Computed from model’s laws of motion. Risk-free rate is in annualized %. Capital in all panels is fixed at the mean capital stock.

A nice feature of the smoothing effects in this model is that they arise endogenously. This contrasts with the risk-free rate smoothing effects of Campbell and Cochrane (1999), which are the result of a parameter choice. To see this, it helps to return to approximation of the risk-free rate (1.28). This expression shows that, if one allows the preference parameter $\lambda$ to vary over time, then one can control the magnitude of the precautionary savings effect. Indeed, Campbell and Cochrane (1999) do just that, and they choose the magnitude of the channel to exactly cancel out the intertemporal substitution effect. This model has no such freedom. The magnitude of this channel comes through the amount of countercyclicality in consumption volatility, which is the result of general equilibrium
effects of investors’ preferences for precautionary savings on firms’ production decisions.

1.6.6 Cash Flow Dynamics and Alternative Dividend Processes

The model has a very simple notion of dividends. Dividends are just the profits (net of investment) of a representative neoclassical firm (see equation (1.13)). This formulation has the advantage of transparency, but it abstracts from a number of issues which affect dividends in the data. For example, the model abstracts from debt and payout policy, as well as labor market frictions. The literature has shown that these abstractions lead to some counterfactual behavior of dividends (Rouwenhorst (1995), Jermann (1998), Kaltenbrunner and Lochstoer (2010)), and my model is not exempted from this issue. Extending the model to incorporate richer financial and production elements would help the model fit dividend dynamics (i.e. Jermann and Quadrini (2012), Kuehn, Petrosky-Nadeau, and Zhang (2012)), at the expense of complexity.

Would the main results still hold under a richer and more data-like dividend process? Table 1.7 suggests that the answer is yes. The table shows basic properties of dividends and excess returns for alternative dividend processes. The alternative processes specify dividends as stochastically related to consumption, which allows one to capture basic empirical dividend moments. This procedure follows the endowment economy literature (Bansal and Yaron (2004), Campbell and Cochrane (1999)). The table shows that, regardless of the dividend process, the model still produces a large and volatile equity premium.

The ‘neoclassical dividend’ column shows results from the baseline model. Consistent with previous studies, this simple technology leads to some counter-
Table 1.7: Alternative Dividend Processes: Basic Statistics

Figures are annual. ‘Neoclassical Dividend’ is a claim on the dividends from the firm (see equation 1.13). ‘Calibrated Dividend’ is a claim on the dividend process (1.29) calibrated to match U.S. data. ‘Consumption Claim’ is a claim on consumption. ∆d is dividend growth, ∆c is consumption growth, and r − rf is excess returns of the CRSP index. Dividend growth is computed using no reinvestment. Details of the data are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th></th>
<th>US 1948-2011</th>
<th>Neoclassical Dividend</th>
<th>Calibrated Dividend</th>
<th>Consumption Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ(∆d) (%)</td>
<td>13.11</td>
<td>6.21</td>
<td>13.10</td>
<td>1.20</td>
</tr>
<tr>
<td>ρ(∆d, ∆c)</td>
<td>0.17</td>
<td>-0.68</td>
<td>0.18</td>
<td>1.00</td>
</tr>
<tr>
<td>E(r − rf) (%)</td>
<td>6.47</td>
<td>6.77</td>
<td>8.01</td>
<td>6.86</td>
</tr>
<tr>
<td>σ(r − rf) (%)</td>
<td>16.07</td>
<td>16.07</td>
<td>20.69</td>
<td>17.71</td>
</tr>
<tr>
<td>AC1(r − rf)</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.06</td>
<td>-0.06</td>
</tr>
</tbody>
</table>

factual dividend statistics. Dividends are too smooth compared to the data, and dividend growth is negatively correlated with consumption, while the data show a mild positive correlation. Despite the negative correlation with consumption growth, a claim on the neoclassical dividend produces a large and volatile equity premium. One might think that, due to this negative correlation, this claim would serve as a hedge for consumption risk and offer low returns. However, the vast majority of the claim’s value comes from its continuation value, not the next year’s dividends. A negative shock leads the firm to cut investment and boost dividends temporarily, but this negative shock is a bad sign for a long future of dividends. This result of a large risk premium with negative consumption-dividend correlation is consistent with results from previous habit models (Jermann (1998)).

The ‘calibrated dividend’ column presents results from a dividend process which matches some basic dividend data moments. The dividend process is a
simple function of consumption

\[ d_{x,t} \equiv \psi c_t + \sigma_d \epsilon_{d,t} \] (1.29)

where \( \psi \) and \( \sigma_d \) are parameters and \( \epsilon_{d,t} \) is a standard normal shock. This approach captures some key statistical features of dividend growth and follows Campbell and Cochrane (1999) and Bansal and Yaron (2004), among others. Specifically, I choose \( \psi = 2 \) and \( \sigma_d = 0.18 \) to hit the correlation between consumption and dividend growth, and the volatility of dividend growth.

Table 1.7 shows the calibrated dividend process matches the volatility of dividend growth and its correlation with consumption growth, as intended. It also generates a data-like equity premium. Excess returns on this claim are large on average, volatile, and mildly negatively autocorrelated. This large risk premium exists despite the low correlation between dividend and consumption growth. Once again, the vast majority of this asset’s value comes from its continuation value, not the next year’s dividends. The ‘consumption claim’ column shows further robustness of the large and volatile equity premium. A claim on consumption also produces data-like stock returns.

Table 1.8 shows that alternative dividend processes help the model better fit the evidence of time-varying excess returns. It shows regressions of future dividend growth and future excess returns on today’s price-dividend ratio, using alternative dividend processes. As seen in Section 1.5, the benchmark model (‘neoclassical dividend’ column) captures the lack of dividend predictability at shorter horizons, although the amount of predictability is somewhat overstated at the 5-year horizon. The calibrated dividend column shows that this mild overstatement is fixed with a more realistic dividend process. The calibrated dividend
Table 1.8: Alternative Dividend Processes: Predictive Regressions

Figures are annual. 'Neoclassical Dividend' is a claim on the dividends from the firm (see equation 1.13). 'Calibrated Dividend' is a claim on the dividend process (1.29) calibrated to match U.S. data. ‘Consumption Claim’ is a claim on consumption. \( \Delta d \) is dividend growth, \( p_t - d_t \) is the log price-dividend ratio, and \( r - r_f \) is excess returns. Standard errors are Newey-West with \( 2(L - 1) \) lags. Details of the data are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th>L</th>
<th>US 1948-2011</th>
<th>Neoclassical Dividend</th>
<th>Exogenous Dividend</th>
<th>Consumption Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.03</td>
<td>-0.00</td>
<td>0.03</td>
<td>-0.04</td>
</tr>
<tr>
<td>( \hat{\beta} )</td>
<td>3 0.01</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>5 0.03</td>
<td>0.13</td>
<td>0.05</td>
<td>-0.01</td>
</tr>
<tr>
<td>SE(( \hat{\beta} ))</td>
<td>1 0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>3 0.07</td>
<td>0.05</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>5 0.09</td>
<td>0.07</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>1 0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>3 0.00</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>5 0.00</td>
<td>0.10</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L</th>
<th>US 1948-2011</th>
<th>Neoclassical Dividend</th>
<th>Exogenous Dividend</th>
<th>Consumption Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td>( \hat{\beta} )</td>
<td>3 -0.27</td>
<td>-0.30</td>
<td>-0.38</td>
<td>-0.38</td>
</tr>
<tr>
<td></td>
<td>5 -0.40</td>
<td>-0.44</td>
<td>-0.56</td>
<td>-0.55</td>
</tr>
<tr>
<td>SE(( \hat{\beta} ))</td>
<td>1 0.05</td>
<td>0.05</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>3 0.08</td>
<td>0.09</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>5 0.12</td>
<td>0.10</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>1 0.09</td>
<td>0.10</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>3 0.19</td>
<td>0.26</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>5 0.26</td>
<td>0.37</td>
<td>0.29</td>
<td>0.30</td>
</tr>
</tbody>
</table>

claim produces essentially zero dividend predictability, with a maximum \( R^2 \) of 0.02 at the 5-year horizon. This result is an intuitive extension of the consumption claim results. As seen in the 'consumption claim' column, consumption
growth is unpredictable at all horizons. Since the calibrated dividend claim is essentially a levered claim on consumption, it inherits this lack of predictability. This lack of dividend predictability shows up in excess return forecasts. Both the calibrated dividend claim and consumption claim produce strongly predictable excess returns.

The apparent lack of cash flow predictability is due to the uninformative-ness of asset prices, rather than a fundamental lack of predictability in the economy itself. This effect of conditioning information is seen in Figure 1.7, which shows plots of expected consumption growth computed using the model's laws of motion, that is, conditioning on all information available in the economy. The left panel shows consumption growth against the price-dividend ratio from the benchmark 'neoclassical' dividend process. The panel shows a cloud of consumption growth observations, and no relationship with asset prices. But the significant vertical dispersion of the cloud indicates that consumption growth is very predictable: it's just that asset prices are not very informative about consumption growth. Note that neither the price-dividend ratio nor the consumption-wealth ratio predicts consumption growth (see the 'consumption claim' column of Table 1.8). On the other hand, the model's state variables are informative. The middle and right panels show that consumption growth varies systematically with capital, surplus consumption, and productivity. Consumption growth is decreasing in capital and productivity, but increasing in surplus consumption. These relationships are more complicated than these panels suggest however, since the state variables are correlated in simulations.
**Figure 1.7: Predictability of Consumption Growth: Effect of the Conditioning Set.** Figures are annualized and computed from the model's laws of motion. Price-dividend ratio is for a claim on the firm's dividends. The left panel shows scatterplots from model simulations. The middle and right panels show consumption growth as a function of state variables.

### 1.6.7 Comparative Statics

The quantitative results come from estimated parameter values, but as with all econometric methods, the point estimates can be sensitive to choices of the econometrician. This section investigates the effect of changing the parameter values. The comparative statics also confirm the intuition developed earlier in the paper.

Table 1.9 shows key moments from these comparative statics exercises. Each column examines moments generated by models where only one of the parameter values is changed from the estimation described by Table 1.1. Three different parameter changes are examined: lower persistence of habit, weaker steady state habit, and lower capital adjustment costs. The magnitude of the perturbations are chosen to be the smallest change that produces a clearly recognizable deviation from the estimated results. This approach helps isolate the direct effect of changing a parameter value from its interaction with other elements of the model.
Table 1.9: Comparative Statics

Figures are annual. ‘Estimated’ represents parameter values from Table 1.1. All other columns use the estimated values but with one parameter changed. ‘Lower persistence of habit’ sets $\rho_s = 0.80$. ‘Weaker steady state habit’ sets $\tilde{S} = 0.12$. ‘Lower adjustment costs’ sets $\phi = 40$. Details of the data are found in Appendix 1.8.2.

<table>
<thead>
<tr>
<th>Identifying Moments</th>
<th>Data 1948-2011</th>
<th>Estimated</th>
<th>Lower persistence of habit</th>
<th>Weaker steady state habit</th>
<th>Lower adjustment costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}(r_f)$ (%)</td>
<td>0.98</td>
<td>0.98</td>
<td>3.31</td>
<td>8.25</td>
<td>4.66</td>
</tr>
<tr>
<td>$\mathbb{E}(R - R_f)/\sigma(R)$</td>
<td>0.48</td>
<td>0.48</td>
<td>0.39</td>
<td>0.31</td>
<td>0.38</td>
</tr>
<tr>
<td>$\sigma(r - r_f)$ (%)</td>
<td>16.07</td>
<td>16.07</td>
<td>24.22</td>
<td>14.77</td>
<td>8.92</td>
</tr>
<tr>
<td>$\sigma(c_{hp})/\sigma(y_{hp})$</td>
<td>0.47</td>
<td>0.47</td>
<td>0.43</td>
<td>0.64</td>
<td>0.39</td>
</tr>
<tr>
<td>$\sigma(y_{hp})$ (%)</td>
<td>1.50</td>
<td>1.50</td>
<td>1.63</td>
<td>1.53</td>
<td>1.49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Untargeted Moments</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}(r - r_f)$ (%)</td>
<td>6.47</td>
<td>6.77</td>
<td>7.46</td>
<td>3.84</td>
<td>3.11</td>
</tr>
<tr>
<td>$\text{AC1}(r - r_f)$</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.12</td>
<td>-0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td>$\sigma(r_f)$ (%)</td>
<td>2.24</td>
<td>2.96</td>
<td>9.20</td>
<td>5.20</td>
<td>2.30</td>
</tr>
<tr>
<td>$\text{AC1}(r_f)$</td>
<td>0.56</td>
<td>0.88</td>
<td>0.66</td>
<td>0.86</td>
<td>0.93</td>
</tr>
<tr>
<td>$\mathbb{E}(p - d)$</td>
<td>3.42</td>
<td>2.67</td>
<td>2.23</td>
<td>2.15</td>
<td>2.63</td>
</tr>
<tr>
<td>$\sigma(p - d)$</td>
<td>0.43</td>
<td>0.46</td>
<td>0.39</td>
<td>0.33</td>
<td>0.39</td>
</tr>
<tr>
<td>$\text{AC1}(p - d)$</td>
<td>0.95</td>
<td>0.90</td>
<td>0.78</td>
<td>0.89</td>
<td>0.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$R^2$ from forecasting $r - r_f$</th>
<th>1-year</th>
<th>0.08</th>
<th>0.10</th>
<th>0.07</th>
<th>0.04</th>
<th>0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>with $p - d$</td>
<td>3-year</td>
<td>0.19</td>
<td>0.26</td>
<td>0.14</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>5-year</td>
<td>0.26</td>
<td>0.37</td>
<td>0.17</td>
<td>0.15</td>
<td>0.25</td>
</tr>
</tbody>
</table>

**Lower persistence of habit**

The third column of Table 1.9 examines a model where the persistence of habit $\rho_s$ is lowered from the estimated value of 0.963 to 0.800. This parameter is important because the strong persistence of habit is one way in which this model deviates from previous models with habit and production. The high estimated persistence means that habit today depends on a very long history of consumption. In contrast, habit in Jermann (1998) and Boldrin, Christiano, and Fisher (2001) depends only on the last quarter’s consumption.
The table shows the high persistence is critical. In the lower persistence model, the volatility of the risk-free rate triples from the estimated value of 2.96% to 9.20%, bringing the model in line with the high risk-free rate volatility of Jermann (1998) and Boldrin, Christiano, and Fisher (2001). Figure 1.8 explains why. The figure plots the risk-free rate decomposition of equation (1.28) for the low persistence of habit model. The figure shows that the intertemporal substitution effect is still countercyclical and the precautionary saving effect is still procyclical. However, the two channels no longer cancel each other out quantitatively. The intertemporal substitution effect is extremely countercyclical and overwhelms precautionary saving effects. This change can traced to the lower persistence of habit. A low persistence of habit means that habit will strongly mean revert tomorrow. This means that, in bad times, there is a pronounced desire to borrow in order to consume today.

Another consequence of faster mean reversion is that precautionary saving effects become weaker. Faster mean reversion means that habit will recover quickly
from bad times, and thus there is less countercyclical uncertainty about savings. This weakening of precautionary savings effects is reflected in a reduction in the amount of time-varying risk premia. The third column of Table 1.9 shows that the amount of time-variation in risk premia falls. $R^2$’s from regressions of future excess returns on the log price-dividend ratio drop significantly.

The persistence of habit has a strong effect on many other asset price moments. The parameter is identified with the volatility of excess returns, and, as expected, the low persistence model strongly overpredicts this volatility. Another effect is that the persistence of the risk-free rate and the price-dividend ratio drops. Intuitively, the persistence of preferences is reflected in the persistence of asset prices.

**Weaker steady state habit**

The fourth column of Table 1.9 raises steady state surplus consumption $\bar{S}$ from the estimated value of 0.063 to 0.120. This comparative static brings the model closer to CRRA utility. Roughly speaking, $\bar{S} = 1.00$ is an economy with no habit, so this model is still very far from the standard model.

Weakening steady state habit lowers the volatility of marginal utility of the household (see equation (1.21)), and thus has the direct effect of lowering the Sharpe ratio from the estimated value of 0.48 to 0.31. This reduced Sharpe ratio has the obvious effect of reducing the equity premium from the estimated value of 6.77% to 3.84%. The Sharpe ratio does not drop as much as one might expect, however, since the volatility of consumption increases. This increase in consumption volatility is due to a weaker general equilibrium consumption smoothing effect (see Section 1.6.4).

Weakening steady state habit has the additional effect of weakening precau-
tionary savings effects. With weaker habit, there is less need for precautionary savings, and the risk-free rate jumps nearly ten-fold. This weakening of precautionary savings effects can be seen in the return forecasting regressions too. The $R^2$’s from regressions of future returns on the price-dividend ratio all drop by more than half.

**Lower adjustment costs**

The final column of Table 1.9 examines the effect of lowering capital adjustment costs. The adjustment cost parameter $\phi$ is lowered from the estimated value of 75.00 to 40.00. This comparative static helps compare my model with Lettau and Uhlig (2000), who also examine external habit in a production economy. My model deviates from theirs in two important ways. The first is that I include capital adjustment costs, while Lettau and Uhlig assume costless adjustment. The second is that I use a non-linear, global solution method, while Lettau and Uhlig linearize the model around the non-stochastic steady state. This comparative static shows that capital adjustment costs make a large impact on the model results and that lowering the costs brings my model closer to Lettau and Uhlig (2000).

The table shows that lowering adjustment costs reduces the relative volatility of consumption to output falls from the estimated value of 0.47 to 0.39. This is consistent with the intuition from Section 1.6.4. Low adjustment costs reduce the incentive for smooth investment and encourages volatile consumption.

This decrease in consumption risk is reflected in a lower Sharpe ratio. The equity premium drops by half. With lower consumption risk, the precautionary motive is weakened and the risk-free rate rises by a factor of five. This decrease in the precautionary motive is also seen in lower $R^2$’s from return forecasting.
1.7 Conclusion

Time-varying uncertainty plays a key role in many theories of aggregate asset prices. Previous papers take time-varying uncertainty as exogenous. This paper shows that time-varying uncertainty is the natural result of a fundamental economic motive: the desire for precautionary savings. These motives make investors very sensitive to shocks in bad times, and lead to countercyclical consumption volatility. This mechanism, when amplified by external habit preferences, quantitatively accounts for the empirical evidence of time-varying risk premia. The model also provides a parsimonious description of asset prices and the real economy. The estimated model matches a long list of facts about the aggregate stock market, safe government bonds, as well as consumption, output, and investment.
1.8 Proofs, Data, and Solution Method

1.8.1 Proofs for the Two-period Model

**Proof of Proposition 1.** A change of variables shows that this model is equivalent to a standard consumption-savings problem. Shift consumption by assigning $C^* = C_1 - H$. Then the date 1 consumption rule can be written as

$$C_1(W_1) = C^*(W_1 - H_1, -H_2) + H_1$$ (1.30)

where $C^*(W, Y)$ solves a simple consumption-savings problem with wealth $W$ and certain future income of $Y$:

$$C^*(W, Y) \equiv \arg \max_C u(C) + \mathbb{E}\{u[R(W - C) + Y]\}$$ (1.31)

and for ease of notation I suppress the subscript 2 on $R$. It turns out that for $Y \neq 0$, $R$ random, and CRRA utility, $C^*(W, Y)$ is strictly convex in $W$. That is, with CRRA utility, rate of return randomness, and the introduction of any (even constant) future income is a sufficient condition for generating strict convexity of the consumption function. This is not one of the sufficient conditions shown in Carroll and Kimball (1996), so I will show that it is sufficient in what follows.

The FOC of the shifted problem (1.31) is

$$u'(C^*(W, Y)) = \phi'(W - C^*(W, Y))$$ (1.32)

Where, for convenience, I’ve defined the function

$$\phi(S) \equiv \mathbb{E}\{u[RS + Y]\}$$
Taking $\frac{\partial}{\partial W}$ of the FOC and rearranging gives

$$\frac{\partial}{\partial W} C^*(W, Y) = \frac{\phi''}{u'' + \phi''}$$

Take another $\frac{\partial}{\partial W}$, do some serious algebra, and we get an expression for the convexity of $C^*$:

$$\frac{\partial^2}{\partial W^2} C^*(W, Y) = \left[ \frac{(u'')^2 (\phi'')^2}{u' \times [u'' + \phi'']^3} \right] \left[ \frac{\phi' \phi''' - u' u'''}{(\phi'')^2} \right]$$

(1.33)

The first bracket is negative simply because $u' > 0$ and $u'' < 0$. To show that the second bracket is (strictly) positive, first note that, due to the CRRA specification, $\frac{u'u'''}{(u'')^2} = 1 + \frac{1}{\gamma}$. I will now show that, due to the non-zero future income $Y$, $\frac{\phi' \phi'''}{(\phi'')^2} > 1 + \frac{1}{\gamma}$. This is an extension of Carroll and Kimball (1996)'s Lemma 4.

Proving $\frac{\phi' \phi'''}{(\phi'')^2} > 1 + \frac{1}{\gamma}$ requires the following technical Lemma.

**Lemma 1.** Let $\Phi_i$ for $i = 1, ..., N$ be $2 \times 2$ symmetric matrices with the following properties:

- the diagonals of each $\Phi_i$ are positive
- the off-diagonals of each $\Phi_i$ are all negative
- for every $i$, $|\Phi_i| = 0$
- $|\sum_{i=1}^N \Phi_i| = 0$

Then for each pair $i, j$ there is some constant $k$ such that

$$\Phi_i = k\Phi_j$$

**Proof.** I will first show this for the case where $N = 2$. I’ll then use the $N = 2$ results
to prove the general case. For ease of notation assign \( \Phi_1 \equiv \begin{pmatrix} p & q \\ q & r \end{pmatrix} \) and \( \Phi_2 \equiv \begin{pmatrix} x & y \\ y & z \end{pmatrix} \). With some algebra, one can show that

\[
|\Phi_1 + \Phi_2| = |\Phi_1| + |\Phi_2| + [\sqrt{prz} - \sqrt{xz}]^2 + 2[\sqrt{prxz} - qy] \tag{1.34}
\]

A few facts will let us simplify this expression dramatically. First \(|\Phi_1| = |\Phi_2| = |\Phi_1 + \Phi_2| = 0\), so those terms all drop out. Then note that, since \(|\Phi_1| = |\Phi_2| = 0\), we have \(pr = q^2\) and \(xz = y^2\), and we can rewrite

\[
\sqrt{prxz} = \sqrt{q^2y^2} = qy
\]

and so the last term in equation (1.34) also drops out. Thus equation (1.34) implies that \(pz = xr\), or

\[
\frac{p}{x} = \frac{r}{z} = k \tag{1.35}
\]

where \(k\) is the conjectured constant of proportionality. We just need to show that

\[
\frac{q}{y} = \frac{p}{x} = k.
\]

To show this, plug \(pr = q^2\) and \(xz = y^2\) into (1.35) and we have

\[
\frac{p}{x} = \frac{q^2/p}{y^2/x} \Rightarrow \frac{p}{x} = \frac{q}{y}
\]

This completes the \(N = 2\) case.

To show the general case, first note that if \(\Phi_i\) and \(\Phi_j\) satisfy the requirements of the lemma, then \(\Phi_i + \Phi_j\) also satisfies those requirements. Thus I can apply the \(N = 2\) results to the general case, where one matrix is \(\Phi_i\) and the other matrix
is $\sum_{j \neq i} \Phi_i$. Moreover, note that if there is a $k$ such that $\Phi_i = k \sum_{j \neq i} \Phi_i$, then there is $m$ such that $\sum_j \Phi_j = m \Phi_i$. Apply this to all $i$ and get the desired result $\Phi_i = k \Phi_j$. 

Now, back to proving the proposition. I want to show that $\phi' \phi'' > 1 + \frac{1}{\gamma}$. Suppose, for contradiction, that $\frac{\phi' \phi''}{(\phi')^2} \leq 1 + \frac{1}{\gamma}$. I can write this expression using the determinant of a $2 \times 2$ matrix by defining

$$\Phi \equiv \mathbb{E} \begin{bmatrix} \phi' & \sqrt{1 + \frac{1}{\gamma} \phi''} \\ \sqrt{1 + \frac{1}{\gamma} \phi''} & \phi''' \end{bmatrix} = \mathbb{E} \begin{bmatrix} Ru'(z) & \sqrt{1 + \frac{1}{\gamma} R^2 u''(z)} \\ \sqrt{1 + \frac{1}{\gamma} R^2 u''(z)} & R^3 u'''(z) \end{bmatrix}$$

Where, for ease of notation, $z \equiv R(W - C^*(W, Y)) + Y$. This expression can now be written compactly as

$$|\Phi| \leq 0$$

Note that $\Phi$ is the weighted sum of many component matrices

$$\begin{bmatrix} Ru'(z) & \sqrt{1 + \frac{1}{\gamma} R^2 u''(z)} \\ \sqrt{1 + \frac{1}{\gamma} R^2 u''(z)} & R^3 u'''(z) \end{bmatrix}$$

and that due to the CRRA specification of $u$, the determinant of each component matrix is zero. Thus $\Phi$ is positive semidefinite, so $|\Phi| \geq 0$. But our assumption for contradiction says $\Phi$ is negative semidefinite, and so it must be that $|\Phi| = 0$.

Now I use Lemma 1. The lemma states that if $|\Phi| = 0$, all of the component matrices must be proportional to one another. This means that for any states $i$ and $j$ the ratio of the diagonal terms of the corresponding matrices is equal, that
is, for any \(i\) and \(j\),

\[
\frac{R_i u'_i}{R_j u'_j} = \left( \frac{R_i}{R_j} \right)^3 \frac{u'''_i}{u'''_j}
\]

\[
\Rightarrow \left( \frac{R_i S + Y}{R_j S + Y} \right)^{-\gamma} = \left( \frac{R_i}{R_j} \right)^2 \left( \frac{R_i S + Y}{R_j S + Y} \right)^{-\gamma^2}
\]

\[
\Rightarrow \frac{R_i S + Y}{R_j S + Y} = \frac{R_i}{R_j}
\]

\[
\Rightarrow S + \frac{Y}{R_i} = S + \frac{Y}{R_j}
\]

\[
\Rightarrow R_i = R_j
\]

which is a contradiction, since \(R\) is random. Note that the presence of a nonzero income \(Y\) is critical because otherwise, I could not move from the fourth line to the fifth line in the equations above.

Therefore, \(\frac{\phi' \phi'''}{(\phi')^2} > 1 + \frac{1}{\gamma}\), and by equation (1.33), \(C^*(W, Y)\) is strictly concave in \(W\), and by equation (1.30), \(C''_1(W) < 0\).

Proof of Proposition 2. I first show that the transformed consumption function satisfies \(\frac{\partial^3}{\partial W^2} C^*(W, Y) > 0\). To show this, note that \(C^*(W, 0)\) is linear, but for any \(\epsilon > 0\), \(C^*(W, \epsilon)\) is strictly concave. Thus we can sign the derivative

\[
\frac{\partial^3}{\partial Y \partial W^2} C^*(W, 0) = \lim_{\epsilon \to 0} \left[ \frac{\partial^2}{\partial W^2} C^*(W, \epsilon) - \frac{\partial^2}{\partial W^2} C^*(W, 0) \right] \epsilon = \lim_{\epsilon \to 0} \left[ \frac{\partial^2}{\partial W^2} C^*(W, \epsilon) \right] \frac{\epsilon}{\epsilon} < 0
\]

Assuming that \(\frac{\partial^3}{\partial Y \partial W^2} C^*(W, 0)\) is continuous, this means that there is some \(\bar{Y} > 0\) such that for any \(\delta < \bar{Y}\), \(\frac{\partial^3}{\partial Y \partial W^2} C^*(W, \delta) < 0\).
Since $C^*(W,Y)$ is HD1, I can take derivatives to show that

$$\frac{\partial^3}{\partial W^3} C^*(W/\delta, 1) = \frac{\delta^3}{W} \left[ - \frac{1}{\delta^2} \frac{\partial^2}{\partial W^2} C^*(W, \delta) - \frac{\partial^3}{\partial Y \partial W^2} C^*(W, \delta) \right] < 0$$

Now note that I am free to choose $W$, so this means, for any $W$, $\frac{\partial^3}{\partial W^3} C^*(W, 1) > 0$. But by homogeneity, $\frac{\partial^3}{\partial W^3} C^*(W, Y) = \frac{1}{Y^2} \frac{\partial^3}{\partial W^3} C^*(W, 1) > 0$, and thus the third derivative with respect to wealth of the transformed consumption function is negative (as long as $Y \neq 0$).

Now to finish proving the proposition. Using the transformed consumption function (1.31) I can relate the desired derivative to the third derivative of the consumption function. First take $\frac{\partial}{\partial W}$ twice:

$$C''_1(W_1) = \frac{\partial^2}{\partial W^2} C^*(W_1 - H_1, -H_2)$$

Then take $\frac{\partial}{\partial H_1}$

$$\frac{\partial}{\partial H_1} C''_1(W_1) = - \frac{\partial^3}{\partial W^3} C^*(W_1 - H_1, -H_2) < 0$$

$\square$

**Proof of Corollary 2.** Let $W_0$ and $\overline{W}_0$ be the minimum and maximum of $W_0$, re-
spectively.

\[
\max_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_1)] - \min_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_1)]
\]

\[
= \sigma_0(\Delta W_1) \left[ \max_{W_0 \in \mathbb{W}_0} C'_1(W_0) - \min_{W_0 \in \mathbb{W}_0} C'_1(W_0) \right]
\]

\[
= \sigma_0(\Delta W_1) \left[ C'_1(W_0) - C'_1(W_0) \right]
\]

\[
= \sigma_0(\Delta W_1) \left[ \int_{\mathbb{W}_0} dW C''_1(W) \right]
\]

\[
= -\sigma_0(\Delta W_1) \left[ \int_{\mathbb{W}_0} dW C''_1(W) \right]
\]

Where the 2nd line uses equation (1.2), the third line uses \(C''_1(W) < 0\).

Then take \(\frac{\partial}{\partial H_1}\)

\[
\frac{\partial}{\partial H_1} \left[ \max_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_1)] - \min_{W_0 \in \mathbb{W}_0} \sigma_0[C_1(W_1)] \right] = -\sigma_0(\Delta W_1) \left[ \int_{\mathbb{W}_0} dW \frac{\partial}{\partial H_1} C''_1(W) \right] > 0
\]

\(\square\)

1.8.2 Data Details

The data span 1948 thru 2011. Macroeconomic quantities are from the BEA NIPA and BEA fixed asset tables. All data is real and per-capita. Output is simply GDP. Consumption is nondurable consumption plus services consumption. Investment and capital are fixed investment plus durable consumption plus government investment. All quantities are constructed by dividing nominal figures (Table 1.1.5 or Table 3.9.5) by the appropriate price index (Table 1.1.4 or Table 3.9.4), and then dividing by the population (from Table 2.1).

Most studies exclude government investment from their definition of invest-
ment, citing the fact that their models do not include government. However, this logic would also imply that one should remove government purchases from GDP, which is not typically done. I choose to keep government purchases in GDP and include government investment in investment as this approach preserves the idea that the data come from a single general equilibrium system. Moreover, this approach is closer to the spirit of Cooley and Prescott (1995). Choice of the empirical definition of investment does not have a significant effect on any of the results.

Asset price data is taken from CRSP. The equity is the CRSP value-weighted index. Price-dividend ratios are computed annually using the stock market reinvestment assumption. As discussed in Cochrane (2011), this aggregation method preserves the Campbell-Shiller identity, which is useful for identifying the source of asset price fluctuations. CRSP returns are deflated using the CPI, also obtained from CRSP. The risk-free rate is computed using a forecast of the ex-post real return of the 90-day T-bill following Beeler and Campbell (2009). Following Beeler and Campbell (2009), the ex-post real return is calculated by deflating the 90-day nominal T-bill yield using seasonally adjusted CPI from the BLS. The forecast is constructed by regressing next quarter's ex-post real return on today's nominal 90-day yield and the mean inflation rate over the previous year.

1.8.3 Simulated Method of Moments Details

This section spells out details of the SMM method. I use a different notation than the brief discussion in the text in order to be more precise.

---

Cooley and Prescott (1995) say that “Our economy is very abstract: it contains no government sector, no household production sector, no foreign sector and no explicit treatment of inventories. Accordingly, the model economy’s capital stock, $K$, includes capital used in all of these sectors plus the stock of inventories. Similar, output, $Y$, includes the output produced by all of this capital.”
Econometric Details

I transform the parameters so that I do not need to be concerned about corner solutions. For example, rather than estimate $\beta \in [0, 1]$, I estimate $\logit(\beta) \in \mathbb{R}$. After the estimation is done, I find the point estimates of the original parameters by inverting the transformation, and the standard errors by the delta method.

To be explicit, let $\zeta \in \Theta$ be the vector of original parameters and $\psi^{-1}(\zeta) = \theta \in \mathbb{R}_K$ by the transformed parameters. Then

$$\sqrt{T}(\hat{\zeta}_T - \zeta_0) = \sqrt{T}[(\hat{\theta}_T) - \psi(\theta_0)] \quad (1.36)$$

$$\approx \sqrt{T}[(\hat{\theta}_0)'(\hat{\theta}_T - \theta_0)] \quad (1.37)$$

so the asymptotic variance of the original parameters can be estimated by

$$\text{Var}[\sqrt{T}(\hat{\zeta}_T - \zeta_0)] = D\psi(\hat{\theta}_T)'\text{Var}[\sqrt{T}(\hat{\theta}_T - \theta_0)]D\psi(\hat{\theta}_T) \quad (1.38)$$

I also use the delta method on the moment errors, because some of the “moments” I use are not moments in the traditional sense (i.e. the Sharpe Ratio) but are more formally described as transformations of moments. That is, let the GMM objective be

$$\hat{\theta}_T = \arg\min_{\theta \in \mathbb{R}} Q_T(\theta) = G_T(\theta)'W_T G_T(\theta)$$

where $G_T(\theta)$ are some moment errors.

$$G_T(\theta) \equiv S^{-1} \sum_{s=1}^{S} h[\tilde{f}_T^s(\theta)] - h[\tilde{f}_T^+]$$

68
where

\[
\tilde{f}_t^s(\theta) \equiv T^{-1} \sum_t f_t^s(\theta) \quad [M \times 1]
\]

is a set of moments from simulation \(s\)

\[
h : \mathbb{R}^M \rightarrow \mathbb{R}^N
\]

transforms the moments into some other statistic

and \(S\) is the total number of simulations. This \(h\) function has no effect on the GMM formula for the asymptotic variance

\[
\hat{\text{Var}}[\sqrt{T}(\hat{\theta}_T - \theta_0)] \equiv [DG_T W_T DG'_T]^{-1}[DG_T W_T \hat{X}_T W_T DG'_T][DG_T W_T DG'_T]^{-1}
\]

where

\[
DG_T \equiv DG_T(\hat{\theta}_T)
\]

\[
\hat{X}_T \equiv \hat{\text{Var}}(\sqrt{T}G_T(\theta_0))
\]

but it does have an effect on the estimator of the variance of the moment errors \(\hat{\text{Var}}(\sqrt{T}G_T(\theta_0))\).

\[
\hat{X}_T \equiv D h(\hat{f}_T^*)' \left(1 + \frac{1}{S} \Sigma_T \right) D h(\hat{f}_T^*)
\]

\[
\Sigma_T \equiv \sum_{j=-k}^{k} \left( \frac{k - |j|}{k} \right) T^{-1} \sum_{t=1}^{T} \hat{e}_t \hat{e}'_{t-j}
\]

\[
\hat{e}_t \equiv f_t^* - T^{-1} \sum_{s=1}^{T} (f_s^*)
\]

That is, we simply adjust the Duffie and Singleton (1993) estimator by applying the delta method to the spectral density. To see how this works, note that the
moment error can be approximated by

$$G_T(\theta_0) = S^{-1} \sum_s h(\tilde{\tilde{f}}^s_T(\theta_0)) - \tilde{\tilde{f}}^s_T$$

$$\approx S^{-1} \sum_s [Dh(\tilde{\tilde{f}}^s_T(\theta_0))][\tilde{\tilde{f}}^s_T(\theta_0) - \tilde{\tilde{f}}^s_T]$$

$$= Dh(\tilde{\tilde{f}}^*_T) \left[ S^{-1} \sum_s \tilde{\tilde{f}}^s_T(\theta_0) - \tilde{\tilde{f}}^*_T \right]$$

Thus we just need to adjust the traditional Newey-West estimator by $Dh(\tilde{\tilde{f}}^*_T)$.

**Numerical Details**

To calculate the moment errors, I simulate the model 1000 times ($S = 1000$). The initial state is set to be the median state for a long simulation close to the estimated parameter values. Derivatives are computed with a two-sided finite difference.

I optimize using Levenberg-Marquardt (LM). I choose this method rather than the more commonly used simulated annealing method for two reasons. The first is that the moment function does not display an extreme number of local minima, which is where simulated annealing has an advantage. With a relatively smooth objective function, a method which uses derivative information is much more efficient. Derivative information is particularly helpful in the GMM setting with small residuals, since the Gauss-Newton method provides a quick positive semi definite approximation of the Hessian. Speed is important as the estimation takes around 24 hours with a very good guess. Another advantage of LM is robustness. Simulated annealing tends to be very sensitive to the choice of the annealing schedule (Press, Teukolsky, Vetterling, and Flannery (1992)).

I begin with an LM parameter of 1000, and use a simple algorithm for adjusting the parameter: if the new function value is a good enough improvement, I
decrease the LM parameter by a factor of 10. Otherwise, I set it back to 1000. “Good enough” is judged by the difference between the improvement in the objective and the predicted improvement according to a quadratic model.

Since I use exact identification, the maximum moment error provides a clean convergence criterion. I consider the algorithm converged if the maximum moment error is less than 1.0E-4.

1.8.4 The Interpretation of Habit

Here I describe how deviating from Campbell and Cochrane (1999)’s specification of $\lambda$ raises some questions regarding the interpretation of habit in the model. The key issue is that the constant $\lambda$ of my model can make habit decrease in response to an increase in consumption. This violates some traditional notions of habit.

This issue can be illustrated by taking the derivative of log habit $h_{t+1}$ with respect to log consumption $c_{t+1}$:

$$\frac{dh_{t+1}}{dc_{t+1}} = 1 - \frac{\lambda}{S_{t+1}^{-1} - 1}$$

Thus if $S_{t+1}$ is large enough, $\frac{dh_{t+1}}{dc_{t+1}}$ will be negative.

The preferences of this paper still preserve the standard notion of habit in that $H_t$ is a geometric average of previous consumption. This can be seen by following the analysis of Campbell and Cochrane (1999) found in Campbell (2003). I can log-linearize the log surplus consumption ratio around the steady state:

$$s_t = \log[1 - \exp(h_t - c_t)] \approx \kappa - \lambda^{-1}(h_t - c_t) \quad (1.39)$$
Plugging this into the definition of the habit process (1.5), I find the link between habit and historical consumption

\[ h_{t+1} \approx (\text{Constants}) + \rho_s h_t + (1 - \rho_s) c_t \]  
\[ = (\text{Constants}) + (1 - \rho_s) \sum_{j=0}^{\infty} \rho_j^s c_{t-j} \]  
(1.40)  
(1.41)

This informal demonstration is verified by simulated data. In the simulated data, habit is highly correlated with consumption. The contemporaneous correlation is 0.978 and the correlation with lagged consumption is 0.983. Habit growth and consumption growth are moderately correlated. The correlation between \( \Delta h_t \) and \( \Delta c_t \) is 0.411.

That habit should move non-negatively with consumption everywhere is not required if one entertains a very slow-moving, historical average of consumption as responsible for our current reference point for consumption. Moreover, Campbell and Cochrane (1999)’s specification is also vulnerable to this this issue. Ljungqvist and Uhlig (2009) show that while habit moves positively with small movements in consumption, it can move negatively with large movements.

The issue illustrated in this section is related to Campbell and Cochrane (1999)’s three requirements on \( \lambda(s_t) \). They require (i) the risk-free rate is constant, (ii) habit is predetermined at the steady state surplus consumption, and (iii) habit is predetermined near the steady state. The first assumption is not critical for making habit move non-negatively with consumption. In my model, (ii) is satisfied, but (iii) is not. (iii), in combination with Campbell and Cochrane (1999)’s specification for \( \lambda(s_t) \) results in \( \frac{dh_{t+1}}{dc_{t+1}} \geq 0 \) for all \( s_t \).
1.8.5 Solution Method Details

**Euler Equation** To be explicit about the firm's Euler equation, let \( \pi_Z(Z_i, Z_j) \) be the transition matrix for the discretized productivity process. The firm's problem is to find capital policy \( K' = G(K; \hat{K}, S, Z_i) \) to solve

\[
V(K; \hat{K}, S, Z_i) = \max_{K', I, N} \left\{ \Pi(K, Z_i, N) - W(\hat{K}, S, Z_i)N - \Phi(I, K) - I \\
+ \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) V(K'; \hat{K}', S', Z_j) \right\}
\]

subject to

\[
K' = I + (1 - \delta)K
\]

The FOC for investment and the envelope condition are:

\[
1 + D_1 \Phi(I, K) = \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) D_1 V(K'; \hat{K}', S', Z_j)
\]

\[
D_1 V(K; \hat{K}, S, Z_i) = D_1 \Pi(K, N, Z_i) + (1 + D_1 \Phi(I, K))(1 - \delta) - D_2 \Phi(I, K)
\]

which together produce the Euler equation

\[
1 + D_1 \Phi(I, K) = \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j)[D_1 \Pi(K', Z_j, N') + (1 + D_1 \Phi(I, K'))(1 - \delta) - D_2 \Phi(I, K')]
\]

Impose the fact that the household does not value leisure and consistency,
and we have

\[
1 = \sum_{Z_j} \pi_{Z_i}(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) \hat{R}_t(\hat{K}, S, Z_i; Z_j) \tag{1.42}
\]

\[
\hat{R}_t(\hat{K}, S, Z_i; Z_j) \equiv \frac{D_1 \Pi(\hat{K}', Z_j, 1) + (1 + D_1 \Phi(\hat{I}_j, \hat{K}'))(1 - \delta) - D_2 \Phi(\hat{I}_j, \hat{K}')}{1 + D_1 \Phi(\hat{I}, \hat{K})}
\]

Where

\[
M(\hat{K}, S, Z_i, Z_j) = \beta \left( \frac{C_j S_j}{C S} \right)^{-\gamma} \tag{1.43}
\]

\[
C_j = \Pi(\hat{K}', Z_j, 1) - \Phi(\hat{I}_j, \hat{K}') - \hat{I}_j
\]

\[
C = \Pi(\hat{K}, Z_i, 1) - \Phi(\hat{I}, \hat{K}) - \hat{I}
\]

\[
\hat{K}' = \hat{G}(\hat{K}, S, Z_i)
\]

\[
\hat{I} = \hat{G}(\hat{K}, S, Z_i) - (1 - \delta) \hat{K}
\]

\[
\hat{I}_j = \hat{G}(\hat{K}', S', Z_j) - (1 - \delta) \hat{K}'
\]

and the evolution of surplus consumption satisfies

\[
s_j = (1 - \rho_s) \bar{s} + \rho_s s + \lambda (c_j - c) \tag{1.44}
\]

The projection algorithm looks for cubic spline coefficients which solve equations (1.42), (1.43), and (1.44).

**Solving for asset prices** To find asset prices, the firm's Bellman equation, with optimal values plugged in, is:

\[
V(K; \hat{K}, S, Z_i) = \Pi(K, Z_i, 1) - W(\hat{K}, S, Z_i) - \Phi(I, K) - I
\]

\[
+ \sum_{Z_j} \pi_{Z_i}(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) V(K'; \hat{K}', S', Z_j)
\]
In equilibrium, \( W(\hat{K}, S, Z_i) = (1 - \alpha)AZ_i\hat{K}^\alpha, K = \hat{K}, I = \hat{I}, \text{and } K' = \hat{K}' \), so

\[
V(\hat{K}; \hat{K}, S, Z_i) = \alpha AZ_i\hat{K}^\alpha - \Phi(\hat{I}, \hat{K}) - \hat{I} + \sum_{Z_j} \pi_Z(Z_i, Z_j) M(\hat{K}, S, Z_i; Z_j) V(\hat{K}', \hat{K}', S', Z_j)
\]

Let’s define \( \hat{V}(\hat{K}, S, Z_i) \equiv V(\hat{K}; \hat{K}, S, Z_i) \). The above equation suggests that \( \hat{V}(\hat{K}, S, Z_i) \) can be found by repeatedly applying the above equation as an operator (using the law of motion for capital \( \hat{G}(\hat{K}, S, Z_i) \)).

**Approximations**

I approximate the autoregressive process for productivity \( z_t \) with a 13 point Markov Chain using the Rouwenhorst method. I approximate the law of motion for capital in the \( K \) and \( S \) directions using a two-dimensional cubic spline. The spline is of 6th degree in the \( K \) direction and 14th degree in the \( S \) direction. The spline breakpoints are log-spaced in both the \( K \) and \( S \) directions. I find that increasing the degree to the 14th in the \( \hat{K} \) direction has no material impact on the quantitative results.

In a projection method, one must define what it means to satisfy the Euler equation. I use the collocation method, which specifies that the Euler equation should hold exactly at a set of points (collocation nodes) in the \( K \) and \( S \) domain. I choose these nodes to be the standard nodes for splines using knot averaging. I consider the algorithm converged if, across these collocation nodes, the maximum absolute Euler equation residual, expressed as \( 1 - E[M'R'_i] \), is less than 1.0E-8. I search for spline coefficients which satisfy this condition by using Broyden’s method.
1.8.6 Adjustment Costs and Accounting

Many papers specify adjustment costs in the following manner:

\[ K' = (1 - \delta)K + \phi(I/K)K \]

(e.g. Jermann (1998), Gourio (2009), Kaltenbrunner and Lochstoer (2010), Guvenen (2009), among others). This formulation preserves the traditional Cobb-Douglas formulation of output:

\[ Y = ZK^\alpha N^{1-\alpha} = C + I \]

This formulation, however, deviates from the standard accounting treatment of investment and capital. The standard treatment specifies that end-of-period capital is beginning-of-period capital plus investment less depreciation. With geometric depreciation, this translates into the standard, adjustment-cost-free formulation of capital accumulation:

\[ K' = (1 - \delta)K + I \]

I choose to preserve this accounting identity. As a result, capital adjustment costs are pushed into output:

\[ Y = ZK^\alpha N^{1-\alpha} - [\text{Adj Cost}] = C + I \]

Fortunately, both choices result in the same capital and consumption alloca-
tions. For example, Gourio (2009) uses

\[ K^{*\prime} = (1 - \delta)K + \phi^* (I^* / K)K \]
\[ \phi^* (x) = x - \frac{\eta}{2} (x - \delta)^2 \]

From these expressions, we have capital evolution and consumption

\[ K^{*\prime} = (1 - \delta)K + I^* - \frac{\eta}{2} \left( \frac{I^*}{K} - \delta \right)^2 K \]
\[ C = Z K^\alpha N^{1-\alpha} - I^* \]

which is identical to my formulation with

\[ I \equiv I^* - \frac{\eta}{2} \left( \frac{I^*}{K} - \delta \right)^2 K \]

1.8.7 Characterization of Equilibrium

This section describes the equilibrium laws of motion.

Figure 1.9 shows the distribution of the model’s state variables. The model has three aggregate states, capital, surplus consumption, and productivity, and this makes describing the distribution more complicated than in other asset pricing models. The univariate descriptions are fairly straightforward. Panel A shows the univariate distribution of each state variable. These distributions are constructed by computing histograms from simulated data. Each state variable is roughly log-normal on its own, and the productivity distribution displays the discreteness of the Markov chain approximation. The state variables interact, however. Panel B shows bivariate distributions as contour plots. This panel shows that all of the state variables are positively correlated. High capital, productivity,
and surplus consumption all represent good times. This high correlation is particularly pronounced for surplus consumption and productivity, which can be seen in the lower right corner. Surplus consumption and productivity move together very closely. This equilibrium relationship is important for understanding the model's laws of motion which follow.

**Figure 1.9: Distribution of State Variables.** Panel A shows simple histograms from simulated data. Panel B shows 2D histograms as contour plots.

Figure 1.10 shows the law of motion for consumption. It shows four three-dimensional plots to provide a complete portrait of this function of three variables. The top two panels hold productivity fixed and show consumption as a function surplus consumption and capital. The bottom two panels hold capital fixed and show consumption as a function surplus consumption and productivity. Consumption is increasing and concave in capital and productivity, which is consistent with the precautionary volatility mechanism from Section 1.2. Consumption is decreasing and slightly convex in surplus consumption. This shape seems inconsistent with the mechanism, however, surplus consumption and productivity are very highly correlated (Figure 1.9). As a result, the do-
Figure 1.10: Law of Motion for Consumption.

main of the bottom panels which is explored in equilibrium runs from the bottom left to the upper right corner of the plots. Along this selection of the plot, consumption is strictly concave. Overall, Figure 1.10 shows that consumption is more sensitive to shocks when the household is poor.

Concave consumption results in countercyclical volatility, which is shown in Figure 1.11. This figure shows the law of motion for consumption volatility as a function of surplus consumption and productivity for two different levels of capital. I choose to focus on the state variables surplus consumption and productivity because most of the variation in consumption volatility occurs along these dimensions. Consumption volatility is high in bad times: either when surplus consumption is low or when productivity is low. This relationship is a natural result of precautionary savings dynamics. In bad times, future consumption and savings policies are particularly uncertain and this uncertainty leads to high volatility.

This countercyclical volatility is reflected in the Hansen-Jagannathan bound.
Figure 1.11: Law of Motion for Consumption Growth Volatility. Consumption growth volatility is the volatility of log first-differenced consumption.

Figure 1.12 plots this bound as a function of surplus consumption and productivity for various levels of capital. The maximum Sharpe ratio declines in both surplus consumption and productivity. In bad times, when surplus consumption or productivity are low, consumption is especially volatile, making marginal utility especially volatile. This behavior results in investors demanding countercyclical returns for assets which are related to this countercyclical risk.
Figure 1.12: Law of Motion for the Hansen Jaganathan Bound.
Chapter 2

Habit, Production, and the Cross Section of Stock Returns

2.1 Introduction

In the decades since the publication of Mehra and Prescott (1985)'s equity premium puzzle, economics and finance have produced a handful of models which provide a quantitative description of aggregate asset prices. Whether these solutions also help us understand the cross-section of stock returns is a relatively unexplored question.

Exploring this question is important because what we ultimately want is not a solution to the equity premium puzzle, but a framework for understanding asset prices in general. A model which explains the return on the CRSP index but cannot address any other asset leaves much to be desired.

In this paper, I show that one solution to the equity premium puzzle, habit formation, provides insight into the cross-section of stock returns. I find that a real business cycle with external habit formation and idiosyncratic firm produc-
tivity generates a data-like value premium. In the model and the data, expected returns are linear in log (B/M). Moreover, the model quantitatively captures the slope of the relationship. In model and data, the slope on log (B/M) is about 5, indicating that a 20% higher B/M implies a 100 b.p. increase in expected returns. These cross-sectional predictions come with good predictions about many equity premium and business cycle facts.

The value premium is due to the temporal distribution of cash flows. In the model, value firms have low productivity. Mean reversion in productivity leads to high productivity growth, and thus high cash flow growth. High cash flow growth means that value firms’ cash flows are distributed late into the future, and are more exposed to the discount rate shocks which characterize the habit model. The value premium is compensation for this higher exposure.

The mechanism implies that value firms have high cash flow growth. This prediction is contrary to the conventional view that value firms have short durations (Dechow, Sloan, and Soliman (2004), Da (2009)). Measuring the duration of an equity claim, however, is a nontrivial task. Measuring duration requires the estimation of a discount rate and a terminal cash flow growth rate, both of which are very difficult to identify. On the other hand, measuring cash flow growth is more straightforward. As recently emphasized by Chen (2012b), by a number of measures, value firms have higher cash flow growth (see also Lakonishok, Shleifer, and Vishny (1994), Bansal, Dittmar, and Lundblad (2005b), Hansen, Heaton, and Li (2008), Chen, Petkova, and Zhang (2008)). In this paper, I present additional empirical evidence supporting the mechanism of the model. While the existing papers focus on dividends or earnings before extraordinary income, I show that value firms’ also have higher cash flow growth when one includes the extraordinary items, adds back depreciation, or subtracts net
Both production and general equilibrium are critical to the model’s results. The importance of production can be seen by comparing the model’s results to that of Santos and Veronesi (2010). Santos and Veronesi study external habit in an endowment economy, and they come to the opposite conclusion of this paper, that is, they find that external habit in general leads to a value discount. Because their model lacks production, they cannot identify value firms in their model with B/M, and instead characterize ‘value’ with D/P. This definition means that value firms are high dividend firms, and mean reversion implies low cash flow growth and low exposure to discount rate shocks. In contrast, a model with production characterizes value with B/M, which is the most powerful value-oriented firm variable (Fama and French (1992)). Value firms are then low productivity firms, and mean reversion implies high cash flow growth and high exposure to discount rate shocks.

To illustrate the importance of general equilibrium, I show that in a partial equilibrium version of the model, the value premium disappears. The partial equilibrium model is a firm’s problem where I use the stochastic discount factor from the general equilibrium model, but reduce the adjustment costs to a value similar to partial equilibrium estimates (i.e. Whited (1992)). I call this model ‘partial equilibrium’ because markets will not clear and consumption volatility is not well defined. In this model, the dispersion in cash flow growth between value and growth firms shrinks dramatically. Intuitively, value firms do not want to have high cash flow growth since it results in high discount rates and reduces their firm value. The low adjustment costs of the partial equilibrium model allow value firms to disinvest aggressively and reduce their cash flow growth. This contrasts with general equilibrium, where the low EIS of habit preferences and
calibration to consumption data pins down high adjustment costs and leads to a value premium.

The paper proceeds as follows. In the remainder of this section I discuss related literature. Section 2.2 presents empirical evidence in support of the model’s mechanism. Section 2.3 presents the model. Section 2.4 calibrates the model to post-war U.S. data. Section 2.5 shows quantitative results regarding both equity premium puzzles and the cross section of returns. Section 2.6 inspects the mechanism for generating the value premium. Section 2.7 concludes.

**Related Literature**  This paper is closely related to Chen (2012a). That paper shows that the external habit formation of Campbell and Cochrane (1999) can be extended into a production economy. In a production economy, external habit leads to countercyclical consumption volatility which endogenizes the habit sensitivity function of Campbell and Cochrane (1999). This paper extends Chen (2012a) to include heterogenous firms and examines the cross-section of stock returns.

Regarding the literature which addresses both the equity premium and the cross-section of returns, the most closely related paper is Santos and Veronesi (2010). Santos and Veronesi (2010) find that, in an endowment economy, the external habit framework tends to imply a value discount, rather than a value premium. This paper differs primarily by considering a production economy. Production is critical, and indeed, the introduction of production completely reverses their conclusion.

More broadly, the literature on both the equity premium and the cross-section is rather limited. This is likely due to the fact that the curse of dimensionality looms large when one considers both aggregate uncertainty and a nontriv-
ial cross-section of firms. To my knowledge, the only other papers which have ventured into this territory are Favilukis and Lin (2011) and Ai and Kiku (2012). Both find that the long-run risk framework of Bansal and Yaron (2004) is consistent with the value premium, although they rely on the use of alternative technologies such as operating leverage. Gabaix (2012) shows that the time-varying disaster framework is qualitatively consistent with the characteristics vs. covariances puzzle (Daniel and Titman (1997)) regarding the value premium.

Regarding the partial equilibrium literature on the cross-section, a closely related paper is Zhang (2005). Zhang (2005) finds that a q-theoretical model with a countercyclical price of risk naturally produces a large value premium. His SDF is consistent with empirical features of aggregate asset returns, but is exogenously specified. As this model also features a countercyclical price of risk, one can consider this paper a general equilibrium foundation for Zhang (2005).

2.2 Empirical Evidence on Value and Cash Flow Growth

The model’s mechanism relies on value firms having higher cash flow growth than growth firms. In this section, I discuss the existing empirical evidence regarding the mechanism, and present some new evidence in support of the mechanism.

The conventional view is that the mechanism is counterfactual, that value firms have low cash flow growth. Evidence for this view comes exclusively from equity duration studies, namely Dechow, Sloan, and Soliman (2004) and Da (2009). Both papers find that value firms have short durations, and thus, low cash flow growth.
Equity duration, however, is very difficult to measure, or even to define. Generally speaking, the duration of an asset is something like

$$\text{Duration} = \sum_{t=0}^{\infty} \frac{PV(CF_t)}{P_0} t$$

Note that measuring duration requires both a discount rate, and a terminal value. While these two are both directly observable for bonds, they are extremely difficult to measure for equities.

Dechow, Sloan, and Soliman (2004) and Da (2009) rely on identifying assumptions which bias them toward finding that value firms have low durations. In particular, Dechow, Sloan, and Soliman (2004) assume that the terminal value is equal to the market value, less some present value of the next 10 years of cash flow. Since value firms have low market value compared to current cash flows, this leads to a low terminal value and, thus, a low duration. Da (2009) assumes that the terminal ROE is equal to the mean of ROE for the first seven years after portfolio formation. Since value firms have low ROE at portfolio formation (Fama and French (1995)), this assumption leads to a low terminal value and, thus, low duration. These biases are pointed out by Chen (2012b).

Measuring cash flow growth is much more straightforward. Indeed, a number of papers focused on other issues happen to provide summary statistics regarding value and cash flow growth (i.e. Lakonishok, Shleifer, and Vishny (1994), Bansal, Dittmar, and Lundblad (2005b), Hansen, Heaton, and Li (2008)). These papers uniformly find that value portfolios have higher cash flow growth.

Unfortunately, even cash flow growth offers multiple methods of measurement. For example, the previously mentioned papers use portfolios which are rebalanced periodically using various methods. Chen (2012b) provides a detailed
examination of cash flow growth and value in buy-and-hold portfolios. For the majority of his methods of analysis, he finds that value firms have high cash flow growth.

This section presents additional empirical evidence regarding value and cash flow growth. In particular, Chen (2012b) focuses on dividends and earnings before extraordinary income. Other definitions of cash flow are arguably more appropriate empirical targets for models. Existing models, including the one in this paper, abstract from dividend policy (Miller and Modigliani (1961)). In most models, dividends are equal to net income plus depreciation less net investment (cash flow from operating and investing activities). Both sides of the equation can be considered as cash flow, and the RHS is arguably a more appropriate empirical target since it captures the real (non-financial) activities of the firm.

I look at four different notions of cash flow. The first is earnings before extraordinary income (ib). The second is earnings (ni). Extraordinary income affects the cash flows of value investors, and it is not obvious why one should exclude this effect. The third measure I examine is earnings plus depreciation (ni + dp). This measure is close to cash flow from operations. The last measure I examine is earnings plus depreciation less net investment (ni + dp - capx + sppe). This measure is closest in spirit to the cash flows of the model.

I use tercile book-to-market sorted portfolios and CRSP and COMPUSTAT data from 1971-2011. The portfolios are buy-and-hold portfolios. Delisted stocks are reinvested in the remaining stocks, following Chen (2012b)’s procedure. The choice of terciles is due to the use of net investment. Net investment is quite volatile, and the use of large portfolios averages out much of this volatility and paints a clearer picture of the typical cash flow dynamics. Additionally, the choice of terciles significantly simplifies the presentation of the results. The 1971 begin-
ning date is limited due to the availability of sales of plant, property, and equipment (sppe). The relatively modern sample period is useful, however, in that one of the only measurements where Chen (2012b) finds that value firms have low cash flow growth is when he looks at dividend growth for buy-and-hold portfolios in the post-1963 sample. I will show that, during a similar sample period, many other definitions of cash flow present the opposite picture for buy-and-hold portfolios.

Table 2.1 shows cash flow levels. It shows the cash flow from a $1 investment in value or growth portfolios, averaged across portfolio formation years. The first thing that jumps out from the table is that value stocks do not have significantly lower cash flows than growth stocks. In the 1st year after portfolio formation, value stocks pay 7 cents per dollar invested while growth stocks pay 6 cents, with respect to earnings before extraordinary income. In fact, net of extraordinary income and investment, value firms pay much less. Using this definition, in the 1st year value pays half a cent per dollar while growth pays an order of magnitude more. The second pattern which emerges from the table is that value has higher cash flow growth. There is little action in the cash flows of growth firms, but the value cash flows exhibit apparent growth.

Table 2.2 shows growth rates of the cash flows from the previous figure. It also considers two additional definitions of cash flow: earnings (after extraordinary income) and earnings plus depreciation. By all definitions of cash flow, value portfolios have much higher cash flow growth than growth portfolios in year 2. Indeed, using the definition closest in spirit to the model (earnings plus depreciation less net investment), value experiences a huge 186% growth in cash flow between year 2 and year 1, while growth gets a meager 5% growth. Cash flow growth is also monotonically increasing in B/M using all definitions. Cash
flow growth of value exceeds that of neutral which exceeds that of growth. An additional pattern which is seen in Table 2.2 is that the growth rates mean revert. Value begins with strikingly high cash flow growth in year 2, but growth slows down quickly. Growth portfolios generally follows the opposite pattern. Both the high cash flow growth of value portfolios and its subsequent mean reversion will be seen in the model.

Analyzing cash flow growth for the longer term faces data limitations. 40 years of data provides only 10 non-overlapping 4-year periods. Thus, it is probably best to focus on the year 2 and year 3 growth rates. Nevertheless, the fact that the cash flow patterns are common across multiple definitions of cash flow is reassuring.

2.3 A General Equilibrium Model with Heterogeneous Firms

Having established that the mechanism is consistent with the data, I now present the model. The model is basically an rbc model with external habit formation, capital adjustment costs, and idiosyncratic firm productivity. It is designed to have the minimal features for both an equity premium and an endogenous cross-section of firms.

Markets are complete, time is discrete and infinite. For the remainder of the paper, lowercase denotes logs, i.e. $c_t \equiv \log C_t$. 

90
2.3.1 Representative Household

A unit measure of identical households \( j \in [0,1] \) chooses asset holdings to maximize lifetime utility

\[
\mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{(C_{j,t} - H_t)^{1-\gamma}}{1-\gamma} \right\}
\]

(2.1)

where \( H_t \), the level of habit, is determined by aggregate consumption, and so is taken as external by the household.

I specify the evolution of habit using surplus consumption, rather than the level of habit itself. That is, let

\[
S_t \equiv \frac{C_t - H_t}{C_t}
\]

(2.2)

be the surplus consumption ratio, where \( C_t \) is aggregate consumption. Then surplus consumption follows an AR1-process in logs

\[
s_{t+1} \equiv (1 - \rho_s) \tilde{s} + \rho_s s_t + \lambda (c_{t+1} - c_t)
\]

(2.3)

This approach is chosen for comparability with the existing literature on external habit (Campbell and Cochrane (1999), Wachter (2006), among others).

Note that the conditional volatility of the surplus consumption process (2.3) \( \lambda \) is a constant. This is a substantial deviation from the literature on external habit models in endowment economies, which specify this conditional volatility as time-varying and countercyclical (e.g. Campbell and Cochrane (1999), Menzly, Santos, and Veronesi (2004)). In a previous paper, I show that the introduction of production results in countercyclical consumption volatility, which is quantitatively very similar to the assumed countercyclical volatility of surplus
consumption typical of endowment economy models. For comparability with Campbell and Cochrane (1999), I fix $\lambda$ at their steady state value

$$\lambda = \frac{1}{S} - 1 \quad (2.4)$$

The fact that markets are complete and habit is external means that the household's side of the model boils down to a simple SDF

$$M_{t,t+1} = \beta \left( \frac{C_{t+1}}{C_t} \frac{S_{t+1}}{S_t} \right)^{-\gamma} \quad (2.5)$$

### 2.3.2 Heterogeneous Firms

There is a unit measure of heterogeneous firms, indexed by $i \in [0,1]$. The firms produce consumption using capital $K_{i,t}$:

$$\Pi(K_{i,t}, B_{i,t}, A_t) = A_t B_{i,t} K_{i,t}^\alpha \quad (2.6)$$

where aggregate productivity $A_t$ and idiosyncratic productivity $B_{i,t}$ are both AR1 processes in logs:

$$a_{t+1} = \rho_a a_t + \sigma_a \epsilon_{a,t+1} \quad (2.7)$$

$$b_{i,t+1} = \rho_b b_{i,t} + \sigma_b \epsilon_{b,i,t+1} \quad (2.8)$$

where $\epsilon_{a,t+1}$ and $\epsilon_{b,i,t+1}$ are both standard normal random variables which are uncorrelated across $t$ and $i$.

All heterogeneity in the models originates from the idiosyncratic productivity process (2.8). This approach is used for three reasons. The first is that it is a very simple way of introducing a cross section of firms. The second is that a large lit-
erature documents substantial heterogeneity in productivity (Syverson (2011)). The third is that this approach is the standard way of modeling firm heterogeneity in both macroeconomics and finance (Hennessy and Whited (2005), Zhang (2005), Khan and Thomas (2008), Bloom (2009)). We will see, however, that this approach has difficulties matching the tremendous heterogeneity which is seen in the data. Matching the heterogeneity in the data with additional sources of heterogeneity is an interesting question for future research, however, it is beyond the scope of this paper.

Capital accumulates according to the usual capital accumulation rule

\[ K_{i,t+1} = I_{i,t} + (1 - \delta)K_{i,t} \]  

(2.9)

and firms face a convex capital adjustment cost

\[ \Phi(I_{i,t}, K_{i,t}) = \frac{\phi}{2} \left( \frac{I_{i,t}}{K_{i,t}} - \delta \right)^2 K_{i,t} \]  

(2.10)

I assume that the adjustment costs are a pure loss. They do not represent payments to labor. Adjustment costs are included because production economies produce a counterfactually smooth Tobin’s Q unless one includes an investment friction. Quadratic costs are chosen for simplicity, but a richer model would incorporate investment frictions by modeling an investment goods sector, as in Boldrin, Christiano, and Fisher (2001), or would feature heterogeneous plants with non-convex costs of adjustment, as in Khan and Thomas (2008).

Because of complete markets, the firm’s objective is standard

\[ \max_{\{I_{i,t}, K_{i,t+1}\}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} M_{0,t} [A_{t}B_{i,t}K_{i,t} - I_{i,t} - \Phi(I_{i,t}, K_{i,t})] \right\} \]  

(2.11)
It chooses investment and capital to maximize future dividends, discounted with the household’s SDF.

2.3.3 Recursive Competitive Equilibrium

Equilibrium is defined recursively. Let $\mu$ represent the distribution of firms over capital $K_i$ and idiosyncratic productivity $B_i$. The aggregate state is the triple of the distribution of firms $\mu$, surplus consumption $S$, and aggregate productivity $A$.

The recursive competitive equilibrium is laws of motion for the distribution of firms $\Gamma(\mu, S, A)$ and aggregate consumption $C(\mu, S, A)$, a capital policy for the firm $G(K_i, B_i; \mu, S, A)$ and value function for the firm $V(K_i, B_i; \mu, S, A)$ such that

1. Firm optimality holds: $G(K_i, B_i; \mu, S, A)$ and $V(K_i, B_i; \mu, S, A)$ solve

$$V(K_i, B_i; \mu, S, A) = \max_{(K_i)'} \left\{ \Pi(K_i, A, B_i) - \Phi(I, K_i) ight. \\
+ \int_{-\infty}^{\infty} dF(\epsilon'_a) \int_{-\infty}^{\infty} dF(\epsilon'_b) M(A'; \mu, S, A) V(K_i', B_i'; \mu', S', A') \right\}$$

(2.12)

where the productivity processes are given by (2.7) and (2.8), capital accumulation is given by (2.9), the SDF is the household’s intertemporal marginal rate of substitution

$$M(A'; \mu, S, A) = \beta \left( \frac{C(\mu', S', A') S'}{C(\mu, S, A) S} \right)^{-\gamma}$$

(2.13)

$S'$ evolves according to (2.3), $\mu'$ is given by $\Gamma(\mu, S, A)$, and $F(\epsilon'_a)$ is the standard normal CDF.
2. Firm decisions are consistent with the law of motion for consumption:

\[ C(\mu, S, A) = \int d\mu(K_i, B_i) \left\{ \Pi(K_i, B_i, Z) - \Phi(I(K_i, B_i; \mu, S, A), K_i) \right\} \] (2.14)

where \( I(K_i, B_i; \mu, S, A) = G(K_i, B_i; \mu, S, A) - (1 - \delta)K_i \).

3. Firm decisions are consistent with the law of motion for the distribution of firms, that is, let \( B \) be the Borel algebra for \( \mathbb{R}^2_+ \). Then \( \mu' = \Gamma(\mu, S, A) \) is given by

\[
\forall (K_1, B_1) \in B,
\mu'(K_1, B_1) = \int_{\{(K_i, B_i) \text{ s.t. } G(K_i, B_i; \mu, S, A) \in K_1\}} d\mu(K_i, B_i) \int_{\{\epsilon'_b \text{ s.t. } \exp(\rho b + \sigma_b \epsilon'_b) \in B_1\}} dF(\epsilon'_b)
\] (2.15)

### 2.3.4 Krusell-Smith Solution Method

I solve the model with a variant of the Krusell and Smith (1998) algorithm, similar to Khan and Thomas (2008). As in Khan and Thomas (2008), I approximate the distribution of firms \( \mu \) with the aggregate capital stock \( K \). Thus, the approximate aggregate state is a triple of aggregate capital, surplus consumption, and aggregate productivity: \((K, S, A)\)

I discretize the aggregate and idiosyncratic productivity processes (2.7) and (2.8) using the Rouwenhorst (1995) method. I then conjecture that the laws of motion for aggregate consumption and capital follow the following log-linear
forms:

\[ c' = \log \tilde{C}(K, S, A_j) = \theta_{0,j}^C + \theta_{1,j}^C k + \theta_{2,j}^C \hat{s} \]

\[ k' = \log \tilde{\Gamma}(K, S, A_j) = \theta_{0,j}^\Gamma + \theta_{1,j}^\Gamma k + \theta_{2,j}^\Gamma \hat{s} \]

where \( j \in \{A_1, ..., A_{N_A}\} \) represents the aggregate productivity state. Note that the use of aggregate productivity dependent coefficients allows for a non-linear relationship between consumption and the aggregate state.

The goal of the Krusell-Smith method is then to find the coefficients to the laws of motion \( \theta_{i,j}^C, \theta_{i,j}^\Gamma \) such that

1. Firms maximize value given the laws of motion \( \theta_{i,j}^C, \theta_{i,j}^\Gamma \)

2. Estimates of (2.16) on simulated data using policies from step 1. produce coefficients which are close to \( \theta_{i,j}^C, \theta_{i,j}^\Gamma \), and produce \( R^2 \)'s which are close to 1.

The most straightforward application of Krusell-Smith searches for this approximate equilibrium by doing a fixed-point iteration using the firm's problem defined in (2.12) and simulation of a distribution of firms. However, there is no theorem which suggests that this fixed-point iteration will converge, and indeed it does not.

To aid in finding equilibrium, I apply the 'equilibrium-in-simulation' method (Krusell and Smith (1997)). That is, I first solve solve the approximate equilibrium version of (2.12). I then plug the resulting value function into the following
problem:

\[ G(K_i, B_i; K, S, A; C) = \arg\max_{\{I, K'\}} \left\{ \Pi(I, A, B_i) - \Phi(I, K_i) \right\} \]

\[ + \int_{-\infty}^{\infty} dF(\epsilon'_a) \int_{-\infty}^{\infty} dF(\epsilon'_b) M^*(A'; K, S, A; C) V(K'_i, B'_i; K', S', A') \right\} \]

\[ (2.17) \]

\[ M^*(A'; K, S, A; C) = \beta \left( \frac{\tilde{C}(K', S', A') S'}{CS} \right)^{-\gamma} \]

\[ (2.18) \]

That is, I introduce today's aggregate consumption as an additional state variable and solve for a new investment policy which accounts for aggregate consumption. I then use this augmented investment policy \( G(K_i, B_i; K, S, A; C) \) in the simulation step. This allows me to find a 'market-clearing' \( C \) at each date in the simulation. That is, at each date, I use a root finder to find the \( C \) that solves equation (2.14). Note that once the equilibrium is found, aggregate consumption from the simulation of the firms and that produced by the law of motion are equal, and so problem (2.17) with market clearing (2.14) and problem (2.12) produce identical choices.

The presence of external habit significantly complicates the computationally demanding Krusell-Smith algorithm. External habit preferences introduce an additional aggregate state variable, surplus consumption, which is completely absent from the standard RBC economy. As a result, using the RBC equilibrium as an initial guess for the Krusell-Smith algorithm will cause the algorithm to fail. To address this problem, I apply a homotopy method. I solve a series of models with the following altered SDF

\[ M' = \beta \left( \frac{C'}{C} \left( \frac{S'}{S} \right)^{\frac{1}{\gamma}} \right)^{-\gamma} \]

\[ (2.19) \]
I begin by solving the model with $\chi = 0$. Here the RBC model serves as a good initial guess. Once the program is fairly close to equilibrium, I increase $\chi$ by 0.100 and use the previous laws of motion as an initial guess for the new model. I repeat this process until $\chi = 1.000$, which is equivalent to the model presented in (2.3).

Surplus consumption also adds the difficulty that it is an endogenous state variable that is not pre-determined. As a result, the habit process equation (2.3) must be solved at every date in the simulation step of the algorithm. Note that the simulation step involves simulating an entire distribution of firms, and so an entire distribution of decision rules must be accounted for in solving equation (2.3). This significantly increases the computational burden of the algorithm.

### 2.4 Calibration to Post-War U.S. Data

The model is calibrated to post-war (post 1947) U.S. Data. This sample period is chosen because the World Wars introduce structural changes which may not be captured by the model. In particular, over the long sample (post 1929) HP-filtered output and investment are essentially uncorrelated.

Aggregate quantities are taken from the BEA. Firm-level data is taken from CRSP/COMPSTAT. B/M sorted portfolio data is taken from Ken French’s website. Aggregate asset price moments are taken from Beeler and Campbell (2009).

Table 2.3 shows the calibration. Preference parameters are chosen as much as possible to fit unconditional moments of asset prices. Time preference $\beta$ is chosen to fit the mean 30-day T-bill return, and the persistence of surplus consumption $\rho_s$ is chosen to fit the persistence of the 30-day T-bill return (using Beeler and Campbell (2009)’s ex-ante measure). In an external habit model, the steady state surplus consumption $\bar{S}$ and utility curvature $\gamma$ jointly control the
risk-aversion of the representative household, and thus are difficult to identify separately. For ease of comparison with the literature on external habit, I choose $\gamma = 2$ to match Campbell and Cochrane (1999)’s value, and then choose $\bar{S}$ to fit the mean Sharpe ratio of the CRSP index.

I choose ‘aggregate’ technology parameters to fit moments of the real economy. The production curvature $\alpha$ and depreciation rate $\delta$ are chosen to fit capital’s share of output and the mean, growth-adjusted investment rate. The volatility of aggregate productivity $\sigma_a$ is chosen to fit the volatility of HP-filtered log GDP (I use a smoothing parameter of 6.25, as argued by Ravn and Uhlig (2002)). The persistence of aggregate productivity is chosen to fit the persistence of the constant-labor Solow residual. A critical parameter of the production technology is the quadratic adjustment cost parameter $\phi$. I choose this parameter value to hit the volatility of aggregate consumption growth (non-durables and services).

The firm level technology parameters are chosen as follows. The persistence of idiosyncratic productivity $\rho_b$ is chosen to fit the persistence of plant-level productivity. The volatility of idiosyncratic productivity $\sigma_b$ is chosen to fit the mean volatility of the individual stock return.

### 2.5 Quantitative Results

Section 2.5.1 shows that the model addresses equity premium puzzles. This section mostly verifies the results of Chen (2012a) and so the discussion will be brief.

Section 2.5.2 contains the main quantitative results. It shows that the model generates a data-like value premium according to Fama-Macbeth regressions.
2.5.1 Matching the Data on Equity Premium Puzzles

Table 2.4 shows aggregate asset price moments. The model produces a large and volatile equity premium, and a low and smooth risk-free rate.

Table 2.5 shows basic business cycle moments. As intended by the calibration, the model produces a low consumption volatility. As in the data, investment is much more volatile than output and consumption is much less volatile. Also as in the data, consumption and investment co-move with GDP.

Table 2.6 shows regressions of future dividend growth and returns on the price-dividend ratio. Because the Campbell and Shiller (1988) identity links dividend growth, returns, and the price-dividend ratio, these regressions form a nice characterization of the drivers of asset price fluctuations (for example, Cochrane (2011) and Koijen and Van Nieuwerburgh (2010)). The table shows that, as in the data, the price dividend ratio has little predictive power for future dividend growth, and has strong predictive power for future excess returns.

Taken together, Tables 2.4, 2.5, and 2.6 show that the model does a good job of addressing the equity premium puzzles.

2.5.2 Matching the Data on the Cross-Section of Stock Returns

We have seen that the model is able to address the equity premium puzzles. External habit combined with production produces a large and volatile equity premium, a low and smooth risk-free rate, and asset price fluctuations which are linked to excess returns rather than the price-dividend ratio. This brings us to the main question of the paper. Is external habit consistent with the cross-section of stock returns?

Table 2.7 examines this question and shows the main result of the paper. It
shows regressions of next year’s returns on today’s log B/M. The regressions are
firm-level and follow the Fama and MacBeth (1973) method. In both the model
and data, the log(B/M)_{i,t} coefficient is positive and highly statistically signifi-
cant. Moreover, the slopes are large and similar in magnitude, with a value of
about 5. This slope means that in both the model and data, a 20% higher B/M
implies a roughly 100 b.p. higher expected return.

Though firm-level regressions provide the most statistical power and offer the
simplest quantitative description of the value premium, the literature often ex-
amines portfolio sorts. Table 2.8 shows summary statistics on decile B/M sorted
portfolios. The expected returns columns show that, in both the model and data,
expected returns are monotonically increasing in B/M. In the model, expected
returns culminate to an economically significant decile 10-1 return of about 2%
per year.

The decile value premium is smaller than the data, but it comes with a much
smaller dispersion in B/M. Recall that all heterogeneity in the model originates
from a simple AR1 idiosyncratic productivity process. For parsimony and to
maintain clarity of the mechanism, the model abstracts from other sources of
heterogeneity such as differences in financial frictions or life-cycle effects. As a
result, the model generates a dispersion in B/M which is significantly less than
the data. In the model, log B/M differs by 0.4 between the high and low deciles. In
the data, this differences is 1.4. This difference in spreads means that the portfo-
lio sorts of the model are not comparable to the data. Reproducing the enormous
B/M dispersion in the data is an interesting question, but is beyond the scope of
this paper.

A more effective way to illustrate the portfolio sort results is to interpret them
as a non-parametric regression (Cochrane (2011)). Figure 2.1 provides this inter-
pretation. It plots the average returns of the 10 B/M sorted portfolios against demeaned log B/M. This figure shows that the model captures the log-linear form seen in the data. The match is not only qualitative, but quantitative too. The slope of the relationship between expected returns and log B/M is similar in both model and data, consistent the results of the Fama-Macbeth regressions (Table 2.7).

2.6 Inspecting the Mechanism

Here, I explain how this value premium works. I begin with by showing “value” as a function of firm state variables. Figure 2.2 plots B/M and expected returns against the two firm state variables, idiosyncratic productivity and capital. The left panel shows that value firms are low productivity firms with high capital. Capital, however, is slow-moving. As a result, value is primarily characterized by low productivity. The right panel shows how value is connected to expected returns. Expected returns decline strongly in idiosyncratic productivity. Overall, we see that value firms are low productivity firms have high expected returns, consistent with the quantitative results of Section 2.5.

2.6.1 Value and Cash Flow Growth

The low productivity of value firms leads to cash flow growth. This relationship is illustrated in figure 2.3 which plots the cash flow growth of portfolios sorted on B/M. The darkest lines show value portfolios. Soon after portfolio formation, value portfolios have high dividend growth, but this growth slows down quickly and eventually reaches the average growth rate of zero (the model abstracts from balanced growth). Intuitively, value firms have low productivity,
but this low productivity is temporary. Mean reversion implies that productivity grows, and so cash flows grow. This contrasts with growth portfolios (the lightest lines) which show the reverse pattern. When growth firms are declared as growth, they have very high productivity. Mean reversion then means that their productivity will fall, so you have low productivity growth and low cash flow growth. Both this initial spread in cash flow growth and its subsequent mean reversion are consistent with the empirical evidence of Section 2.2.

The exact relationship between value and cash flow growth is, of course, the result of competitive equilibrium and optimal choices (Section 2.3.3), as well as the calibration (Section 2.4). Cash flow is the result of both production and investment. Low productivity affects both. Regarding production, low productivity leads to mean reversion and tends to imply high cash flow growth. Regarding investment, we have the opposite. Low productivity encourages disinvestment which tends to imply low cash flow growth. Which channel dominates depends on the parameter choices.

In fact, the model without habit produces the counterfactual prediction that value firms have low cash flow growth. This is seen in figure 2.4, which plots cash flows for B/M portfolios in a model without habit. As with the habit model, the adjustment costs in this model are chosen to match the volatility of consumption growth. The figure shows exactly the opposite pattern of figure 2.3, with growth firms (dotted line) beginning with high cash flow growth and eventually mean reverting to the steady state of no growth. Here, the investment channel dominates the production channel. Intuitively, removing habit increases the EIS of the household, and requires that the adjustment costs be lowered in order to fit the volatility of consumption. With low costs of adjusting capital, that is low costs of investing, the investment channel becomes dominant.
As discussed in the Section 2.2, the model’s prediction that value firms have high cash flow growth is inconsistent with the conventional wisdom from equity duration studies, but is consistent with direct measurement of cash flow growth.

### 2.6.2 Cash Flow Growth and Expected Returns

So far, I have shown that value is characterized by low productivity and high cash flow growth. To finish the story, I need to show that high cash flow growth leads to high expected returns. This last link is due to the large discount rate shocks which drive the external habit model. High cash flow growth means that cash flows are distributed far into the future, and thus are more exposed to large discount rate shocks. Investors then demand high returns in exchange for bearing this higher exposure. A number of previous papers discuss this link (Cornell (1999), Lettau and Wachter (2007), Santos and Veronesi (2010), Chen (2012b)), but here I provide a new sketch of the intuition. This sketch uses simple Investments 101 formulas and so steps outside of the GE model.

Consider a growing perpetuity

\[
P_0 = \frac{D_1}{\kappa_0 - g}
\]

where \(\kappa_0\) is the discount rate and \(g\) is the growth rate of cash flows. Now suppose that the discount rate gets hit by a completely unexpected shock \(\Delta\kappa\). The price next period is then

\[
P_1 = \frac{D_1(1 + g)}{\kappa_0 + \Delta\kappa - g}
\]
Taking a 1st-order Taylor approximation of the definition of the return gives us

\[
R_1 \equiv \frac{D_1 + P_1}{P_0} \approx (1 + \kappa_0) - \left( \frac{1 + g}{\kappa_0 - g} \right) \Delta \kappa
\]  
(2.22)

If discount rates suddenly go up, the stock price takes a hit, and so we have a negative sign on the second term. Notice also that the second term is increasing in the cash flow growth rate \( g \). Discount rate shocks hit high cash flow growth assets particularly hard. Intuitively, high cash flow growth means that most of the cash flows will occur in the distant future, and these distant cash flows are hit multiple times by a persistent shock to discount rates.

Informally,\(^1\) the law of one price implies that

\[
\mathbb{E}_0 [R_1 - R_f] \approx \left( \frac{1 + g}{\kappa_0 - g} \right) \text{Cov}_t (-\Delta \kappa, -M_1) \frac{\sigma_0(M_1)}{\mathbb{E}_0(M_1)}
\]  
(2.23)

Provided that the discount rate shock is positively correlated with the SDF, this higher exposure to discount rate shocks commands a risk premium. In the model, this correlation is indeed positive. Intuitively, a positive technology shock tomorrow implies that consumption will be high tomorrow and thus marginal utility tomorrow \((M_1)\) will be low. Since marginal utility is low tomorrow, the conditional volatility of the SDF is also low and the discount rate shock is low (negative).

Though this link between the conditional volatility of the SDF and marginal utility is a standard assumption of external habit models in endowment economies (Campbell and Cochrane (1999)), the link in this production economy comes through the precautionary savings dynamics. In a previous paper (Chen (2012a)),

\(^1\)Formally, there is no uncertainty at date 0 and so the expected return should be \(1/\mathbb{E}_0[M_1]\). A more formal illustration would involve shocks to the volatility of the SDF but precludes the use of simple formulas.
I study a representative firm version of the model of this paper and provide further discussion of these precautionary savings dynamics.

This theoretical link between cash flow growth and expected returns is very similar to that which comes from the standard duration formula

\[
\%\Delta P_t \approx -[\text{Duration}]\Delta \kappa
\]  

(2.24)

High cash flow growth firms have long durations and thus are more sensitive to discount rate shocks. However, I choose to focus on cash flow growth because it is much easier to define and measure.

### 2.6.3 Other Potential Mechanisms

The model does not have fixed operating costs, irreversible investment, nor asymmetric adjustment costs on capital. Thus, the mechanism is distinct from the operating leverage channel of Carlson, Fisher, and Giammarino (2005) and the inflexibility channel of Zhang (2005) (see also Gala (2010)).

There is still one important channel to exclude: the cyclicality of cash flows. It could be that value firms have higher returns due to the fact that their cash flows are more ‘procyclical.’ Figure 2.5 shows that this is not the case.

The left panels of Figure 2.5 plot the cash flows of value and growth firms against the two state variables which represent the business cycle in this model: surplus consumption and aggregate productivity. For both value and growth firms, cash flows decline in surplus consumption. Since high surplus consumption represents a good state, in this respect both value and growth firms have countercyclical cash flows. Regarding the magnitude of the countercyclicality, there is no apparent difference. On the other hand, in terms of aggregate pro-
ductivity, growth firms are clearly more procyclical. Value firm cash flows are generally invariant to aggregate productivity, but, growth firm cash flows clearly increase. The cyclicality of cash flows itself would then lead to a value discount, not a value premium.

The right panels of Figure 2.5 shows that this result is intuitive. These panels show net investment (investment net of depreciation) for value and growth firms. In bad times, that is, in states with low surplus consumption or low aggregate productivity, value firms are disinvesting. These are times when the household really values consumption, and since value firms are unproductive, it is efficient for the value firms to discard their capital and provide cash flows to the household. This behavior leads to countercyclical cash flows for value firms. On the other hand, growth firms are investing in bad states. The household wants consumption, but since growth firms are so productive, it is efficient for the firm to give the household less consumption so that it can invest for the future. This behavior leads to procyclical and riskier cash flows for growth firms.

Of course, the risk of holding a stock is not just the risk of its cash flow next period. Every cash flow into the infinite future affects the risk of the stock. Both the temporal distribution and the short-term cyclical of a firm's cash flows affect its risk and return. In net, the high cash flow growth of value outweighs the lower cyclicality of its cash flows.

### 2.6.4 The Role of General Equilibrium

The model is general equilibrium (GE), and GE has many important implications for the results. One important role is that it pins down difficult-to-observe investment frictions. These investment frictions have a significant effect on the model's cross-sectional asset pricing results.
To show this, I conduct a ‘partial equilibrium’ (PE) experiment. First I take the laws of motion for consumption and aggregate capital which come from the calibrated GE model (2.16). I then plug these laws of motion into the firm's problem (2.12) and solve for firm investment policies, but I change the adjustment costs for the firm's problem to be 1/20th of their calibrated value. These lower adjustment costs are in line with ‘partial equilibrium estimates’ which use a constant SDF (e.g. Whited (1992)). Lastly I simulate a panel of firms using these PE investment policies (updating aggregates using the GE laws of motion).

This procedure mimics that used in the large literature on partial equilibrium dynamic firm models (e.g. Zhang (2005), Carlson, Fisher, and Giammarino (2005), Hennessy and Whited (2005)). I am conjecturing an SDF and then solving for optimal firm behavior given this SDF, but I do not go on to check that the SDF is consistent with firm behavior. Note that in this example markets will not clear, that is, equation (2.14) does not hold. Indeed, consumption is not clearly defined since I can calculate consumption either from the conjectured law of motion or by aggregating in the panel simulation.

Table 2.9 shows that in this PE model, the value premium disappears. It shows Fama-Macbeth regressions of next year's returns on today's log B/M ratio. While the GE model matches the data quite nicely, in the GE model, the slope on log B/M becomes tiny and statistically insignificant.

Figure 2.6 explains why the value premium goes away. It shows the cash flow growth of value and growth firms, comparing the GE model to the PE model. In the GE model, there is a large spread in cash flow growth, but in PE, the spread is tiny. Intuitively, a value firm does not want to have high cash flow growth since it raises its discount rate and lowers its firm value. It tries to reduce its discount rate by shifting its cash flows from the future to the present, that is, by disinvesting.
The low adjustment costs of the PE model reduce the costs of this disinvestment, and thus result in a lower value premium.

Note that the low elasticity of intertemporal substitution (EIS) implied by external habit preferences and the need to match aggregate consumption volatility are critical to the quantitative effects in this discussion. This low EIS means that the household has a strong desire to smooth consumption across time, and, through the SDF, the firm has a strong incentive to smooth cash flows. This strong smoothing motive combined with the volatility of consumption growth seen in U.S. data then imply large investment frictions. This stands in contrast to long run risk and disaster models, which typically imply a large EIS, thus small investment frictions.

2.7 Conclusion

I show that external habit formation is consistent with one important aspect of the cross-section of stock returns. A real business cycle model, extended to include external habit preferences and idiosyncratic productivity generates a data-like value premium. The result is robust, and requires neither operating leverage, nor investment irreversibilities. Rather, the value premium arises as a result of the temporal distribution of cash flows. Value firms are temporarily low productivity firms. Mean reversion implies that they have high cash flow growth. These late-distributed cash flows are more exposed to the discount rate shocks which drive the external habit model. The value premium is compensation for this exposure.
2.8 Figures

Figure 2.1: Expected Returns as a Function of B/M. Figures are annual

Figure 2.2: B/M and Expected Returns as a Function of Firm States. Figures are annual
Figure 2.3: Cash Flow Growth of Book-to-Market Sorted Portfolios. Figures are annual.

Figure 2.4: Cash Flow Growth of Book-to-Market Sorted Portfolios: No Habit. Figures are annual.
Figure 2.5: The Cyclicality of Value and Growth Cash Flows. Value firm plots are calculated using the median capital and productivity of firms in the 10th decile of B/M-sorted portfolios. Growth firm plots are from the 1st decile. Net investment is investment net of depreciation. Aggregate capital is fixed at its mean.
Figure 2.6: Partial Equilibrium Experiment: Cash Flow Growth of Book-to-Market Sorted Portfolios. ‘GE’ uses calibrated parameter values from Table 2.3 and equilibrium laws of motion. ‘PE’ uses equilibrium aggregate laws of motion, but firm-level decision rules consistent with adjustment costs which are 1/20th of the value from Table 2.3.
2.9 Tables

Table 2.1: Mean Cash Flow of Value and Growth

‘E before Extraordinary’ is earnings before extraordinary income (ib). ‘E + Dep - Net Inv’ is earnings plus depreciation less capital expenditures plus sales of plant, property, and equipment (ni + dp - capx + sppe). ‘Year’ is year after portfolio formation. ‘Growth,’ ‘Neutral,’ and ‘Value’ are buy-and-hold value-weighted tercile portfolios sorted on B/M. Cash flow is averaged across portfolio formation years.

<table>
<thead>
<tr>
<th>Year</th>
<th>E Before Extraordinary</th>
<th>E + Dep - Net Inv</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth</td>
<td>Neutral</td>
</tr>
<tr>
<td>1</td>
<td>0.063</td>
<td>0.084</td>
</tr>
<tr>
<td>2</td>
<td>0.067</td>
<td>0.093</td>
</tr>
<tr>
<td>3</td>
<td>0.072</td>
<td>0.102</td>
</tr>
</tbody>
</table>
Table 2.2: Growth Rates of Mean Cash Flow for Value and Growth

‘E before Extraordinary’ is earnings before extraordinary income (ib). ‘E’ is earnings (ni), ‘Dep’ is depreciation (dp), ‘Net Inv’ is capital expenditures less sales of plant, property, and equipment (capx-sppe). ‘Year’ is year after portfolio formation. ‘Growth,’ ‘Neutral,’ and ‘Value’ are buy-and-hold value-weighted tercile portfolios sorted on B/M. The growth rate in year $t$ is the difference between the log of mean cash flows for year $t$ and the log of mean cash flows for year $t - 1$.

<table>
<thead>
<tr>
<th>Year</th>
<th>E Before Extraordinary</th>
<th>E</th>
<th>E + Dep</th>
<th>E + Dep - Net Inv</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth</td>
<td>Neutral</td>
<td>Value</td>
<td>Growth</td>
</tr>
<tr>
<td>2</td>
<td>6.6</td>
<td>9.7</td>
<td>29.7</td>
<td>6.2</td>
</tr>
<tr>
<td>3</td>
<td>6.4</td>
<td>9.0</td>
<td>13.7</td>
<td>6.4</td>
</tr>
<tr>
<td>4</td>
<td>8.1</td>
<td>7.4</td>
<td>6.7</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>E + Dep</td>
<td>Neutral</td>
<td>Value</td>
<td>Growth</td>
</tr>
<tr>
<td>2</td>
<td>8.5</td>
<td>8.9</td>
<td>15.9</td>
<td>5.0</td>
</tr>
<tr>
<td>3</td>
<td>8.3</td>
<td>8.0</td>
<td>9.0</td>
<td>11.3</td>
</tr>
<tr>
<td>4</td>
<td>8.7</td>
<td>9.8</td>
<td>6.9</td>
<td>13.5</td>
</tr>
</tbody>
</table>
The model is annual, and all parameter values and empirical moments are annual. Consumption is real non-durable goods and services consumption. The volatility of GDP and relative volatility of consumption are logged and HP-filtered with a smoothing parameter of 6.25.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Empirical Target</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unconditional Asset Price Moments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>Time Preference</td>
<td>0.89</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>Persistence of Surplus Consumption</td>
<td>0.87</td>
</tr>
<tr>
<td>$\bar{S}$</td>
<td>Steady-State Surplus Consumption</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Long-Run Growth Moments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Production Curvature</td>
<td>0.35</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Depreciation Rate</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Unconditional Business Cycle Moments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_a$</td>
<td>Volatility of TFP</td>
<td>0.03</td>
</tr>
<tr>
<td>$\rho_a$</td>
<td>Persistence of TFP</td>
<td>0.94</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Adjustment Cost</td>
<td>23</td>
</tr>
<tr>
<td><strong>Firm Level Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>Vol of idio prod</td>
<td>0.60</td>
</tr>
<tr>
<td>$\rho_b$</td>
<td>Persist of idio prod</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>Chosen Outside of the Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Utility Curvature</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Table 2.4: Unconditional Aggregate Asset Price Moments

All figures are annual. Data moments are taken from Beeler and Campbell (2009) and correspond to 1947-2008. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

<table>
<thead>
<tr>
<th>US Data</th>
<th>Model</th>
<th>mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibrated Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E(r_f)$ (%)</td>
<td>0.89</td>
<td>0.87</td>
<td>-3.00</td>
<td>0.05</td>
<td>9.50</td>
</tr>
<tr>
<td>$AC1(r_f)$</td>
<td>0.84</td>
<td>0.79</td>
<td>0.45</td>
<td>0.82</td>
<td>0.97</td>
</tr>
<tr>
<td>$E(R_m - R_f)/\sigma(R_m)$</td>
<td>0.44</td>
<td>0.47</td>
<td>0.26</td>
<td>0.46</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Untargeted Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E(r_m - r_f)$ (%)</td>
<td>6.36</td>
<td>6.84</td>
<td>3.67</td>
<td>7.06</td>
<td>8.59</td>
</tr>
<tr>
<td>$\sigma(r_m - r_f)$ (%)</td>
<td>16.52</td>
<td>16.23</td>
<td>10.59</td>
<td>16.03</td>
<td>23.19</td>
</tr>
<tr>
<td>$AC1(r_m - r_f)$</td>
<td>0.08</td>
<td>-0.04</td>
<td>-0.19</td>
<td>-0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sigma(r_f)$ (%)</td>
<td>1.82</td>
<td>2.85</td>
<td>0.57</td>
<td>1.83</td>
<td>8.84</td>
</tr>
<tr>
<td>$E(p_m - d_m)$</td>
<td>3.36</td>
<td>2.68</td>
<td>1.97</td>
<td>2.74</td>
<td>3.16</td>
</tr>
<tr>
<td>$\sigma(p_m - d_m)$</td>
<td>0.45</td>
<td>0.34</td>
<td>0.20</td>
<td>0.33</td>
<td>0.54</td>
</tr>
<tr>
<td>$AC1(p_m - d_m)$</td>
<td>0.87</td>
<td>0.87</td>
<td>0.76</td>
<td>0.87</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table 2.5: Business Cycle Moments

All figures are annual. Data moments correspond to 1947-2011. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

<table>
<thead>
<tr>
<th>US Data</th>
<th>Model</th>
<th>mean</th>
<th>5%</th>
<th>50%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Calibrated Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma(y_{hp})$ (%)</td>
<td>1.50</td>
<td>1.64</td>
<td>1.04</td>
<td>1.58</td>
<td>2.21</td>
</tr>
<tr>
<td>$\sigma(c_{hp})$ (%)</td>
<td>1.32</td>
<td>1.27</td>
<td>0.85</td>
<td>1.22</td>
<td>1.80</td>
</tr>
<tr>
<td><strong>Untargeted Moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma(c_{hp})/\sigma(y_{hp})$</td>
<td>0.49</td>
<td>0.50</td>
<td>0.42</td>
<td>0.48</td>
<td>0.59</td>
</tr>
<tr>
<td>$\sigma(i_{hp})/\sigma(y_{hp})$</td>
<td>2.68</td>
<td>3.34</td>
<td>3.01</td>
<td>3.33</td>
<td>3.73</td>
</tr>
<tr>
<td>$\rho(y_{hp}, c_{hp})$</td>
<td>0.84</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\rho(y_{hp}, i_{hp})$</td>
<td>0.58</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$AC1(\Delta c)$</td>
<td>0.52</td>
<td>0.04</td>
<td>-0.14</td>
<td>0.02</td>
<td>0.24</td>
</tr>
<tr>
<td>$E(Adj Cost/ Y)$ (%)</td>
<td>0.75</td>
<td>0.46</td>
<td>0.73</td>
<td>1.24</td>
<td>1.24</td>
</tr>
<tr>
<td>$E(Adj Cost/ I)$ (%)</td>
<td>4.37</td>
<td>2.20</td>
<td>4.25</td>
<td>8.61</td>
<td>8.61</td>
</tr>
</tbody>
</table>
Table 2.6: Predicting Dividend Growth and Excess Returns with the Price-Dividend Ratio

All figures are annual. Data moments are taken from Beeler and Campbell (2009) and correspond to 1947-2008. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

### Panel A: Predicting Dividend Growth

\[ \sum_{j=1}^{L} \Delta d_{m,t+j} = \alpha + \hat{\beta}(p_{m,t} - d_{m,t}) + \epsilon_{t+L}. \]

<table>
<thead>
<tr>
<th>L</th>
<th>US Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>mean 5% 50% 95%</td>
</tr>
<tr>
<td>1</td>
<td>0.003</td>
<td>-0.002 -0.014 -0.001 0.006</td>
</tr>
<tr>
<td>3</td>
<td>0.012</td>
<td>-0.005 -0.024 -0.003 0.013</td>
</tr>
<tr>
<td>5</td>
<td>0.044</td>
<td>-0.011 -0.043 -0.009 0.019</td>
</tr>
<tr>
<td>t-stat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.112</td>
<td>-0.128 -1.219 -0.110 1.142</td>
</tr>
<tr>
<td>3</td>
<td>0.193</td>
<td>-0.385 -2.376 -0.270 1.499</td>
</tr>
<tr>
<td>5</td>
<td>0.482</td>
<td>-0.646 -2.816 -0.711 1.557</td>
</tr>
<tr>
<td>(\hat{\beta})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.12</td>
<td>-0.149 -0.315 -0.148 -0.035</td>
</tr>
<tr>
<td>3</td>
<td>-0.27</td>
<td>-0.281 -0.568 -0.315 -0.081</td>
</tr>
<tr>
<td>5</td>
<td>-0.42</td>
<td>-0.422 -0.757 -0.472 -0.114</td>
</tr>
<tr>
<td>(R^2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.009 0.000 0.005 0.023</td>
</tr>
<tr>
<td>3</td>
<td>0.001</td>
<td>0.019 0.000 0.011 0.061</td>
</tr>
<tr>
<td>5</td>
<td>0.011</td>
<td>0.036 0.000 0.022 0.116</td>
</tr>
</tbody>
</table>

### Panel A: Predicting Excess Returns

\[ \sum_{j=1}^{L} (r_{m,t+j} - r_{f,t+j}) = \alpha + \hat{\beta}(p_{m,t} - d_{m,t}) + \epsilon_{t+L}. \]

<table>
<thead>
<tr>
<th>L</th>
<th>US Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>mean 5% 50% 95%</td>
</tr>
<tr>
<td>1</td>
<td>-1.12</td>
<td>-1.49 -3.15 -1.148 -0.035</td>
</tr>
<tr>
<td>3</td>
<td>-0.27</td>
<td>-0.281 -0.568 -0.315 -0.081</td>
</tr>
<tr>
<td>5</td>
<td>-0.42</td>
<td>-0.422 -0.757 -0.472 -0.114</td>
</tr>
<tr>
<td>t-stat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.63</td>
<td>-2.039 -3.349 -2.083 -0.484</td>
</tr>
<tr>
<td>3</td>
<td>-3.19</td>
<td>-2.872 -5.117 -2.842 -0.709</td>
</tr>
<tr>
<td>5</td>
<td>-3.37</td>
<td>-3.697 -7.730 -3.596 -1.084</td>
</tr>
<tr>
<td>(\hat{\beta})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.63</td>
<td>-2.039 -3.349 -2.083 -0.484</td>
</tr>
<tr>
<td>3</td>
<td>-3.19</td>
<td>-2.872 -5.117 -2.842 -0.709</td>
</tr>
<tr>
<td>5</td>
<td>-3.37</td>
<td>-3.697 -7.730 -3.596 -1.084</td>
</tr>
<tr>
<td>(R^2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.09</td>
<td>0.089 0.007 0.085 0.178</td>
</tr>
<tr>
<td>3</td>
<td>0.19</td>
<td>0.164 0.019 0.153 0.337</td>
</tr>
<tr>
<td>5</td>
<td>0.26</td>
<td>0.249 0.043 0.275 0.520</td>
</tr>
</tbody>
</table>
Table 2.7: Regressions of Future Returns on Book-to-Market

All figures are annual. Data moments correspond to 1947-2012. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

<table>
<thead>
<tr>
<th>Dependent Var: $R_{i,t+1}$</th>
<th>US Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>5%</td>
</tr>
<tr>
<td>intercept</td>
<td>18.62</td>
<td>16.63</td>
</tr>
<tr>
<td>t-stat</td>
<td>5.74</td>
<td>3.92</td>
</tr>
<tr>
<td>$\log(B/M)_{i,t}$</td>
<td>5.70</td>
<td>5.17</td>
</tr>
<tr>
<td>t-stat</td>
<td>4.88</td>
<td>3.48</td>
</tr>
</tbody>
</table>

Table 2.8: Summary Statistics from 10 Book-to-Market Sorted Portfolios

All figures are annual. Data moments correspond to 1947-2012. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

<table>
<thead>
<tr>
<th>port</th>
<th>US Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E(R_{port})$ (%)</td>
<td>$\sigma(R_{port})$ (%)</td>
</tr>
<tr>
<td>Low</td>
<td>7.75</td>
<td>20.86</td>
</tr>
<tr>
<td>2</td>
<td>8.00</td>
<td>17.13</td>
</tr>
<tr>
<td>3</td>
<td>8.08</td>
<td>16.65</td>
</tr>
<tr>
<td>4</td>
<td>8.60</td>
<td>17.62</td>
</tr>
<tr>
<td>5</td>
<td>9.52</td>
<td>18.33</td>
</tr>
<tr>
<td>6</td>
<td>9.72</td>
<td>17.64</td>
</tr>
<tr>
<td>7</td>
<td>9.81</td>
<td>19.27</td>
</tr>
<tr>
<td>8</td>
<td>11.65</td>
<td>21.06</td>
</tr>
<tr>
<td>9</td>
<td>11.86</td>
<td>20.38</td>
</tr>
<tr>
<td>High</td>
<td>13.29</td>
<td>25.67</td>
</tr>
</tbody>
</table>
Table 2.9: Partial Equilibrium Experiment: Regressions of Future Returns on Book-to-Market

All figures are annual. Data moments correspond to 1947-2012. The model columns show means and percentiles from 500 simulations of the same length as the empirical sample.

<table>
<thead>
<tr>
<th>Dependent Var: $R_{i,t+1}$</th>
<th>US data</th>
<th>GE model</th>
<th>PE model</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>18.62</td>
<td>16.63</td>
<td>11.31</td>
</tr>
<tr>
<td>t-stat</td>
<td>5.74</td>
<td>3.92</td>
<td>3.96</td>
</tr>
<tr>
<td>log($B/M$)$_{i,t}$</td>
<td>5.70</td>
<td>5.17</td>
<td>0.70</td>
</tr>
<tr>
<td>t-stat</td>
<td>4.88</td>
<td>3.48</td>
<td>1.30</td>
</tr>
</tbody>
</table>
Chapter 3

Implications of Security Market Data for Production Technologies

3.1 Introduction

Stocks are claims on real assets, yet formally modeling the relationship between asset prices and the real economy has proven to be difficult. These difficulties span a wide range of models (Cochrane (2008b)), suggesting the existence of a set of general principles. Indeed, Hansen and Jagannathan (1991) provide a general restriction which a broad array of models must satisfy in order to be consistent with asset prices. Using a basic prediction of optimal consumption, Hansen and Jagannathan (1991) show that the volatility of the intertemporal marginal rate of substitution of the investor must exceed the Sharpe ratios of assets available to the investor. This restriction applies regardless of the completeness of markets, the production technology, or the stochastic structure of shocks.

This paper expands the set of general principles of asset price modeling. I
show that as long as the firm's value function is “factorable,” a firm's stock return equals its investment return (a.k.a. marginal rate of transformation). A sufficient condition for factorability is that the production technology is homogenous of any degree. This condition nests the linear homogeneity condition of Cochrane (1991) and Restoy and Rockinger (1994). Factorability is also the result of a first-order approximation to a general value function. In numerical experiments, I show that this first-order approximation forms a good approximation of a nonlinear solution.

The generality of the investment-return stock-return equality means that this equality provides a succinct and intuitive description of the way asset prices are anomalous with respect to optimal investment behavior. The high volatility of stock returns means that investment returns must be very volatile. This restriction holds in a broad range of models, regardless of the completeness of markets, preferences, beliefs, or the stochastic structure of shocks. This generality means that the restriction forms a powerful diagnostic test. Using this equality, one can quickly identify whether a model can match asset prices by simply looking at the production technology.

In the vast majority of models, the volatility of investment returns stems from capital adjustment costs or investment-specific technology (IST) shocks. The magnitude of these two channels is typically small, however, leading to stock volatility which is far less than that of the data. This statement can be made quantitative, as shown in Figure 3.1. The figure shows admissible parameter values for IST volatility and capital adjustment costs implied by U.S. macroeconomic data and the investment return volatility restriction. Both IST volatility and adjustment costs are restricted to be very high. IST volatility must be at least 10% per year, and the elasticity of the investment rate with respect to marginal Q

122
must be less than 0.5. Since this restriction applies to a broad range of models, it provides a succinct description of the literature. The figure also plots parameter values from various papers in the literature. The admissible region neatly separates models which match the equity premium (in green) and models which do not (in red). This clean division applies to a broad array of mechanisms, including habit (Jermann (1998)), long-run risk (Kaltenbrunner and Lochstoer (2010)), disaster risk (Gourio (2012)), IST shocks (Papanikolaou (2011)), labor operating leverage (Danthine and Donaldson (2002)) and limited participation (Guvenen (2009)).

The intuition behind these results is closely analogous to the intuition behind the Hansen and Jagannathan (1991) bounds. The difficulty in satisfying the Hansen and Jagannathan (1991) bounds is that consumption volatility is very small compared to the Sharpe ratio. Thus generating a large Sharpe ratio requires amplification through utility function curvature (risk aversion) or an additional source of volatility (i.e. concerns about wealth). Analogously, the difficulty in matching the bounds of this paper is that the volatility of changes in the investment rate is far smaller than the volatility of stock returns. Thus, generating a large stock return volatility requires amplification mechanism through production curvature (capital adjustment costs), or an additional source of volatility (i.e. IST shocks).

This production-based restriction complements the consumption-based restrictions. Satisfying Hansen and Jagannathan (1991)’s consumption-based restrictions only means that the model can match the large Sharpe ratio seen in the data. A large Sharpe ratio does not imply a large equity premium, however. Indeed, since 1991 a number of models have been built which satisfy the Hansen and Jagannathan (1991) bounds but fail to generate a large equity pre-
mium (Tallarini (2000), Kaltenbrunner and Lochstoer (2010), Gourio (2012)). The production-based bounds explain why. As seen in Figure 3.1, these models have neither large IST volatility nor large capital adjustment costs.

The intuition that adjustment costs help models fit asset prices has been discussed in previous papers (i.e. Jermann (1998), Danthine and Donaldson (2002)). The previous papers, however, only discuss adjustment costs within specific parametric models. This paper broadens the class of models which require high adjustment costs as much as possible. It also shows that habit formation, labor operating leverage, and limited participation ultimately affect asset prices by allowing for a higher adjustment cost. These different mechanisms are in a sense alternative foundations for the same reduced form.

The investment return - stock return equality has been empirically studied in a number of papers (Cochrane (1991), Restoy and Rockinger (1994), Liu, Whited, and Zhang (2009)), but its generality has been overlooked. Indeed, the generality of this equality has been seen as a weakness, since verifying the equality in the data does not provide evidence for efficient markets, or any other specific notions of market behavior (Kogan and Papanikolaou (2012)). The generality is a strength, however, in that failure to satisfy the implications of this equality means that a model will be rejected by asset price data, regardless of an innumerable other details of the model. The goal of seeking restrictions also leads me to focus on other aspects of this equality. While previous papers focus on the first moment of investment returns in order to highlight the success of formal asset price modeling, this paper focuses on the second moment in order to characterize the difficulties.

The paper proceeds as follows. Section 3.2 presents a semi-parametric model which nests a broad array of models in the literature. Section 3.3 shows how
this model can be restricted using data. Section 3.4 examines these restrictions.
Section 3.6 shows how the results can be extended beyond constant returns to scale. Section 3.6 concludes.

### 3.2 A Semi-Parametric Model

This section lays out a semi-parametric model. The model features a representative firm but is otherwise minimal on assumptions. The specification nests many models and mechanisms in the literature, including traditional RBC (Rouwenhorst (1995)), internal habit (Jermann (1998)), external habit (Chen (2012a)), Epstein-Zin preferences (Tallarini (2000)), long-run risk (Kaltenbrunner and Lochstoer (2010)), IST shocks (Papanikolaou (2011)), long-run productivity risk (Croce (2010)), limited participation (Guvenen (2009)), extrapolative expectations (Hirshleifer and Yu (2011)), and disappointment aversion (Campanale, Castro, and Clementi (2010)). The model also captures operating leverage as in Danthine and Donaldson (2002) and Favilukis and Lin (2011), though these models are not exactly nested.

Time is discrete and the horizon is infinite. A representative firm solves

\[
V(k_t, s_t) = \max_{(k_{t+1}, n_t)} \Pi(k_t, n_t, k_{t+1}, s_t) + \tilde{E}_t[M(s_t, s_{t+1})V(k_{t+1}, s_{t+1})]
\]  

(3.1)

where \(\Pi(k_t, n_t, k_{t+1}, s_t)\) is a profit function, \(k_t\) is capital, \(n_t\) is labor, and \(s_t\) is the aggregate state. The aggregate state is abstract, and can contain anything including productivity, the volatility of productivity, disaster probability, surplus consumption, etc. The firm discounts future profits using some stochastic discount factor \(M(s_t, s_{t+1})\). This discount factor need not be unique, and can accommodate any sort of preferences. The firm uses subjective expectations \(\tilde{E}\) which could
potentially deviate from objective probabilities. Note that the stochastic structure of the underlying shocks to the model remains unspecified and can also accommodate a wide variety of processes.

Firm profits are parameterized in order to allow for a number of technological features from the literature

\[
\Pi(k_t, n_t, k_{t+1}, s_t) = z(s_t)[d(s_t)k_t]^\alpha n_t^{1-a} - w(s_t)n_t
\]

Operating Profit

\[
- \frac{1}{\mu(s_t)} \psi\left(\frac{k_{t+1}}{d(s_t)k_t}\right) d(s_t)k_t
\]

Capital Expenditures

Operating profits are revenues less payments to labor. The firm makes revenues from neutral technology \(z(s_t)\), undestroyed capital \(d(s_t)k_t\), and labor \(n_t\). \(d(s_t)\) is a depreciation state which allows for disasters in the style of Gourio (2012). The exact neutral technology process remains unspecified, and so the model nests many mechanisms in the literature such as long-run productivity risk (Croce (2010)), disaster productivity risk (Gourio (2012)), and uncertainty shocks (Bloom (2009)). Revenues are constant returns, as is assumed in most of the literature (Jermann (1998), Kaltenbrunner and Lochstoer (2010), among others). This assumption means that investment returns will exactly equal stock returns, but Section 3.5 shows that this relationship will hold approximately under reasonable deviations from constant returns.

The firm pays its hired labor at a wage \(w(s_t)\) which can be any function of the aggregate state. This generality means that the model nests labor-based operative leverage as in Danthine and Donaldson (2002) and Favilukis and Lin (2011).

Capital expenditures feature IST shocks \(\mu(s_t)\) as well as capital adjustment costs \(\psi\left(\frac{k_{t+1}}{d(s_t)k_t}\right)\). The IST shocks \(\mu(s_t)\) captures innovation in producing capital-
specific goods, but can also be interpreted as the state of financial frictions (Justiniano, Primiceri, and Tambalotti (2011)). I introduce these shocks because Papanikolaou (2011) shows that they may play a key role in determining asset prices.

The capital adjustment cost features a constant elasticity of the investment rate with respect to Tobin’s Q

\[
\psi\left(\frac{k_{t+1}}{d(s_t)k_t}\right) = \left[\frac{1 - \phi}{a_1} \left(\frac{k_{t+1}}{d(s_t)k_t} - 1 + \delta - a_2\right)\right]^{1/(1-\phi)}
\]

where \(a_1\) and \(a_2\) are chosen to make the non-stochastic steady state invariant to the adjustment cost \(\phi\). One can show that \(1/\phi\) is the elasticity of the investment rate. This formulation is equivalent to the adjustment costs in Jermann (1998), Guvenen (2009), Kaltenbrunner and Lochstoer (2010).

This concludes the model description. Note that a large number of features of the model remain unparameterized and unrestricted: productivity, depreciation, wages, preferences, as well as beliefs. Previous papers describe this as a limitation of the investment-return stock-return equality (i.e. Kogan and Papanikolaou (2012)). Indeed, the generality of this model means that verifying its predictions cannot provide evidence of rational markets or test any other notions of investor behavior. This generality has a strength, however, in that restrictions implied by this model apply to a broad array of models.

### 3.3 Restricting the Model with Data

Despite the generality of the model in the previous section, the model has a powerful prediction which can be used to restrict the model’s parameters. The prediction is that investment returns equal stock returns, and thus the moments
of these two restrictions are equal. This section formally describes this prediction and shows how I use data to construct investment returns. Readers eager to see results should skip to Section 3.4.

The model’s production technology is linearly homogenous. Thus, the model satisfies the sufficient conditions of (Restoy and Rockinger (1994)) and the following equality holds

$$R_{t+1}^I = R_{t+1}$$

(3.2)

where $R_{t+1}$ is the stock return, and $R_{t+1}^I$ is the investment return and is defined as

$$R_{t+1}^I \equiv \frac{\Pi_1(k_{t+1}, n_{t+1}, k_{t+2}, s_{t+1})}{-\Pi_3(k_t, n_t, k_{t+1}, s_t)}$$

(3.3)

and the subscript 1 on $\Pi$ indicates the derivative with respect to the 1st argument. Intuitively, the stock return tomorrow is high if the marginal profitability of capital tomorrow is high (numerator is large) or the cost of creating capital tomorrow is low (denominator is small). An implication of this equality is that all of the moments match, and in particular

$$\mathbb{E}(R_{t+1}^I) = \mathbb{E}(R_{t+1})$$

(3.4)

$$\sigma(R_{t+1}^I) = \sigma(R_{t+1})$$

(3.5)

All higher moments match too, but we will see that the volatility restriction is already quite challenging for models to satisfy. Note that this expectation and volatility do not have a tilde, indicating that they use objective probability measures and may differ from the subjective expectation in the firm’s objective (3.1).
I restrict the model by repeating the following many times. I first pick some parameter values. I then combine the parameter values with macroeconomic data to construct investment returns. Last I check if the parameter values are admissible, that is, I check if they are consistent with equation (3.4).

Crunching out the derivatives in the definition of investment returns (3.3) yields

$$R_{t+1}^I = \frac{d_{t+1} \left\{ \alpha \frac{1}{d_{t+1}} \frac{y_{t+1}}{k_{t+1}} + \frac{1}{\mu_{t+1}} \left[ \frac{1}{\psi_{t+1}^{-1}} \left( \frac{1}{d_{t+1}} \frac{i_{t+1}}{k_{t+1}} \right) (1 - \delta + \psi^{-1} \left( \frac{1}{d_{t+1}} \frac{i_{t+1}}{k_{t+1}} \right)) - \frac{1}{d_{t+1}} \frac{i_{t+1}}{k_{t+1}} \right] \right\}}{\frac{1}{\mu_{t+1}} \psi_{t+1}^{-1} \left( \frac{1}{d_{t+1}} \frac{i_{t+1}}{k_{t+1}} \right)}$$

(3.6)

The task now is to identify as many components of the above as possible using data.

The central dynamic variables of the above equation are the output-to-capital ratio $y_t/k_t$ and the investment to capital ratio $i_t/k_t$. Macroeconomic data provide these variables. Note that taking these values from the data implicitly assumes that we require that the model to match the data in terms of output and investment. Since these are the two central variables in business cycles, demanding that models match these data series is intuitive. I take these variables from the BEA’s NIPA accounts. Output is real, per capita GDP. Investment is also real and per capita, and is defined in a broad sense. Investment is defined as durable consumption plus private investment in structures, equipment, and residences, plus government fixed investment. This broad definition is consistent with the use of a representative firm: in principle the output and inputs of this firm represents all sectors of the economy. Alternative approaches include excluding government or residential investment, or using Compustat data, however the main results should hold regardless of the measurement of investment. I use very loose
restrictions on investment return moments to compensate for alternative measurement of the real variables. Moreover, the main results only require that aggregated investment is much less volatile than stock returns, and this holds regardless of the investment measure.

The capital destruction state \( d_{t+1} \) is identified with the assumption that there has been no destruction of aggregate capital between 1947 and 2010 in the United States. Disastrous destruction of capital should be observable. One may argue that the financial crisis of 2008 is the destruction of intangible capital, but the main results should not be sensitive to the exclusion of post 2007 data. Note that this identifying assumption does not say that I assume that the probability of disaster is constant.

The IST state \( \mu_t \) is unobserved. In many models, \( \mu_t \) is identified by the relative price of investment goods (Greenwood et al. (1997)), however, it is not difficult to write down a model in which \( \mu_t \) deviates from this relative price (Justiniano, Primiceri, and Tambalotti (2011)). I assume that \( \mu_t \) follows

\[
\Delta \log \mu_t = g_{\mu} + \frac{\sigma_{\mu}}{\sigma(\Delta \log i_t)} [\Delta \log i_t - \mathbb{E}(\Delta \log i_t)]
\]

(3.7)

where \( \sigma_{\mu} \) is a parameter which controls the volatility of IST shocks and \( g_{\mu} \) is chosen so that the mean growth rate of investment shocks is zero. I also initialize \( \mu_0 = 1 \) though the results are not sensitive to this initialization as long as the growth rate of \( \mu_t \) is approximately 0. This assumption means that IST shocks are perfectly correlated with investment growth, consistent with theoretical predictions (Justiniano, Primiceri, and Tambalotti (2011)).

The remaining parameters are restricted by firm optimality and asset prices. These remaining parameters are capital’s share of output \( \alpha \), the capital adjust-
ment cost parameter $\phi$, and the volatility of IST shocks $\mu_t$. Firm optimality means that the mean and volatility of investment returns equals the mean and volatility of stock returns. Asset prices then provide us the mean and volatility of stock returns. Importantly these same restrictions apply to a broad range of models, including RBC models, internal habit, external habit, long run risk, disasters, and more.

### 3.4 The Restrictions

This section contains the main results. I construct investment returns using data as described in the previous section and then use firm optimality and asset price data to generate regions of admissible parameters. I proceed in two cases. Section 3.4.1 shuts down IST shocks since many models do not have this feature. It then identifies restrictions on the capital share and adjustment costs which apply to the broad range of models which do not feature IST shocks. Section 3.4.2 turns on the IST shocks to extend the set of restricted models.

Throughout this section I make only use very loose bounds imposed by asset prices. That is, I assume that the data indicate that the expected stock return should be between 4 and 12%, and that the volatility of stock returns should be between 10 and 30%. These bounds are loose because papers frequently adjust the aggregate stock market return for financial leverage. Others might argue that the measured aggregate return deviates from the model's return due to taxes or differences among private and public firms. These loose bounds should encompass all of these deviations. I avoid precise measurement of these bounds in order to focus the results on the investment return moments.
3.4.1 Case 1: No Investment-Specific Shocks

Here I shut down IST volatility ($\sigma_\mu = 0$) and examine restrictions on the capital share $\alpha$ and adjustment cost $\phi$.

Figure 3.2 shows mean investment returns as a function of the capital share and the adjustment cost. Mean investment returns are monotonically increasing in the capital share but have little relationship with adjustment costs. The shaded region represents restrictions implied by firm optimality and stock market data. The figure shows that the expected investment return restricts capital’s share of output to about 0.36, consistent with the capital share implied by measurement of labor income (Cooley and Prescott (1995)). On the other hand, since adjustment costs have little relationship with the mean investment return, the mean investment return places no restrictions on adjustment costs.

Figure 3.3 shows the volatility of investment returns as a function of the capital share and adjustment cost. The capital share has no effect on investment return volatility, while investment return volatility is monotonically increasing in the adjustment cost. Once again, the shaded region shows restrictions implied by stock market data. The restriction is quite severe, despite the fact that the bounds entertain stock return volatility of as low as 10% per year. To satisfy this bound, the adjustment cost parameter $\phi$ must be at least 2, implying an elasticity of the investment rate with respect to Tobin’s Q of $1/\phi = 0.5$. This value is much lower than the value used in the literature, which is on the order of 30 (i.e. Whited (1992)).

Both the mean and volatility restrictions are necessary for consistency with basic asset price moments. Thus the admissible capital share and adjustment cost but lie in the intersection of the previous figures as shown in Figure 3.4. The figure shows that capital share must be between 0.3 and 0.5, while the adjust-
ment cost must be at least 2. This restriction applies regardless of the completeness of markets, preferences, beliefs, or the stochastic structure of shocks. As a result, the figure provides a quick interpretation of the literature. The figure also plots parameter values used by various papers in the literature. Models which fit the equity premium are shown in green, while models which cannot are shown in red. The production based restrictions cleanly separate the two groups, despite the fact that these models use a broad array of mechanisms from time-varying disaster risk (Gourio (2012)) to limited participation (Guvenen (2009)).

The role of adjustment costs can be seen in the following approximation of the investment return

\[ R_{t+1} = R_{t+1}^I \approx \alpha \frac{y_{t+1}}{k_{t+1}} + \frac{\phi}{\mathbb{E}[i_{t+1}/k_{t+1}]} [\Delta(i_{t+1}/k_{t+1})] - \frac{\phi \delta}{\mathbb{E}[i_{t+1}/k_{t+1}]} (i_{t+1}/k_{t+1}) + \text{Constants} \]

(3.8)

where \( y_t/k_t \) is the output to capital ratio, \( i_t/k_t \) is the investment rate, \( \Delta(i_t/k_t) \) is its first difference, \( \alpha \) is the capital share, and \( \phi \) is the adjustment cost parameter. The trouble with satisfying this equation is that stock return volatility is far higher than the volatility of the components of the investment return on the RHS. Table 3.1 shows summary statistics on these variables. While stock return has a huge volatility of about 17% per year, the output-to-capital ratio has a tiny volatility of 3.5%. The investment rate and its first difference have even tinier volatilities on the order of 0.1% per year. Since \( \alpha \) must be within (0, 1), it can have only a limited role in amplifying these volatilities, in addition to the fact that it is pinned down by the mean investment return. Capital adjustment costs, on the other hand are not so restricted and amplify the volatilities of both the growth in the investment rate and the investment rate itself.

There are a number of ways to interpret the link between adjustment costs
and volatility. The standard Q-theoretical interpretation is that capital adjustment costs mean that installed capital has value. The higher the adjustment cost, the more “stuck” is this capital and the more it is exposed to shocks. An alternative interpretation is that adjustment costs drive a wedge between stock prices and investment, and in particular make investment more sluggish a variable. Yet another interpretation is that the firm would like to minimize the amount of discounting suffered by its dividends. Allocating production in a way which minimizes the volatility of its stock price would be one way to achieve this goal. The production technology, and in particular, adjustment costs restrict the firm’s ability to minimization its risk. These multiple interpretations suggest that capital adjustment costs are a robust way of understanding the disconnect between asset price and investment volatility.

The qualitative results from this section are consistent with Cochrane (1991). Cochrane finds with a similar technology that the marginal product of capital is controls the mean stock return while the adjustment cost parameter controls the volatility of stock returns. Rather than pin down the adjustment cost using stock return volatility, Cochrane fits the standard deviation of fitted values of a regression of stock returns on leaded and lagged investment rates with the standard deviation of a similar regression of investment returns. This approach is well-suited for an investment-based alternative to consumption-based models of stock returns. The approach used here of directly fitting volatilities is designed to restrict consumption-based models and interpret the production-asset-pricing literature which has blossomed since the publication of Cochrane (1991).
3.4.2 Case 2: Investment-Specific Shocks

Here I fix the capital share at $\alpha = 0.36$ to be consistent with the expected stock return as well as the measured labor share of income. I then examine restrictions on the volatility of the IST shock and the adjustment cost.

Figure 3.5 shows the mean investment return as a function of IST volatility and adjustment costs. The figure shows that both IST vol and adjustment costs have very little effect on the expected investment return. The shaded region is admissible according to loose bounds implied by expected stock returns. The entire figure is shaded in. Expected stock returns do not restrict either parameter.

Figure 3.6 shows the volatility of investment returns. Both parameters affect investment return volatility. Higher adjustment costs or high IST volatility imply higher investment return volatility, with level curves radiating out from the origin. A striking aspect of the figure is the magnitudes. The volatility of stock returns restricts IST volatility or adjustment costs to be very high. It is not until an IST vol of 10% or an adjustment cost of 2 do we reach an admissible investment return volatility of 10% per year. To place the IST volatility in perspective, the volatility of the relative price of investment is only about 5% per year Fernández-Villaverde and Rubio-Ramírez (2007). Moderate IST volatility close to the directly measured value does not help much with lowering the extreme requirement on adjustment costs from the no IST shock case. Indeed, since the level curves are circular, an IST volatility of 5% means that adjustment costs need to be even larger than the no IST shock case.

Intersecting the mean and volatility figures produces the figure from the introduction, Figure 3.1. Since expected returns do not restrict these parameter values, the admissible region is the same as that from the volatility figure. Once again, a broad range of models must be in this admissible region. The inclu-
sion of IST volatility allows this figure to include an additional, important paper in the field of asset pricing. Papanikolaou (2011) is one of the few asset pricing production economy papers which can match the equity premium, and we see from Figure 3.1 that his success can also be placed neatly into this structure. Papanikolaou (2011) succeeds on the equity premium because he uses a very large value of IST shocks.

IST shocks differ from other kinds of productivity shocks because they are less restricted by the data. This “invisibility” can be seen in an approximation of the investment return

\[
R_{t+1} = R_{t+1}^I \approx az_{t+1} \left( \frac{n_{t+1}}{k_{t+1}} \right)^{1-\alpha} - \Delta \mu_{t+1} + \frac{\phi}{E[i_{t+1}/k_{t+1}]} \left[ \frac{1}{\mu_{t+1}} \frac{i_{t+1}}{k_{t+1}} - \frac{1}{\mu_t} \frac{i_t}{k_t} \right]
\]

(3.9)

+ Constants

(3.10)

Investment returns come from three terms: one due to neutral productivity shocks \(z_{t+1}\), one due to IST shocks \(\Delta \mu_{t+1}\), and another related to growth in the investment rate. The IST shock has a direct effect on the investment return. A high IST state tomorrow means that capital is cheap to produce tomorrow. Cheap capital means that firm value is low, implying a low realized stock return.

A key aspect of the IST shock is that it is not directly limited by observables. In some models, \(\mu_t\) is equal to the relative price of investment goods (Greenwood, Hercowitz, and Krusell (1997)). In general, however, it is not observable (Justiniano, Primiceri, and Tambalotti (2011)). This contrasts with shocks that affect neutral productivity \(z_{t+1}\). The firm’s revenue function means that \(z_{t+1} \left( \frac{n_{t+1}}{k_{t+1}} \right)^{1-\alpha} = y_{t+1}/k_{t+1}\), and the RHS is pinned down by the data. This observability strongly limits the effect of many mechanisms, including the standard RBC TFP shock, time-varying disaster productivity shocks (Gourio (2012)), long-run productivity...
risk (Croce (2010)), and uncertainty shocks (Bloom (2009)).

Equation (3.9) also shows that, in this class of models, investor preferences and beliefs only affect stock returns indirectly. To be more explicit, alternative preferences or beliefs can only increase stock return volatility to the extent that they allow for more volatile IST shocks or higher adjustment costs. This invariance is related to the perfect corporate control assumed by the firm’s objective (3.1). This perfect corporate control means that the investment return shares the same preference / belief distortions as that of the stock return.

These production-based restrictions neatly complement the consumption-based restrictions of Hansen and Jagannathan (1991). Hansen and Jagannathan (1991) provide an upper bound for the Sharpe ratio

$$\frac{\mathbb{E}[R_{t+1} - R_{t+1}^f]}{\sigma(R_{t+1})} \leq \sigma(\log M_{t+1})$$  (3.11)

while this paper restricts the volatility of returns

$$\sigma(R_{t+1}) \approx \sigma(R_{t+1}^f)$$  (3.12)

Both of these restrictions are necessary in a broad range of models. The consumption-based restriction is silent on the volatility of stock returns. Meanwhile the production-based restrictions are silent on the risk-free rate.

### 3.4.3 Additional Restrictions

The previous sections establish that, for a broad class of models, security prices imply that either adjustment costs are very high or the volatility of IST shocks must be very high. Not all models are consistent however with high ad-
justment costs or high IST volatility. In other words, these technological restrictions place restrictions on other aspects of the model.

This section investigates what kind of beliefs or preferences are consistent with high adjustment cost. Extending the restrictions requires some additional structure. I must take a stand on the stochastic structure of shocks, preferences, as well as beliefs. In particular, I assume AR1 neutral productivity with a persistence parameter of 0.95 per quarter. I also assume Epstein-Zin preferences and rational expectations. The model structure is the same as Kaltenbrunner and Lochstoer (2010)’s temporary shock models.

I focus on the elasticity of intertemporal substitution (EIS) and restrict this parameter by requiring that the model approximately match the relative volatility of consumption (consumption volatility/ GDP volatility). To accommodate alternative measurement of consumption and GDP, I make the loose requirement that this ratio should be between 0.4 and 0.6. I leave the risk aversion parameter at 6 and the volatility of the neutral shock at 0.015 per quarter. These parameters have little effect on the relative volatility of consumption (Schmitt-Grohe and Uribe (2004), Van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramirez (2012)).

The model is solved using projection methods (Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012)) and simulated for 2000 quarters. Figure 3.7 shows relative consumption volatility as a function of the EIS and adjustment costs. Both parameters have a strong effect on consumption volatility, and only a small region of the plot is admissible. A low adjustment cost permits many EIS values, however, for adjustment costs larger than 0.5, the EIS is strongly restricted and must be less than 0.5. Recall that stock return volatility restricts adjustment costs

---

1 One can avoid taking a stand by simply solving many models. This alternative approach is interesting for future work.
to be greater than 2 (Figure 3.3). This technological production-based restriction then implies that the EIS must be very low, less than 0.1.

Though this EIS restriction is derived in a special case, it is consistent with results from the literature on a broader set of models. Long run risk and disaster models require a large EIS. These models have had trouble matching the volatility of stock returns and tend to have low adjustment costs (Kaltenbrunner and Lochstoer (2010), Croce (2010), Gourio (2012), Kung and Schmid (2011)). Habit models and limited participation models imply a low EIS, do a good job matching stock return volatility, and also tend to have very high adjustment costs (Jermann (1998), Guvenen (2009), Chen (2012a)).

### 3.5 Generalizing Beyond Constant Returns

I have shown that for a broad class of models, the technology must feature either large adjustment costs or volatile IST shocks in order to be consistent with basic features of asset markets. The previous analysis assumes however that the profit function is homogenous of degree 1. While this is a common assumption and nests a wide range of mechanisms, one might be concerned that the results are sensitive to this assumption. In this section, I show that the results are robust to decreasing returns. I establish this both theoretical and numerical results.

Consider the general firm problem

\[
V(k, s) = \max_{(k', n)} \{\pi(k, n, k', s) + \mathbb{E}[M(s, s') V(k', s')]\}
\]

Define stock and investment returns in the standard way and the investment re-
Under this more general structure, we have equality of investment returns and stock returns as long as firm value is “factorable.”

**Proposition 3.** If there exist real-valued functions $p(k)$ and $q(s)$ such that

$$V(k, s) = p(k)q(s)$$

then for any $k, s, s'$

$$R(k, s, s') = R^I(k, s, s')$$

**Proof.** $V(k', s') = r(k')q(s')$, so stock return can be written without $k'$

$$R(k, s, s') = \frac{V(k', s')}{V(k, s) - \Pi(k, n, k', s)} = \frac{V(k', s')}{\mathbb{E}_s[M(s, s')V(k', s')] - p(k')q(s')} = \frac{V(k', s')}{\mathbb{E}_s[M(s, s')p(k')q(s')]}$$

The investment return can also be written without

$$R^I(k, s, s') = \frac{D_1(k', n', k'', s')}{D_2(k', n, n, k', s)} = \frac{D_1 V(k', s')}{\mathbb{E}_s[M(s, s')D_1 V(k', s')] - D_3 p(k')q(s')} = \frac{D_1 V(k', s')}{\mathbb{E}_s[M(s, s')D p(k')q(s')]} = \frac{D_1 V(k', s')}{\mathbb{E}_s[M(s, s')q(s')] - q(s')}$$

This proposition says that as long as firm value is “factorable,” then invest-
ment returns equal stock returns. But what does it mean that firm value is factorable? The following proposition provides an illustration

**Proposition 4.** If $\Pi(k, n, k', s)$ is HDv in $(k, n, k')$, then

$$R(k, s, s') = R^I(k, s, s')$$

*Proof.* The corollary to the Contraction Mapping Theorem shows that $V(k, s)$ is HDv in $k$. Thus we can rewrite firm value $V(k, s) = \frac{\nu}{p(k)} \frac{V(1, s)}{q(s)}$.

Proposition 4 extends the linear homogeneity case of Cochrane (1991) and Restoy and Rockinger (1994).

Factorable firm value can also be understood as an approximation of an arbitrary value function. Define the log value

$$v(k, s) \equiv \log V(k, s)$$

Taylor expand around $\bar{k}, \bar{s}$

$$v(k, s) \approx \bar{v} + D_1 v(\bar{k}, \bar{s})(k - \bar{k}) + D_2 v(\bar{k}, \bar{s})(s - \bar{s})$$

$$V(k, s) \approx e^{\bar{v} + D_1 v(\bar{k}, \bar{s})(k - \bar{k})} e^{D_2 v(\bar{k}, \bar{s})(s - \bar{s})}$$

To first order, any value function is separable. More generally, a log-Taylor expansion implies that firm value can be written in the following form

$$V(k, s) = p(k) q(s) H(k, s)$$
where

\[ \log F(k) = D_1 v(\bar{k}, \bar{s})(k - \bar{k}) + (1/2)D_1^2 v(\bar{k}, \bar{s})(k - \bar{k})^2 + ... \]
\[ \log G(k) = D_2 v(\bar{k}, \bar{s})(s - \bar{s}) + (1/2)D_2^2 v(\bar{k}, \bar{s})(s - \bar{s})^2 + ... \]
\[ \log H(k, s) = (1/2)D_1 D_2 v(\bar{k}, \bar{s})(k - \bar{k})(s - \bar{s}) + ... \]

Factorability does not require log-linearity, rather, it requires only that the cross terms in \( H(k, s) \) are small. This result suggests that factorability and the investment return stock return equality holds approximately in an even broader range of models outside of the constant returns case.

I examine this suggestion in a specific case. I add decreasing returns to the model used to impose additional restrictions (Section 3.4.3). As a reminder, this model is essentially an RBC model with Epstein-Zin preferences. I alter this model to include decreasing returns in production, that is, the firm's revenue function is

\[ y = z(k^\alpha n^{1-\alpha})^\eta \]

\( \eta \) controls the returns to scale of the firm. I continue to assume that the capital accumulation and expenditures has constant returns. Indeed, Proposition 4 shows that if both revenues, expenditures, and capital accumulation share the same returns to scale, investment returns equal stock returns.

I solve this model using projection methods, simulate it, and then compare the investment returns and stock returns. The results show that the first two moments of investment and stock returns are approximately equal for reasonable decreasing returns specifications. Figure 3.8 shows the results which indicate that even under strong decreasing returns of 0.8, the investment return volatility
remains within a few percentage points of the stock return volatility. The figure shows volatilities as a function of adjustment costs to show a range of stock return volatilities. For each adjustment cost, I choose the EIS in order to match the relative volatility of consumption to GDP of about 0.5. The left panel shows results where returns to scale are 1, and as predicted by Q theory stock return and investment return volatility are exactly equal. The middle panel shows returns to scale of 0.9 and we see some deviation between the two volatilities, though both volatilities are monotonically increasing in adjustment costs. Moreover, investment return volatility deviates from stock return volatility at most 1% per year. At severe decreasing returns of 0.8, the same pattern exists, and the volatility deviation maxes out at about 2%. Figure 3.9 shows that mean returns are also approximately equal under decreasing returns. Mean returns are not appreciably different under moderately decreasing returns of 0.9. Only when decreasing returns are severe and adjustment costs are high does a difference appear, and even in this extreme case mean investment returns only deviate from mean stock returns by about half a percent per year.

Taken together, these theoretical and numerical results show that the admissible regions produced in the main results (Section 3.4) are robust. Even under severe decreasing returns, investment returns approximately equal stock returns. If one is very concerned about severe decreasing returns, he can increase the bounds from Section 3.4 by 2%. Even in this case, adjustment costs must be very high or IST costs must be very volatile.
3.6 Conclusion

Asset prices place restrictions on production technologies. The high volatility of stock returns implies that the volatility of investment returns must be very high. This result provides an intuition for the failure of standard models. In order to match basic asset price facts, the technology must have very large adjustment costs or volatile IST shocks. This restriction applies to a broad range of models, including habit, long-run risk, disasters, extrapolative expectations, limited participation, and more.

3.7 Tables and Figures
Figure 3.1: Admissible Region  The shaded region shows admissible regions for the volatility of the investment-specific technology shock $\sigma_\mu$ and adjustment cost $\phi$ restricted by the mean and volatility of stock returns. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal } Q$. x’s and literature references represent parameter values used in the respective papers. Models in green are able to fit the equity premium. KL ’10 represents the benchmark permanent shock model from Kaltenbrunner and Lochstoer (2010). Gourio ’13 represents the unlevered equity return from his model.
Figure 3.2: Expected Investment Return: No Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows mean investment returns as a function of capital share $\alpha$ and adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi$ = the elasticity of the investment rate with respect to marginal Q. The shaded area shows the admissible region implied by expected stock returns.
Figure 3.3: Volatility of Investment Return: No Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows the standard deviation of investment returns as a function of capital share $\alpha$ and adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal } Q$. The shaded area shows the admissible region implied by stock return volatility.
Figure 3.4: Admissible Region: No Investment-Specific Technology Shocks. The shaded region shows admissible regions for the capital share $\alpha$ and adjustment cost $\phi$ restricted by the mean and volatility of stock returns. The adjustment costs are constant elasticity with $1/\phi$ = the elasticity of the investment rate with respect to marginal Q. $\chi$'s and literature references represent parameter values used in the respective papers. Models in green are able to fit the equity premium. KL ’10 represents the benchmark permanent shock model from Kaltenbrunner and Lochstoer (2010). Gourio ’13 represents the unlevered equity return from his model.
Figure 3.5: Expected Investment Return: Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows mean investment returns as a function of the volatility of the investment-specific technology shock $\sigma_\mu$ and the adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi$ = the elasticity of the investment rate with respect to marginal Q. The shaded area shows the admissible region implied by expected stock returns.
Figure 3.6: Volatility of the Investment Return: Investment-Specific Technology Shocks. Figures are annualized. This contour plot shows the volatility of investment returns as a function of the volatility of the investment-specific technology shock $\sigma_\mu$ and the adjustment cost $\phi$. The adjustment costs are constant elasticity with $1/\phi = \text{the elasticity of the investment rate with respect to marginal Q}$. The shaded area shows the admissible region implied by the volatility of stock returns.
Figure 3.7: Additional Restrictions on the Elasticity of Intertemporal Substitution. The contour plot shows the relative volatility of consumption growth as a function of the EIS and adjustment cost. Results are from a standard rbc model with Epstein-Zin preferences (Section 3.4.3). The shaded region represents admissible parameter values implied by the measured volatility of consumption growth.

Figure 3.8: Investment Return and Stock Return Volatility Under Decreasing Returns. All figures are annualized. The plot shows stock return and investment return volatility under various returns to scale and adjustment cost parameters.
Figure 3.9: Mean Investment Return and Stock Return Under Decreasing Returns. All figures are annualized. The plot shows stock return and investment return volatility under various returns to scale and adjustment cost parameters.
Table 3.1: Summary Statistics of Stock Returns and Components of the Investment Return

Figures are annual. $R_t$ is the return on the value-weighted CRSP index. $y_t/k_t$ is the output to capital ratio. $i_t/k_t$ is the investment to capital ratio, and $\Delta(i_t/k_t)$ is its first difference.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>stdev</th>
<th>AC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_t$</td>
<td>0.085</td>
<td>0.167</td>
<td>7.65E-05</td>
</tr>
<tr>
<td>$y_{t+1}/k_{t+1}$</td>
<td>0.565</td>
<td>0.035</td>
<td>0.986</td>
</tr>
<tr>
<td>$\Delta(i_{t+1}/k_{t+1})$</td>
<td>-9.40E-05</td>
<td>0.001</td>
<td>0.074</td>
</tr>
<tr>
<td>$i_{t+1}/k_{t+1}$</td>
<td>0.071</td>
<td>0.003</td>
<td>0.872</td>
</tr>
</tbody>
</table>
Bibliography


Francesco Bianchi, Cosmin Ilut, and Martin Schneider. Uncertainty shocks, asset supply and pricing over the business cycle. 2012.


Emmanuel Farhi and Xavier Gabaix. Rare disasters and exchange rates. 2013.


