Spring 2011

Essays in cross-sectional asset pricing

Scott Hogeland Cederburg
University of Iowa

ESSAYS IN CROSS-SECTIONAL ASSET PRICING

by

Scott Hogeland Cederburg

An Abstract

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Business Administration
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Professor Paul A. Weller

ABSTRACT

In this dissertation, I study the performance of asset-pricing models in explaining the cross section of expected stock returns. The finance literature has uncovered several potential failings of the Capital Asset Pricing Model (CAPM). I investigate the ability of additional risk factors, which are not considered by the CAPM, to explain these problems. In particular, I examine intertemporal risk and long-run risk in the cross section of returns. In addition, I develop a firm-level test to refine and reassess the cross-sectional evidence against the CAPM.

In the first chapter, I test the cross-sectional implications of the Intertemporal CAPM (ICAPM) of Merton (1973) and Campbell (1993, 1996) using a new firm-level approach. I find that the ICAPM performs well in explaining returns. Consistent with theoretical predictions, investors require a large positive premium for taking on market risk and zero-beta assets earn the risk-free rate. Moreover, investors accept lower returns on assets that hedge against adverse shifts in the investment opportunity set. The ICAPM explains more cross-sectional variation in average returns than either the CAPM or Fama–French (1993) model. I also investigate whether the SMB and HML factors of the Fama–French model proxy for intertemporal risk and find little evidence in favor of this conjecture.

In the second chapter, we propose an intertemporal asset-pricing model that simultaneously resolves the puzzling negative relations between expected stock return and analysts' forecast dispersion, idiosyncratic volatility, and credit risk. All three effects emerge in a long-run risk economy accommodating a formal cross section of firms characterized by mean-reverting expected dividend growth. Higher cash flow duration firms exhibit higher exposure to economic growth shocks while they are less sensitive to firm-specific news. Such firms command higher risk premiums but exhibit lower measures of idiosyncratic risk. Empirical evidence broadly supports our model's predictions, as higher dispersion, idiosyncratic volatility, and credit risk firms display lower exposure to long-run risk along with higher firm-specific risk.

Lastly, in the third chapter, we examine asset-pricing anomalies at the firm level. Portfolio-level tests linking CAPM alphas to a large number of firm characteristics suggest that the CAPM fails across multiple dimensions. There are, however, concerns that underlying firm-level associations may be distorted at the portfolio level. In this paper we use a hierarchical Bayes approach to model conditional firm-level alphas as a function of firm characteristics. Our empirical results indicate that much of the portfolio-based evidence against the CAPM is overstated. Anomalies are primarily confined to small stocks, few characteristics are robustly associated with CAPM alphas out of sample, and most firm characteristics do not contain unique information about abnormal returns.

Abstract Approved: ________________________________
Thesis Supervisor

____
Title and Department

____
Date

ESSAYS IN CROSS-SECTIONAL ASSET PRICING

by

Scott Hogeland Cederburg

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Business Administration
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Professor Paul A. Weller

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Scott Hogeland Cederburg

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Business Administration at the May 2011 graduation.

Thesis Committee: ______________________________
Paul A. Weller, Thesis Supervisor

____
David S. Bates

____
Matthew T. Billett

____
B. Ravikumar

____
Ashish Tiwari

ACKNOWLEDGEMENTS

I would like to thank my advisor, Paul A. Weller, for his support and guidance. I also thank my other dissertation committee members: David S. Bates, Matthew T. Billett, B. Ravikumar, and Ashish Tiwari. Each member provided valuable help and insights while I prepared my dissertation.

In addition to my committee members, several other faculty members, fellow PhD students, and coauthors were instrumental throughout the dissertation process. In particular, I thank Doron Avramov, Phil Davies, Redouane Elkamhi, Satadru Hore, Mike O'Doherty, and Tong Yao. Finally, I would like to thank Bev Berg and Renea Jay for their outstanding help during my time at the University of Iowa.

ABSTRACT

In this dissertation, I study the performance of asset-pricing models in explaining the cross section of expected stock returns. The finance literature has uncovered several potential failings of the Capital Asset Pricing Model (CAPM). I investigate the ability of additional risk factors, which are not considered by the CAPM, to explain these problems. In particular, I examine intertemporal risk and long-run risk in the cross section of returns. In addition, I develop a firm-level test to refine and reassess the cross-sectional evidence against the CAPM.

In the first chapter, I test the cross-sectional implications of the Intertemporal CAPM (ICAPM) of Merton (1973) and Campbell (1993, 1996) using a new firm-level approach. I find that the ICAPM performs well in explaining returns. Consistent with theoretical predictions, investors require a large positive premium for taking on market risk and zero-beta assets earn the risk-free rate. Moreover, investors accept lower returns on assets that hedge against adverse shifts in the investment opportunity set. The ICAPM explains more cross-sectional variation in average returns than either the CAPM or Fama–French (1993) model. I also investigate whether the SMB and HML factors of the Fama–French model proxy for intertemporal risk and find little evidence in favor of this conjecture.

In the second chapter, we propose an intertemporal asset-pricing model that simultaneously resolves the puzzling negative relations between expected stock return and analysts' forecast dispersion, idiosyncratic volatility, and credit risk. All three effects emerge in a long-run risk economy accommodating a formal cross sec-

tion of firms characterized by mean-reverting expected dividend growth. Higher cash flow duration firms exhibit higher exposure to economic growth shocks while they are less sensitive to firm-specific news. Such firms command higher risk premiums but exhibit lower measures of idiosyncratic risk. Empirical evidence broadly supports our model's predictions, as higher dispersion, idiosyncratic volatility, and credit risk firms display lower exposure to long-run risk along with higher firm-specific risk.

Lastly, in the third chapter, we examine asset-pricing anomalies at the firm level. Portfolio-level tests linking CAPM alphas to a large number of firm characteristics suggest that the CAPM fails across multiple dimensions. There are, however, concerns that underlying firm-level associations may be distorted at the portfolio level. In this paper we use a hierarchical Bayes approach to model conditional firm-level alphas as a function of firm characteristics. Our empirical results indicate that much of the portfolio-based evidence against the CAPM is overstated. Anomalies are primarily confined to small stocks, few characteristics are robustly associated with CAPM alphas out of sample, and most firm characteristics do not contain unique information about abnormal returns.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1
INTERTEMPORAL RISK AND THE CROSS SECTION OF EXPECTED STOCK RETURNS

1.1 Introduction

Empirical shortcomings of the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965) have elicited both theoretical and empirical responses. Theoretically, Merton (1973) and Campbell (1993, 1996) relax the strong assumption of single-period or myopic investors to derive the Intertemporal CAPM (ICAPM), wherein investors faced with time-varying investment opportunities may be willing to pay a premium for assets which hedge exposure to adverse shifts in the investment opportunity set. Empirically, multifactor models such as the Fama–French (1993) three-factor model have supplemented or even replaced the CAPM as tools to forecast or explain the cross section of returns. The ICAPM is often used as a theoretical justification for multifactor models with additional factors conjectured to be related to intertemporal hedging motives.[1] However, relatively little work has directly investigated whether intertemporal risk is priced in the manner predicted by the ICAPM.

The ICAPM has sharp implications for the cross section of returns. Similar to the static CAPM, the ICAPM predicts that zero-beta assets should earn the risk-free rate and that the market portfolio will carry a positive price of risk.

---

[1] For example, Fama and French (1995) state their results regarding the Fama–French model "are consistent with a multifactor version of Merton's (1973) intertemporal asset-pricing model in which size and BE/ME proxy for sensitivity to risk factors in returns." See also Fama and French (1996), Liew and Vassalou (2000), Vassalou (2003), and Petkova (2006) for discussions and tests of the links between the Fama–French model and the ICAPM.

Intertemporal risk may also be rewarded in equilibrium, so state variables that forecast investment opportunities may be priced in the cross section. Past empirical work motivates many additional factors on this basis, yet Campbell (1996) emphasizes that many studies do not enforce the ICAPM's theoretical implication that state variables can only be priced to the extent they forecast investment opportunities. The asset-pricing literature testing the ICAPM also often ignores the model's theoretical predictions about how intertemporal risk should be rewarded in equilibrium.[2] Campbell (1993, 1996) develops an intertemporal asset-pricing model and shows that a representative investor with a coefficient of relative risk aversion, γ, greater than one should be willing to pay a premium for assets which hedge against a decrease in expected future market returns. Campbell's model therefore allows us to investigate whether compensation for intertemporal risk is consistent with ICAPM theory.

In this paper, I use a new approach to test the cross-sectional predictions of the ICAPM. Specifically, I develop a Bayesian hierarchical model to test the ICAPM in the spirit of Fama and MacBeth (1973), with the advantage that risk factor loadings and prices of risk are simultaneously estimated. This approach, similar to the test of the CAPM by Davies (2010), mitigates the measurement error biases that typically plague cross-sectional tests. As a result, I can effectively test the ICAPM at the firm level. I therefore base inferences on the full cross section of

[2] Black (1993) also raises this issue: "The people who use data often think of the factors as rationally priced, because the factors represent risks that people care about, as in Merton's intertemporal asset pricing model. But they rarely tell us *how* the factors should be priced; they usually don't even predict the *signs* of the factor expected excess returns" (Black (1993), p. 36, his emphasis).

returns, rather than testing among portfolios where results are known to be sensitive to the choice of test assets (e.g., Ahn, Conrad, and Dittmar (2009) and Lewellen, Nagel, and Shanken (2010)). Further, testing asset-pricing models among individual firms generally achieves higher power (Litzenberger and Ramaswamy (1979) and Ang, Liu, and Schwarz (2010a)) and avoids various problems that can occur when aggregating firm returns into portfolios, such as wasting or distorting information in firm returns.[3]

Perhaps more importantly, I directly include estimated innovations in the market risk premium and real interest rate as additional factors while testing the ICAPM. The ICAPM is commonly tested by specifying a multifactor model with several macroeconomic variables as additional factors. Often, these variables are shown to predict investment opportunities, either individually or in combination. However, as noted by Campbell (1996), the ICAPM implies that the prices of risk for these state variables depends on their ability to jointly forecast investment opportunities, while most studies leave the prices of risk unconstrained. This approach does not satisfy Campbell's (1996) critique since the components of the macroeconomic variables that are orthogonal to investment opportunities are allowed to be priced. Further, examining the ICAPM's predictions about how intertemporal risks should be priced in equilibrium is difficult using this approach. In contrast, by directly including shifts in investment opportunities as additional factors, I explicitly enforce Campbell's critique and can examine the theoretical implications of the

[3]See, for example, Roll (1977), Lo and MacKinlay (1990), Kandel and Stambaugh (1995), Conrad, Cooper, and Kaul (2003), Kan (2004), Daniel and Titman (2006b), and Fama and French (2008) for potential problems with portfolios.

ICAPM by relating estimates from my approach to Campbell's (1993, 1996) model.

Over the period 1963 to 2008, I find that the ICAPM performs well when explaining firm returns. Zero-beta assets earn returns equal to the risk-free rate and market risk is significantly positively rewarded, consistent with the predictions of the ICAPM. The intertemporal risk factors based on innovations in the market risk premium and real interest rate also help explain the cross section of returns, as the ICAPM achieves a cross-sectional R^2 of 35% compared to 20% for the CAPM. Moreover, I examine whether risks are being priced in accordance with ICAPM theory. I show that the compensation for intertemporal risk is consistent with equilibrium asset pricing in an economy where the representative investor has $\gamma > 1$. Specifically, investors require a positive premium to hold assets that perform poorly when the market risk premium decreases, while they accept lower returns on assets that hedge against these unfavorable shocks. Similarly, investors require a higher premium for stocks that covary positively with changes in the real interest rate. Overall, the ICAPM performs well in explaining firm returns and several theoretical implications of the model are supported in the data.

While the results support the ICAPM and provide evidence about the economic determinants of expected returns, predicting firm returns using the ICAPM is challenging. The intertemporal risk factors have relatively low volatility and are imperfectly measured, contributing to uncertainty in estimates of expected returns. As Cochrane (2008b) notes, a factor model formed with mimicking portfolios may perform better than the true model under these conditions. The prior literature has proposed an intertemporal hedging explanation for the SMB and HML factors of the Fama–French (1993) model, suggesting these factors may serve as mimicking portfolios for the underlying intertemporal risks. I investigate this conjecture by examining the Fama–French model in the context of the ICAPM.

Perhaps surprisingly, the Fama–French model underperforms the ICAPM when explaining firm returns. The SMB and HML factors are significantly priced in the cross section, but the ICAPM explains more variation in average returns than the Fama–French model. More importantly, the components of SMB and HML that are orthogonal to investment opportunities are supplying the pricing ability of the model. There is little relation between the SMB and HML factors and the intertemporal risk factors. Additionally, alternative intertemporal risk factors formed using state variables based on SMB and HML to forecast investment opportunities perform relatively poorly and are dominated in pricing ability by the SMB and HML factors themselves. Overall, the conjecture that SMB and HML proxy for intertemporal risks is not supported in the empirical tests.

This paper contributes to the asset-pricing literature testing the ICAPM. In contrast to many previous studies, I enforce Campbell's critique and examine whether risks are priced in accordance with ICAPM theory. While Campbell (1996), Hodrick, Ng, and Sengmueller (1999), and Chen (2003) are notable exceptions to these criticisms, they approach the ICAPM using a VAR approach that necessitates the use of a small number of test assets. As such, their results may be sensitive to using alternative sets of assets. In contrast, I develop a regression-based approach that allows me to test the ICAPM across the full cross section of firm returns. This paper also contributes to the literature searching for an economic explanation of the Fama–French model's empirical success. Past studies supporting an ICAPM explanation for the Fama–French model tend to fall subject to the criticisms outlined above. While enforcing the theoretical implications of the ICAPM, I find little evidence that SMB and HML are related to intertemporal risk.

The remainder of the paper is organized as follows. Section 1.2 discusses the cross-sectional predictions of the ICAPM. Section 1.3 introduces the model and estimation procedure. Section 1.4 contains a description of the data and estimates of the determinants of the investment opportunity set. Section 1.5 tests the ICAPM and examines the Fama–French model in the context of the ICAPM. Section 1.6 concludes.

1.2 ICAPM Theory and Related Literature

In this section, I discuss the theory of the ICAPM as it relates to cross-sectional asset pricing and review the associated literature. Section 1.2.1 presents the cross-sectional predictions of the ICAPM. Section 1.2.2 discusses the findings in related research.

1.2.1 Cross-Sectional Implications of the ICAPM

Merton (1973) develops the ICAPM and draws several implications for the cross section of returns. Most importantly, Merton shows that investors may wish to hedge against shifts in the investment opportunity set. As a result, state variables that forecast investment opportunities may appear as priced risk factors in the cross section.[4] Extending this literature, Campbell (1993, 1996) develops an intertemporal asset-pricing model in discrete time, relying on a log-linear approximation to solve the model. Campbell's framework is particularly helpful for understanding the underlying intuition of the ICAPM and developing predictions about how factors should be priced in equilibrium, so I build on his model in this paper.

Following Campbell (1993, 1996), I consider an economy where every asset is tradable (including human capital) and there exists a representative agent. The representative investor has non-separable recursive utility proposed by Epstein and Zin (1989, 1991) and Weil (1989),

$$U_t = \left[(1-\beta)C_t^{(1-\gamma)/\theta} + \beta\left(E_t U_{t+1}^{1-\gamma}\right)^{1/\theta}\right]^{\theta/(1-\gamma)}, \tag{1.1}$$

where C_t is consumption, γ is the coefficient of relative risk aversion, β is a time-preference parameter, σ is the elasticity of intertemporal substitution, and $\theta = (1-\gamma)/[1-(1/\sigma)]$. The investor optimizes utility subject to the budget constraint,

$$W_{t+1} = R_{m,t+1}(W_t - C_t), \tag{1.2}$$

where W_t is wealth and $R_{m,t}$ is the gross return on the market portfolio.[5]

---

[4]While Merton (1969, 1971) and Fama (1970) show the CAPM holds in a multiperiod setting if agents' preferences and investment opportunities are not state dependent, there is ample evidence that the investment opportunity set is stochastic. For example, several variables forecast market returns, including the market dividend yield (Campbell and Shiller (1988) and Fama and French (1988)), the term and default spreads (Fama and French (1989)), book-to-market (Kothari and Shanken (1997)), and the consumption-to-wealth ratio (Lettau and Ludvigson (2001b)).

[5]Campbell (1993, 1996) provides a log-linearized approximation to the solution of the representative investor's problem under the assumption of homoskedastic asset returns and consumption growth. Within Campbell's framework, Chen (2003) examines the implications of heteroskedasticity on asset prices. Exposure to changes in market volatility affects equilibrium expected returns in this case, but Chen finds that the quantitative impact is small. I concentrate on the homoskedastic case for simplicity.

Within this economy, Campbell (1996) develops an intertemporal asset-pricing equation with no reference to consumption. Specifically,

$$E_t r^e_{i,t+1} + \frac{V_{ii}}{2} = \gamma V_{im} + (\gamma - 1) V_{ih}, \tag{1.3}$$

where lower-case letters are logs, $r^e_{i,t+1}$ denotes the excess return on stock i, V_{ii} is the variance of stock i's returns and the $V_{ii}/2$ term arises from Jensen's Inequality with log returns, the covariance between firm returns and market returns is

$$V_{im} = Cov_t \left(r^e_{i,t+1}, r^e_{m,t+1} \right), \tag{1.4}$$

and V_{ih} is a term related to intertemporal risk defined as

$$V_{ih} = Cov_t \left(r^e_{i,t+1}, (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \right). \tag{1.5}$$

The V_{ih} term is stock i's covariance with expected discounted changes in future market returns, which captures the stock's exposure to changes in the investment opportunity set.

Changes in expected market returns may be a priced risk factor according to equation 1.3. The sign of the effect on expected returns depends on the level of investor risk aversion. Investors face a tradeoff when considering the impact of future market returns on their investment decisions. Suppose a particular stock tends to earn high returns when investment opportunities improve and low returns when opportunities worsen. On one hand, investors enjoy earning high returns and having more wealth to invest during periods when wealth is most productive. On the other hand, investors dislike experiencing low returns and being poorer when investment opportunities are degrading. This tradeoff leads to some ambiguity about the premium investors require for owning a stock whose returns are correlated with shocks to expected market returns.

Investor preferences determine the optimal response to intertemporal risk. Fama (1970), Campbell (1993, 1996), and others note that investors behave myopically if $\gamma = 1$, so these investors ignore changes in the investment opportunity set. Agents with a coefficient of relative risk aversion less than one, $\gamma < 1$, prefer to be wealthy when they have the greatest opportunity to profit from their wealth. All else equal, these investors are willing to pay a higher price for stocks that pay off when investment opportunities are improving. Conversely, more risk averse investors with $\gamma > 1$ wish to hedge their exposure to changes in investment opportunities. Due to this hedging motive, these investors are willing to pay a premium for stocks which earn high returns when investment opportunities worsen, which results in less uncertainty about their consumption stream. Equilibrium risk premiums depend on the risk aversion of the representative investor.

Campbell (1993, 1996) writes the asset-pricing equation 1.3 in terms of the gross expected market return. As (Cochrane, 2005, Ch. 1) notes, however, "Our economic understanding of interest rate variation turns out to have little to do with our understanding of risk premia, so it is convenient to separate the two by looking at interest rates and excess returns separately." I take this approach by closely approximating the covariance given by equation 1.5 as the sum of two covariances,

$$V_{ih} \approx V_{i\overline{m}} + V_{ir}, \tag{1.6}$$

where

$$V_{i\overline{m}} = Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}\right), \tag{1.7}$$

$$V_{ir} = Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{f,t+j}\right), \tag{1.8}$$

$r_{f,t}$ is the risk-free rate known at the end of period t and earned during period $t+1$, and $r^e_{m,t+1}$ is the excess market return earned during period $t+1$.[6]

Working with this approximation, the asset-pricing equation 1.3 becomes

$$E_t r^e_{i,t+1} + \frac{V_{ii}}{2} = \gamma V_{im} + (\gamma - 1)V_{i\overline{m}} + (\gamma - 1)V_{ir}. \tag{1.9}$$

This equation has several immediate implications for the cross section of stock returns. First, exposure to market risk should be positively priced in the cross section, similar to the implication of the static CAPM. Second, exposures to changes in the investment opportunity set proxied by the real risk-free rate and the market risk premium may be priced. It is commonly assumed in the asset-pricing literature that $\gamma > 1$, so the reward for V_{ir} and $V_{i\overline{m}}$ should likely be positive.[7] Finally, after controlling for exposures to market risk and changes in the investment opportunity set, additional risk factors or characteristics should not be related to expected returns.

[6] See Appendix A.1.1 for further details.

[7] For example, Campbell and Cochrane (1999), Gomes, Kogan, and Zhang (2003), Bansal and Yaron (2004), and Menzly, Santos, and Veronesi (2004) assume relative risk aversion is greater than one. Relatively high risk aversion is likely needed to match the observed level of the equity premium (e.g., Mehra and Prescott (1985)). Moreover, Campbell and Cochrane (1999) contend that risk aversion needs to be relatively high to simultaneously produce high Sharpe ratios with low consumption volatility at short and long horizons.

1.2.2 Literature Review

Several papers have investigated the pricing of macroeconomic state variables in the cross section of stock returns, often citing the ICAPM for theoretical motivation. As Campbell (1996) points out, however, most of these papers do not impose the theoretical restriction from the ICAPM that state variables only be priced to the extent that they forecast investment opportunities.[8] There are some notable exceptions to this criticism: Campbell (1996), Hodrick, Ng, and Sengmueller (1999), Chen (2003), and Brennan, Wang, and Xia (2004).

As shown above, Campbell (1993, 1996) establishes an insightful framework for exploring the effect of intertemporal hedging on asset prices. He develops a model with time-varying expected market returns and labor income that is not fully captured by the proxy for aggregate wealth. Campbell (1996) uses a VAR approach to empirically implement his equilibrium multifactor model in the cross section using 10 size, 12, industry, and 3 bond portfolios. Generally, Campbell finds that the static CAPM provides a good first-order approximation to the cross section of expected returns. This occurs because innovations in both expected market returns and labor income are highly negatively correlated with realized market returns in Campbell's model. As a result, the market factor picks up a large portion of the other two sources of risk.

[8]Chen, Roll, and Ross (1986), Vassalou (2003), Hahn and Lee (2006), and Sohn (2009), among others, measure the prices of risk for macroeconomic variables without restricting parameters in accordance with their ability to forecast returns. Ferson and Harvey (1999) do impose a constraint that state variables only be priced to the extent that they are related to future returns. However, their focus is somewhat different and they include levels of the state variables in the cross-sectional regression rather than the unanticipated changes in these state variables which should be priced according to ICAPM theory.

Hodrick, Ng, and Sengmueller (1999) extend Campbell's (1993, 1996) model to an international context. They find evidence in favor of several ICAPM predictions, as the model explains the cross section of G7 country equity index returns and the estimated coefficient of relative risk aversion is reasonable at around five. However, the ICAPM is unable to explain the returns of international portfolios based on book-to-market.

Chen (2003) incorporates stochastic market volatility into Campbell's (1993, 1996) framework. In equilibrium, investors require compensation for taking on exposure to changes in market volatility. Empirically, Chen finds that expected returns for three long-short portfolios based on size, book-to-market, and momentum do not appear to be fully explained by an intertemporal hedging motive. Further, expected returns are not substantially affected by the introduction of volatility to the set of state variables for investment opportunities.

The VAR approach implemented by Campbell (1996), Hodrick, Ng, and Sengmueller (1999), and Chen (2003) has advantages and disadvantages. This methodology permits a researcher to ensure that state variables are only priced to the extent they forecast investment opportunities. Further, the method allows for an examination of the compensation for various risks to determine whether the estimates are consistent with ICAPM theory. However, these positive features do not come without a cost. The VAR approach must be implemented among a small set of test assets. Since asset-pricing model tests are sensitive to the set of test assets, we may have poor inferences about a model when the set of test assets is restricted. An alternative and more common way to test asset-pricing models is to use a two-step regression approach (e.g., Fama and MacBeth (1973)). In Section 1.3, I develop a Bayesian hierarchical model based on the asset-pricing equation (1.9) and take a regression-based approach similar to Fama–MacBeth while avoiding several problems associated with the two-step method. Using this approach, I can test the ICAPM across the full cross section of firms while maintaining the advantages noted above.

Brennan, Wang, and Xia (2004) take a regression-based approach to test the ICAPM. They posit a stochastic discount factor that is linear in market returns as well as innovations in the maximal Sharpe ratio and real interest rate. This model implies that exposure to market risk and changes in the investment opportunity set will be priced in the cross section. They filter the maximal Sharpe ratio from government bond yield data according to the theoretical relation between duration and risk premiums that arises in their model and estimate the real interest rate using bond yield and expected inflation data. Brennan, Wang, and Xia then estimate the ICAPM factor exposures of 25 size and book-to-market portfolios and 30 industry portfolios. Finally, they test whether the estimated exposures are priced in the cross section and find significant prices of risk. Similar to my approach, Brennan, Wang, and Xia implement a regression-based test of the ICAPM that enforces Campbell's (1996) critique. However, they do not examine whether rewards for risk are consistent with ICAPM theory.

A related literature examines the Fama–French (1993) three-factor model in the context of the ICAPM. Fama and French (1996) discuss the possibility that the SMB and HML factors proxy for ICAPM state variables. They suggest that systematic distress risk may underly the empirical success of the factors. Liew and Vassalou (2000) show that SMB and HML contain some information about future GDP growth. Similar in spirit, Lettau and Ludvigson (2001a) and Vassalou (2003) show that the pricing abilities of SMB and HML are at least partially shared with macroeconomic variables. Petkova (2006) points out that these studies do not directly examine the Fama–French model in an ICAPM context, since they do not consider Campbell's (1996) critique that factors can only be priced to the extent they forecast investment opportunities. Petkova includes unexpected changes in macroeconomic variables that have been shown to forecast market returns as additional factors in her specification. However, this approach does not fully satisfy Campbell's critique since components of the macroeconomic variables that are orthogonal to investment opportunities may account for the non-zero prices of risk. In contrast, I strictly impose this restriction by directly including market risk premium and real interest rate shocks as additional factors.

1.3 Model

In this section, I present my empirical framework for testing the cross-sectional implications of the ICAPM. Section 1.3.1 introduces the model and Section 1.3.2 describes the estimation procedure. Section 1.3.3 discusses the advantages and disadvantages of my methodology for testing the ICAPM.

1.3.1 Model Outline

Motivated by the asset-pricing equation (1.9), I develop a regression-based approach to test the cross-sectional implications of the ICAPM. For each asset i in each three-year period y, I estimate

$$r^e_{i,t,y} = \alpha_{i,y} + \beta^m_{i,y} r^e_{m,t,y} + \beta^{\overline{m}}_{i,y} \eta_{\overline{m},t,y} + \beta^r_{i,y} \eta_{r,t,y} + \epsilon_{i,t,y}, \quad \epsilon_{i,t,y} \sim N(0, \sigma^2_{i,y}), \qquad (1.10)$$

where $r^e_{i,t,y}$ and $r^e_{m,t,y}$ are log real excess returns in month t for asset i and the market, respectively, and $\eta_{\overline{m},t,y}$ and $\eta_{r,t,y}$ are the unexpected changes in the market risk premium and real interest rate during month t. Factor loadings for each asset i are estimated using 36 months of data and are allowed to change every three years, so I am testing conditional versions of the factor models. The model can also be easily extended to accommodate additional risk factors or characteristics.

In order to examine whether the risk factors are priced within each period, I estimate

$$\overline{r}^e_{i,y} + \frac{s^2_{i,y}}{2} = \lambda_{0,y} + \lambda_{m,y}\beta^m_{i,y} + \lambda_{\overline{m},y}\beta^{\overline{m}}_{i,y} + \lambda_{r,y}\beta^r_{i,y} + \epsilon_{i,y}, \quad \epsilon_{i,y} \sim N(0, \sigma^2_y), \qquad (1.11)$$

where $\overline{r}^e_{i,y}$ and $s^2_{i,y}$ are the mean and variance of returns for asset i in period y.[9] The $s^2_{i,y}/2$ term is a Jensen's Inequality adjustment analogous to that in the asset-pricing equation (1.9). Equations (1.10) and (1.11) measure the relation between factor loadings and average returns in each three-year period. If a particular risk is priced in the cross section, however, exposure to this risk should be systematically rewarded over time. Therefore, I aggregate the price of risk estimates across the entire sample period. In particular, I assume that λ_y is centered around the full-

[9]Since $\eta_{\overline{m},t,y}$ and $\eta_{r,t,y}$ are untraded factors, $\alpha_{i,y}$ in equation (1.10) should take the form $\alpha_{i,y} = \alpha^*_{i,y} + \lambda_{\overline{m},y}\beta^{\overline{m}}_{i,y} + \lambda_{r,y}\beta^r_{i,y}$ ((Campbell, Lo, and MacKinlay, 1996, Ch. 6)), where $\alpha^*_{i,y}$ is the stock's abnormal return relative to the ICAPM and $\lambda_{\overline{m},y}\beta^{\overline{m}}_{i,y} + \lambda_{r,y}\beta^r_{i,y}$ is the expected compensation in month t of period y for exposure to intertemporal risks. I have explored setting the prior mean of $\alpha_{i,y}$ to $\lambda_{\overline{m},y}\beta^{\overline{m}}_{i,y} + \lambda_{r,y}\beta^r_{i,y}$, but I specify a large prior variance on $\alpha_{i,y}$ so the effect of this change is trivial.

period price of risk $\overline{\lambda}$,

$$\lambda_y = \overline{\lambda} + \epsilon_y, \quad \epsilon_y \sim N(0, V_\lambda). \tag{1.12}$$

If a component of the vector $\overline{\lambda}$ is different from zero, there is evidence that the corresponding risk factor is systematically priced in the cross section. In the remainder of the paper, I refer to $\overline{\lambda}$ when discussing a price of risk.

Given test asset and market returns, as well as time series of unexpected changes in the market risk premium and real interest rate ($\eta_{\overline{m}}$ and η_r, respectively), equations (1.10) to (1.12) examine whether ICAPM risk factors are priced in the cross section. However, the market risk premium and real interest rate are unobservable, so $\eta_{\overline{m}}$ and η_r are not readily available. I now discuss the procedure for estimating these factors.

I estimate $\eta_{\overline{m}}$ using the predictive systems framework of Pástor and Stambaugh (2009). Using this methodology, market returns, the market risk premium, and a set of predictive variables x (typically macroeconomic variables) obey a VAR,

$$r^e_{m,t^*} = \overline{r}^e_{m,t^*-1} + \eta_{m,t^*}, \tag{1.13.1}$$

$$\overline{r}^e_{m,t^*} = (1 - \phi_m)E_m + \phi_m \overline{r}^e_{m,t^*-1} + \eta_{\overline{m},t^*}, \tag{1.13.2}$$

$$x_{t^*} = (I - \phi_x)E_x + \phi_x x_{t^*-1} + \eta_{x,t^*}, \tag{1.13.3}$$

$$\eta_{t^*} \sim N(0, \Sigma), \tag{1.13.4}$$

where $t^* = (y-1)T + t$ re-indexes time for ease of notation. The parameters E_m and E_x denote the long-run means of the market risk premium and predictive variable processes, while ϕ_m and ϕ_x are parameters for the speed of mean reversion of these processes. In contrast to commonly-used predictive regressions which specify the market risk premium to be a perfect linear function of predictive variables, predictive systems allow for imperfect prediction of the market risk premium. Information from the predictive variables enters the estimate of the market risk premium through (potentially) correlated errors in equations (1.13.2) and (1.13.3). Furthermore, if the market risk premium is not perfectly predicted by the state variables, there may be correlation between the errors in the market return and market risk premium (equations (1.13.1) and (1.13.2)) which can be captured by the model.

The error in equation (1.13.2), $\eta_{\overline{m}}$, is the unexpected change in the market risk premium which enters into the regression equation (1.10). This term corresponds to the $V_{i\overline{m}}$ term in the asset-pricing equation (1.9), which can be written,[10]

$$V_{i\overline{m}} = \frac{1}{\phi_m(1-\phi_m\rho)} Cov_t\left(r^e_{i,t+1}, \eta_{\overline{m},t+1}\right). \tag{1.14}$$

ICAPM theory therefore implies that asset i's risk premium may depend on the covariance of its returns with $\eta_{\overline{m}}$, which enters directly into the regression equation (1.10) as an additional factor.

Finally, I estimate η_r using a Forward Filtering, Backward Sampling (FFBS) technique.[11] Specifically, I assume that the real interest rate and expected inflation

---

[10]See Appendix A.1.2 for derivation.

[11]FFBS was developed by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Also see West and Harrison (1997) for additional information about FFBS.

obey

$$r_{n,t^*} = r_{f,t^*} + E_{t^*}[\pi_{t^*+1}] + \epsilon_{r,t^*+1}, \tag{1.15.1}$$

$$\pi_{t^*+1} = E_{t^*}[\pi_{t^*+1}] + \epsilon_{\pi,t^*+1}, \tag{1.15.2}$$

$$r_{f,t^*} = \phi_r r_{f,t^*-1} + \eta_{r,t^*}, \tag{1.15.3}$$

$$E_{t^*}[\pi_{t^*+1}] = \phi_\pi E_{t^*-1}[\pi_{t^*}] + \eta_{\pi,t^*}, \tag{1.15.4}$$

$$\epsilon_{t^*+1} \sim N(0, V), \tag{1.15.5}$$

$$\eta_{t^*} \sim N(0, W), \tag{1.15.6}$$

where $r_{n,t}$ is the nominal interest rate known at the end of month t and earned during month $t+1$, $r_{f,t}$ is the real risk-free rate, and π_t is realized inflation during month t. In this system, the nominal interest rate is the sum of the real interest rate and expected inflation plus a small measurement error. I further assume that expected inflation is an unbiased estimate of realized inflation. Finally, the real interest rate and expected inflation follow (potentially) correlated AR(1) processes. The unexpected change in the real interest rate that enters into the regression equation (1.10) is the error term from equation (1.15.3), η_r, since[12]

$$V_{ir} = \frac{1}{\phi_r(1-\phi_r\rho)} Cov_t\left(r^e_{i,t+1}, \eta_{r,t+1}\right). \tag{1.16}$$

The asset-pricing equation (1.9) implies that, assuming $\gamma > 1$, the expected return on asset i should be positively related to the covariance of its returns with changes in the risk-free rate and market risk premium. We can recover estimates of the coefficients on V_{im}, $V_{i\overline{m}}$, and V_{ir} from equation (1.12) of the regression approach.

[12]See Appendix A.1.2 for derivation.

Rewriting equation (1.9) to allow these coefficients to be free parameters and allow for a non-zero intercept, we have

$$E_t r^e_{i,t+1} + \frac{V_{ii}}{2} = b_0 + b_m V_{im} + b_{\overline{m}} V_{i\overline{m}} + b_r V_{ir}. \tag{1.17}$$

Estimates of the b coefficients in equation (1.17) can be made using $\overline{\lambda}$ from equation (1.12). Specifically,[13]

$$b_0 = \overline{\lambda}_0 \tag{1.18}$$

and

$$\begin{bmatrix} b_m \\ b_{\overline{m}} \\ b_r \end{bmatrix} = (\Sigma Z Z')^{-1} \begin{bmatrix} \overline{\lambda}_m \\ \overline{\lambda}_{\overline{m}} \\ \overline{\lambda}_r \end{bmatrix}, \tag{1.19}$$

where Σ is the covariance matrix of the market excess returns and intertemporal risk factors and Z is a scaling factor,

$$Z = \begin{bmatrix} 1 \\ \sqrt{\phi_m(1-\phi_m\rho)} \\ \sqrt{\phi_r(1-\phi_r\rho)} \end{bmatrix}. \tag{1.20}$$

I set $\rho = 0.9949$ following Campbell (1996). In equilibrium, we expect $b_0 = 0$, $b_m > 1$, $b_{\overline{m}} > 0$, and $b_r > 0$ when $\gamma > 1$.

1.3.2 Model Estimation

The model in equations (1.10) to (1.13) and equation (1.15) involves a high-dimensional parameter space since firm-specific parameters must be estimated for thousands of firms each period. As such, using maximum likelihood estimation is

[13]See Appendix A.1.3 for further details.

computationally challenging. An alternative estimation strategy is the generalized method of moments (GMM). An advantage of GMM is that it does not require strong distributional assumptions on parameters. However, Ferson and Foerster (1994) show that GMM has rather poor finite-sample properties, especially when the number of parameters and test assets is large.

I adopt a Bayesian hierarchical approach to estimate the model.[14,15] The principle advantages of using a Bayesian framework are exact finite-sample inference and a complete accounting of parameter uncertainty. For example, uncertainty about $\eta_{\overline{m}}$ and η_r as well as assets' loadings on these factors is fully reflected in the posterior distribution of $\overline{\lambda}$.

As a result of taking a Bayesian approach, I must specify explicit priors. I specify proper and diffuse priors for all model parameters with the following exceptions. I adopt Pástor and Stambaugh's (2009) approach by introducing some prior information about the nature of the market risk premium, including information regarding the long-run average market return, the speed of mean reversion, and the correlation between shocks to the market risk premium and current market returns. Further discussion on these prior parameter choices is included in Appendix A.2

---

[14]Several papers have used Bayesian techniques to examine asset-pricing models. McCulloch and Rossi (1991) and Geweke and Zhou (1996) develop Bayesian analyses of the Arbitrage Pricing Theory (APT), while Shanken (1987), Harvey and Zhou (1990), Kandel, McCulloch, and Stambaugh (1995), and Cremers (2006) propose Bayesian tests for the mean-variance efficiency of a given portfolio. Ang and Chen (2007) use Bayesian methods to examine whether the conditional CAPM can explain the value premium. Cosemans, Frehen, Schotman, and Bauer (2009), Davies (2010), and Cederburg, Davies, and O'Doherty (2010) use Bayesian approaches to test the CAPM.

[15]See (Rossi, Allenby, and McCulloch, 2005a, Ch. 5) for a discussion of Bayesian hierarchical models.

and in Pástor and Stambaugh (2009). Inferences are not sensitive to reasonable alterations to the prior parameters.

Within a Bayesian framework, I estimate my model using a Markov chain Monte Carlo (MCMC) technique. Within each iteration, I first draw time series of $\eta_{\overline{m}}$ and η_r from the systems of equations (1.13) and (1.15), respectively. Drawing new time series of these factors in each iteration ensures that my posterior distributions fully reflect the uncertainty about these unobserved sequences.[16] I then use these draws as factors in equation (1.10) when estimating equations (1.10) to (1.12). In contrast to the traditional Fama–MacBeth approach which estimates these equations sequentially, I estimate this system of equations simultaneously as a Bayesian hierarchical model. More specifically, equation (1.12) is a hierarchical prior for the λ_y parameters in equation (1.11), while equation (1.11) acts as a hierarchical prior for the $\beta_{i,y}$ parameters in equation (1.10). My approach to estimating equations (1.10) to (1.12) is similar to that of Davies (2010), who tests the cross-sectional implications of the CAPM using a one-step procedure.

1.3.3 Model Discussion

My approach to testing the ICAPM has several advantages. From an economic perspective, I examine whether risks are priced in accordance with ICAPM theory while imposing the theoretical restrictions implied by the model. From a methodological perspective, I use a one-step procedure to estimate factor loadings

[16] In this paper, I assume that the cross section of returns contains no additional information about $\eta_{\overline{m}}$ and η_r. I include details in Appendix A.2 about how to incorporate this extra information.

and prices of risk to mitigate biases from measurement error, test the ICAPM at the firm level, and implement Pástor and Stambaugh's (2009) predictive systems approach to obtain better estimates of the market risk premium. Below, I discuss these features of my methodology.

While testing the ICAPM, I directly include $\eta_{\overline{m}}$ and η_r as risk factors in the regression equation (1.10). A common alternative for testing the ICAPM is to show that a set of state variables has some ability to forecast investment opportunities and then form a multifactor model using these state variables as additional factors. However, this approach is susceptible to Campbell's critique. The multifactor model could describe the cross section of returns for reasons unrelated to intertemporal risk. In particular, the components of the state variables that contain no incremental information about investment opportunities may provide the pricing ability of the model. By directly including $\eta_{\overline{m}}$ and η_r in equation (1.10), state variables are only allowed to explain returns through the channel of the intertemporal risk factors which satisfies Campbell's critique. In addition, examining $\eta_{\overline{m}}$ and η_r simplifies the task of determining whether intertemporal risks are priced in accordance with the theoretical predictions of the ICAPM outlined in Section 1.3.1. In contrast, determining the anticipated sign for several state variables that may or may not predict investment opportunities is a more difficult task, and consequently often ignored.

Regarding the estimation methodology, the first important aspect of my approach is the simultaneous estimation of equations (1.10) to (1.12). As Davies (2010) notes, this one-step approach allows the researcher to retain several desirable features of the Fama–MacBeth (1973) methodology while reducing or eliminating several issues with a two-step approach. Both methods allow firm factor loadings to change over time, allowing the researcher to test conditional versions of factor models. Further, the one-step approach implicitly reflects any heteroskedasticity and cross-correlations in the cross section of firm returns, similar to this feature of Fama–MacBeth shown by Shanken (1992). Therefore, I can test a model using a large number of test assets without the need to estimate a large variance-covariance matrix.

While maintaining the positive features of the Fama–MacBeth (1973) procedure, the one-step approach makes important improvements. Notably, this method mitigates the errors-in-variables (EIV) problem that commonly plagues cross-sectional tests of asset-pricing models. The EIV problem arises during the second step of Fama–MacBeth tests, a cross-sectional regression of average returns on estimated factor loadings. This issue is avoided when the model is estimated in one step. Davies (2010) shows that measurement errors produce strongly biased estimates of prices of risk in tests of the CAPM, and inferences about the model change substantially when the one-step approach is used. Measurement error is likely to be an even larger problem when testing the ICAPM using the traditional approach. The relatively low volatility of $\eta_{\overline{m}}$ and η_r reduces the precision of factor loading estimates, which would result in an understatement of the importance of intertemporal risks.

The characteristics of the one-step approach lead to the second advantage of my approach, which is the ability to effectively use individual firms as test assets. As noted by Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973), among others, the EIV problem in two-step tests is exacerbated by the low precision of firm factor loading estimates. In response, researchers typically turn to portfolios in tests of asset-pricing models. However, several potential problems arise from using portfolios. Litzenberger and Ramaswamy (1979) and Ang, Liu, and Schwarz (2010a) argue that valuable information is lost when forming portfolios, leading to lower power in asset-pricing tests. Further, information about firm-level returns can be distorted when forming portfolios, potentially resulting in poor inferences about a model (e.g., Roll (1977), Kandel and Stambaugh (1995), and Fama and French (2008)).[17] Consistent with these assertions about problems with portfolios, inferences about asset-pricing models appear to be quite sensitive to the choice of test assets (e.g., Daniel and Titman (2006b), Ahn, Conrad, and Dittmar (2009), and Lewellen, Nagel, and Shanken (2010)). Directly using individual firms avoids these issues, and the one-step approach is well suited for testing models using the full cross section of firms.

The third key aspect of my approach lies in the estimation of the innovations in the market risk premium. Estimates of the price of risk for $\eta_{\overline{m}}$ will generally be more precise the better is the estimate of the premium. The market risk premium can be estimated using a variety of techniques.[18] I work within the predictive systems framework of Pástor and Stambaugh (2009). Unlike predictive regressions

---

[17]For other concerns, see Lo and MacKinlay (1990), Conrad, Cooper, and Kaul (2003), and Kan (2004).

[18]Methods for estimating the market risk premium include predictive regressions (e.g., Fama and Schwert (1977) and Campbell (1987)), valuation models (e.g., Fama and French (2002)), and the implied cost of capital (Pástor, Sinha, and Swaminathan (2008)).

that make the unrealistic assumption that the market risk premium is perfectly predicted by a set of state variables, Pástor and Stambaugh's methodology allows for imperfect prediction. In this context, additional information about the market risk premium can be taken from market returns through correlated errors.[19] Pástor and Stambaugh (2009) present evidence that the predictive systems outperform predictive regressions when predicting market returns.

ICAPM investors may require compensation for taking on exposure to changes in the investment opportunity set. In this paper, I take the market risk premium and real interest rate as state variables that forecast investment opportunities. However, investors may be interested in additional state variables, such as market volatility. Chacko and Viceira (2005) develop a model with stochastic volatility in incomplete markets and find that market volatility is negatively priced by investors. In practice, however, they contend that volatility is not variable and persistent enough to generate large intertemporal hedging demands. Similarly, Chen (2003) finds that hedging demand related to market volatility does not cause statistically or economically significant differences in expected returns. Therefore, I concentrate on the market risk premium and real interest rate to test the ICAPM.

[19]The correlation between market return and market risk premium errors is likely negative. While negative correlation between current market returns and the market risk premium leads to a potential explanation of the excess volatility puzzle of LeRoy and Porter (1981) and Shiller (1981) (e.g., Cochrane (1992)), positive correlation would deepen the puzzle (Pástor and Stambaugh (2009)). For empirical evidence of negative correlation, see Campbell (1991), Campbell and Ammer (1993), and van Binsbergen and Koijen (2010).

1.4 Data and Preliminary Estimation

In this section, I discuss the data used in testing the ICAPM and show estimates of the market risk premium and real interest rate. Section 1.4.1 concentrates on $\eta_{\overline{m}}$ and η_r, while Section 1.4.2 examines the other risk factors included in the asset-pricing tests. Section 1.4.3 discusses the test assets used for testing the ICAPM.

1.4.1 Investment Opportunity Set

Changes in the investment opportunity set are proxied by unexpected shocks to the market risk premium and real interest rate. I estimate $\eta_{\overline{m}}$ over the period July 1963 to June 2008 using a predictive system. Following Petkova (2006), I use the aggregate dividend-to-price ratio, term spread, risk-free rate, and the default spread as macroeconomic state variables.

Figure 1.1 shows estimates of the market risk premium using both a predictive system and a predictive regression. While the estimates from the two methods appear to be closely related, there are several instances where the approaches disagree about the level and changes in the market risk premium. Mean draws from the posterior distribution of the predictive system parameters in equation (1.13) appear in Panel A of Table 1.1. Consistent with the prior literature, the four state variables appear to provide information about market returns. Additional information about the market risk premium is provided through the correlated errors between expected and realized market returns, while the predictive regression ignores this information.

Consistent with Pástor and Stambaugh (2009), I find that the predictive system forecasts stock market returns more accurately than the predictive regression. Market risk premium estimates from a predictive system have an adjusted R^2 of 5.7% in predicting the next month's excess return on the value-weighted CRSP index, while a predictive regression achieves an adjusted R^2 of 5.4%. All else equal, a more accurate market risk premium estimate should lead to a more precise estimate of the price of risk for $\eta_{\overline{m}}$.

To estimate the real interest rate, I use information about Treasury bill yields and realized inflation. I use the one-month Treasury bill rate from Kenneth French's website as the nominal interest rate. The personal consumption expenditures (PCE) price index serves as my proxy for realized inflation.[20] PCE is a chained index which adjusts for consumer substitution across goods in response to price changes, in contrast to the consumer price index (CPI) which is a fixed-basket index. I use core inflation data to eliminate the effects of transitory movements in food and energy prices which are less likely to influence expected inflation.

Figure 1.2 graphs the estimates of the nominal interest rate and expected inflation during the sample period. The difference between the nominal interest rate and expected inflation is the real interest rate, shown in Figure 1.3. The real interest rate tends to be fairly persistent during normal economic periods, and is generally punctuated by increases prior to NBER recessions and sharp decreases during recessionary periods. The mean parameter draws for the system of equations (1.15) appear in Panel B of Table 1.1.

---

[20] PCE data is available at http://research.stlouisfed.org/fred2/.

1.4.2 Risk Factors

In the previous section, I estimate time series of $\eta_{\overline{m}}$ and η_r, which may be priced in the cross section according to ICAPM theory. I therefore include these variables as factors in the cross-sectional tests. I annualize the $\eta_{\overline{m}}$ and η_r factors by multiplying monthly estimates by 12. I assume the excess return on the aggregate wealth portfolio is a linear function of the excess return on the value-weighted CRSP stock market index. Factor return data for the Fama–French (1993) three-factor model is from Kenneth French's website.

Table 1.2 shows summary statistics of the risk factors. Panel A reports averages and standard deviations. The series of $\eta_{\overline{m}}$ and η_r are less volatile than the returns on the market, SMB, and HML portfolios. Panel B of Table 1.2 contains the correlation matrix for the risk factors. The $\eta_{\overline{m}}$ factor is strongly negatively correlated with market returns as expected, while it has lower correlations with SMB and HML. The η_r factor, on the other hand, is nearly uncorrelated with market returns, SMB, and HML.

1.4.3 Test Assets

I conduct tests of the ICAPM using individual firms as test assets. I include all firm-periods on the NYSE, Amex, or NASDAQ exchanges with the required data available on CRSP and Compustat. I exclude financial firms (SIC code 6000–6999) from the sample. I use three-year periods and monthly subperiods. Since the sample period is July 1963 to June 2008, there are 15 three-year periods in the sample. I require that each firm has all monthly returns during the period for the firm-period to be included in the sample. The final sample includes $37,648$ firm-period observations. I use real returns in all tests.

Firm size and book-to-market are calculated for each firm year using data from CRSP and Compustat. Let j denote the first year of period y. Size is the natural log of price per share times the number of outstanding shares at the end of June of year j. I calculate the natural log of the ratio of book value of equity to market value of equity. Following Fama and French (2008), the book value of equity is total assets (at), minus total liabilities (lt), plus balance sheet deferred taxes and investment tax credits (txditc) if available, minus the book value of preferred stock if available. Depending on availability, I use liquidating value (pstkl), redemption value (pstkrv), or carrying value (upstk) for the book value of preferred stock. The market value of equity is price per share times the number of shares outstanding at the end of December of year $j-1$. Firms with negative book equity are excluded from the sample. Panel C of Table 1.2 reports summary statistics for the test assets during the sample period.

1.5 Asset-Pricing Model Tests

In this section, I test the cross-sectional implications of the ICAPM and examine the Fama–French (1993) three-factor model in the context of the ICAPM. Section 1.5.1 presents the results of testing the CAPM and ICAPM. Section 1.5.2 investigates whether the empirical success of the Fama–French model can be attributed to the intertemporal hedging motives of the ICAPM.

To test the ICAPM, I examine the intercept and prices of risk for the market factor and intertemporal risk factors from equation (1.12). If the ICAPM holds, the intercept should equal zero and the market factor should carry a positive price of risk. The $\eta_{\overline{m}}$ and η_r factors may also have non-zero prices of risk according to the ICAPM. I further estimate the elements of b in equation (1.17) to examine whether the prices of risk are consistent with ICAPM theory.[21]

To help assess model performance, I report the R^2 from the cross-sectional regression in equation (1.11). I average the cross-sectional R^2 across periods to measure performance over the full sample period. The cross-sectional R^2 is based on Gelman and Pardoe (2006), who describe a method to calculate explained variance for each level of a hierarchical model. I concentrate on the cross-sectional R^2 from equation (1.11) as an informal measure of model performance to achieve comparability with the previous literature (e.g., Jagannathan and Wang (1996), Lettau and Ludvigson (2001a), Lustig and Van Nieuwerburgh (2005), and Li, Vassalou, and Xing (2006)). Lewellen, Nagel, and Shanken (2010) caution against using cross-sectional R^2 as the sole measure of model performance among the 25 portfolios sorted by size and book-to-market of Fama and French (1993), but I examine the performance of models among individual firms which are not plagued by the same problems as these portfolios. In this case, the cross-sectional R^2 should provide useful information about a model's ability to explain returns.[22]

---

[21] Statistical significance in the empirical analysis is based upon Bayesian credible intervals. I also refer to the mean of the draws from the posterior distribution as the parameter estimate.

[22] In unreported results, also examined the Akaike Information Criterion–Monte Carlo ($AICM$) and Bayesian Information Criterion–Monte Carlo ($BICM$) model fit measures of Raftery, Newton, Satagopan, and Krivitsky (2007), which are Bayesian analogues to the

1.5.1 Testing the ICAPM

Table 1.3 shows the results from testing the ICAPM, with the CAPM used as a baseline model. Beginning with the CAPM, the model appears to work reasonably well. As the CAPM predicts, market risk is strongly positively rewarded, as over 99% of the posterior distribution of $\overline{\lambda}_m$ lies above zero and the posterior mean is 0.73%. This evidence is consistent with the findings of Davies (2010) for individual firms and Ahn, Conrad, and Dittmar (2009) using basis assets formed on desirable statistical properties. Moreover, the posterior distribution of $\overline{\lambda}_0$ is centered very near zero at 0.04%, so zero-beta assets are earning returns near the risk-free rate. The evidence suggests the CAPM's predictions about the intercept and price of market risk hold reasonably well among individual firms.

Once intertemporal risk factors are included, I find that the ICAPM outperforms the CAPM when explaining the cross section of firm returns. The ICAPM intercept estimate is near zero (0.14%) and market risk carries a large positive premium (0.69%), consistent with the predictions of the ICAPM. The intertemporal risk factors also help explain average returns. The posterior distribution of $\overline{\lambda}_r$ is centered at 0.06% and over 90% of the posterior lies above zero. The estimate of $\overline{\lambda}_{\overline{m}}$ is -0.14%. The precision on this estimate is relatively low and only about 73% of the posterior distribution lies below zero. This result is partially explained by the correlation of -0.55 between the market and $\eta_{\overline{m}}$ factors, which contributes to the difficultly of estimating $\overline{\lambda}_{\overline{m}}$ with a high degree of precision. However, the estimated

---

Akaike (1974) and Schwarz (1978) information criteria. Generally, the ICAPM provides a worse fit than the CAPM among individual firms, which is unsurprising due to the uncertainty surrounding loadings on intertemporal hedging factors.

price of risk is economically large and $\eta_{\overline{m}}$ contributes to pricing the cross section of returns. The cross-sectional R^2 for the ICAPM specification is about 35%, compared to only 20% for the CAPM. The improvement in pricing is reflected in Panels A and B of Figure 1.4.[23] Differences in fitted returns for the CAPM and ICAPM are most pronounced for firms with more extreme values of average returns.

Perhaps more interesting than the prices of risk discussed above are the estimated coefficients of the asset-pricing equation (1.17). With these estimates, we can investigate whether intertemporal risks are priced in accordance with ICAPM theory. Consistent with ICAPM predictions, the intercept b_0 is nearly zero at 0.14. Further, covariance between firm and market returns is positively rewarded with a b_m estimate of 4.64. Finally, covariances between firm returns and both $\eta_{\overline{m}}$ and η_r are positively related to average returns. The $b_{\overline{m}}$ estimate associated with $\eta_{\overline{m}}$ is 1.47, while b_r is estimated to be 0.89. The posterior probability that the three coefficients are jointly positive is nearly 98%. Therefore, consistent with the theoretical implications of the ICAPM, investors appear to be requiring compensation for taking on exposure to market risk as well as exposures to adverse shocks to the investment opportunity set. More specifically, investors require higher returns for stocks that do poorly when the market risk premium or real interest rate decrease, while they accept lower returns on assets that provide a hedge against worsening investment opportunities. This behavior is consistent with the equilibrium actions

[23]Figure 1.4 contains scatter plots of fitted versus average returns of firms, where the returns of firms are aggregated by equally weighting firms within each of the 25 size and book-to-market classifications. The size and book-to-market quintile breakpoints are from Kenneth French's website.

of a representative agent with $\gamma > 1$. There are, however, some implications of Campbell's (1996) model that do not appear to hold perfectly. Equation (1.9) implies that $b_m = b_{\overline{m}} + 1 = b_r + 1$, while the posterior probability that $b_m > b_{\overline{m}} + 1$ and $b_m > b_r + 1$ is nearly 99%. Overall, however, there is substantial empirical support for intertemporal risk as an important economic determinant of expected returns.

The firm-level results in this paper add to Campbell's (1996) portfolio-level inferences about the ICAPM. As noted by Campbell, the CAPM will suffice for cross-sectional asset-pricing under three conditions: (i) γ is exactly one, (ii) no stocks are exposed to intertemporal risks, or (iii) the factor loadings on the market are perfectly cross-sectionally correlated with loadings on the intertemporal risk factors. Similar to Campbell, I find that the first two possibilities are not likely to hold. First, the estimated b_m coefficient, which translates to γ in Campbell's model, for the CAPM is 3.76 with over 99% of the posterior distribution lying above one. Therefore, investors should not behave myopically and ignore intertemporal risks. Second, there is strong evidence in this paper and in the preceding literature that investment opportunities are time varying, and firms carry non-zero exposures to this risk. The third possibility implies that the CAPM market factor perfectly captures exposure to intertemporal risks because of the relation between the risk factors. Campbell finds that this relation holds reasonably well among portfolios. In contrast, I find that intertemporal risks contribute to cross-sectional asset pricing among individual firms. While there is a fairly large negative correlation between market returns and $\eta_{\overline{m}}$ (about -0.55), the intertemporal risk factors contain information about the cross section of returns that is not captured by the static CAPM. Specifically, $\eta_{\overline{m}}$ and η_r

are able to capture much of the variation in average returns left unexplained by the CAPM and improve cross-sectional asset pricing.

I find substantial evidence in favor of the ICAPM's predictions. As suggested by Berk (1995) and Jagannathan and Wang (1998), however, including firm characteristics in the model allows the researcher to detect model misspecification. I therefore include firm size and book-to-market as characteristics in equation (1.11) and rerun the asset-pricing model tests. Table 1.4 contains the results. Size and book-to-market are related to returns even after controlling for exposure to the CAPM and ICAPM factors. Over 99% of the posterior distribution of $\overline{\lambda}_{B/M}$ lies above zero for both models, while over 90% of the posterior of $\overline{\lambda}_{Size}$ is below zero. The two characteristics are strongly related to average returns.

After controlling for the ICAPM factors, no additional factors or characteristics should be priced. Therefore, the non-zero relations between average returns and the size and book-to-market characteristics does provide evidence that the ICAPM is misspecified in some way. There are several potential explanations of this result. First, important components of the investment opportunity set may be omitted in my specification of the ICAPM. Second, similar to the Roll (1977) critique of the CAPM, the market portfolio may not be a linear function of the CRSP value-weighted stock index. Third, the ICAPM may be an incomplete description of reality and may omit important risk factors such as human capital (e.g., Mayers (1972)). On the other hand, data mining may produce a large number of characteristics – perhaps including size and book-to-market – that appear to be unexplained by an asset-pricing model even when the model works well, leading to spurious relations between these variables and average returns when the same data is used to test the model.

In sum, the main implications of the ICAPM hold reasonably well in the cross section. Market risk and intertemporal risks are related to average returns as predicted by ICAPM theory. Investors require higher returns for stocks that do poorly when aggregate wealth declines and when investment opportunities are degrading. Further, the $\eta_{\overline{m}}$ and η_r factors contribute to a substantially higher R^2 for the ICAPM relative to the CAPM.

1.5.2 The ICAPM and the Fama–French Model

The ICAPM's predictions are generally supported, suggesting that both market risk and intertemporal risks are important economic determinants of the cross section of expected returns. Despite its theoretical attractiveness and success in cross-sectional tests, however, the ICAPM is difficult to apply in many applications. For example, using the ICAPM may not be optimal for determining the expected returns of individual firms for portfolio choice or cost of capital purposes, since the uncertainty surrounding the intertemporal risk factor prices and loadings may lead to substantial errors in predicting the returns of any single firm.

In this context, we may be able to improve performance by choosing a factor model that augments the market portfolio with factor portfolios that mimic the intertemporal risk factors. As Cochrane (2008b) notes, "if we have the perfect model of the marginal utility of wealth, then a portfolio formed by its regression on to asset returns will work just as well." Further, this new factor model will have more frequently observed data which will improve performance in empirical applications relative to the true model. In line with this argument, Campbell and Cochrane (2000) demonstrate that the CAPM outperforms the consumption CAPM in asset-pricing model tests even when the consumption CAPM is a perfect conditional asset-pricing model. A similar argument applies to the ICAPM when intertemporal risk factors are imperfectly or infrequently measured.

Motivated by the empirical advantages of mimicking portfolio models, I turn to investigating the Fama–French (1993) three-factor model in the context of the ICAPM. The ICAPM has often been suggested as an underlying theoretical motivation for the Fama–French model. Fama and French (1996) discuss their model in the context of the ICAPM and present some tests of whether the ICAPM or APT explain the SMB and HML factors. Liew and Vassalou (2000), Vassalou (2003), and Petkova (2006) find links between the SMB and HML factors and macroeconomic variables, contending that the Fama–French factors are proxying for ICAPM state variables. Identifying intertemporal risks as the economic cause of the empirical success of the Fama–French model would suggest the additional factors are likely to continue to help price the cross section of assets and provide evidence in favor of using the Fama–French model in applications requiring predictions of expected returns. I therefore examine the impact of jointly including the $\eta_{\overline{m}}$ and η_r factors along with SMB and HML and investigate whether the pricing ability of SMB and HML arises through an ICAPM channel.

Table 1.5 presents tests of several asset-pricing models. The results of the CAPM and ICAPM tests in Table 1.3 are reprinted for ease of comparison. In addition, the table reports estimated prices of risk for the Fama–French three-factor model and a model with both the ICAPM intertemporal risk factors and the Fama–French factors.

The Fama–French model performs about as well as the ICAPM at explaining firm returns. The estimate of $\overline{\lambda}_0$ is 0.18%, which is still indistinguishable from zero in a statistical sense, and market risk is strongly positively rewarded with over 99% of the posterior lying above zero. The SMB and HML factors also contribute to pricing the cross section. The price of risk for HML is highly significant, while about 90% of posterior for the corresponding estimate for SMB lies above zero. The cross-sectional R^2 for the Fama–French model is 34% compared to 21% for the CAPM, so the SMB and HML factors are contributing to explaining firm returns. The ICAPM does slightly outperform the Fama–French model, with the models achieving an R^2 of 35% and 34%, respectively. There is some cause for concern, however, for the Fama–French model as the estimate of $\overline{\lambda}_{HML}$ is negative. Since HML is a traded portfolio, its price of risk should be equal to its average return (e.g., (Campbell, Lo, and MacKinlay, 1996, Ch. 6)) which is strongly refuted by the data. The evidence of a negative $\overline{\lambda}_{HML}$ is consistent with Ahn, Conrad, and Dittmar (2009) who also test the Fama–French model using alternative test assets.

I now investigate whether the SMB and HML factors are related to intertemporal risk. The evidence does not support an ICAPM explanation for the additional factors. When all factors are included in the model, $\overline{\lambda}$ estimates are similar to those form the base models. Further, the cross-sectional R^2 improves relative to both the ICAPM and Fama–French model, suggesting that the $\eta_{\overline{m}}$ and η_r factors contain largely different information than SMB and HML. The non-zero prices of risk for SMB and HML do provide some evidence against the ICAPM, since no additional factors should be priced. Overall, the evidence suggests that the Fama–French model and the ICAPM are working through different channels when pricing individual stocks.[24]

Table 1.5 provides some evidence that intertemporal hedging motives may not explain the empirical success of the Fama–French model. I now move to an alternative test that allows SMB and HML to act as predictive variables. Suppose an empirically constructed factor, such as SMB or HML, is a factor-mimicking portfolio proxying for an unobserved state variable. Then the factor portfolio returns have two components, one of which is perfectly correlated with the underlying state variable and another orthogonal noise component. In the ICAPM, exposure to the orthogonal component should be unpriced. In the following tests, I allow state variables related to SMB and HML to forecast investment opportunities by acting as predictive variables for the market risk premium. Any remaining exposure to SMB and HML after controlling for this predictive ability should not be priced.

I construct predictive variables using information from SMB and HML. As Brennan, Wang, and Xia (2001) note, SMB and HML returns are technically inappropriate for use as state variables since the returns follow a geometric Brownian motion in the limit. Discretely measured SMB and HML returns can be viewed

[24]In unreported results, I investigate whether the connection between the Fama–French model and the ICAPM using 25 portfolios sorted by size and book-to-market as test assets. While the Fama–French model performs well in explaining the returns of these portfolios, the pricing ability of the model appears to be largely unrelated to the intertemporal risk factors of the ICAPM.

as noisy signals of the information underlying the expected returns of these factor portfolios, but using state variables related to the relative prices of the SMB and HML component portfolios is likely to lead to better return predictions. Briefly, these price-based measures are constructed as the log of the book-to-market of a portfolio with equal investments in the small-growth, small-neutral, and small-value portfolios minus the log of the analogously constructed book-to-market for the big portfolios. Therefore, the SMB-based state variable is large when big stocks are expensive relative to small stocks. Similarly, I construct an HML-based state variable that is high when growth stocks are relatively expensive. See Appendix A.1.4 for additional details on state variable construction.[25]

Table 1.6 contains the results of testing the ICAPM with the SMB and HML price factors as predictive variables. The ICAPM using the Fama–French factors as predictive variables underperforms the ICAPM in Table 1.5 that utilizes macroeconomic variables, achieving a cross-sectional R^2 of only 0.30 compared to 0.35 for the base ICAPM specification. There is some weak initial evidence that the intertemporal risk factors may be priced as predicted by the ICAPM, since there is a 77% chance that the b coefficients are positive. However, adding the raw SMB and HML factors to the model degrades the performance of the intertemporal risk factors. The $\eta_{\overline{m}}$ and η_r factors using SMB and HML state variables share explanatory power

[25]Portfolio rebalancing at the end of each June creates the potential for changes in the state variables that are unrelated to underlying economic changes. I have explored several methods of adjusting for rebalancing as described in Appendix A.1.4. For the base case, I include the state variables described in the text as well as two additional variables which capture the changes in the state variables resulting from rebalancing and take the value of zero when no rebalancing occurs. Test results are nearly identical across specifications so I show only results for the base case.

with the SMB and HML factors. The increase in cross-sectional R^2 from adding $\eta_{\overline{m}}$ and η_r to the Fama–French model is small. The $\overline{\lambda}_{\overline{m}}$ and $\overline{\lambda}_r$ estimates are influenced by the inclusion of SMB and HML, while $\overline{\lambda}_{SMB}$ and $\overline{\lambda}_{HML}$ are nearly unchanged from the base Fama–French model estimates. The evidence suggests that the components of SMB and HML that are orthogonal to intertemporal risks may be driving the explanatory power of these factors. Further, after including SMB and HML, the b_r estimate is negative, inconsistent with the predictions of the ICAPM. Overall, the evidence does not support an ICAPM explanation for the Fama–French model.

The results in this section suggest that the empirical successes of the Fama–French model are not attributable to ICAPM hedging concerns. Several potential alternative explanations for the SMB and HML factors exist, including additional priced risks not captured by the ICAPM (e.g., Mayers (1972)), the APT (e.g., Fama and French (1996)), and data snooping bias (e.g., Lo and MacKinlay (1990) and Ferson, Sarkissian, and Simin (1999)). Determining the true economic explanation of the additional factors is left for future research.

1.6 Conclusion

I develop a regression-based approach to test the cross-sectional implications of the ICAPM with several advantages including (i) the strict enforcement of Campbell's (1996) critique of ICAPM tests, (ii) the mitigation of measurement error biases by using a one-step approach similar to Davies (2010), (iii) the ability to test the model across the full cross section of returns, and (iv) the use of Pástor and Stambaugh's (2009) predictive systems framework. In addition to studying the results from the regression approach, I estimate the parameters of Campbell's (1993, 1996) intertemporal asset-pricing model to determine whether risks are being priced in accordance with ICAPM theory.

I find that the ICAPM performs well in explaining the cross section of firm returns. Investors require a premium for taking on market risk, while zero-beta assets earn a return equal to the risk-free rate. In addition, investors accept lower returns on assets that hedge against adverse changes in investment opportunities. The ICAPM explains more of the variation in average firm returns than either the CAPM or Fama–French model. Overall, the predictions of the ICAPM are largely borne out in the data.

The Fama–French (1993) three-factor model has often been conjectured to be an ICAPM model. I test this explanation and find little evidence that SMB and HML are related to intertemporal risks. Instead, the components of SMB and HML that are orthogonal to intertemporal risk appear to provide most of the pricing power of the factors. The economic interpretation of the Fama–French model is still an open question.

Table 1.1: Parameter Estimates

Panel A: Predictive System Parameter Estimates

E_x		E_m
DEF	0.786	0.378
$TERM$	0.897	
DY	2.190	ϕ_m
RF	3.549	0.951

ϕ_x	DEF	$TERM$	DY	RF
DEF	0.950	-0.007	0.003	0.005
$TERM$	0.162	0.942	-0.011	-0.013
DY	-0.074	0.008	0.971	0.020
RF	-0.214	0.035	0.028	1.009

Σ	r_m^e	DEF	$TERM$	DY	RF	$\bar{r}_m^e$
r_m^e	18.167	-0.055	-0.064	-0.575	0.146	-0.674
DEF	-0.055	0.010	0.006	0.002	-0.014	0.023
$TERM$	-0.064	0.006	0.081	0.002	-0.116	0.036
DY	-0.575	0.002	0.002	0.036	-0.004	0.022
RF	0.146	-0.014	-0.116	-0.004	0.229	-0.100
$\bar{r}_m^e$	-0.674	0.023	0.036	0.022	-0.100	0.115

Panel B: Interest Rate/Inflation System Parameter Estimates

ϕ_r	ϕ_π
0.972	0.996

V	r_n	π	W	r_f	$E[\pi]$
r_n	0.001	-0.000	r_f	0.003	-0.001
π	-0.000	0.009	$E[\pi]$	-0.001	0.002

Note: This table presents parameter estimates for the market risk premium predictive system in equation (1.13) and the interest rate/inflation system in equation (1.15). Panel A shows parameters of the predictive system of Pástor and Stambaugh (2009) using the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables. Panel B reports parameters for the system for the real interest rate and expected inflation. Mean draws from the posterior distribution of each parameter are shown. The sample period is July 1963 to June 2008.

Table 1.2: Summary Statistics

Panel A: Factor Return Summary Statistics					
	Market	SMB	HML	$\eta_{\overline{m}}$	η_r
Mean	0.349	0.184	0.389	0.009	0.043
Standard Deviation	4.407	3.166	2.869	2.723	0.648

Panel B: Factor Return Correlation Matrix					
	Market	SMB	HML	$\eta_{\overline{m}}$	η_r
Market	1.000				
SMB	0.308	1.000			
HML	-0.405	-0.277	1.000		
$\eta_{\overline{m}}$	-0.547	-0.179	0.210	1.000	
η_r	-0.008	-0.005	-0.055	-0.025	1.000

Panel C: Test Asset Summary Statistics				
	Excess Return	ln(Size)	ln(B/M)	Average N per Period
Firms	0.940	18.421	-0.516	2509.867
	(13.276)	(2.076)	(0.928)	
	(2.043)			

Note: This table presents summary statistics for factors and test assets. Panel A shows the average and standard deviation of monthly log real excess returns of the factors, while Panel B shows a correlation matrix for these returns. The value-weighted CRSP index is used to proxy for the market portfolio. Unexpected changes in the market risk premium are denoted by $\eta_{\overline{m}}$ and are estimated using the predictive systems approach of Pástor and Stambaugh (2009) with the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables in the system of equations (1.13). Unexpected changes in the real interest rate, η_r, are estimated from the system of equations (1.15). For the statistics involving $\eta_{\overline{m}}$ and η_r, I report the mean of the statistic across the 5,000 draws from the MCMC chain. Panel C reports summary statistics for the test assets. Standard deviations are reported in parentheses below the means. For excess returns, both the average time-series standard deviation of returns and the average cross-sectional standard deviation of three-year averages of returns are reported. The sample period is July 1963 to June 2008.

Table 1.3: Estimated Prices of Risk for ICAPM Exposures

Panel A: Regression System Parameter Estimates

$$\underline{E_t r_i^e + \frac{V_{ii}}{2} = \overline{\lambda}_0 + \overline{\lambda}_m \beta_i^m + \overline{\lambda}_{\overline{m}} \beta_i^{\overline{m}} + \overline{\lambda}_r \beta_i^r}$$

Model	$\overline{\lambda}_0$	$\overline{\lambda}_m$	$\overline{\lambda}_{\overline{m}}$	$\overline{\lambda}_r$	R^2
CAPM	0.036	0.730			0.203
	(0.197)	(0.205)			
ICAPM	0.141	0.686	-0.148	0.055	0.347
	(0.186)	(0.199)	(0.248)	(0.039)	

Panel B: Asset-Pricing Equation Parameter Estimates

$$\underline{E_t r_i^e + \frac{V_{ii}}{2} = b_0 + b_m V_{im} + b_{\overline{m}} V_{i\overline{m}} + b_r V_{ir}}$$

Model	b_0	b_m	$b_{\overline{m}}$	b_r	$Pr(b_0 > 0)$	$Pr(b_{-0} > 0)$
CAPM	0.036	3.760			0.575	0.999
	(0.197)	(1.053)				
ICAPM	0.141	4.636	1.467	0.889	0.779	0.978
	(0.186)	(1.540)	(0.873)	(0.648)		

Note: This table reports the prices of risk for exposures to market risk and changes in the investment opportunity set. The value-weighted CRSP index is used to proxy for the market portfolio. The ICAPM specifications use estimated shocks to the market risk premium and real interest rate as additional factors. Unexpected changes in the market risk premium are estimated using the predictive systems approach of Pástor and Stambaugh (2009) with the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables. Unexpected changes in the real interest rate are estimated using FFBS. The $\eta_{\overline{m}}$ and η_r factors are annualized by multiplying monthly estimates by 12. The reported R^2 is the time-series average of cross-sectional R^2. The sample period is July 1963 to June 2008. The standard deviations of draws from the posterior distribution are listed in parentheses.

Table 1.4: Detecting Model Misspecification Using Firm Characteristics

	$E_t r_i^e + \frac{V_{ii}}{2} = \overline{\lambda}_0 + \overline{\lambda}_f' \beta_i + \overline{\lambda}_z' Z_i$						
Model	$\overline{\lambda}_0$	$\overline{\lambda}_m$	$\overline{\lambda}_{\overline{m}}$	$\overline{\lambda}_r$	$\overline{\lambda}_{Size}$	$\overline{\lambda}_{B/M}$	R^2
CAPM	0.036	0.730					0.203
	(0.197)	(0.205)					
CAPM with Size & B/M	-0.022	0.806			-0.130	0.285	0.268
	(0.201)	(0.189)			(0.088)	(0.100)	
ICAPM	0.141	0.686	-0.148	0.055			0.347
	(0.186)	(0.199)	(0.248)	(0.039)			
ICAPM with Size & B/M	0.088	0.788	0.013	0.041	-0.138	0.278	0.394
	(0.204)	(0.189)	(0.234)	(0.039)	(0.101)	(0.109)	

Note: This table reports estimates of the relations between average returns and the size and book-to-market characteristics after controlling for exposures to market risk and changes in the investment opportunity set. The value-weighted CRSP index is used to proxy for the market portfolio. The ICAPM specifications use estimated shocks to the market risk premium and real interest rate as additional factors. Unexpected changes in the market risk premium are estimated using the predictive systems approach of Pástor and Stambaugh (2009) with the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables. Unexpected changes in the real interest rate are estimated using FFBS. Size and book-to-market are included in the cross-sectional regression to detect model misspecification. The $\eta_{\overline{m}}$ and η_r factors are annualized by multiplying monthly estimates by 12. The reported R^2 is the time-series average of cross-sectional R^2. The sample period is July 1963 to June 2008. The standard deviations of draws from the posterior distribution are listed in parentheses.

Table 1.5: The ICAPM and Fama–French Model – Standard ICAPM Factors

Panel A: Regression System Parameter Estimates

Model	$\overline{\lambda}_0$	$\overline{\lambda}_m$	$\overline{\lambda}_{\overline{m}}$	$\overline{\lambda}_r$	$\overline{\lambda}_{SMB}$	$\overline{\lambda}_{HML}$	R^2
CAPM	0.036	0.730					0.203
	(0.197)	(0.205)					
ICAPM	0.141	0.686	-0.148	0.055			0.347
	(0.186)	(0.199)	(0.248)	(0.039)			
Fama–French	0.177	0.594			0.251	-0.358	0.336
	(0.167)	(0.184)			(0.205)	(0.145)	
All Factors	0.199	0.588	-0.063	0.023	0.251	-0.334	0.393
	(0.179)	(0.189)	(0.245)	(0.028)	(0.212)	(0.154)	

Panel B: Asset-Pricing Equation Parameter Estimates

Model	b_0	b_m	$b_{\overline{m}}$	b_r	b_{SMB}	b_{HML}	$Pr(b_0 > 0)$	$Pr(b_{-0} > 0)$
CAPM	0.036	3.760					0.575	0.999
	(0.197)	(1.053)						
ICAPM	0.141	4.636	1.467	0.889			0.779	0.978
	(0.186)	(1.540)	(0.873)	(0.648)				
Fama–French	0.177	2.136			0.900	-2.743	0.860	0.088
	(0.167)	(1.324)			(2.438)	(2.301)		
All Factors	0.199	3.186	1.348	0.586	1.213	-2.364	0.876	0.139
	(0.179)	(1.754)	(0.885)	(0.668)	(2.538)	(2.411)		

Note: This table contains tests of the CAPM, ICAPM, and Fama–French (1993) three-factor model. The value-weighted CRSP index is used to proxy for the market portfolio. The ICAPM specifications use estimated shocks to the market risk premium and real interest rate as additional factors. Unexpected changes in the market risk premium are estimated using the predictive systems approach of Pástor and Stambaugh (2009) with the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables. Unexpected changes in the real interest rate are estimated using FFBS. The SMB and HML factors of the Fama–French model are also included in the model. The $\eta_{\overline{m}}$ and η_r factors are annualized by multiplying monthly estimates by 12. The reported R^2 is the time-series average of cross-sectional R^2. The sample period is July 1963 to June 2008. The standard deviations of draws from the posterior distribution are listed in parentheses.

Table 1.6: The ICAPM and Fama–French Model – ICAPM Factors Based on SMB and HML

Panel A: Regression System Parameter Estimates							
Model	$\overline{\lambda}_0$	$\overline{\lambda}_m$	$\overline{\lambda}_{\overline{m}}$	$\overline{\lambda}_r$	$\overline{\lambda}_{SMB}$	$\overline{\lambda}_{HML}$	R^2
CAPM	0.036	0.730					0.203
	(0.197)	(0.205)					
ICAPM	0.112	0.665	-0.565	0.034			0.295
	(0.197)	(0.203)	(0.209)	(0.045)			
Fama–French	0.177	0.594			0.251	-0.358	0.336
	(0.167)	(0.184)			(0.205)	(0.145)	
All Factors	0.207	0.579	-0.380	0.016	0.247	-0.369	0.366
	(0.177)	(0.192)	(0.185)	(0.031)	(0.219)	(0.155)	

Panel B: Asset-Pricing Equation Parameter Estimates								
Model	b_0	b_m	$b_{\overline{m}}$	b_r	b_{SMB}	b_{HML}	$Pr(b_0 > 0)$	$Pr(b_{-0} > 0)$
CAPM	0.036	3.760					0.575	0.999
	(0.197)	(1.053)						
ICAPM	0.112	5.949	3.023	0.314			0.722	0.773
	(0.197)	(2.792)	(2.091)	(0.491)				
Fama–French	0.177	2.136			0.900	-2.743	0.860	0.088
	(0.167)	(1.324)			(2.438)	(2.301)		
All Factors	0.207	3.795	2.603	-0.009	1.074	-2.987	0.890	0.073
	(0.177)	(2.322)	(1.633)	(0.522)	(2.628)	(2.453)		

Note: This table contains tests of the CAPM, ICAPM, and Fama–French (1993) three-factor model, where the intertemporal risk factors of the ICAPM use variables based on SMB and HML as predictive variables. The value-weighted CRSP index is used to proxy for the market portfolio. The ICAPM specifications use estimated shocks to the market risk premium and real interest rate as additional factors. Unexpected changes in the market risk premium are estimated using the predictive systems approach of Pástor and Stambaugh (2009) with variables based on SMB and HML as predictor variables. Unexpected changes in the real interest rate are estimated using FFBS. The SMB and HML factors of the Fama–French model are also included in the model. The $\eta_{\overline{m}}$ and η_r factors are annualized by multiplying monthly estimates by 12. The reported R^2 is the time-series average of cross-sectional R^2. The sample period is July 1963 to June 2008. The standard deviations of draws from the posterior distribution are listed in parentheses.

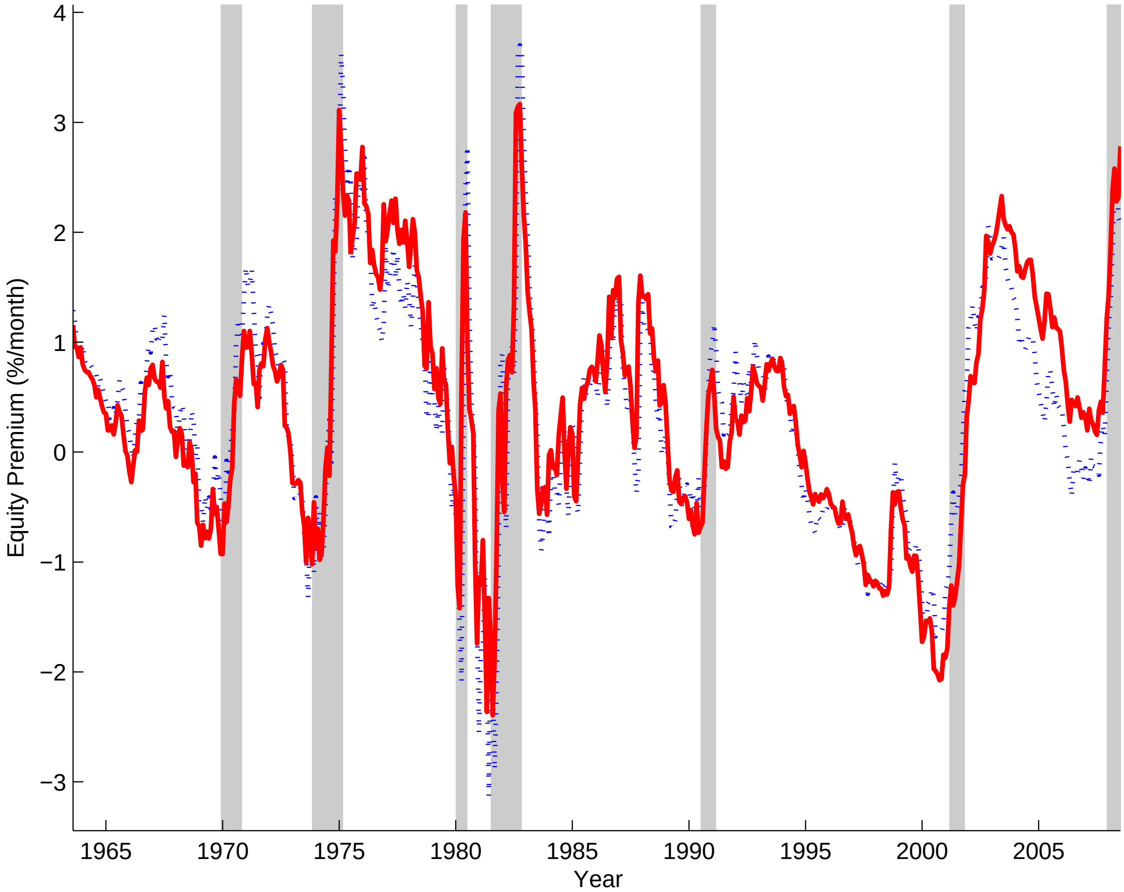

Figure 1.1: Market Risk Premium Estimate

Note: This figure shows estimates of the market risk premium using Pástor and Stambaugh's (2009) predictive systems approach given by equation (1.13) and a predictive regression approach. The dividend-to-price ratio, term spread, risk-free rate, and default spread are used jointly as predictive variables. The predictive systems estimate is the solid line, and the dotted line is the estimate from a predictive regression. The market risk premium is expressed in percent per month. NBER recessions are shaded.

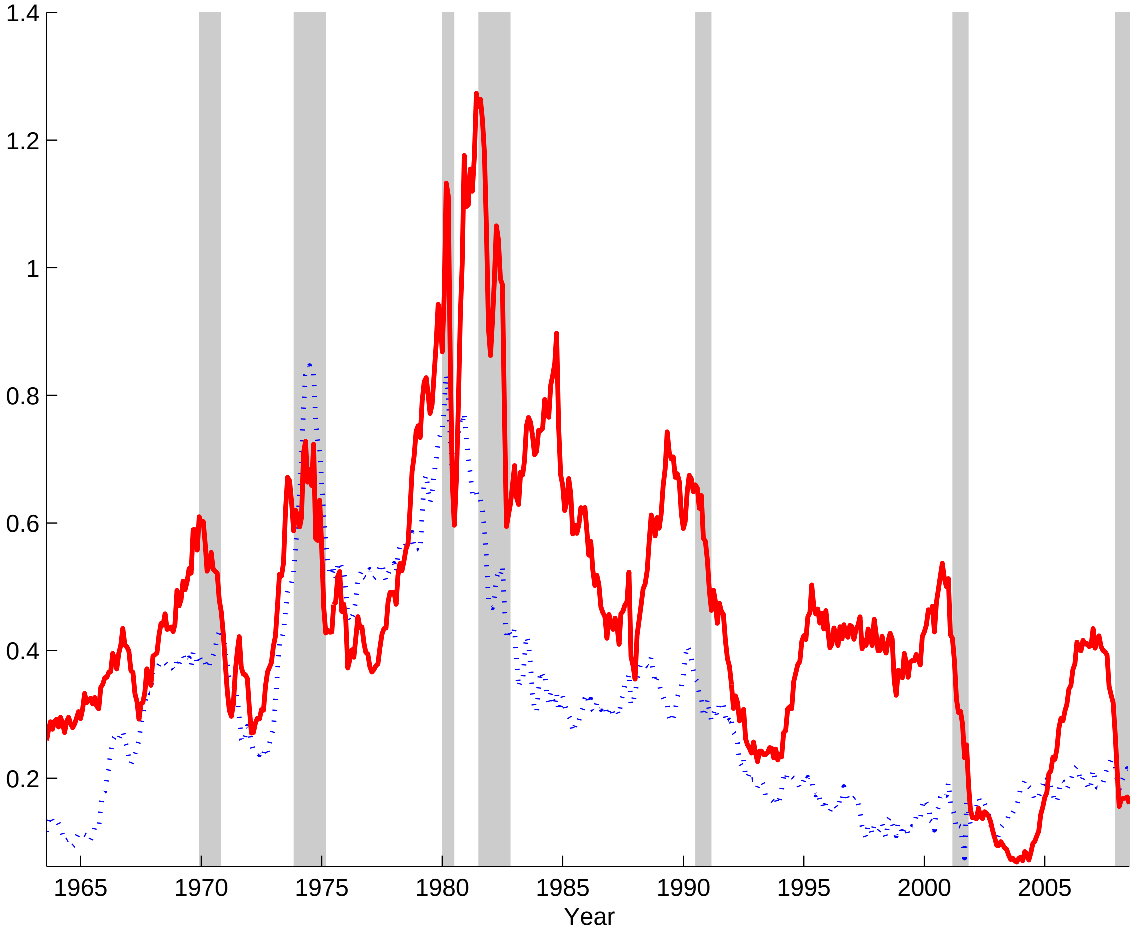

Figure 1.2: Nominal Interest Rate and Expected Inflation

Note: This figure shows estimates of the nominal interest rate and expected inflation. I use FFBS to estimate the system of equations (1.15). The nominal interest rate is the solid line and the dotted line is expected inflation. The nominal interest rate and expected inflation are expressed in percent per month. NBER recessions are shaded.

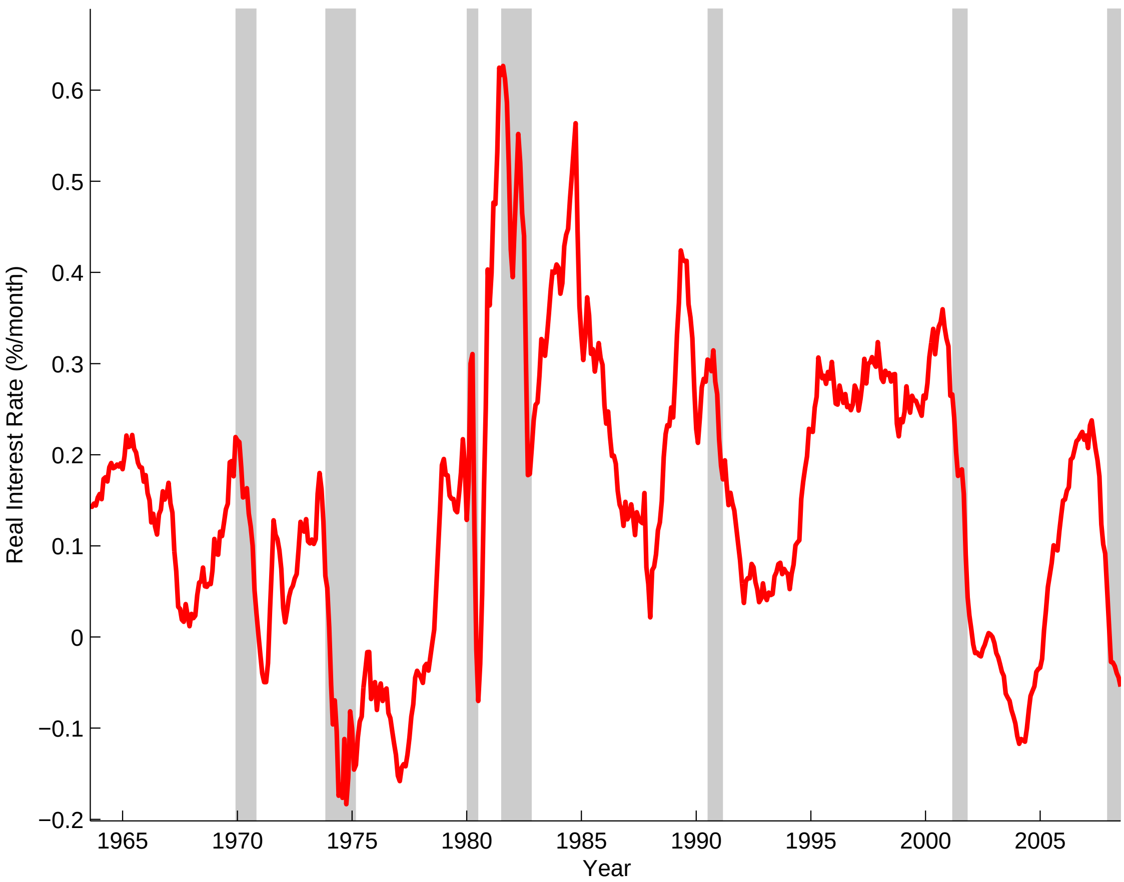

Figure 1.3: Real Interest Rate

Note: This figure shows an estimate of the real interest rate. The real interest rate is estimated from the system of equations (1.15) using FFBS with nominal interest rate and inflation data. The real interest rate is expressed in percent per month. NBER recessions are shaded.

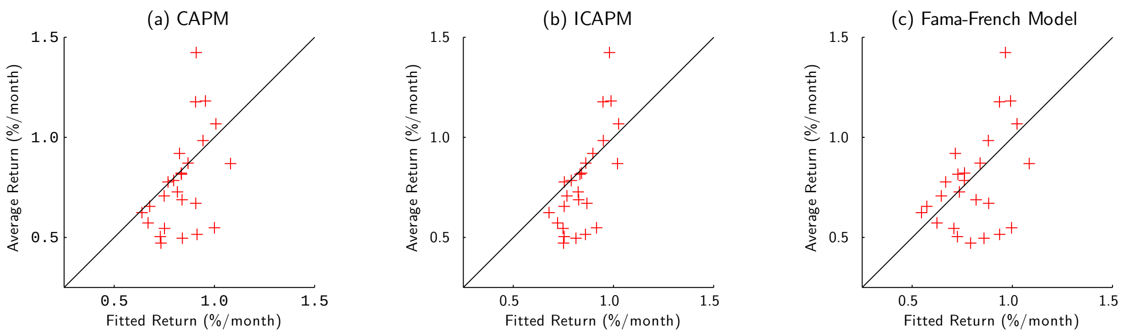

Figure 1.4: Fitted Returns versus Average Returns of Individual Firms

Note: This figure shows fitted returns and average returns for individual firms. The CAPM, the ICAPM with the dividend-to-price ratio, term spread, risk-free rate, and default spread as predictor variables, and the Fama–French (1993) model are used to explain returns. Firms are grouped by equally weighting returns of firms in each size and book-to-market category based on NYSE quintiles. Returns are expressed in percent per month.

CHAPTER 2
CROSS-SECTIONAL ASSET PRICING PUZZLES: AN EQUILIBRIUM PERSPECTIVE

2.1 Introduction

Recent empirical works document counterintuitive relations in the cross section of average stock returns. Specifically, Diether, Malloy, and Scherbina (2002) demonstrate that firms with more uncertain earnings (higher forecast dispersion) underperform lower dispersion firms. Ang, Hodrick, Xing, and Zhang (2006) document a negative relation between average return and idiosyncratic volatility (IV). Dichev (1998) and Campbell, Hilscher, and Szilagyi (2008) find evidence that financially distressed stocks deliver abnormally low returns, and Avramov, Cederburg, and Hore (2009) show that high credit rating firms considerably outperform their lower-rated counterparts. These three effects are anomalous because, while investors are expected to discount uncertainty about firm fundamentals (e.g., Merton (1987)), they appear to pay a premium for bearing such uncertainty. Indeed, all three effects are unexplained by traditional asset pricing models such as the Sharpe–Lintner CAPM and the Fama–French (1993) three-factor model. Past work offers potential explanations for the puzzling effects. Johnson (2004) interprets forecast dispersion as a proxy for unpriced information risk, and shows that in the presence of leverage, equity value rises (and expected return falls) as unpriced risk increases. Garlappi, Shu, and Yan (2008) attribute the credit risk effect to shareholders' ability to extract rents through strategic renegotiation of debt obligations, while George and Hwang (2010) propose that the endogenous choice of leverage explains the credit risk effect.

This paper proposes a unified resolution of the three puzzling effects within a long-run risk equilibrium. Two ingredients pertaining to the dynamics of aggregate and firm fundamentals underly our paradigm. The aggregate economy is formulated based on Bansal and Yaron (2004) in that long-run risk arises through the interaction of the stochastic differential utility of Duffie and Epstein (1992) with persistent consumption and dividend growth rates. In the cross section, we employ a tractable mean-reverting process to formulate firm dividend share, which allows us to solve for cross-sectional asset pricing quantities while ensuring that the cross section of firms aggregate to form the economy. Ultimately, we successfully merge the long-run risk literature with the shares-based cross section literature advocated by Menzly, Santos, and Veronesi (2004) and Santos and Veronesi (2006). Moreover, we develop an intertemporal asset pricing model in the spirit of Merton (1973), as a firm's expected return is affected by its beta with respect to an economic growth hedge portfolio.

In a long-run risk economy, the cross section of expected returns is determined by firms' cash flow duration. Long-run cash flows have pronounced sensitivities to the persistent economic growth rate and thus command large risk premiums. Therefore, high duration firms have high expected returns due to their reliance on long-run cash flows. The duration of a firm in the cross section is positively related to its expected dividend growth, which is summarized by a characteristic labeled relative share. Relative share – a term introduced by Menzly, Santos, and Veronesi (2004) – is the long-run expected dividend share of a firm as a proportion of its current dividend share, with dividend share being the fraction of the dividend paid by the firm relative to the aggregate dividend.

We show that the negative relations between expected return and dispersion, IV, and credit risk are a manifestation of cross-sectional patterns in exposures to idiosyncratic and systematic risks. Low relative share firms exhibit low duration as they derive their values primarily from short-run cash flows. Whereas such firms display low expected returns, they have high sensitivities to firm-specific shocks leading to high dispersion, IV, and credit risk. In contrast, high relative share firms display high duration and are particularly sensitive to aggregate shocks but less susceptible to firm-specific dividend shocks.

Indeed, we establish theoretical links between the dispersion, IV, and credit risk measures and expected stock returns. Low relative share firms deliver low expected returns. These firms exhibit high levels of dispersion proxied by uncertainty about firm cash flow growth. Moreover, the values of low relative share firms are highly sensitive to firm-specific dividend shocks, which leads to high IV. Finally, the majority of default events tend to occur among low relative share firms as a result of negative idiosyncratic shocks, inducing high measures of financial distress for this subset of firms.

The empirical evidence broadly supports our model predictions. First, we do document mean reversion in dividend share as well as establish, based on model selection criteria, the dominance of our formulation for firm-level dividend growth relative to other specifications. Next, dispersion, IV, and credit risk measures exhibit high cross-sectional correlation, and portfolios formed on each of these characteristics display similar patterns in composition and behavior. More importantly, we find empirical support for a long-run risk explanation of these effects, as portfolios of high (low) dispersion, IV, and credit risk firms have low (high) estimated relative share as predicted by our model. Furthermore, the expected returns implied by our model are highly correlated with realized average returns for these portfolios.

In sum, recent work utilizes the long-run risk framework to offer explanations for aggregate and cross-sectional asset pricing puzzles including the equity premium, risk-free rate, and excess volatility puzzles (Bansal and Yaron (2004)), as well as the value premium (e.g., Bansal, Kiku, and Yaron (2007), Hansen, Heaton, and Li (2008), and Ai and Kiku (2010)) and the momentum effect (Zurek (2007) and Avramov and Hore (2008)). We enhance the long-run risk literature in both methodology and substantive findings. In particular, we derive formal cross-sectional asset pricing restrictions and are able to express our model through an intuitive ICAPM relationship. Moreover, our model simultaneously resolves three apparently counterintuitive regularities – the dispersion, IV, and credit risk effects.

The remainder of the paper is organized as follows. In Section 2.2, we introduce the aggregate economy, develop our formulation of the cross section of firms, and establish asset pricing results. We examine the dispersion, IV, and credit risk effects within our framework in Section 2.3 and provide empirical evidence about these relations in Section 2.4. Section 2.5 concludes.

2.2 The Model

Two essential ingredients pertaining to the dynamics of aggregate and firm fundamentals underly our equilibrium setup. The aggregate economy is based on a particularly simple version of Bansal and Yaron's (2004) long-run risk framework in which an economic agent with recursive preferences is exposed to persistent consumption and dividend growth rates.[1] This economy generates an equity premium, market volatility, and a risk-free rate that match the data. Further, the interaction between recursive preferences and persistent growth introduces additional volatility into the pricing kernel, which is essential to satisfy the Hansen–Jagannathan (1991) bound. Within the aggregate economy, we formulate a cross section of firms whose dividends sum to the aggregate dividend. Firms are differentiated by expected cash flow timing. In a long-run risk equilibrium, investors require larger risk premiums to hold firms weighted toward long-run cash flows. Ultimately, firms with higher dispersion, IV, and credit risk measures exhibit lower cash flow duration, thus they deliver lower returns.

2.2.1 The Aggregate Economy

The representative investor is endowed with the stochastic differential utility of Duffie and Epstein (1992), a continuous-time equivalent to Epstein–Zin (1989) preferences, with

$$J_t = E_t \left[\int_t^{\infty} f(C_s, J_s) ds \right], \tag{2.1}$$

[1]Empirically, Bansal, Kiku, and Yaron (2009) find substantial evidence for a predictable component in consumption growth (see also Bansal and Yaron (2004), Kiku (2006), and Hansen, Heaton, and Li (2008)). Yang (2009) finds strong evidence of persistence in durable consumption growth. Theoretically, Hansen and Sargent (2010) show that investors operate as if the economy has persistent consumption growth when faced with uncertainty about the true nature of the consumption growth process, while Kaltenbrunner and Lochstoer (2008) find that long-run risk in consumption naturally arises through the optimal consumption smoothing of Epstein–Zin (1989) investors.

where J_t is the value function and $f(C_s, J_s)$ is a normalized aggregator of current consumption and continuation utility. Assuming that the elasticity of intertemporal substitution (EIS) equals one, the normalized aggregator takes the form

$$f(C,J) = \beta(1-\gamma)J\left[\log C - \frac{\log((1-\gamma)J)}{1-\gamma}\right], \tag{2.2}$$

where γ is the coefficient of relative risk aversion and β is the time preference parameter.[2]

Next, we formulate aggregate consumption and dividend growth rates as[3]

$$\frac{dC_t}{C_t} = (\mu_C + \lambda X_t)dt + \sigma_C dW_C, \tag{2.3}$$

$$\frac{dD_t}{D_t} = (\mu_D + X_t)dt + \sigma_D dW_D, \tag{2.4}$$

where X_t follows the mean-reverting dynamics

$$dX_t = -\kappa X_t dt + \sigma_X dW_X. \tag{2.5}$$

The time-varying component of economic growth, X_t, reverts to zero, with the speed of reversion governed by κ. Shocks to X_t propagate through several periods of

[2]Empirically, there is a wide debate in the macroeconomics literature about the value of EIS. Kandel and Stambaugh (1991) focus on EIS close to zero. Hansen and Singleton (1982), Attanasio and Weber (1989), Vissing-Jorgensen (2002), and Guvenen (2006) argue that EIS is greater than 1, while Hall (1988) and Campbell and Mankiw (1989) argue that EIS is less than 1. Yogo (2004) estimates the EIS parameter in several countries and cannot reject EIS=1 in 11 of 13 countries.

[3]Consumption and dividend growth are the continuous-time versions of the discrete-time processes employed by Bansal and Yaron (2004). We adopt a model with homoskedastic shocks to dividend growth, consumption growth, and the aggregate growth rate. Bansal and Yaron (2004) investigate both homoskedastic and heteroskedastic cases. The homoskedastic model is simpler and sufficient for our purposes. Without altering the spirit of the model dynamics from the Bansal and Yaron (2004) economy, we multiply the stochastic component of expected consumption growth by λ instead of scaling the time-varying portion of expected dividend growth.

consumption and dividend growth rates due to slow mean reversion, which combined with recursive preferences gives rise to long-run risk. Since expected consumption growth is less volatile than expected dividend growth we allow for a differential effect of X_t on consumption and dividend growth rates, with λ controlling the relative strength of shocks. In the data we find $\lambda < 1$, which is consistent with Bansal and Yaron (2004) and Constantinides and Ghosh (2008), among others, and we impose this restriction in our analysis. Thus consumption growth is smoother than dividend growth, capturing the insight of Abel's (1999) levered economy. Following Bansal and Yaron (2004), we assume the processes dW_C, dW_D, and dW_X are uncorrelated for simplicity of presentation.[4]

We present asset pricing quantities for the aggregate setup in the proposition below and refer the interested reader to Hore (2008) for complete derivation and Appendix A.3.1 for additional details.[5]

Proposition 1. *The price-dividend ratio is*

$$\frac{P_t}{D_t} = G(X_t) = \int_t^\infty S(X_t, \tau) ds \qquad [\tau = s - t], \tag{2.6}$$

where the $S(X_t, \tau)$ function is described in Appendix A.3.1. The risk premium is

$$\mu_t = \frac{\lambda(\gamma - 1)}{\kappa + \beta} \sigma_X^2 \frac{G_X}{G}. \tag{2.7}$$

[4]Introducing correlation between the consumption and dividend processes leads to an additional term in expected returns taking the form $\rho\sigma_C\sigma_D\gamma$, which is the familiar expected return in an *i.i.d.* growth economy with CRRA preferences. This term is economically small for reasonable values of risk aversion (e.g. Mehra and Prescott (1985)). We assume zero correlation to simplify the model and focus on the effects of long-run risk. Further, firms in our economy share equal exposures to this source of risk, so our cross-sectional inferences are ultimately unaffected by this assumption.

[5]We impose the following parameter restrictions: $\gamma > 1$, $\lambda < 1$, $\beta > 0$, $\sigma_X > 0$, and $\kappa > 0$.

The price-dividend ratio in equation (2.6) does not take an exponentially affine form, which leads to stochastic volatility in the market return and thus a time-varying risk premium. In discrete-time solutions, the price-dividend ratio is typically posited to be exponentially affine, which produces constant risk premiums. Hence, to generate time-varying risk premiums, a second risk channel – stochastic volatility of consumption growth – is often employed.

The essential feature of our aggregate economy is the term structure of risk premiums, which is formulated in the following theorem.

Theorem 1. *Expected return is increasing in duration for any asset that pays a non-negative portion of the aggregate dividend at all times.*

Proof. See Appendix A.3.1. □

Theorem 1 establishes that longer-run cash flows are riskier, thus commanding higher risk premiums. Figure 2.1 exhibits the representative investor's required discount rate on cash flows as a function of duration. Risk premiums on low duration cash flows are negligible while high duration cash flows command large risk premiums.

Why do agents require higher risk premiums on longer-run cash flows? Investors endowed with Duffie–Epstein preferences prefer early resolution of uncertainty if $\gamma > 1$ with unit EIS ($\psi = 1$), or if $\gamma > \frac{1}{\psi}$ when ψ is unrestricted. Even with persistent consumption and dividend growth dynamics, shocks to growth rates have little effect on short-run cash flows. Short-run cash flows are thus relatively safe. On the other hand, persistent growth heightens uncertainty about the magnitudes of long-run dividends and consumption. Much of this uncertainty remains unresolved for a long period. Thus, investors discount long-run cash flows heavily giving rise to a long-run risk premium.

Before we proceed, we note that related work examines equity duration implications for asset pricing in a different context. Menzly, Santos, and Veronesi (2004) and Wachter (2006) show a positive duration–expected return relation in an economy with the external habit formation preferences of Campbell and Cochrane (1999). Da (2009) develops a model based on cash flow covariance and duration factors, exhibiting empirical results that are consistent with long-run risk predictions. Lettau and Wachter (2007) and Croce, Lettau, and Ludvigson (2009) argue that the positive duration–expected return relation is inconsistent with the value premium. In contrast, Ai and Kiku (2010) show that the value premium does emerge in a long-run risk economy accommodating growth options. Empirically, Bansal, Dittmar, and Lundblad (2005) find that value firms have higher cash flow betas and Bansal, Dittmar, and Kiku (2009) demonstrate that value firms display higher exposures to long-run consumption risk, both consistent with a long-run risk explanation.

2.2.2 The Cross Section of Firms

We consider a cross section of n firms that aggregates to the economy derived above. Explicitly modeling the cross section ensures the reasonable long-term evolution of firm cash flows, which is particularly important in a long-run risk framework. Our tractable framework facilitates drawing cross-sectional inferences by pricing a single firm.

In the cross section, firm i contributes a time-varying portion of the aggregate dividend, defined as the firm's dividend share θ_t^i. Thus, if the aggregate dividend is D_t, firm i contributes $D_t^i = \theta_t^i D_t$, whereas the remaining firms contribute a total of $(1-\theta_t^i)D_t$. The dividend share is formulated using the mean-reverting Wright–Fisher (WF) process[6]

$$d\theta_t^i = \alpha(\bar{\theta}^i - \theta_t^i)dt + \delta\sqrt{(1-\theta_t^i)\theta_t^i}dW_{\theta^i}. \tag{2.8}$$

Moreover, we parameterize long-run dividend shares such that $\sum_{i=1}^n \bar{\theta}^i = 1$ and assume the covariance structure of the $(n \times 1)$ vector dW_θ satisfies

$$\rho_t(dW_{\theta^i}, dW_{\theta^j}) = -\sqrt{\frac{\theta_t^i \theta_t^j}{(1-\theta_t^i)(1-\theta_t^j)}} \text{ for all } i,\ j. \tag{2.9}$$

The negative correlation between dividend share shocks naturally arises from the aggregation identity $\sum_{i=1}^n D_t^i = D_t$. Essentially, as one firm's dividend share increases through a firm-specific shock, dividend shares of all other firms are "crowded out" to maintain proper aggregation. In a typical economy, dividend shares of all firms are small so $\rho_t(dW_{\theta^i}, dW_{\theta^j})$ is negligible (e.g. correlation of -0.01 between firms each paying 1% of the aggregate dividend).

Several features of the dividend share process are desirable. First, the dividend share of all firms is bounded between zero and one.[7] Second, as we formally

[6] A WF process for dividend share arises endogenously in an endowment economy with two trees, as shown by Cochrane (2008b). WF processes originate in genetics literature by Fisher (1930) and Wright (1931). See Crow and Kimura (1970) for examples and further discussion of WF processes. We impose the parameter restrictions $\alpha > 0$ and $\delta > 0$.

[7] Assuming that the initial dividend share of firm i is between zero and one, i.e. $\theta_0^i \in (0,1)$, then the boundary points zero and one are unattainable within finite time if $2\alpha\bar{\theta}^i > \delta^2$ (see Karlin and Taylor (1981), pp 239–241). If the restriction $2\alpha\bar{\theta}^i > \delta^2$ is violated, we propose a reflecting barrier at each of the boundaries. The cross section of firms continues to aggregate properly given the assumption of reflecting barriers.

show below, $\sum_{i=1}^{n} \theta_t^i = 1$ for all t, so the cross section of firms does aggregate to form the economy described above. Third, firm dividend shares revert towards their long-run means. Indeed, mean reversion in dividend (or consumption) share is a common assumption in the shares-based literature (e.g. Menzly, Santos, and Veronesi (2004), Santos and Veronesi (2006), and Da (2009)). Da (2009) finds empirical support for a similar mean-reverting cash flow share process for book-to-market, size, and reversal portfolios. Also common are the assumptions of mean reversion in expected dividend growth (e.g. Campbell and Shiller (1988)) as well as long-run convergence to an economy-wide steady state dividend growth (e.g. Pástor, Sinha, and Swaminathan (2008) and Da and Warachka (2009)). Mean reversion ensures that no firm eventually dominates the economy, and it captures the intuitive notion that expectations of future dividends are revised less than current dividends in the presence of a contemporaneous shock to the firm's dividend. Empirically, we find support for our dividend share process. As shown in Section 2.4, the evidence from portfolios formed based on dispersion, IV, and credit risk measures is supportive of our specification, and our share process outperforms several reasonable alternative specifications.

Following Menzly, Santos, and Veronesi (2004), we refer to $\frac{\bar{\theta}^i}{\theta_t^i}$ as the "relative share" of the firm. The relative share is the long-run expected dividend share of the firm as a proportion of its current dividend share. Some of our model implications depend on dividend share θ_t^i through the relative share $\frac{\bar{\theta}^i}{\theta_t^i}$. To briefly illustrate the dividend share and relative share characteristics, suppose that a firm is currently paying 1% of the total market dividend but is expected to exhibit high dividend growth so that it eventually pays 5% of the market dividend. The dividend share of this firm is 0.01 and the relative share is 5 $\left(\frac{5\%}{1\%}\right)$. Consider another firm that is currently paying 1% of the market dividend but is expected to eventually pay only 0.2% of the market dividend. This firm has a dividend share of 0.01 and a relative share of 0.2 $\left(\frac{0.2\%}{1\%}\right)$.

We now turn to demonstrating how dividend shares and relative shares aggregate at the portfolio level. Theorem 2 provides a convenient aggregation result where the sum of any two dividend shares also follows a WF process.

Theorem 2. *Assume that the dividend share of each firm $i = 1, ..., n$ follows a WF process, as in equation (2.8), with the correlation structure described in equation (2.9). Then the dividend share of a portfolio of any two firms i and j follows a WF process,*

$$d\theta_t^p = \alpha(\bar{\theta}^p - \theta_t^p)dt + \delta\sqrt{(1-\theta_t^p)\theta_t^p}dW_{\theta^p}, \tag{2.10}$$

where $\bar{\theta}^p = \bar{\theta}^i + \bar{\theta}^j$ and $\theta_t^p = \theta_t^i + \theta_t^j$.

Proof. See Appendix A.3.2. □

Theorem 2 leads to proper aggregation of the cross section. Since the sum of any two WF processes also follows a WF process under the correlation structure in equation (2.9), it is trivial to show that the dividend shares of the n firms in the economy sum to one. Hence, firm dividends also sum to the aggregate dividend. Further, our model's implications for firm-level asset pricing carry over to the portfolio level. We can demonstrate how the exposure of portfolios to long-run risk is determined from the exposures of firms held within the portfolios. Specifically, a value-weighted portfolio's relative share is the dividend-weighted average of firm relative shares,

$$\frac{\bar{\theta}^p}{\theta_t^p} = \sum_{i\in\mathcal{P}} \frac{D_t^i}{\sum_{i\in\mathcal{P}} D_t^i}\frac{\bar{\theta}^i}{\theta_t^i}, \tag{2.11}$$

where $\mathcal{P}$ is the set of firms held in the portfolio. See Appendix A.3.2 for additional details on aggregation.

We assume throughout that $Cov\left(d\theta_t^i, \frac{dC_t}{C_t}\right) = Cov\left(d\theta_t^i, dX_t\right) = 0$, so the dividend share does not covary with the pricing kernel. Hence, no asset in this economy hedges against transient consumption shocks from dW_C. In the parlance of Menzly, Santos, and Veronesi (2004) and Santos and Veronesi (2006), cash flow betas of all securities are zero. Even then, as we show below, firm relative shares interact with the persistent component of consumption growth to play a powerful role in determining time-varying exposures to systematic risk for each firm.

Next, firm dividend growth, obtained by applying Itô's Lemma to $D_t^i = \theta_t^i D_t$, is given by

$$\frac{dD_t^i}{D_t^i} = \left(\alpha\left(\frac{\bar{\theta}^i}{\theta_t^i} - 1\right) + \mu_D + X_t\right) dt + \delta\sqrt{\frac{1-\theta_t^i}{\theta_t^i}} dW_\theta^i + \sigma_D dW_D. \tag{2.12}$$

Firm-level dividend growth absorbs the expected aggregate dividend growth drift component $\mu_D + X_t$ as well as the aggregate dividend shock $\sigma_D dW_D$. In better economic conditions, when X_t is high, expected dividend growth is higher. Firm relative share $\frac{\bar{\theta}^i}{\theta_t^i}$ also affects firm dividend growth. To illustrate, if $\bar{\theta}^i > \theta_t^i$, then the dividend share is likely to increase towards $\bar{\theta}^i$. In this case, $\frac{\bar{\theta}^i}{\theta_t^i} > 1$ and expected dividend growth is higher than aggregate dividend growth. On the other hand, if $\bar{\theta}^i < \theta_t^i$, the dividend share is expected to decrease, leading to low expected dividend growth relative to the aggregate dividend. Finally, if $\bar{\theta}^i = \theta_t^i$ then the expected dividend growth rate is equal to the expected aggregate dividend growth rate. Still, firm dividend growth is more volatile due to the firm-specific shock dW_θ^i.

2.2.3 Cross-Sectional Asset Pricing

The cross section of returns is driven by the interaction between two properties of the economy: (i) long-run cash flows carry a higher risk premium than short-run cash flows and (ii) firms have different expected cash flow timing, reflected by the relative share characteristic. Firms with dividends that are concentrated in the short-run primarily derive their value from relatively safe cash flows, while firms with primarily long-run dividends are risky investments. Relative share is positively related to expected dividend growth and cash flow duration. Hence, relative share is positively related to systematic risk exposure and expected return.

Below, we formalize asset pricing properties at the firm level.

Theorem 3. *The firm price-dividend ratio is*

$$\begin{aligned}\frac{P_t^i}{D_t^i} \equiv G^i(X_t, \theta_t^i; \bar{\theta}^i, \alpha) &= \int_t^\infty S(X_t, \tau) E_t[\theta_s^i] ds \qquad [\tau = s - t] \\ &= \frac{P_t}{D_t} + \left(\frac{\bar{\theta}^i}{\theta_t^i} - 1\right) \int_t^\infty S(X_t, \tau)\left(1 - e^{-\alpha\tau}\right) ds, \qquad (2.13)\end{aligned}$$

where $S(X_t, \tau)$ is defined in Proposition 1.[8] *The firm price-dividend ratio is increasing in relative share $\frac{\bar{\theta}^i}{\theta_t^i}$. Next, the firm-level risk premium is*

$$\mu_t^i = \frac{\lambda(\gamma - 1)}{\kappa + \beta} \sigma_X^2 \frac{G_X^i}{G^i}, \qquad (2.14)$$

and μ_t^i is increasing in relative share. Finally, the firm-level instantaneous variance is

$$\sigma_{i,t}^2 = \sigma_D^2 + \left(\frac{G_X^i}{G^i}\right)^2 \sigma_X^2 + \left(1 + \frac{\theta_t^i G_\theta^i}{G^i}\right)^2 \delta^2 \frac{1 - \theta_t^i}{\theta_t^i}. \qquad (2.15)$$

Proof. See Appendix A.3.3. □

The firm-level price-dividend ratio is affected by aggregate and firm-specific conditions. In particular, all firms are affected by expected economic growth, and all

[8] Integrability of (2.13) is trivially satisfied due to the transversality condition of $S(X_t, \tau)$ with $\alpha > 0$.

price-dividend ratios increase when X_t increases. If $\theta_t^i = \bar{\theta}^i$ (i.e. the dividend share is at its long-run mean), then the firm price-dividend ratio equals that of the market. If $\frac{\bar{\theta}^i}{\theta_t^i} > 1$ ($\frac{\bar{\theta}^i}{\theta_t^i} < 1$), then firm expected dividend growth is higher (lower) than aggregate expected dividend growth, which pushes the firm price-dividend ratio higher (lower) than the aggregate price-dividend ratio. However, high relative share also implies that firm dividends are weighted towards the long-run and agents exposed to long-run risk discount these cash flows more heavily. The higher discount rate for these cash flows moderates the increase in firm price-dividend ratio that arises from an increase in relative share.

Finally, observe from equation (2.15) that firm return variance has both systematic and idiosyncratic components. Shocks to the aggregate dividend and growth rate are market-wide risks to which all firms are exposed. The systematic portion of firm return volatility varies in the cross section only to the extent that firms have differential exposures to systematic risk through $\frac{G_X^i}{G^i}$. On the other hand, the $\left(1 + \frac{\theta_t G_\theta^i}{G^i}\right)^2 \delta^2 \frac{1-\theta_t^i}{\theta_t^i}$ component arises solely from firm-specific shocks to dividend share. Later, we discuss properties of IV as well as describe the cross-sectional relation between expected return and IV.

2.2.3.1 A Parsimonious Beta Pricing Model

Comparing the aggregate and firm risk premiums in equations (2.7) and (2.14), respectively, we are able to formulate a conditional single-factor model. The firm expected return is related to an economic growth hedge portfolio through the beta pricing relation

$$\mu_t^i = \beta_t^i \mu_t, \tag{2.16}$$

where

$$\beta_t^i \left(X_t, \frac{\bar{\theta}^i}{\theta_t^i} \right) = \frac{G_X^i}{G^i} \frac{G}{G_X} \tag{2.17}$$

is the firm (time-varying) beta with respect to the hedging portfolio. Here, the hedging portfolio is constructed to be exposed solely to economic growth shocks through dW_X and to have expected return μ_t. Additional information about the hedging portfolio is available in Appendix A.3.3. Firm beta is positively related to relative share as illustrated in Figure 2.2. In fact, beta and duration are closely linked in the cross section. Long-run cash flows are more sensitive to growth rate shocks, leading to higher exposures to hedging portfolio returns for high duration firms. Therefore, firms with higher (lower) than average duration have betas greater (less) than one. The positive relative share–beta relation translates to a positive relation between relative share and expected returns, shown by Figure 2.2. There is an economically significant spread in expected returns in the cross section between high and low relative share firms arising from differential exposures to long-run risk.

This convenient beta pricing model in the spirit of Merton's (1973) ICAPM is novel in the asset pricing literature studying recursive preferences. Exposure to a single factor hedging against economic growth shocks exclusively determines cross-sectional dispersion in expected returns. Typically, the ICAPM is expressed as a multifactor model with the aggregate wealth portfolio and additional intertemporal hedge portfolios. However, we zero out the correlation between dividend and consumption growth shocks to concentrate on long-run risk implications, thereby eliminating exposure to the total wealth portfolio.

In the presence of correlation between consumption and dividend growth shocks, the expected firm return takes the form

$$\mu_t^i = \gamma \sigma_C \sigma_D \rho + \frac{\lambda(\gamma - 1)}{\kappa + \beta} \sigma_X^2 \frac{G_X^i}{G^i}, \tag{2.18}$$

$$= \beta_{TW}^i \mu^{TW} + \beta_t^i \mu_t, \tag{2.19}$$

where $\mu^{TW} = \sigma_C^2 \gamma$ and $\beta_{TW}^i = \frac{\sigma_D \sigma_C \rho}{\sigma_C^2}$ are the expected return and firm beta on the total wealth portfolio, while μ_t and β_t^i are the risk premium and firm beta on the economic growth hedge portfolio. Notice that β_{TW}^i is common across all assets, suggesting that reintroducing correlation between consumption and dividend growth has no effect on cross-sectional results. Further relaxing assumptions to include non-zero correlation between consumption and dividend shares would cause firms to have differing exposures to this risk, but the effects would be rather small since μ^{TW} is relatively small for reasonable parameter values. Meanwhile, long-run risk produces considerable variation in the cross section of expected returns. Investors are willing to realize lower returns on assets that hedge against unfavorable economic growth shocks, while they require large premiums on assets whose returns strongly covary with aggregate growth shocks.

2.2.3.2 Asset Pricing with Leverage

While the dispersion and IV effects can be analyzed whether or not firms employ leverage, the credit risk effect is exclusively studied in the context of leverage. We adopt a simple specification for firm debt from Merton (1974). In this framework, a firm takes on zero-coupon debt with a face value of B payable at time T. The firm defaults if its value at time T falls below the face value of debt. If the firm does default, all its value is distributed to debtholders. Otherwise, shareholders receive the residual firm value, the value remaining after the principal payment to debtholders. Despite its simplicity, Merton's (1974) framework achieves empirical success in forecasting firm defaults. Duffie, Saita, and Wang (2007) establish that distance to default, a key determinant of credit risk in the Merton model, is an important predictor of bankruptcies and defaults.

To formalize, the value of levered equity is given by

$$V_t^i = E_t \int_t^T \frac{\Lambda_s}{\Lambda_t} D_s^i ds + E_t \left[\frac{\Lambda_T}{\Lambda_t} \max\left(P_T^i - B, 0\right)\right]. \tag{2.20}$$

The first component is the expected discounted dividend stream paid prior to bond maturity, while the second is the expected discounted payoff at bond maturity. Closed-form solutions for equity value, beta, and expected return are unavailable in the presence of leverage. However, these quantities can be estimated through simulation. Appendix A.4.4 describes the simulation details.

In the presence of leverage, equity and debt expected returns still follow a conditional ICAPM.[9] Simulations show that equity value is more sensitive to economic growth shocks than debt value. Since the influence of economic growth shocks on equity is amplified by leverage, levered equity beta is generally higher than

[9] Firm levered equity return follows the process $dR_t^{V_i} = \mu_t^{V_i} dt + \phi_X^i dW_X + O(\sqrt{dt})$, where ϕ_X^i is the elasticity of firm equity value with respect to aggregate growth shocks. The $O(\sqrt{dt})$ term contains all other diffusion terms (e.g. dW_θ and dW_D) that are uncorrelated with growth shocks. Then $\mu_t^{V_i} = \frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_X \phi_X^i = \phi_X^i \frac{G}{\sigma_X G_X}\mu_t \equiv \beta_t^{V_i}\mu_t$. We do not have an analytical expression for ϕ_X^i, which prevents us from deriving a closed-form solution for $\beta_t^{V_i}$.

unlevered beta. Figure 2.3 exhibits estimated conditional equity betas for firms with one year to debt maturity and a market debt ratio of 0.5, confirming that levered equity betas are indeed higher than their unlevered counterparts (Figure 2.3 versus Figure 2.2). Therefore, expected excess equity returns for levered firms are higher than for unlevered firms (Figure 2.3 versus Figure 2.2), while the positive expected return–relative share relation persists among levered firms.

We are now ready to develop our model's implications for the cross-sectional associations between expected return and (i) dispersion, (ii) IV, and (iii) credit risk. We ultimately confirm that our single-factor model provides a unified resolution of these three apparently puzzling negative relations.

2.3 Three Cross-Sectional Effects

2.3.1 Dispersion

Based on the firm dividend growth process formulated in equation (2.12), dispersion is proxied by

$$Dispersion_t^i = \left| \frac{\sqrt{\delta^2 \frac{1-\theta_t^i}{\theta_t^i} + \sigma_D^2}}{\alpha\left(\frac{\bar{\theta}^i}{\theta_t^i} - 1\right) + \mu_D + X_t} \right|. \tag{2.21}$$

Our dispersion proxy measures the uncertainty about future dividend growth, defined as the standard deviation of future dividend growth normalized by expected dividend growth. Diether, Malloy, and Scherbina (2002) empirically examine the dispersion of analysts' earnings forecasts, which can be viewed as an alternative proxy for uncertainty about the underlying cash flow dynamics. Diether, Malloy, and Scherbina (2002) show a positive relation between dispersion in earnings forecasts and cash flow variability but dismiss a risk-based explanation based on cash flow uncertainty due to the apparently puzzling negative relationship between expected return and dispersion. Instead, they attribute the dispersion effect to equity mispricing. However, Barron, Stanford, and Yu (2009) decompose dispersion into earnings uncertainty and information asymmetry components using the model of Barron, Kim, Lim, and Stevens (1998). Cash flow uncertainty is significantly negatively associated with future returns, while information asymmetry carries a significant positive relation with future returns. That the dispersion effect is driven by earnings uncertainty is inconsistent with the mispricing explanation of Diether, Malloy, and Scherbina (2002) but is consistent with our own rational explanation. Finally, Zhang (2006) and Huang (2009) directly demonstrate negative empirical relations between expected return and measures of cash flow variability closely related to ours.

Theorem 4 establishes sufficient conditions for a negative relation between relative share and dispersion.

Theorem 4. *Dispersion decreases in relative share under the restrictions* $\alpha\left(\frac{\bar{\theta}^i}{\theta^i_t}-1\right)+\mu_D+X_t>0$ *and* $\alpha\left(1-\bar{\theta}^i\left(1-\frac{\sigma_D^2}{\delta^2}\right)\right)>\mu_D+X_t$.

Proof. See Appendix A.3.4. □

In words, the first condition in Theorem 4 ensures that a firm has positive expected dividend growth. The second condition holds quite generally, since the term in parentheses is close to one and estimates of α are much larger than μ_D for typical parameter estimates. Overall, the two conditions for the negative relation between dispersion and relative share emerge quite generally for the large majority of firms in the cross section.[10] Simulations show that dispersion and relative share exhibit a negative relation across the full cross section under a wide variety of conditions. Further discussion appears in Appendix A.3.4.

Theorem 4 establishes a negative cross-sectional relation between dispersion and expected return. Since expected return is increasing in relative share, our model implies a negative expected return–dispersion relation in the cross section. Figure 2.4 illustrates this negative relation among simulated firms.

2.3.2 Idiosyncratic Volatility

While the IV effect seems to be at odds with economic theory, it naturally arises within our setup. We first note that the dispersion effect augments the IV effect. In particular, low relative share firms display higher cash flow risk, portending larger idiosyncratic shocks among these firms. This is the first factor triggering the IV effect. The second contributing factor is the cross-sectional pattern in price sensitivity to dividend share shocks. Low relative share firms are particularly vulnerable to firm-specific shocks due to their heavy dependence on short-run cash flows. In contrast, the values of high relative share firms are closely tied to systematic shocks to economic growth while they are more immune to cash flow shocks, leading to high expected returns matched with low IV. This second mechanism is stronger than the first, and a substantial IV effect would arise in our economy even in the absence of a dispersion effect.

[10] The relation between dispersion and relative share is positive among firms with negative expected dividend growth. Since the large majority of firms have positive expected dividend growth, the cross-sectional relation holds in general.

To formalize, the idiosyncratic component of instantaneous return volatility formulated in equation (2.15) is

$$IV_t^i = \delta\sqrt{\frac{1-\theta_t^i}{\theta_t^i}}\left(1+\frac{\theta_t^i G_\theta^i}{G^i}\right). \tag{2.22}$$

The elasticity of firm value with respect to dividend share, $\frac{\theta_t^i P_\theta^i}{P^i}$, can be written

$$\begin{aligned}\frac{\theta_t^i P_\theta^i}{P^i} &= 1+\frac{\theta_t^i G_\theta^i}{G^i} \\ &= \frac{\int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds}{\int_t^\infty S(X_t,\tau)\left(e^{-\alpha\tau}+\frac{\bar{\theta}^i}{\theta_t^i}\left(1-e^{-\alpha\tau}\right)\right)ds},\end{aligned} \tag{2.23}$$

so the term in parentheses from equation (2.22) simply reflects firm value's sensitivity to firm-specific dividend share shocks. This elasticity is always between zero and one and sharply decreases in relative share. Therefore, the value of any firm increases with a positive shock to dividend share, but the increase is less than one-to-one. More importantly, the values of low relative share firms have greater sensitivities to firm-specific dividend share shocks. Meanwhile, the values of high relative share firms are relatively inelastic with respect to changes in dividend share. This pattern in price responses to firm-specific shocks implies that a negative relative share–idiosyncratic volatility relation holds quite generally.

From a technical perspective, idiosyncratic volatility decreases in relative share under the following sufficient condition.

Theorem 5. *Idiosyncratic volatility decreases in relative share under the restriction on the $(\bar{\theta},\alpha)$-plane $\left(\frac{\bar{\theta}^i}{\theta_t^i}-2\bar{\theta}^i\right)(\sqrt{2\alpha}M-1)>1$ where $M=\frac{|S(X_t,\tau)|_1}{|S(X_t,\tau)|_2}$ is independent of α, and $|\cdot|_i$ represents the i-th norm.*

Proof. See Appendix A.3.4. □

A condition for Theorem 5 to hold is that $\theta_t^i < 1/2$, which occurs almost surely given reasonable firm-level parameters and must obtain for at least $n-1$ of the n firms in the cross section. For a given $\theta_t^i < 1/2$, any $(\bar{\theta}^i, \alpha)$ that satisfy the restriction imposed by Theorem 5 will give rise to a negative relation between relative share and IV. In fact, since $S > 0$, the 1-norm of S is much greater than the 2-norm and $M >> 1$, such that the above condition flexibly holds under a wide set of parameter values.

Theorem 5 establishes a negative relation between relative share and IV. This relation in turn implies a negative relation between IV and expected return. Figure 2.4 provides simulation evidence of this negative relation.

In sum, the apparently puzzling dispersion and IV effects are attributable to cross-sectional differences in exposure to systematic risk. Our resolution for the dispersion and IV effects is substantially different from that of Johnson (2004). In his option-pricing setup, increasing IV leads to higher current levered equity values, which in turn leads to lower expected returns as future expected payoffs remain unchanged. While Johnson's explanation arises from exposure to idiosyncratic risk, our explanation is rooted in exposure to time-varying priced as well as unpriced sources of risk. Moreover, Johnson's model relies on leverage to produce the dispersion and IV effects. Our model produces the dispersion and idiosyncratic volatility effects for both levered and unlevered firms, consistent with the empirical evidence of Diether, Malloy, and Scherbina (2002) and Ang, Hodrick, Xing, and Zhang (2006).

2.3.3 Credit Risk

Empirical work has uncovered a negative relation between credit risk and expected return (e.g., Dichev (1998), Campbell, Hilscher, and Szilagyi (2008), and

Avramov, Cederburg, and Hore (2009)). This evidence initially seems quite surprising, as high credit risk firms are commonly perceived to be riskier. The literature has typically offered non-risk-based explanations for the credit risk effect. However, we show that the credit risk effect arises naturally in our model due to cross-sectional differences in exposures to systematic and idiosyncratic risks.

From an intuitive perspective, the large majority of a firm's return volatility is idiosyncratic (e.g. Campbell, Lettau, Malkiel, and Xu (2001)). Therefore, one would expect idiosyncratic shocks to be the primary cause of defaults.[11] Low relative share firms with high levels of idiosyncratic risk are most likely to experience this type of default. Meanwhile, high relative share firms have low IV and do not often default for firm-specific reasons. Default risk models designed to predict bankruptcies and defaults on a year-to-year basis (e.g., Altman (1968), Ohlson (1980), and Campbell, Hilscher, and Szilagyi (2008)) are likely to identify firms with high IV as high default risk firms. However, in equilibrium these firms have low systematic risk exposures leading to a negative relation between expected return and distress risk estimates.

In contrast to the relatively frequent idiosyncratic defaults, occasional large negative shocks to expected economic growth influence all firms and may lead to a "default wave" with many firms in the cross section simultaneously experiencing financial distress. The default waves produced our model are consistent with the empirical evidence of Chen (2009). When the economy experiences a negative

[11] Consistent with this intuition, Campbell and Taksler (2003) find that the IV of firm equity is as effective as credit rating in explaining cross-sectional variation in credit spreads. Additionally, the time series of credit spreads generally matches up with average IV, further suggesting IV plays an important role in defaults.

systematic shock, high relative share firms with large economic growth betas are most affected. As such, many of the firms caught in a default wave are those which appeared to have high credit quality ex ante. Investors require additional compensation for holding high relative share firms and taking on exposure to default waves despite their low probability.

In the Merton (1974) framework, default occurs if the firm value drops below the face value of bonds at maturity, that is, if $P_T < B$. We do not have a closed-form solution for the probability of default. Instead, we estimate the default probability by simulating the firm value through the bond maturity date and then calculating the percentage of simulations in which the firm defaults. In the simulation, we assume each firm has zero-coupon debt with one year to maturity and a debt ratio of 50%.[12] The simulation procedure is explained in more detail in Appendix A.4.4.

Our model produces a robust negative relation between expected return and credit risk in simulations. Low relative share firms tend to default more frequently than high relative share firms. This effect translates to a negative credit risk–expected return relation, illustrated by Figure 2.4. The default events among high relative share firms in the simulation are infrequent but concentrated in times of poor economic conditions, when the expected economic growth rate has experienced a large negative shock. Meanwhile, low relative share firms default relatively

[12]For simplicity, we assume that all firms have a face value of debt equal to 50% of the firm value at time 0. This assumption biases us against finding a negative relation relative to the case where all firms have the same market debt ratio, since the debt of the high credit risk firms has lower value so these firms have lower leverage. Additionally, George and Hwang (2010) show that firms with potentially high distress costs (a close parallel to the high relative share firms in our model) will optimally choose lower levels of leverage, while in our simulation these firms have higher leverage.

frequently and more uniformly across economic conditions. In general, our model produces a robust negative relation between expected return and credit risk when parameters are set to match the empirical regularity that idiosyncratic volatility is higher than systematic volatility for most firms in the cross section.

In sum, we have demonstrated theoretical connections between the dispersion, IV, and credit risk effects. Cross-sectional differences in cash flow timing tend to produce a negative relation between firm-specific risk and exposure to systematic risk in the cross section. Expected returns are negatively related to each of the three firm characteristics as a result. We next turn to empirical evidence on the three anomalies.

2.4 Empirical Evidence

This section presents empirical evidence about the dispersion, IV, and credit risk effects based the sample period July 1981 through June 2008. Our model predicts that the dispersion, IV, and credit risk measures are related across firms, as each measure is driven by the same relative share characteristic. Our model also implies that all three puzzling effects are mutually related as they are all explained by the same common factor, namely economic growth. The empirical evidence broadly supports our model's predictions. In particular, the three anomalies are indeed tightly positively linked in the data. Moreover, higher dispersion, IV, and credit risk firms display lower exposures to long run risk thereby commanding lower risk premiums. Finally, the expected returns implied by our model generally match the empirical patterns in average returns.

2.4.1 Relating the Anomalies

Prior research suggests positive relations between dispersion and IV (e.g., Barron, Stanford, and Yu (2009)), dispersion and credit risk (e.g., Avramov, Chordia, Jostova, and Philipov (2009)), and IV and credit risk (e.g., Campbell and Taksler (2003)). Here, we simultaneously examine the relations between all three measures. The analysis is based on the dispersion measure of Diether, Malloy, and Scherbina (2002), the IV measure of Ang, Hodrick, Xing, and Zhang (2006), and the financial distress measure of Campbell, Hilscher, and Szilagyi (2008). Additional details about data construction are available in Appendix A.4.1.

Panel A of Table 2.1 exhibits the time-series average of the cross-sectional Spearman rank correlations between dispersion, IV, and financial distress within each month. The evidence shows that all three characteristics are positively related. A firm with high IV is quite likely to have high distress risk as evidenced by the correlation of 0.41 between the two measures. Similarly, the dispersion–IV correlation is 0.31 and the dispersion–default probability correlation is also 0.31.

Panel B of Table 2.1 examines the characteristics of firms in three groups of decile portfolios sorted on dispersion, IV, and distress risk. It displays the mean dispersion, IV, and probability of default across firms in each portfolio. In the ten dispersion portfolios, both IV and the probability of default are nearly monotonically increasing as the dispersion ranking increases. Similar results hold for the IV and credit risk portfolios, with all three characteristics having nearly monotonic patterns across the portfolios.

More specifically, moving from the lowest to highest dispersion portfolio, the mean IV increases from 2.40% to 3.67%, and the mean default probability rises from 0.08% to 0.30%. Similarly, moving from the lowest to highest IV portfolio, the mean dispersion rises from 0.07 to 0.49, and the mean default probability advances from 0.05% to 0.97%. Finally, moving from the lowest to highest distress risk portfolio, the mean dispersion rises from 0.09 to 0.65, and the mean IV increases from 2.25% to 5.80%.

Overall, the evidence exhibited in Table 2.1 illustrates that a strong empirical link exists between the dispersion, IV, and credit risk measures. The positive relations between these measures suggest that portfolios formed on the basis of each characteristic will be similar. Indeed, we find that a stock in the top decile of the IV and credit risk characteristics has a 39% chance of lying in the top dispersion decile and 73% chance of being in the top three deciles of dispersion.

To establish the relation between the dispersion, IV, and credit risk anomalies, we examine the properties of returns of value-weighted decile portfolios for each characteristic. Portfolio formation details are in Appendix A.4.1. Table 2.2 exhibits statistics summarizing portfolio returns.

Panel A merely displays evidence already documented in past work. In particular, there are statistically significant and economically large negative relations between average returns and dispersion, IV, and credit risk. To illustrate, buying low dispersion (IV) [default probability] firms and selling high dispersion (IV) [default probability] firms yields an investment payoff of 77 (190) [117] basis points per month. All figures are highly significant.

In Panel B, we investigate the joint properties of extreme decile portfolio returns for each of the anomalies. The correlation matrix shows that the returns of the low dispersion, IV, and credit risk portfolios are highly correlated (all pairwise correlations are at least 0.83). Similarly, the high dispersion, IV, and credit risk portfolio returns have substantial comovement, with correlations of at least 0.81. In contrast, cross correlations of low and high anomaly portfolios are relatively low, with several correlations less than 0.50. In sum, portfolios formed on the three anomaly variables behave similarly.

Indeed, we have established an empirical link between the dispersion, IV, and credit risk anomalies. The three variables are related in the cross section of stocks, and portfolios formed on the basis of these characteristics are similar. The remaining task is to show that all three effects are explained by exposures to long run risk. We establish this result below.

2.4.2 Model Estimation

We estimate aggregate economy parameters based on the processes in equations (2.3), (2.4), and (2.5). To estimate portfolio-level parameters, we construct a quarterly dividend time series for each portfolio following Bansal, Dittmar, and Lundblad (2005). Dividend share is calculated as the portfolio dividend divided by the sum of dividends across all ten portfolios. Overall, we consider 30 portfolios including ten dispersion, ten IV, and ten credit risk portfolios. Aggregate parameters as well as portfolio dividend share parameters are estimated in a Bayesian framework. Full details of the estimation methodology are available in Appendix A.4.2.

We first examine the WF dividend share process formulated in equation (2.8). The evidence supports mean reversion in dividend shares. In particular, we reject the hypothesis of a unit root at a 5% level for 19 of 30 portfolios. We then compare our dividend share process to several alternative specifications using the Akaike Information Criterion–Monte Carlo (AICM) and Bayesian Information Criterion–Monte Carlo (BICM) model selection criteria of Raftery, Newton, Satagopan, and Krivitsky (2007), which are posterior simulation-based versions of the AIC and BIC criteria. The alternative processes considered are

$$d\theta_t^i = \alpha^i(\bar{\theta}^i - \theta_t^i)dt + \delta^i\sqrt{(1-\theta_t^i)\theta_t^i}dW_{\theta^i}, \tag{2.24}$$

$$d\theta_t^i = \alpha(\bar{\theta}^i - \theta_t^i)dt + \sigma^i dW_{\theta^i}, \tag{2.25}$$

$$d\theta_t^i = \mu^i dt + \sigma^i dW_{\theta^i}. \tag{2.26}$$

Specification (2.24) is a WF process which allows for firm-specific α and δ parameters, specification (2.25) eliminates the dependence of the diffusion term on θ_t^i, and specification (2.26) is a simple Brownian motion with firm-specific drift and diffusion terms. Our dividend share process compares favorably, outperforming all alternatives for the IV and credit risk portfolios. For dispersion portfolios, our process dominates formulations (2.25) and (2.26), and is only slightly outperformed by specification (2.24). Overall, our dividend share process is parsimonious and provides a good fit to the data. Additional details of the dividend share process tests appear in Appendix A.4.3.

Table 2.3 reports the annualized mean draws from the posterior distributions of parameters underlying the aggregate economy as well as cross section of portfolios.

Panel A exhibits the estimates for aggregate parameters. The table also reports our choices for the risk aversion γ (7.5) and time preference β (0.01) parameters. Panel B of Table 2.3 exhibits the mean draws of relative share and the long-run dividend share for each portfolio, as well as the dividend share process parameters α and δ. The evidence shows that, for each of the three sets of portfolios, there exists a strong relation between relative share and portfolio rank. The lowest dispersion, IV, and distress risk portfolios display the highest relative share estimates, given by 5.81, 1.32, and 1.54, respectively. Meanwhile, high dispersion, IV, and distress risk portfolios display quite low relative share estimates of 0.53, 0.02, and 0.03, respectively. In each of the three cases, there is a substantial spread in relative share across the decile portfolios.

Our estimates of relative share reflect differential dividend growth rates across the decile portfolios. In fact, the average dividend growth of low dispersion, IV, and distress risk portfolios is much higher than that of the corresponding high measure portfolios. Indeed, relative share estimates of the portfolios imply that each of the three characteristics should be negatively related to expected returns.

Table 2.4 reports the model-implied expected returns for the dispersion, IV, and credit risk portfolios. The patterns in expected returns generally match those of realized returns, suggesting that exposure to unexpected changes in economic growth may explain much of the cross-sectional variation in expected returns of these portfolios. The model generates large spreads in expected returns, with long-short portfolios earning expected annual returns of 2.8% (dispersion), 4.7% (IV), and 5.0% (credit risk). Further, the model is able to capture this variation in returns despite forming expected return estimates that use only aggregate consumption, aggregate dividend, and portfolio-level dividend data.

However, it should be noted that our expected return estimates tend to be "too flat" relative to realized returns. There are at least two potential explanations of the result that realized returns vary more than expected returns. First, in this paper we make simplifying assumptions to achieve tractability and establish theoretical results on the relations between expected return and dispersion, IV, and credit risk. The magnitudes of cross-sectional spreads in expected returns may be matched by relaxing such assumptions as zero correlations between dividend shares and aggregate variables or unit EIS. In support of this explanation, the model-implied expected returns achieve a high rank correlation with realized average returns, suggesting the model may be capturing the underlying economic mechanism. Alternatively, long-run risk may provide a reasonably good, albeit imperfect, explanation of market anomalies, allowing our model to capture only a fraction of the variation in expected returns across the portfolios. In general, however, long-run risk does provide a qualitative and intuitive resolution of the dispersion, IV, and credit risk effects.

2.5 Conclusion

This paper develops and applies an intertemporal asset pricing model in a long-run risk economy with a formal cross section of firms. Expected returns in the economy are positively related to cash flow duration, while firms in the cross section are characterized by expected cash flow timing. Firms with high relative share, which corresponds to high expected dividend growth, are more sensitive to shocks to the persistent economic growth rate and hence command higher risk premiums.

Our model suggests that the puzzling negative cross-sectional relations between expected stock returns and analysts' forecast dispersion, idiosyncratic volatility, and credit risk emerge as an equilibrium response to long-run risk. While high duration firms are highly exposed to systematic shocks, low duration firms are particularly sensitive to firm-specific dividend shocks. As a result, firms with high measures of idiosyncratic risk tend to have low systematic risk and low expected returns, which explains the observed patterns.

The model predictions are broadly supported in the data. For one, the evidence shows that the dividend shares of dispersion, IV, and credit risk portfolios mean revert, and our dividend share process dominates alternative specifications based on model selection criteria. Moreover, dispersion, IV, and credit risk measures are highly positively related in the data, and portfolios formed based on these three characteristics display similar attributes. More importantly, high dispersion, IV, and credit risk firms exhibit low cash flow duration, thereby command relatively low risk premiums.

This paper contributes to the growing long-run risk literature, which offers explanations for a variety of aggregate and cross-sectional asset pricing puzzles, including the equity premium, risk-free rate, and excess volatility puzzles, as well as the momentum and value effects. Our contributions are in both methodology and substantive findings. We are the first to derive formal cross-sectional asset pricing restrictions in a long-run risk setting summarized by an intuitive conditional beta representation. Moreover, we demonstrate that the analysts' forecast dispersion, idiosyncratic volatility, and credit risk effects are inherently linked through cross-sectional exposures to idiosyncratic and systematic sources of risk. Overall, our model generates aggregate and firm-level return characteristics that largely match the empirical evidence and offers risk-based explanations of puzzling cross-sectional effects.

Table 2.1: Relation between Dispersion, Idiosyncratic Volatility, and Distress Risk

Panel A: Spearman Rank Correlation	Dispersion	Idiosyncratic Volatility	Default Probability
Dispersion	1.00	0.31	0.31
Idiosyncratic Volatility		1.00	0.41
Default Probability			1.00

Panel B: Average Characteristics of Firms in Portfolios	Low	2	3	4	5	6	7	8	9	High
Dispersion Portfolios										
Dispersion	0.00	0.01	0.02	0.03	0.04	0.06	0.08	0.12	0.23	1.53
IV	2.40	1.95	2.07	2.19	2.34	2.50	2.69	2.93	3.25	3.67
Default Probability	0.08	0.05	0.06	0.07	0.09	0.11	0.15	0.20	0.26	0.30
IV Portfolios										
Dispersion	0.07	0.10	0.13	0.17	0.20	0.23	0.29	0.34	0.41	0.49
IV	0.77	1.27	1.63	1.98	2.36	2.80	3.32	4.01	5.07	8.43
Default Probability	0.05	0.05	0.06	0.08	0.10	0.15	0.22	0.32	0.52	0.97
Distress Risk Portfolios										
Dispersion	0.09	0.09	0.10	0.13	0.18	0.20	0.27	0.40	0.57	0.65
IV	2.25	2.25	2.30	2.41	2.53	2.65	2.97	3.58	4.50	5.80
Default Probability	0.02	0.02	0.03	0.04	0.05	0.06	0.07	0.10	0.18	2.52

Note: This table reports statistics on the relation between dispersion, idiosyncratic volatility (IV), and distress risk (Default Probability). Dispersion is the standard deviation of earnings forecasts divided by the absolute value of the mean forecast. IV is the standard deviation of the residual from a Fama–French (1993) three-factor regression using daily returns from the month prior to portfolio formation. The probability of default is based on the Campbell, Hilscher, and Szilagyi (2008) measure. Panel A reports the correlations between dispersion, IV, and default probability. The reported correlations are time-series averages of the monthly Spearman cross-sectional correlations between the three measures. Panel B shows the average dispersion, IV, and default probability for firms appearing in decile portfolios sorted on each of the three variables. The dispersion and IV portfolios are rebalanced monthly, while the distress risk portfolio is rebalanced annually in July. The sample period is July 1981 to June 2008.

Table 2.2: Anomaly Portfolio Return Characteristics

Panel A: Average Returns of Anomaly Portfolios

	Low	2	3	4	5	6	7	8	9	High	High-Low
Dispersion	0.99	0.56	0.46	0.62	0.46	0.72	0.47	0.60	0.28	0.22	-0.77
	(4.20)	(2.47)	(1.94)	(2.42)	(1.70)	(2.50)	(1.67)	(1.83)	(0.82)	(0.59)	(-2.64)
IV	0.73	0.71	0.57	0.61	0.55	0.71	0.24	0.05	-0.37	-1.17	-1.90
	(3.73)	(3.19)	(2.23)	(2.18)	(1.77)	(1.98)	(0.60)	(0.11)	(-0.74)	(-2.32)	(-4.13)
Def Prob	0.78	0.66	0.65	0.62	0.63	0.44	0.41	-0.08	0.06	-0.39	-1.17
	(3.43)	(2.94)	(2.81)	(2.57)	(2.34)	(1.56)	(1.24)	(-0.20)	(0.13)	(-0.73)	(-2.74)

Panel B: Anomaly Portfolio Return Correlations

	Low Dispersion	Low IV	Low Def Prob	High Dispersion	High IV	High Def Prob
Low Dispersion	1.00	0.85	0.84	0.64	0.48	0.49
Low IV		1.00	0.83	0.61	0.41	0.47
Low Def Prob			1.00	0.75	0.61	0.63
High Dispersion				1.00	0.81	0.81
High IV					1.00	0.84
High Def Prob						1.00

Note: This table reports average excess returns and correlations for portfolios sorted on analysts' forecast dispersion, idiosyncratic volatility (IV), and distress risk (Def Prob). Dispersion is the standard deviation of earnings forecasts divided by the absolute value of the mean forecast. IV is the standard deviation of the residual from a Fama–French (1993) three-factor regression using daily returns from the month prior to portfolio formation. The probability of default is based on the Campbell, Hilscher, and Szilagyi (2008) measure. Portfolios are value weighted. The dispersion and IV portfolios are rebalanced monthly, while the distress risk portfolio is rebalanced annually in July. Panel A shows average real monthly excess returns and t-stats in parentheses. Panel B shows the correlations between the real monthly returns of the extreme decile portfolios. The sample period is July 1981 to June 2008.

Table 2.3: Parameter Estimates

Panel A: Aggregate Parameters

γ	β	μ_C	μ_D	σ_C	σ_D	σ_X	λ	κ
7.500	0.010	0.009	0.041	0.010	0.061	0.025	0.253	0.051

Panel B: Portfolio Parameters

	Dispersion	Idiosyncratic Volatility	Distress Risk
		Relative Share $\left(\frac{\bar{\theta}^i}{\theta_0^i}\right)$	
Low	5.811	1.322	1.540
2	0.620	1.180	0.998
3	0.468	0.964	1.261
4	0.679	0.996	1.114
5	0.340	0.855	0.964
6	1.584	0.951	0.718
7	0.420	0.374	0.653
8	0.772	0.169	0.140
9	0.287	0.040	0.167
High	0.528	0.017	0.034
		Long-Run Mean Dividend Share $\left(\bar{\theta}^i\right)$	
Low	0.390	0.172	0.213
2	0.101	0.276	0.220
3	0.071	0.224	0.148
4	0.073	0.144	0.141
5	0.062	0.105	0.127
6	0.195	0.061	0.089
7	0.031	0.013	0.053
8	0.048	0.003	0.006
9	0.010	0.001	0.003
High	0.019	0.000	0.000
		Process Parameters (α and δ)	
α	0.055	0.312	0.224
δ	0.057	0.076	0.061

Note: This table reports aggregate and portfolio parameter estimates. Panel A shows our chosen preference parameters γ and β and estimates of the aggregate parameters from equations (2.3), (2.4), and (2.5). The parameters are estimated using annual US real per capita consumption and dividend data from 1948 to 2008. We use a Bayesian Markov chain Monte Carlo (MCMC) procedure that appears in Appendix A.4.2.1 to estimate the parameters. Panel B reports portfolio parameter estimates from equation (2.8) for portfolios sorted on dispersion, idiosyncratic volatility, and credit risk. Portfolios are value weighted and the sample period is July 1981 to June 2008. We follow Bansal, Dittmar, and Lundblad (2005) to calculate quarterly portfolio dividends to obtain a time series of dividend share. The MCMC procedure used to estimate the portfolio parameters is in Appendix A.4.2.2. All figures are the mean of the posterior distribution of the parameter and are annualized.

Table 2.4: Model-Implied Expected Returns of Anomaly Portfolios

	L	2	3	4	5	6	7	8	9	H	H-L
Dispersion	0.90	0.74	0.70	0.74	0.67	0.82	0.67	0.74	0.61	0.67	-0.23
	(0.14)	(0.12)	(0.12)	(0.12)	(0.11)	(0.12)	(0.12)	(0.13)	(0.13)	(0.13)	(0.11)
IV	0.81	0.78	0.76	0.78	0.77	0.79	0.66	0.57	0.44	0.42	-0.39
	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.13)	(0.10)	(0.10)	(0.09)
Def Prob	0.84	0.76	0.80	0.77	0.75	0.71	0.71	0.52	0.54	0.43	-0.41
	(0.12)	(0.11)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.10)	(0.08)

Note: This table reports model-implied expected excess returns for portfolios sorted on analysts' forecast dispersion, idiosyncratic volatility (IV), and distress risk (Def Prob). The returns of the value-weighted decile portfolios are from the sample period July 1981 to June 2008. Expected returns are given by equation (2.14). We report the mean draw of expected excess return from 20,000 iterations of the simulation, with the aggregate and portfolio parameters drawn from their posterior distributions in each iteration. See Appendix A.4.2.1 for details on aggregate parameter estimation, Appendix A.4.2.2 for portfolio parameter estimation, and Table 1.3 for the mean draws of parameters. The standard deviation of posterior draws of expected returns are reported in parentheses.

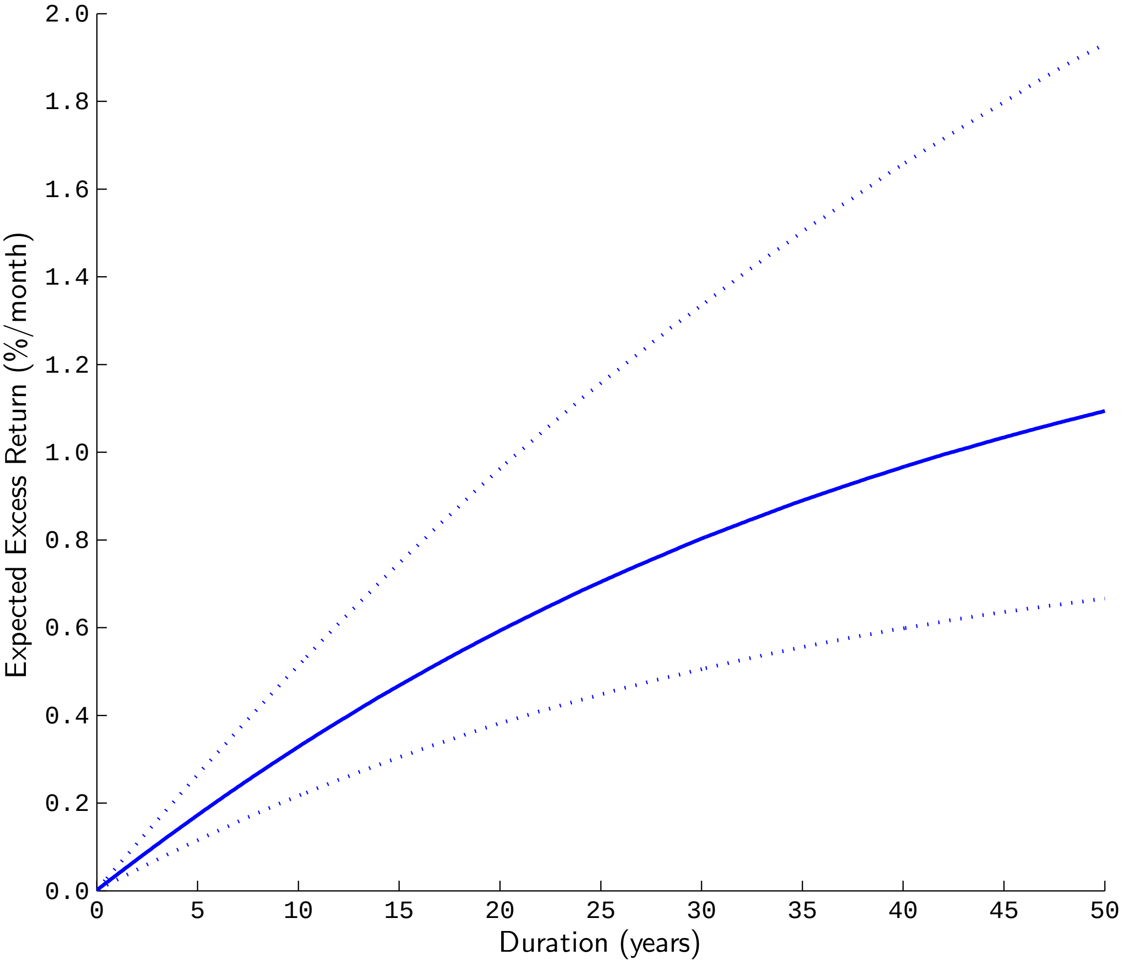

Figure 2.1: Cash Flow Duration and Expected Returns

Note: This figure shows the expected monthly excess return of an asset as a function of its cash flow duration. The asset considered pays a non-negative portion of the aggregate dividend at all times t. The dotted lines show the 90% confidence band for expected excess returns given the posterior distribution of aggregate parameters. See Appendix A.4.2.1 for details on aggregate parameter estimation and Table 2.3 for the mean draws of aggregate parameters.

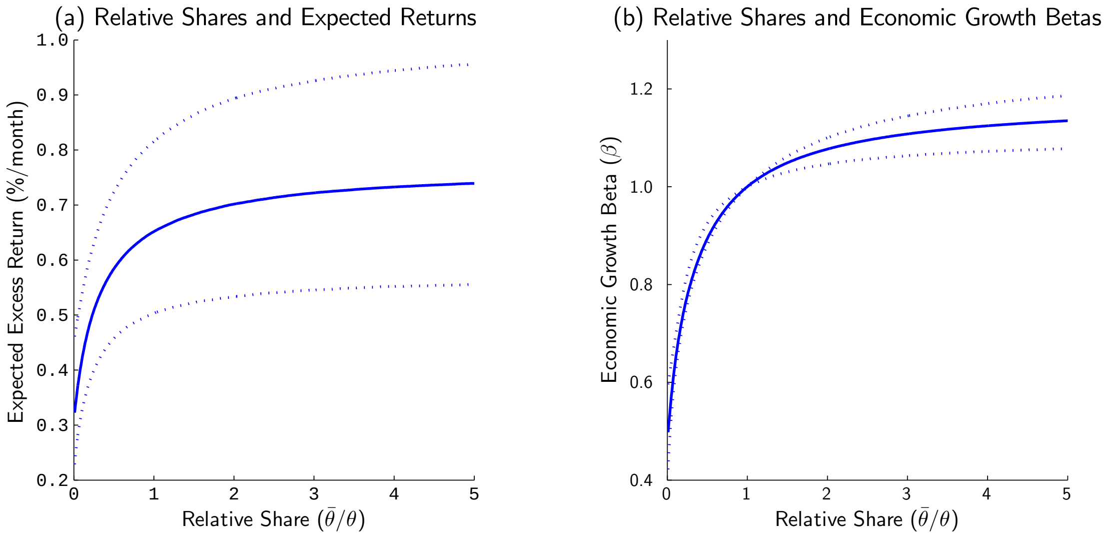

Figure 2.2: Expected Excess Returns and Economic Growth Betas for Unlevered Firms

Note: This figure shows expected excess returns and economic growth betas as a function of firm relative share for an unlevered firm. The first figure graphs the expected monthly excess return against relative share. The expected excess return of firm i is given by equation (2.14). The second figure plots the firm's beta relative to an economic growth hedge portfolio, given by equation (2.16), as a function of relative share. The dotted lines show the 90% confidence bands given the posterior distribution of aggregate and portfolio parameters. See Appendix A.4.2.1 for details on aggregate parameter estimation, Appendix A.4.2.2 for portfolio parameter estimation, and Table 2.3 for the mean draws of parameters.

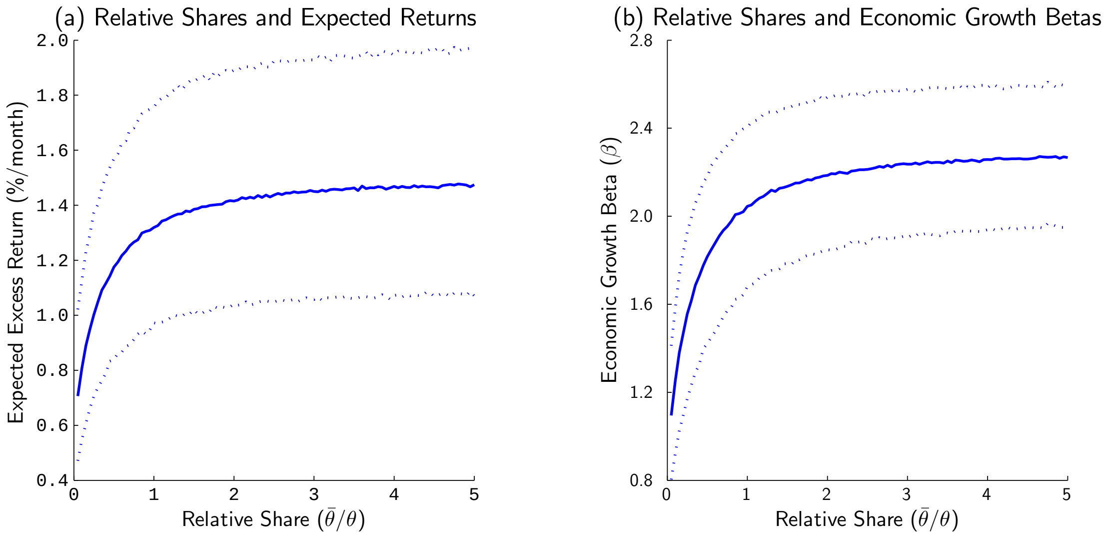

Figure 2.3: Expected Excess Returns and Economic Growth Betas for Levered Firms

Note: This figure shows expected excess returns and economic growth betas as a function of firm relative share for a levered firm. The firm is assumed to have a market debt ratio of 0.5 with one year to debt maturity. Following the Merton (1974) framework, a firm is assumed to default if its firm value at debt maturity is less than the face value of its debt. The first figure graphs the expected monthly excess return against relative share. The second figure plots the firm's beta relative to an economic growth hedge portfolio as a function of relative share. The dotted lines show the 90% confidence bands of simulated values from 5,000 iterations of the simulation, where the aggregate parameters in each iteration are drawn from their posterior distribution. See Appendix A.4 for details on aggregate and portfolio parameter estimation and the simulation procedure.

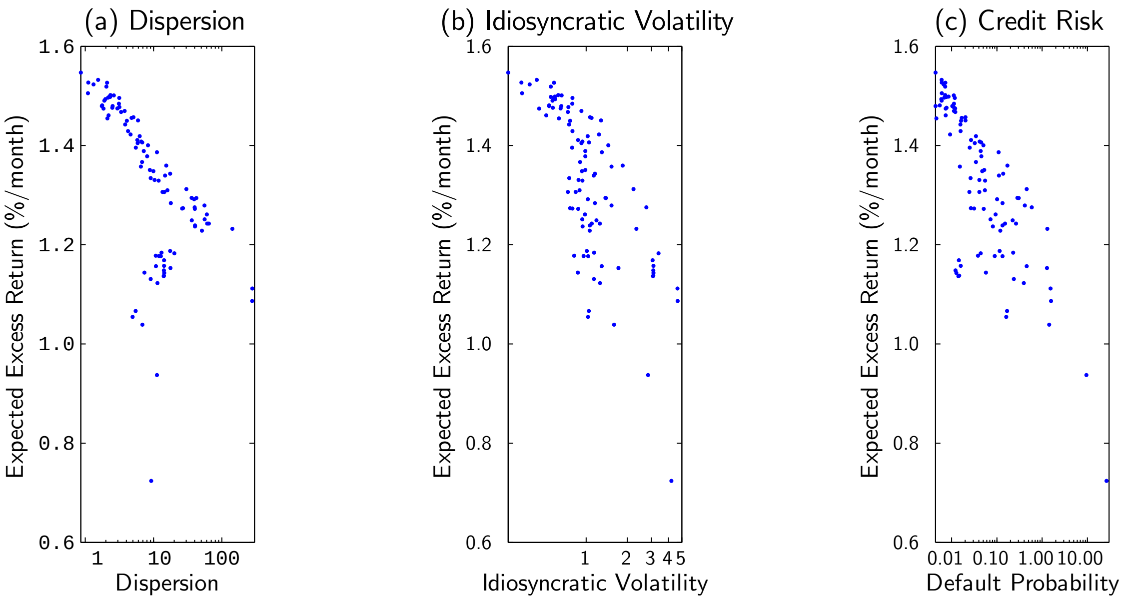

Figure 2.4: Expected Returns and Firm Characteristics

Note: This figure plots the cross-sectional relations between expected excess return and each of dispersion, idiosyncratic volatility (IV), and credit risk. The plots are based on a cross section of 100 simulated firms. For each firm, dispersion and IV are calculated based on equations (2.21) and (2.22), respectively. Credit risk is estimated from the percentage of simulated draws in which the firm defaults. The firm is assumed to have a market debt ratio of 0.5 with one year to debt maturity. Following the Merton (1974) framework, a firm is assumed to default if its value at debt maturity is less than the face value of its debt. See Appendix A.4.4 for full details on the simulation procedure.

CHAPTER 3
ASSET-PRICING ANOMALIES AT THE FIRM LEVEL

3.1 Introduction

An anomaly is a pattern in average stock returns that is inconsistent with the predictions of the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965). Anomalies are commonly identified using a portfolio-based approach. The researcher sorts stocks on a firm characteristic and constructs a zero-cost hedge portfolio by taking long and short positions in the extreme groups. If the hedge portfolio earns abnormal returns relative to the CAPM, the sorting characteristic is classified as an anomaly. Over the past three decades, a large number of anomalies have been uncovered, suggesting the CAPM is unable to explain much of the cross-sectional variation in average stock returns.

There are, however, growing concerns in the literature about the use of portfolios to identify anomalies and, more generally, to test asset-pricing models. These arguments are centered around the idea that grouping firms into portfolios and aggregating returns wastes and potentially distorts valuable information about cross-sectional patterns in abnormal returns.[1] One way to avoid the concerns with portfolios is to use firm-level data. To examine anomalies at the firm level, however, the

[1] For example, Litzenberger and Ramaswamy (1979) and Ang, Liu, and Schwarz (2010b) consider the loss in efficiency from using portfolios rather than individual firms in asset-pricing tests, while Roll (1977), Kandel and Stambaugh (1995), and Fama and French (2008) discuss how patterns in firm-level pricing errors can be distorted at the portfolio level. Lo and MacKinlay (1990) highlight the data-snooping biases inherent in portfolio-based asset-pricing tests. Ahn, Conrad, and Dittmar (2009) and Lewellen, Nagel, and Shanken (2010) show inferences in asset-pricing tests are remarkably sensitive to the choice of test portfolios. For other issues, see Conrad, Cooper, and Kaul (2003), Kan (2004), and Daniel and Titman (2006b).

researcher has to relate firm characteristics to abnormal returns. Abnormal returns are not directly observable, so the researcher must model and estimate the evolution of betas – a challenging problem, especially at the firm-level.

Two recent firm-level studies adopt contrasting approaches to control for market risk. Avramov and Chordia (2006) model market risk as an exact linear function of firm size, book-to-market, and macroeconomic variables, while Fama and French (2008) argue that market risk should not be related to firm characteristics removing the need to examine abnormal returns. Both approaches are problematic. Specifying betas as an exact linear function of covariates is only valid if the researcher knows the full set of variables associated with variation in betas, while tests examining the relation between firm characteristics and raw returns will overstate the CAPM's failings if firm-level betas are associated with firm characteristics.

In this paper, we develop a hierarchical Bayes approach to explore anomalies at the firm level. Specifically, we simultaneously estimate (1) conditional CAPM model parameters for each firm using an approach similar to Lewellen and Nagel (2006) which specifies short time periods and avoids the need for conditioning information, (2) the cross-sectional relation between conditional alphas and firm characteristics in each time period, and (3) the systematic association between alphas and firm characteristics across the entire sample period. Our approach has several desirable features relative to the prior literature. We put little structure on the dynamics of conditional betas, thereby minimizing potential model mis-specification. Our one-step methodology eliminates a measurement error problem encountered in traditional two-step approaches (e.g., Brennan, Chordia, and Subrahmanyam (1998) and Avramov and Chordia (2006)). We also implicitly control for cross-sectional heteroskedasticity and cross-correlations among stocks.

We use this approach to examine nine anomalies over the period 1963 to 2008: size, book-to-market, momentum, reversal, profitability, asset growth, net stock issues, accruals, and financial distress.[2] Studying each anomaly separately, we find that firm-level associations are distorted at the portfolio level for four of the nine anomalies. For example, the traditional portfolio approach suggests size and reversal are associated with abnormal returns, but using information from the entire cross section of stocks there is no evidence of a robust relation between either of these variables and firm alphas. Further analysis suggests the portfolio-level results for size and reversal are driven by a small subset of stocks with extreme values for these characteristics. Nevertheless, the initial firm-level evidence still paints a bleak picture for the CAPM. Seven of the nine characteristics are significantly associated with alphas, suggesting that the CAPM does indeed fail across multiple dimensions. These results, however, may be misleading for three reasons.

First, a disadvantage of using the entire cross section of firms to study anoma-

[2]Previous papers show a positive relation between average returns and book-to-market equity (Rosenberg, Reid, and Lanstein (1985), Chan, Hamao, and Lakonishok (1991), and Fama and French (1992)), stock return momentum (Jegadeesh (1990) and Jegadeesh and Titman (1993, 2001)), and profitability (Haugen and Baker (1996) and Cohen, Gompers, and Vuolteenaho (2002)). There is a negative relation between average returns and size (Banz (1981) and Fama and French (1992)), stock return reversal (DeBondt and Thaler (1985) and Chopra, Lakonishok, and Ritter (1992)), asset growth (Fairfield, Whisenant, and Yohn (2003), Titman, Wei, and Xie (2004), and Cooper, Gulen, and Schill (2008)), net stock issues (Loughran and Ritter (1995), Ikenberry, Lakonishok, and Vermaelen (1995), Daniel and Titman (2006a), and Pontiff and Woodgate (2008)), accruals (Sloan (1996), Collins and Hribar (2000), and Xie (2001)), and financial distress (Dichev (1998) and Campbell, Hilscher, and Szilagyi (2008)).

lies is that inferences may be heavily influenced by small, illiquid stocks. It is possible that the anomalous patterns are being driven primarily by firms which represent only a tiny fraction of the total market capitalization. We investigate this issue by allowing associations between alphas and firm characteristics to vary across micro, small, and big stocks.[3] We find the associations are strongest in terms of statistical and economic magnitude for micro and small stocks. For big stocks, alphas are significantly associated with only three of the nine characteristics - asset growth, net stock issues, and accruals.

Second, anomalies could be the result of the temporary market mis-pricing or data snooping by researchers. These explanations suggest established deviations from the CAPM should weaken over time. To examine this issue, we consider whether the relation between alphas and each firm characteristic attenuates or persists after the anomaly is established in the asset-pricing literature. Anomalies that persist post publication are more likely to reflect a fundamental failure of the CAPM. Of the seven anomaly variables with sufficient post-publication sample periods only two – book-to-market and accruals – are significantly related to abnormal returns after publication. Further, these relations are driven by micro stocks for which transaction costs and liquidity concerns diminish investors' ability to exploit anomalies and correct mis-pricings (e.g., Jensen (1978)). Among big stocks, no firm characteristic is significantly associated with abnormal returns post publication.

Third, firm characteristics could be correlated with each other and offer lit-

---

[3]Following Fama and French (2008), we classify stocks into three size groups - micro, small, and big. The breakpoints are based on the 20th and 50th percentiles of market capitalization for NYSE stocks at the end of June each year.

tle unique information about abnormal returns. Asset-pricing tests that consider each firm characteristic in isolation are likely to suffer from an omitted variable bias that will result in the importance of an anomaly being overstated. Traditional portfolio approaches are unable to adequately address this omitted variable problem. Researchers typically rely on multi-dimensional sorts to isolate the effects of a particular characteristic, but controlling for more than one or two characteristics simultaneously is infeasible. In contrast, our approach is particularly well suited to assess which anomalies contain unique information; we simply specify conditional alphas as a function of multiple firm characteristics. Our results suggest that univariate tests do indeed suffer from a pronounced omitted variable bias. Considering all characteristics simultaneously we find that size, momentum, reversal, asset growth, and financial distress do not contain significant incremental information about abnormal returns, in contrast to the corresponding portfolio-level results.

Taken together, the results suggest that while the CAPM does not perfectly explain firm returns, the anomaly-based evidence against the CAPM is generally overstated. Relations between firm characteristics and conditional firm-level alphas are primarily focused among micro and small stocks and tend not to persist after the anomalies are first documented. Furthermore, few of the firm characteristics associated with alphas actually contain unique information.

The paper is organized as follows. Section 3.2 develops our econometric model for testing asset-pricing anomalies and discusses the advantages and disadvantages of the proposed approach. Section 3.3 describes the data. Section 3.4 presents the empirical results. Section 3.5 concludes.

3.2 Methodology

This section develops our firm-level approach for identifying anomalies relative to the CAPM. The Sharpe–Lintner version of the CAPM states that

$$E\left[r_{i,t}\right] = \beta_i E\left[r_{m,t}\right], \tag{3.1}$$

where $E\left[r_{i,t}\right]$ denotes the expected excess return on stock i at time t, $E\left[r_{m,t}\right]$ is the market risk premium, and $\beta_i = \frac{Cov(r_{i,t}, r_{m,t})}{Var(r_{m,t})}$ captures stock i's exposure to market risk. The Sharpe–Lintner CAPM relates the unconditional expectations of firm and market returns. In reality, as a firm grows and evolves, its exposure to market risk will change. Similarly, the market risk premium is likely to vary depending on the state of the economy and the risk tolerance of investors. In the presence of time-varying risk exposures and risk premiums, a conditional version of the CAPM,

$$E_{t-1}\left[r_{i,t}\right] = \beta_{i,t} E_{t-1}\left[r_{m,t}\right], \tag{3.2}$$

may hold even if the unconditional CAPM does not (Jagannathan and Wang (1996)).

The conditional CAPM implies that the expected conditional alpha, defined as

$$E_{t-1}[\alpha_{i,t}] = E_{t-1}\left[r_{i,t}\right] - \beta_{i,t} E_{t-1}\left[r_{m,t}\right], \tag{3.3}$$

should equal zero for all stocks. A common way of testing this prediction is to examine whether alphas can be forecasted by firm characteristics. Many existing tests in the literature rely on portfolio-based approaches. However, grouping firms into portfolios and aggregating returns has adverse effects. Specifically, valuable information is discarded while averaging across firms and cross-sectional patterns in firm returns may be distorted as a result of the portfolio formation procedure. An alternative approach involves testing the CAPM's prediction that alphas are not forecastable using the full cross section of firm returns by examining the cross-sectional relation,

$$\alpha_{i,t} = \delta_0 + \delta_x x_{i,t-1} + \epsilon_{i,t}, \tag{3.4}$$

where $x_{i,t-1}$ is a firm characteristic that is observable at time $t-1$. The conditional CAPM implies that $\delta_x = 0$ in a cross-sectional regression based on equation (3.4).

Analysis of this cross-sectional regression is complicated by the fact that the dependent variable, $\alpha_{i,t}$, is a latent variable. As such, a model for the latent alphas is necessary to examine the relation in equation (3.4). A test using such a model would ideally have two features. First, the specification should not introduce a spurious relation between $\alpha_{i,t}$ and $x_{i,t-1}$ through the model for the latent alphas. Second, given the structure of the problem, the posterior precision of δ_x should be maximized.

With these considerations in mind, we develop a firm-level test of the CAPM's implication that alphas are not predictable. Specifically, we estimate a system of simultaneous equations,

$$r_{i,t,y} = \alpha_{i,y} + \beta_{i,y} r_{m,t,y} + \epsilon_{i,t,y},\ \epsilon_{i,t,y} \sim N\left(0, \sigma^2_{i,y}\right), \tag{3.5}$$

$$\alpha_{i,y} = X_{i,y}\delta_y + \eta_{i,y},\ \eta_{i,y} \sim N\left(0, \sigma^2_{\alpha,y}\right), \tag{3.6}$$

$$\delta_y = \overline{\delta} + \nu_y,\ \nu_y \sim MVN\left(0, \mathbf{V}\right), \tag{3.7}$$

where $r_{i,t,y}$ denotes the excess return on stock i in subperiod t of time period y, $r_{m,t,y}$ is the excess market return, and $X_{i,y}$ is a matrix including a constant and firm characteristics observable at the beginning of period y. In the primary model specification, we use monthly subperiods (t) and annual periods (y). We therefore allow firm alphas and betas to change each year, utilizing the short-window regression approach of Lewellen and Nagel (2006) to test the conditional CAPM. In equation (3.6), δ_y measures the year-by-year relations between alphas and firm characteristics. In a given year, however, abnormal returns may be related to characteristics purely by chance. To examine whether there is a systematic relation between firm characteristics and alphas throughout the entire sample period, we assume that the parameter vectors, $\{\delta_y\}_{y=1}^{Y}$, in equation (3.6) are drawn from the multivariate normal distribution specified in equation (3.7). If an element of $\overline{\delta}$ is focused away from zero, there is evidence of an anomaly that persists through time. In our empirical analysis, we analyze $\overline{\delta}$ when assessing the importance of firm characteristics in forecasting alphas.

We estimate the system of equations (3.5) to (3.7) simultaneously as a hierarchical Bayes model.[4,5] The model structure and estimation technique provides important benefits when examining the relations between alphas and characteristics. In particular, we minimize the potential for specification issues while modeling the

[4]Several papers have used Bayesian techniques to examine asset-pricing models. McCulloch and Rossi (1991) and Geweke and Zhou (1996) develop Bayesian analyses of the Arbitrage Pricing Theory (APT), while Shanken (1987), Harvey and Zhou (1990), Kandel, McCulloch, and Stambaugh (1995), and Cremers (2006) propose Bayesian tests for the mean-variance efficiency of a given portfolio. Ang and Chen (2007) use Bayesian methods to examine whether the conditional CAPM can explain the value premium. Davies (2010) and Cederburg, Davies, and O'Doherty (2010) test the CAPM and the ICAPM, respectively, using Bayesian approaches.

[5]See (Rossi, Allenby, and McCulloch, 2005a, Ch. 5) for a discussion of hierarchical Bayes models.

latent alphas and maximize the precision of $\overline{\delta}$.

Relative to existing approaches, our methodology is unlikely to find spurious relations between alphas and characteristics. We make limited assumptions about the evolution of betas over time, only assuming that betas are relatively stable within each year. In contrast, the conditional CAPM is often tested by allowing betas to vary as a function of state variables. Avramov and Chordia (2006) take this approach and model firm betas as an exact linear function of size, book-to-market, and macroeconomic variables. However, such an approach requires the econometrician to know the "right" state variables (e.g., Harvey (1989), Shanken (1990), Jagannathan and Wang (1996), and Lettau and Ludvigson (2001b)). Further, misspecification of the process for betas may introduce spurious relations between measured alphas and firm characteristics. If betas are related to other firm characteristics, such as profitability or leverage, that are not included in the model, firm betas will be systematically mismeasured and a researcher may generate incorrect inferences about $\overline{\delta}_x$.

Rather than taking an approach which relies on conditioning information, we directly examine firm alphas and betas within each year, which Fama and French (2006) note is less vulnerable to specification issues. Fama and French (2008) also avoid complex dynamics for betas by regressing raw returns on firm characteristics to examine anomalies, implicitly assuming that all stocks have betas of one. However, even in the absence of a relation between alphas and firm characteristics, this approach will find $\overline{\delta}_x \neq 0$ if $Corr(\beta_{i,t}, x_{i,t-1}) \neq 0$. There is ample theoretical and empirical evidence that betas are related to firm characteristics (e.g., Karolyi (1992), Gomes, Kogan, and Zhang (2003), and Avramov and Chordia (2006)), so properly adjusting for market risk is important while testing whether alphas are forecastable.

Given that our model design is unlikely to produce spurious relations between alphas and characteristics, we turn to developing estimates of $\overline{\delta}$ which are as precise as possible conditional on the data. Our approach of simultaneously estimating equations (3.5) to (3.7) maximizes the precision of $\overline{\delta}$. In contrast, a common approach in the literature is to estimate the relations between alphas and characteristics in two steps (e.g., Brennan, Chordia, and Subrahmanyam (1998) and Avramov and Chordia (2006)). In the first step, the latent alphas are estimated for each firm. The second step involves a cross-sectional regression of estimated alphas on the firm characteristics,

$$\hat{\alpha}_{i,t} = \delta_0 + \delta_x x_{i,t-1} + \epsilon_{i,t}. \tag{3.8}$$

However, a two-step approach introduces a measurement error problem which leads to an understatement of the evidence against the conditional CAPM. Alphas are measured with error in the first step, and the variance of each firm's estimated alpha is greater than the posterior variance of the firm alpha. The variance of $\overline{\delta}$ is increasing in the variance of the alphas used as dependent variables, so the measurement error problem decreases the precision of $\overline{\delta}$. As a result, $\overline{\delta}_x$ may not be different from zero and alphas may appear to be unforecastable even when a significant relation exists in the data. Our methodology eliminates this measurement error problem by simultaneously estimating equations (3.5) to (3.7) to maximize the precision of $\overline{\delta}$. In particular, while a two-step approach uses only time-series information about the latent alphas, the simultaneous estimation methodology utilizes both time-series and cross-sectional information to make inferences about alphas.

An additional feature of the model specified in equations (3.5) to (3.7) is that cross-sectional heteroskedasticity and cross-correlations among firms are implicitly taken into account. Cross-sectional heteroskedasticity and cross-correlations will influence the precision of δ_y in each period. By allowing the relation between firm characteristics and alphas to vary over time, these features of the cross section of returns will be reflected in the posteriors of $\overline{\delta}$ and $\mathbf{V}$ (Shanken and Zhou (2007)). Thus, a large number of test assets can be considered without requiring the estimation of a variance-covariance matrix.

Estimating equations (3.5) to (3.7) simultaneously is, however, a challenging problem. The model involves a high-dimensional parameter space since firm-specific parameters must be estimated for thousands of firms in each year. Moreover estimation is further complicated by the fact that the latent variables $\alpha_{i,y}$ and δ_y appear in multiple equations within the system. Fortunately, the problem can be greatly simplified by recognizing the hierarchical structure of the model. Equation (3.7) is a hierarchical prior for δ_y in equation (3.6), while equation (3.6) is a hierarchical prior for $\alpha_{i,y}$ in equation (3.5). Thus, we adopt a hierarchical Bayes approach to estimate equations (3.5) to (3.7) simultaneously. In addition to greatly reducing the computational burden relative to using maximum likelihood estimation or the generalized method of moments, the Bayesian approach provides a complete accounting of parameter uncertainty and exact finite sample inference.

The Bayesian approach does require the researcher to specify explicit priors and hyperparameters for all model parameters. We specify the prior for the parameter vector of interest, $\overline{\delta}$, to be

$$\overline{\delta} \sim MVN\left(0, 100\mathbf{I}\right). \tag{3.9}$$

The prior mean of zero implies that firm-level alphas are not associated with firm characteristics, which is not consistent with the considerable empirical evidence to the contrary. However, the informativeness of the prior depends on the prior variance. We specify a large prior variance indicating that we have little prior information about $\overline{\delta}$, so our prior has little effect on the posterior distribution of $\overline{\delta}$. In unreported results, we considered non-zero prior means for each firm characteristic based on the evidence in the asset-pricing literature, but the impact on the posterior distributions was minimal due to the large prior variance.

We specify the prior for firm-level betas as

$$\beta_{i,y} \sim N\left(1, 10\right). \tag{3.10}$$

We use a prior mean equal to one because the average beta of firms in the market must equal one. We set the prior variance at 10, so the prior mean should have little impact on the posterior distribution of betas for most firms. For comparison, Vasicek (1973) recommends a prior variance of 0.25, which has a much stronger effect of shrinking firm betas toward one.[6]

It is also necessary to specify priors for $\left\{\sigma^2_{i,y}\right\}$, $\left\{\sigma^2_{\alpha,y}\right\}$, and $\mathbf{V}$. We model $\left\{\sigma^2_{i,y}\right\}$ and $\left\{\sigma^2_{\alpha,y}\right\}$ using the Inverse Gamma distribution and $\mathbf{V}$ with the Inverse

[6] We also considered a hierarchical model structure for firm betas, similar to the model specified in equations (3.6) to (3.7) for firm-level alphas. However, we found that the posterior distributions for the parameter vector of interest, $\overline{\delta}$, are almost identical using either the hierarchical prior or the prior specified above so we opt for the more parsimonious specification.

Wishart distribution. The hyperparameters for these distributions are chosen to ensure that they have minimal influence on the posterior distributions. Our results are not sensitive to either doubling or halving the hyperparameter values.

We estimate the model specified in equations (3.5) to (3.7) using standard Markov chain Monte Carlo (MCMC) techniques. We draw directly from the conditional posterior distributions for all model parameters using a Gibbs sampler. The algorithm converges quickly. For our empirical applications, we run the chain for 5,000 iterations and discard the first 2,500 as a burn-in period. To test whether the algorithm has converged, we initially ran the chain for 20,000 iterations and found that the posterior distributions characterized using iterations 2,500 to 5,000 were nearly identical to those based on iterations 17,500 to 20,000.

A detailed description of the estimation algorithm and the prior distributions and associated hyperparameters is provided in Appendix A.5.1. We also conduct a series of simulation experiments to demonstrate the validity of the estimation approach as well as the robustness of inferences to various features of the cross section of firm returns. A summary of these results is provided in Appendix A.5.2.

3.3 Data

This section outlines the sample construction and data requirements for estimating the model described in equations (3.5) to (3.7). We obtain accounting data from the Compustat Fundamentals Annual files and stock return data from CRSP. The sample includes all NYSE, Amex, and NASDAQ ordinary common stocks with the data required to compute at least one of the following firm characteristics: size

(M), book-to-market (BM), momentum (MOM), reversal (REV), profitability (ROA), asset growth (AG), net stock issues (NS), accruals (ACC), and financial distress (OS).

Following Fama and French (1992), year y runs from July of calendar year y through June of calendar year $y+1$. The characteristics are measured at the end of June in each calendar year y. The variables are matched to monthly returns from July of calendar year y to June of calendar year $y+1$. We exclude financial firms (SIC codes between 6000 and 6999) and firms with negative book equity. Based on Fama and French (2008), we classify firms into micro, small, and big categories using the 20th and 50th percentiles of market capitalization for NYSE stocks at the end of June of calendar year y.

The model described in Section 3.2 requires alphas and betas to be estimated for each firm-year observation. For a firm to be included in the estimation sample in a given year, we require 12 months of return data during that year. The final sample includes 163,603 firm-years of data from July 1963 to June 2008. We use the CRSP value-weighted stock market index as the proxy for the unobserved market portfolio. Monthly excess returns on the CRSP value-weighted stock market index, the risk-free rate, and size breakpoints are from Kenneth French's website.[7] See Appendix A.6 for a detailed description of variable definitions and data construction.

[7]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. We thank Kenneth French for making this data available.

3.4 Results

This section applies the methodology developed in Section 3.2 to explore cross-sectional anomalies. Section 3.4.1 presents the main firm-level results from the estimation of the model described in equations (3.5) to (3.7) and contrasts these results with those from traditional portfolio-level tests. Section 3.4.2 takes a more detailed look at CAPM anomalies at the firm level.

3.4.1 Firm-Level Tests

Panel A of Table 1.1 summarizes the posterior distribution of $\overline{\delta}$ in equation (3.7), which measures the systematic relation between alphas and firm characteristics over the entire sample period. Initially we examine each firm characteristic in isolation. Following Avramov and Chordia (2006) and Fama and French (2008) we assume a linear relation between conditional alphas and firm characteristics.

Panel A shows that seven of the nine firm characteristics are significantly associated with firm-level alphas. Alphas are positively associated with book-to-market, momentum, and profitability and negatively associated with asset growth, net stock issues, accruals, and financial distress. In terms of economic significance, a one-standard-deviation change in any of the seven variables is associated with a change in alpha ranging in magnitude from 16 basis points (bps) per month for momentum (0.51 x 0.32) to in excess of 20 bps per month for book-to-market, profitability, asset growth, and net stock issues.

For comparison, in Panel B of Table 1.1 we report results based on the traditional portfolio approach that is commonly used to identify anomalies. For each firm characteristic, we sort stocks into deciles each year at the end of June and then form hedge portfolios that are long the highest decile and short the lowest decile of stocks. The portfolios are equally weighted and rebalanced annually. Panel B presents the average conditional CAPM alphas. The conditional alphas are computed following the short-window regression methodology in Lewellen and Nagel (2006). Specifically, we estimate a separate CAPM regression each year using monthly data to obtain a time series of non-overlapping conditional portfolio alphas. The standard errors reported in Panel B are based on the time-series variability of the estimated conditional alphas.

The hedge portfolios formed from sorts on size, book-to-market, momentum, reversal, asset growth, net stock issues, and accruals have CAPM alphas that are significantly different from zero at the 1% level. We find no evidence of significant abnormal returns for the hedge portfolios formed on profitability and financial distress.[8] The results in Table 1.1 provide evidence that underlying firm-level associations can be obscured at the portfolio level. Comparing the firm-level results in Panel A to the portfolio-based tests in Panel B, we find that inferences differ for four of the nine firm characteristics: size, reversal, profitability, and financial distress.

One clear difference between the firm-level and portfolio-level tests is that the portfolio approach only considers firms in deciles one and ten, ignoring information contained in the remaining 80% of stocks. The firm-level approach, on the other

[8]The results for profitability are consistent with those of Fama and French (2008) for portfolio sorts. Chava and Purnanandam (2010) document that the financial distress anomaly is specific to the post-1980 period used by Dichev (1998) and Campbell, Hilscher, and Szilagyi (2008).

hand, utilizes information from the entire cross section. Thus, the portfolio-level analysis could be unduly influenced by a small number of outlier observations in the extreme deciles. To investigate this possibility, in Panel A of Table 1.2 we specify alphas as a function of a constant, the firm characteristic, and two dummy variables identifying whether a particular firm lies in the top or bottom decile for that characteristic. The results suggest that the portfolio-level associations for size and reversal are driven primarily by the extremes. For example, there is no linear relation between alphas and size, but firms in the smallest decile earn alphas that are nearly 0.4% per month higher than firms in the largest decile. For all other firm characteristics, inferences are not substantially altered by the introduction of dummy variables.[9]

When conducting firm-level tests of the CAPM it is also important to consider the potential impact of non-synchronous returns. Our initial model specification uses monthly returns and assumes that all stocks are traded frequently. If trading is infrequent, betas measured by relating firm returns to contemporaneous market returns will tend to understate exposure to market risk. This issue is particularly relevant for our analysis if the extent to which a firm has non-synchronous returns is associated with a given firm characteristic. To control for non-synchronicities we follow Dimson (1979) and include the lagged excess market return as an additional factor in equation (3.5) to correct for any downward bias in measured betas. Panel B of Table 1.2 shows that allowing for non-synchronicities has little impact on the

[9]In results not reported, we considered other non-linear specifications, including the addition of squared and cubed terms for each characteristic, but our inferences were unchanged.

relations between alphas and firm characteristics.

Although there is evidence in Tables 1.1 and 1.2 that firm-level associations between alphas and firm characteristics are distorted at the portfolio level for four out of nine characteristics, the firm-level analysis nonetheless finds substantial evidence against the conditional CAPM. Seven of the nine firm characteristics are significantly associated with conditional CAPM alphas even after allowing for the possibility of non-linearities and non-synchronous returns. In the next section we take a more detailed look at the empirical shortcomings of the conditional CAPM.

3.4.2 A Closer Look at CAPM Anomalies

Given our main results in Panel A of Table 1.1, it is tempting to conclude that the CAPM provides a poor characterization of stock returns. However, in order to properly evaluate the performance of the CAPM we must consider the performance of the model across three dimensions. First, from an economic perspective, it is important to know whether anomalous patterns in returns are market-wide or limited to illiquid stocks that represent a small portion of the total market capitalization. Second, anomalies could arise due to temporary mis-pricing or data snooping by researchers and, as such, would be unlikely to persist over time. Third, it is important to examine to what extent firm characteristics identified as anomalies contain unique information about abnormal returns. If multiple firm characteristics contain the same information then tests that consider each firm characteristic in isolation are likely to suffer from an omitted variable bias that will result in the importance of an anomaly being overstated.

To examine whether anomalies are pervasive across size groups, we repeat the firm-level analysis from Table 1.1, but allow $\overline{\delta}$ to vary across micro, small, and big stocks.[10] The posterior distributions are presented in Figure 1 for each firm characteristic. Of the nine characteristics considered, seven are significantly related to the conditional CAPM alphas of micro stocks based on 95% credible intervals. In contrast, only three anomaly variables – asset growth, net stock issues, and accruals – are significantly associated with the abnormal returns of big stocks. Moreover, the economic magnitude of the relations is greatly reduced among big stocks relative to micro stocks. For example, a one-standard-deviation shock in asset growth has a 32 bps per month impact on micro stocks compared to just 12 bps for big stocks. The results in Figure 1 suggest that the CAPM provides a much more effective characterization of the returns of big stocks, which constitute over 90% of the total market capitalization.

In Table 1.3 we examine the extent to which firm-level relations between firm characteristics and alphas persist after each firm characteristic is first documented as an anomaly.[11] We re-estimate the model in equations (3.5) to (3.7), but unlike

---

[10] The robustness of anomalies across size subgroups is an active area of interest. For example, Loughran (1997) argues that the value effect is restricted to small stocks, while Fama and French (2006) show Loughran's (1997) results are specific to the value/growth indicator, the sample period, and US stocks. Several other papers documenting individual anomalies conduct double sorts on size and a particular anomaly variable, with mixed results. Fama and French (2008) take a more comprehensive approach by analyzing the relations between returns and several firm characteristics within size subgroups.

[11] In prior research regarding the persistence of anomalies, Schwert (2003) finds that the size and book-to-market effects appear to have attenuated after the anomalies were documented, while the momentum anomaly has persisted. Jegadeesh and Titman (2001) also find that the momentum anomaly appears to have persisted throughout the 1990s.

Panel A of Table 1.1, in which $\overline{\delta}$ is constant across the whole sample period, we allow $\overline{\delta}$ to vary across the pre- and post-publication periods.[12] In pre-publication periods the results in Table 1.3 show that alphas are positively related to book-to-market, momentum, and profitability, and negatively related to accruals and financial distress. In the post-publication periods, only book-to-market and accruals remain significantly associated with firm alphas. Moreover, the results for book-to-market and accruals are driven by micro stocks. Among big stocks, there is no evidence of any robust relations between firm characteristics and conditional alphas post publication.

In Figure 2 we highlight the relation between accruals and firm alphas before and after the initial publication by Sloan in 1996. Pre-publication there is a robust relation between accruals and alphas across stocks of all sizes. Post-publication, the negative relation persists among micro stocks, diminishes for small stocks, and disappears in big stocks. This pattern is consistent with market participants attempting to exploit the anomaly to earn abnormal returns. Among big stocks, where transaction costs are lowest and there are few, if any, short selling constraints, deviations from CAPM pricing are quickly eliminated. In contrast, investors appear to be unable to trade away the anomaly among micro stocks, where transaction costs are high, liquidity is low, and short selling is often difficult to implement (e.g., Jensen

[12]We use publication dates based on the following papers: Banz (1981) (size), Rosenberg, Reid, and Lanstein (1985) (book-to-market), Jegadeesh (1990) (momentum), DeBondt and Thaler (1985) (reversal), Haugen and Baker (1996) (profitability), Sloan (1996) (accruals), and Dichev (1998) (financial distress). The asset growth (Cooper, Gulen, and Schill (2008)) and net stock issues (Daniel and Titman (2006a)) anomalies were only recently uncovered so we do not include these characteristics in our analysis.

(1978)).

The pre-post analysis in Table 1.3 and Figure 2 provides little evidence that anomalies persist after they are first documented, especially among big firms. As such, our evidence is more consistent with the hypothesis that anomalies arise in the data either due to market participants making a mistake which they later correct or due to data snooping by researchers.

Thus far our analysis has focused on the relation between conditional alphas and each firm characteristic in isolation. If firm characteristics are correlated with each other and offer little unique information about alphas then studying each characteristic in isolation will overstate the failings of the conditional CAPM. The traditional portfolio approach is unable to adequately address this omitted variable problem. Researchers typically rely on two- or possibly three-dimensional sorts to isolate the effects of a particular characteristic. Controlling for more than one or two characteristics simultaneously, however, is infeasible and inferences are sensitive to both the sorting technique and the sorting sequence (e.g., Conrad, Cooper, and Kaul (2003)). In contrast, our approach is particularly well suited to assess which anomalies contain unique information; we simply specify firm-year alphas in equation (3.6) as a function of all nine firm characteristics in Table 1.1.

In Figure 3 we compare the posterior distributions from two analyses – one in which each firm characteristic is considered in isolation and one in which all characteristics are considered simultaneously.[13] Momentum, asset growth, and fi-

[13]In results not reported we also considered a model specification in which conditional alphas were modeled as a function of multiple firm characteristics and the relations were allowed to vary across micro, small, and big stocks. As in Figure 1 the relations between

nancial distress are significantly associated with CAPM alphas when considered in isolation, but as Figure 3 highlights, none of these characteristics contain significant incremental information when all characteristics are considered simultaneously. The only firm characteristics that are significantly related to firm-level alphas when multiple characteristics are considered simultaneously are book-to-market, profitability, net stock issues, and accruals. Our analysis therefore suggests that univariate tests provide a low hurdle for firm characteristics to be classified as anomalies.

3.5 Conclusion

In this paper, we use a hierarchical Bayes framework to examine asset-pricing anomalies, modeling firm-year alphas as a function of one or more firm characteristics. We investigate nine anomalies – size, book-to-market, momentum, reversal, profitability, asset growth, net stock issues, accruals, and financial distress – over the period 1963 to 2008. Studying each anomaly separately we find robust evidence that CAPM alphas are positively associated with book-to-market, momentum, and profitability. Alphas are negatively associated with asset growth, net stock issues, accruals, and financial distress.

These initial results imply the failings of the CAPM are widespread. A deeper investigation of anomalies, however, suggests that while the CAPM may not perfectly explain firm returns, the anomaly-based evidence against the CAPM is greatly overstated. Relationships between firm characteristics and conditional firm-level alphas are primarily focused among micro and small stocks and tend not

---

characteristics and alphas are generally driven by micro and small stocks.

to persist after the anomaly is first documented. Among large firms there is no evidence of any persistent anomalies. Furthermore, few of the firm characteristics associated with alphas actually contain unique information.

Table 3.1: Firm Characteristics and CAPM Alphas, 1963-2008

M	BM	MOM	REV	ROA	AG	NS	ACC	OS
Panel A: Base Specification								
Posterior Mean for the Aggregate-Level Parameters, $\overline{\delta}$								
0.05	0.24**	0.51**	-0.06	1.79**	-0.45**	-1.36**	-1.27**	-1.20*
(0.06)	(0.08)	(0.19)	(0.08)	(0.51)	(0.09)	(0.21)	(0.22)	(0.46)
Average Cross-Sectional Standard Deviation of Firm Characteristics								
1.89	0.86	0.32	0.78	0.15	0.58	0.16	0.13	0.14
Panel B: Performance of Hedge Portfolios								
Average Conditional CAPM Alpha, $\hat{\alpha}^{CAPM}$								
-1.08**	1.37**	0.61**	-0.78**	0.17	-1.23**	-1.14**	-0.54**	-0.18
(0.33)	(0.20)	(0.20)	(0.23)	(0.29)	(0.17)	(0.16)	(0.12)	(0.27)

Note: Panel A presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate-level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across the entire sample period. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero. Panel B reports average conditional alphas for hedge portfolios that are long the highest decile of stocks and short the lowest decile for each variable. Following Lewellen and Nagel (2006), the conditional CAPM alphas are estimated annually using monthly data. Standard errors are in parentheses. An * (**) indicates significance at the 5% (1%) level using a two-tailed test. The firm characteristics are described in Appendix A.6.

Table 3.2: Alternative Model Specifications, 1963-2008

	M	BM	MOM	REV	ROA	AG	NS	ACC	OS
Panel A: Nonlinear Specification									
Posterior Mean for the Aggregate-Level Parameters, $\overline{\delta}$									
$\overline{\delta}$, Linear	0.08	0.23**	0.50*	-0.01	1.74**	-0.42**	-0.91**	-1.44**	-0.79*
	(0.07)	(0.09)	(0.24)	(0.09)	(0.51)	(0.10)	(0.23)	(0.28)	(0.33)
$\overline{\delta}$, Decile 1	0.19	-0.07	-0.42**	-0.15	-0.15	-0.20	0.17*	-0.46**	-0.01
	(0.12)	(0.11)	(0.14)	(0.15)	(0.17)	(0.14)	(0.08)	(0.13)	(0.09)
$\overline{\delta}$, Decile 10	-0.18*	-0.02	-0.31*	-0.27*	-0.15	-0.08	-0.16	-0.21*	-0.33*
	(0.08)	(0.11)	(0.13)	(0.11)	(0.10)	(0.12)	(0.11)	(0.10)	(0.16)
Panel B: Sum Betas									
Posterior Mean for the Aggregate-Level Parameters, $\overline{\delta}$									
$\overline{\delta}$	0.09	0.23**	0.48*	-0.07	2.03**	-0.47**	-1.49**	-1.42**	-1.48**
	(0.06)	(0.08)	(0.22)	(0.09)	(0.54)	(0.09)	(0.24)	(0.23)	(0.47)

Note: The table presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate-level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across the entire sample period. Panel A shows estimates from a nonlinear specification including a linear component and dummy variables for firms with characteristic values in the top or bottom deciles. Panel B shows estimates using sum betas. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero.

Table 3.3: Firm Characteristics and CAPM Alphas Pre- and Post-Publication, 1963-2008

	M	BM	MOM	REV	ROA	ACC	OS
Publication Date	1981	1985	1990	1985	1996	1996	1998
Pre-Publication – Posterior Means for the Aggregate-Level Parameters, $\overline{\delta}$							
All	-0.00	0.24*	0.64**	-0.19	1.91**	-1.32**	-1.18*
	(0.09)	(0.11)	(0.24)	(0.11)	(0.58)	(0.25)	(0.52)
Micro	-0.03	0.31*	0.52*	-0.17	1.48**	-1.31**	-1.18**
	(0.12)	(0.13)	(0.20)	(0.13)	(0.46)	(0.22)	(0.40)
Small	0.12	0.24	0.53	-0.19	2.53**	-1.35**	-2.25**
	(0.13)	(0.15)	(0.29)	(0.14)	(0.60)	(0.38)	(0.60)
Big	-0.08	0.09	0.57	-0.12	1.16	-1.31**	0.39
	(0.11)	(0.14)	(0.37)	(0.15)	(0.93)	(0.39)	(0.84)
Post-Publication – Posterior Means for the Aggregate-Level Parameters, $\overline{\delta}$							
All	0.08	0.23*	0.33	0.07	1.29	-0.97*	-1.19
	(0.08)	(0.11)	(0.33)	(0.11)	(0.98)	(0.43)	(0.94)
Micro	0.01	0.41**	0.22	-0.02	0.99	-1.35**	-0.92
	(0.10)	(0.13)	(0.26)	(0.13)	(0.75)	(0.40)	(0.72)
Small	0.11	0.24	0.39	0.04	1.67	-0.50	-1.91**
	(0.11)	(0.15)	(0.37)	(0.14)	(0.95)	(0.64)	(0.74)
Big	-0.01	0.12	0.36	-0.01	0.49	0.11	-1.80
	(0.09)	(0.14)	(0.47)	(0.16)	(1.58)	(0.70)	(1.05)

Note: The table presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior mean and standard deviation for the aggregate level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across time. We allow for different aggregate-level parameters in the pre- and post-publication periods. We estimate two models for each anomaly, one in which $\overline{\delta}$ is restricted to be the same across all firms (All) and one in which $\overline{\delta}$ varies across micro, small, and big stocks. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include zero.

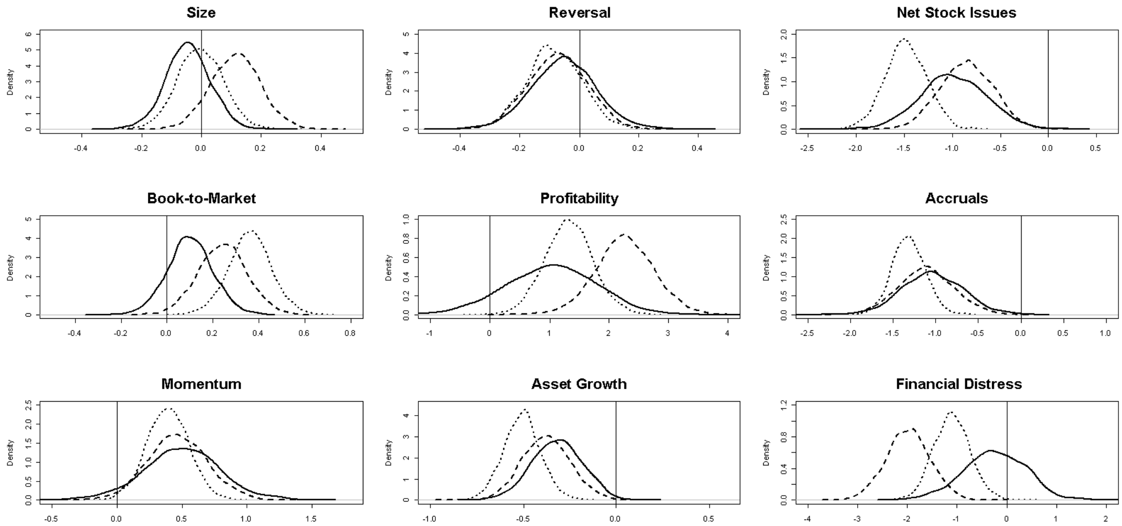

Figure 3.1: Firm Characteristics and CAPM Alphas by Size Group

Note: The figure presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and each firm characteristic separately. We report the posterior distributions for the aggregate-level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across the entire sample period. We estimate a model for each anomaly in which the aggregate-level parameters ($\overline{\delta}$) vary across micro (dotted), small (dashed), and big (line) stocks.

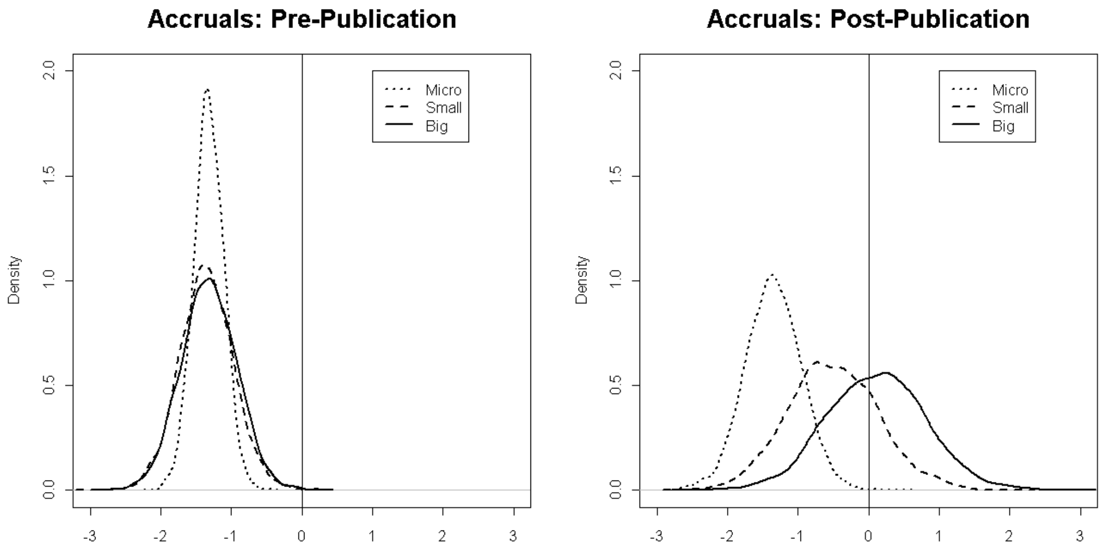

Figure 3.2: The Accruals Anomaly Pre- and Post-Publication

Note: The figure presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and accruals. We report the posterior distributions for the aggregate-level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across time. We allow for different aggregate-level parameters in the pre- and post-publication periods and also allow the aggregate-level parameters to vary across micro (dotted), small (dashed), and big (line) stocks.

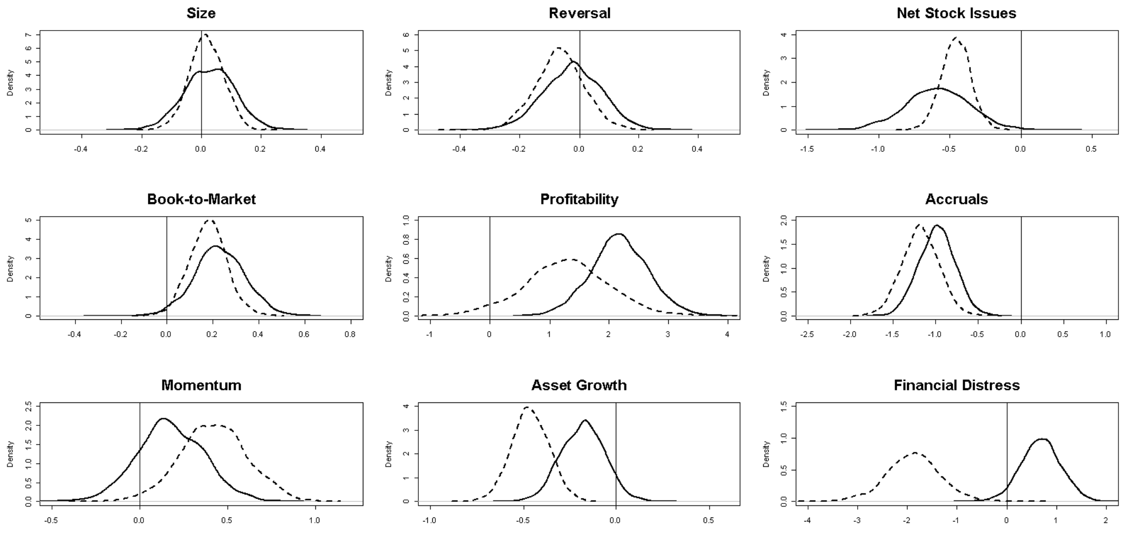

Figure 3.3: Individual and Multiple Anomaly Variables

Note: The figure presents the results from the estimation of the model described in equations (3.5) to (3.7) examining the cross-sectional relation between firm alphas and multiple firm characteristics simultaneously. We report the posterior distributions (line) for the aggregate-level parameters, $\overline{\delta}$, which provide information about the relation between alphas and firm characteristics across the entire sample period. For comparison, for each anomaly variable we also present the posterior distribution (dashed) of $\overline{\delta}$ from estimation of the model described in equations (3.5) to (3.7) for each characteristic in isolation using the same data sample.

APPENDIX
MODEL DEVELOPMENT AND METHODOLOGY

A.1 ICAPM Model Development

A.1.1 Equations (1.6) to (1.8)

Equations (1.6) to (1.8) are derived as follows,

$$V_{ih} = Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{m,t+1+j}\right) \tag{A.1}$$

$$\approx Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j (r_{f,t+j} + r^e_{m,t+1+j})\right) \tag{A.2}$$

$$\equiv V_{i\overline{m}} + V_{ir}. \tag{A.3}$$

Simulations show that the approximation from line (A.1) to line (A.2) is close under reasonable parameter values for generating returns.

A.1.2 Equations (1.14) and (1.16)

Equation (1.14) is derived as

$$V_{i\overline{m}} \equiv Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r^e_{m,t+1+j}\right) \tag{A.4}$$

$$= Cov_t\left(r^e_{i,t+1}, \sum_{j=1}^{\infty}\rho^j (E_{t+1} - E_t)\overline{r}^e_{m,t+j}\right) \tag{A.5}$$

$$= Cov_t\left(r^e_{i,t+1}, \sum_{j=1}^{\infty}\rho^j \phi_m^{j-1}(\overline{r}^e_{m,t+1} - (1-\phi_m)E_m - \phi_m \overline{r}^e_{m,t})\right) \tag{A.6}$$

$$= \frac{1}{\phi_m(1-\phi_m\rho)} Cov_t\left(r^e_{i,t+1}, \eta_{\overline{m},t+1}\right), \tag{A.7}$$

where the first equality reflects the expectation of the market return in equation (1.13.1), the second equality comes from the AR(1) structure of the market risk premium in equation (1.13.2), and the last equality is from the definition of $\eta_{\overline{m},t+1}$ in equation (1.13.2). Similarly, equation (1.16) can be derived using equation (1.15.3),

$$V_{ir} \equiv Cov_t\left(r^e_{i,t+1}, (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{f,t+j}\right) \tag{A.8}$$

$$= Cov_t\left(r^e_{i,t+1}, \sum_{j=1}^{\infty}\rho^j \phi_r^{j-1}(r_{f,t+1} - \phi_r r_{f,t})\right) \tag{A.9}$$

$$= \frac{1}{\phi_r(1-\phi_r\rho)} Cov_t\left(r^e_{i,t+1}, \eta_{r,t+1}\right). \tag{A.10}$$

A.1.3 Relating Equations (1.9) and (1.12)

We can relate the $\overline{\lambda}$ coefficients from equation (1.12) to the coefficients on V_{im}, $V_{i\overline{m}}$, and V_{ir} from equation (1.9). Rewriting equation (1.9) to allow these coefficients to be free parameters and allow for a non-zero intercept, we have equation (1.17) which is rewritten here,

$$E_t r^e_{i,t+1} + \frac{V_{ii}}{2} = b_0 + b_m V_{im} + b_{\overline{m}} V_{i\overline{m}} + b_r V_{ir}.$$

We can transform the covariances into betas and adjust for scaling from equations (1.14) and (1.16). Defining Σ as the covariance matrix of the market and intertemporal risk factors and a vector Z as in equation (1.20),

$$Z = \begin{bmatrix} 1 \\ \sqrt{\phi_m(1-\phi_m\rho)} \\ \sqrt{\phi_r(1-\phi_r\rho)} \end{bmatrix},$$

we can rewrite the right-hand side of equation (A.1.3),

$$b_0 + b_m V_{im} + b_{\overline{m}} V_{i\overline{m}} + b_r V_{ir} = b_0 + \begin{bmatrix} b_m \\ b_{\overline{m}} \\ b_r \end{bmatrix}' \Sigma Z Z' \beta_i, \tag{A.11}$$

where β_i is the vector of coefficients from a regression of excess stock returns of firm i on a constant and the risk factors. Equations (1.11) and (1.12) imply

$$E_t r^e_{i,t+1} + \frac{V_{ii}}{2} = \overline{\lambda}_0 + \overline{\lambda}_m \beta_i^m + \overline{\lambda}_{\overline{m}} \beta_i^{\overline{m}} + \overline{\lambda}_r \beta_i^r, \tag{A.12}$$

and we can rewrite the right-hand side of equation (A.12),

$$\overline{\lambda}_0 + \overline{\lambda}_m \beta_i^m + \overline{\lambda}_{\overline{m}} \beta_i^m + \overline{\lambda}_r \beta_i^r = \overline{\lambda}_0 + \begin{bmatrix} \overline{\lambda}_m \\ \overline{\lambda}_{\overline{m}} \\ \overline{\lambda}_r \end{bmatrix}' \beta_i. \tag{A.13}$$

Equating the right-hand sides of equations (A.1.3) and (A.12), we have

$$b_0 + \begin{bmatrix} b_m \\ b_{\overline{m}} \\ b_r \end{bmatrix}' \Sigma Z Z' \beta_i = \overline{\lambda}_0 + \begin{bmatrix} \overline{\lambda}_m \\ \overline{\lambda}_{\overline{m}} \\ \overline{\lambda}_r \end{bmatrix}' \beta_i \tag{A.14}$$

for all β_i. Therefore, we arrive at equation (1.18),

$$b_0 = \overline{\lambda}_0,$$

and equation (1.19),

$$\begin{bmatrix} b_m \\ b_{\overline{m}} \\ b_r \end{bmatrix} = (\Sigma Z Z')^{-1} \begin{bmatrix} \overline{\lambda}_m \\ \overline{\lambda}_{\overline{m}} \\ \overline{\lambda}_r \end{bmatrix}.$$

A.1.4 SMB and HML State Variables

I form state variables based on SMB and HML to use in a predictive system. The state variables are ratios of the book-to-markets of component portfolios, similar to Brennan, Wang, and Xia (2001). The state variable for SMB is

$$z_{SMB,t} = log\left(\frac{B_{S,t}/M_{S,t}}{B_{B,t}/M_{B,t}}\right). \tag{A.15}$$

In equation (A.15), $B_{S,t}$ and $M_{S,t}$ are the book and market values for the small firm portfolio. When constructing the SMB factor portfolio, the small-growth, small-neutral, and small-value portfolios are equally weighted. Therefore, I calculate $B_{S,t}/M_{S,t}$ as the book-to-market of the equally weighted portfolio of these three component portfolios by adjusting the book and market values of each portfolio prior to summing values across portfolios.[14] Book-to-market for the big firm portfolio is constructed in the same way. The state variable for HML is

$$z_{HML,t} = log\left(\frac{B_{H,t}/M_{H,t}}{B_{L,t}/M_{L,t}}\right), \tag{A.16}$$

which is calculated with a portfolio that equally weights the small-value and big-value portfolios and a portfolio that holds the small-growth and big-growth portfolios.

Notice that the state variables in equations (A.15) and (A.16) are logs of the ratio of book-to-markets. The state variables can be rewritten as

$$z_{SMB,t} = (m_{B,t} - b_{B,t}) - (m_{S,t} - b_{S,t}) \tag{A.17}$$

and

$$z_{HML,t} = (m_{L,t} - b_{L,t}) - (m_{H,t} - b_{H,t}), \tag{A.18}$$

where m and b are log market and book values. The state variable for SMB (HML)

[14]Data on the book and market values of each portfolio is from Kenneth French's website.

is therefore high when big (growth) stocks are expensive relative to small (value) stocks.

The component portfolios of SMB and HML are rebalanced annually, which may lead to jumps in the state variables that do not reflect economically important information. For example, in June of year y, the portfolios are based on the weights from the end of June of year $y-1$ determined from book values from the end of December of year $y-2$ and market values from the end of June of year $y-1$. In July of year y, however, the portfolios are rebalanced based on book values from December of year $y-1$ and market values from June of year y. Therefore, the state variable has potential for rebalancing errors when transitioning from June to July. I check several alternative specifications to ensure that portfolio rebalancing does not affect inferences. For the primary specification, I form June book-to-market for the portfolios after they have been rebalanced at the end of June. June book-to-market is based on the December book value and the June market value found by discounting the July market value by the realized return from July. I then include an additional state variable defined in June as the difference between the book-to-markets of the original portfolio and the rebalanced portfolio, with the state variable taking a value of zero in all other months. The model can therefore take information from both the original and rebalanced portfolios when predicting July returns. Inferences are robust to using the state variables in equations (A.15) and (A.16) with no adjustment or using modified versions of these state variables that replace the June book-to-market of the original portfolio with the book-to-market of the rebalanced portfolio.

A.2 ICAPM Model Estimation

Draws from the posterior distribution of the parameters can be obtained using Markov chain Monte Carlo (MCMC) techniques. Steps 1 to 6 are designed to draw from the posterior distribution of the price of risk, and are a Gibbs sampler equivalent to the Metropolis–Hastings approach of Davies (2010). Steps 7 to 10 draw from the posterior distribution of the market risk premium and are closely related to the approach of Pástor and Stambaugh (2009). Steps 11 to 14 draw from the posterior distribution of the real interest rate using Forward Filtering, Backward Sampling (FFBS).

1. Draw $\beta_{i,y}|\sigma^2_{i,y}, \{\eta_{\overline{m},t,y}\}_{t=1}^T, \{\eta_{r,t,y}\}_{t=1}^T, \lambda_y, \sigma^2_y, \underline{\beta}, \underline{\mathbf{V}}_\beta \sim N(\overline{\beta}_{i,y}, \overline{\mathbf{V}}_{\beta,i,y})$ for $i = 1, ..., N_y$ and $y = 1, ..., Y$, where N_y is the number of assets in period y, $\beta_{i,y} = [\alpha_{i,y} \quad \beta^m_{i,y} \quad \beta^{\overline{m}}_{i,y} \quad \beta^r_{i,y}]'$,

$$\overline{\beta}_{i,y} = \overline{\mathbf{V}}_{\beta,i,y}\left(\underline{\mathbf{V}}_\beta^{-1}\underline{\beta} + \sigma^{-2}_{i,y}X'_{1,y}R^e_{i,y} + \sigma_y^{-2}X'_{2,y}(\overline{R}^e_{i,y} - \lambda_{0,y})\right), \tag{A.19}$$

$$\overline{\mathbf{V}}_{\beta,i,y} = \left(\underline{\mathbf{V}}_\beta^{-1} + \sigma^{-2}_{i,y}X'_{1,y}X_{1,y} + \sigma_y^{-2}X'_{2,y}X_{2,y}\right)^{-1}, \tag{A.20}$$

$X_{1,y} = \left[\iota_T \quad \{r^e_{m,t,y}\}_{t=1}^T \quad \{\eta_{\overline{m},t,y}\}_{t=1}^T \quad \{\eta_{r,t,y}\}_{t=1}^T\right]$, $R^e_{i,y} = \{r^e_{i,t,y}\}_{t=1}^T$, and $X_{2,y} = [0 \quad \lambda_{m,y} \quad \lambda_{\overline{m},y} \quad \lambda_{r,y}]$. I set the prior parameters of $\beta_{i,y}$ to $\underline{\beta} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and

$$\underline{\mathbf{V}}_{\beta} = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix}$$ to specify proper but diffuse priors.

2. Draw $\sigma^2_{i,y}|\beta_{i,y}, \{\eta_{\overline{m},t,y}\}_{t=1}^{T}, \{\eta_{r,t,y}\}_{t=1}^{T}, \underline{\nu}, \underline{s}^2 \sim$ Inverse Gamma$(\frac{\overline{s}^2}{2}, \frac{\overline{\nu}}{2})$ for $i = 1, ..., N_y$ and $y = 1, ..., Y$, where $\overline{\nu} = \underline{\nu} + T$, $\overline{s}^2 = \underline{\nu}\underline{s}^2 + s^2$, and

$$s^2 = \sum_{t=1}^{T} \left(r^e_{i,t,y} - \alpha_{i,y} - \beta^m_{i,y} r^e_{m,t,y} - \beta^{\overline{m}}_{i,y}\eta_{\overline{m},t,y} - \beta^r_{i,y}\eta_{r,t,y}\right)^2. \tag{A.21}$$

The prior parameters are set to $\underline{\nu} = 1$ and $\underline{s}^2 = 100$.

3. Draw $\lambda_y|\{\beta_{i,y}\}_{i=1}^{N_y}, \sigma^2_y, \overline{\lambda}, \mathbf{V}_\lambda \sim N(\hat{\lambda}_y, \hat{\mathbf{V}}_{\lambda y})$ for $y = 1, ..., Y$, where

$$\hat{\lambda}_y = \hat{\mathbf{V}}_{\lambda y}(\mathbf{V}_\lambda^{-1}\overline{\lambda} + \sigma_y^{-2} X'_{1,y}\overline{r}^e_y), \tag{A.22}$$

$$\hat{\mathbf{V}}_{\lambda y} = (\mathbf{V}_\lambda^{-1} + \sigma_y^{-2} X'_{1,y} X_{1,y})^{-1}, \tag{A.23}$$

$\lambda_y = [\lambda_{y,0} \quad \lambda_{y,m} \quad \lambda_{y,\overline{m}} \quad \lambda_{y,r}]'$, $X_{1,y} = \left[\iota_{N_y} \quad \{\beta^m_{i,y}\}_{i=1}^{N_y} \quad \{\beta^{\overline{m}}_{i,y}\}_{i=1}^{N_y} \quad \{\beta^r_{i,y}\}_{i=1}^{N_y}\right]$, $\overline{r}^e_y = \{\overline{r}^e_{i,y} + \frac{s^2_{i,y}}{2}\}_{i=1}^{N_y}$, and $s^2_{i,y}$ is the sample return variance for firm i in period y.

4. Draw $\sigma^2_{\lambda y}|\lambda_y, \{\beta_{i,y}\}_{i=1}^{N_y}, \underline{\nu}_\lambda, \underline{s}^2_\lambda \sim$ Inverse Gamma$(\frac{\overline{s}^2_\lambda}{2}, \frac{\overline{\nu}_\lambda}{2})$ for $y = 1, ..., Y$, where $\overline{\nu}_\lambda = \underline{\nu}_\lambda + N_y$, $\overline{s}^2_\lambda = \underline{\nu}_\lambda \underline{s}^2_\lambda + s^2_\lambda$, and $s^2_\lambda = \sum_{i=1}^{N_y}\left(\overline{r}^e_{i,y} + \frac{s^2_{i,y}}{2} - \lambda_{0,y} - \lambda_{m,y}\beta^m_{i,y} - \lambda_{\overline{m},y}\beta^{\overline{m}}_{i,y} - \lambda_{r,y}\beta^r_{i,y}\right)^2$. The prior parameters are set to $\underline{\nu}_\lambda = 1$ and $\underline{s}^2_\lambda = 1$.

5. Draw $\mathbf{V}_\lambda|\overline{\lambda}, \{\lambda_y\}_{y=1}^{Y}, g, G \sim$ Inverse Wishart$(Y+g, \sum_{y=1}^{Y}(\lambda_y - \overline{\lambda})(\lambda_y - \overline{\lambda})' + G)$, where $g = 3$ and $G = g\mathbf{I}$.

6. Draw $\overline{\lambda}|\{\lambda_y\}_{y=1}^{Y}, \mathbf{V}_\lambda, \underline{V}, \underline{\lambda} \sim N(\tilde{\lambda}, \tilde{\mathbf{V}}_\lambda)$, where

$$\tilde{\lambda} = \tilde{\mathbf{V}}_\lambda \left((\mathbf{V}_\lambda/Y)^{-1} \sum_{y=1}^{Y} \frac{\lambda_y}{Y} + \underline{V}^{-1}\underline{\lambda} \right), \tag{A.24}$$

$$\tilde{\mathbf{V}}_\lambda = \left((\mathbf{V}_\lambda/Y)^{-1} + \underline{V}^{-1} \right)^{-1}, \tag{A.25}$$

$\underline{V} = 0.25\mathbf{I}$, and $\underline{\lambda} = [0 \quad 0 \quad 0 \quad 0]'$.

The following four steps draw a sequence of the market risk risk premium and related parameters from the system of equations (1.13). These steps are closely related to those of Pástor and Stambaugh (2009) with an extra lag of the market risk premium built into the system. The following steps are generalized to the case where information from the cross section of returns is incorporated into the estimates of the market risk premium. For the base case in the current version of the paper, this source of information is omitted while estimating the market risk premium. Adjustments for the base case are noted below. For notational convenience, define $t^* \equiv 12(y-1) + t$ and $T^* = TY$. For simplicity, I also refer to parameters that are constant throughout each period y (e.g., $\beta_{i,y}$ and $\sigma^2_{i,y}$) using the t^* notation.

7. Draw $\{\bar{r}^e_{m,t^*}\}_{t^*=1}^{T^*}$, $\{\eta_{\overline{m},t^*}\}_{t^*=1}^{T^*}$ $|\{\{\beta_{i,t^*}\}_i^{N_{t^*}}\}_{t^*=1}^{T^*}$, $\{\{\sigma^2_{i,t^*}\}_{i=1}^{N_{t^*}}\}_{t^*=1}^{T^*}$, E_m, E_x, ϕ_m, ϕ_x, Σ using a FFBS step. All parameter distributions in this step are conditioned on the previous draws of $\{\{\beta_{i,t^*}\}_i^{N_{t^*}}\}_{t^*=1}^{T^*}$, $\{\{\sigma^2_{i,t^*}\}_{i=1}^{N_{t^*}}\}_{t^*=1}^{T^*}$, E_m, E_x, ϕ_m, ϕ_x, and Σ, but the conditioning is suppressed in the notation for simplicity. The set of historical and current market returns, state variables, and firm returns observable at time t^* is denoted by D_{t^*}. Equation (1.10) in period t^* contains the market risk premium from times t^*-1 and t^*. In order to

incorporate the information from this equation into draws of the time series of $\bar{r}^e_m$, I include the lagged value of $\bar{r}^e_m$ in the vector autoregression (VAR):

$$\begin{bmatrix} r_{m,t^*} - E_m \\ x_{t^*} - E_x \\ \bar{r}^e_{m,t^*} - E_m \\ \bar{r}^e_{m,t^*-1} - E_m \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & \phi_x & 0 & 0 \\ 0 & 0 & \phi_m & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r^e_{m,t^*-1} - E_m \\ x_{t^*-1} - E_x \\ \bar{r}^e_{m,t^*-1} - E_m \\ \bar{r}^e_{m,t^*-2} - E_m \end{bmatrix} + \begin{bmatrix} u_{t^*} \\ v_{t^*} \\ w_{t^*} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} u_{t^*} \\ v_{t^*} \\ w_{t^*} \end{bmatrix} \sim N(\mathbf{0}, \Sigma). \tag{A.26}$$

The structure of the evolution matrix ensures that the current period market risk premium becomes next period's lagged premium in this VAR. Define $\tilde{r}^e_{m,t^*} = [\bar{r}^e_{m,t^*} \quad \bar{r}^e_{m,t^*-1}]'$, $\tilde{\phi}_m = \begin{bmatrix} \phi_m & 0 \\ 1 & 0 \end{bmatrix}$, and $\tilde{\phi}_x = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & \phi_x & 0 & 0 \\ 0 & 0 & \phi_m & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$. Also define

$$\tilde{\Sigma} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \equiv \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} & \Sigma_{uw} & 0 \\ \Sigma_{vu} & \Sigma_{vv} & \Sigma_{vw} & 0 \\ \Sigma_{wu} & \Sigma_{wv} & \Sigma_{ww} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.27}$$

so $\tilde{\Sigma}_{ww} = \begin{bmatrix} \Sigma_{ww} & 0 \\ 0 & 0 \end{bmatrix}$, $\tilde{\Sigma}_{uw} = \begin{bmatrix} \Sigma_{uw} & 0 \end{bmatrix}$, and $\tilde{\Sigma}_{vw} = \begin{bmatrix} \Sigma_{vw} & 0 \end{bmatrix}$.

(a) *Forward Filtering*: The filtering step is used to find sequences for time $t^* = 1, ..., T^*$ of the parameters of the distributions

$$\tilde{r}^e_{m,t^*}|D_{t^*-1} \sim N(a_{t^*}, P_{t^*}) \tag{A.28}$$

and

$$\tilde{r}^e_{m,t^*}|D_{t^*} \sim N(b_{t^*}, Q_{t^*}). \tag{A.29}$$

The prior distribution of $\tilde{r}^e_{m,t^*}$ before observing data from time t^* is given by equation (A.28). There are two additional sources of information about $\tilde{r}^e_{m,t^*}$ available at time t^*. First, realizations of the market return and state variables give information about the market risk premium. Second, there is information about $\tilde{r}^e_{m,t^*}$ from the pricing errors in equation (1.10). For example, if firms with high sensitivity to changes in the market risk premium (large $\beta^{\overline{m}}_{i,t^*}$) perform abnormally well in time t^*, there is evidence of a positive shock to $\bar{r}^e_{m,t^*}$, leading to an upward revision of beliefs about $\bar{r}^e_{m,t^*}$ and a downward revision of beliefs about $\bar{r}^e_{m,t^*-1}$. After observing these two sources of information, the distribution of $\bar{r}^e_{m,t^*}$ is updated to equation (A.29). Define $z_{t^*} = [r^e_{m,t^*} \quad x'_{t^*}]'$, $E_z = [E_m \quad E'_x]'$, and

$$\begin{aligned}
a_{t^*} &= E(\tilde{r}^e_{m,t^*}|D_{t^*-1}), & b_{t^*} &= E(\tilde{r}^e_{m,t^*}|D_{t^*}), \\
P_{t^*} &= Var(\tilde{r}^e_{m,t^*}|D_{t^*-1}), & Q_{t^*} &= Var(\tilde{r}^e_{m,t^*}|D_{t^*}), \\
f_{t^*} &= E(z_{t^*}|D_{t^*-1}), & R_{t^*} &= Var(z_{t^*}|\tilde{r}^e_{m,t^*}, D_{t^*-1}) \\
S_{t^*} &= Var(z_{t^*}|D_{t^*-1}), & G_{t^*} &= Cov(z_{t^*}, \tilde{r}^e_{m,t^*}|D_{t^*-1}).
\end{aligned} \tag{A.30}$$

Also define V as the unconditional variance of $[z \quad \tilde{r}^e_m]'$. Then

$$V = \begin{bmatrix} V_{rr} & V_{rx} & V_{rm} \\ V_{xr} & V_{xx} & V_{xm} \\ V_{mr} & V_{mx} & V_{mm} \end{bmatrix} \tag{A.31}$$

can be calculated as $\mathrm{vec}(V) = [I - (\tilde{\phi}_x \otimes \tilde{\phi}_x)]^{-1}\mathrm{vec}(\tilde{\Sigma})$. Let $V_{zz} = \begin{bmatrix} V_{rr} & V_{rx} \\ V_{xr} & V_{xx} \end{bmatrix}$.

To begin drawing the time series of r^e_m, first note that $b_0 = [E_m \quad E_m]'$ and $Q_0 = V_{mm}$. Also,

$$\tilde{r}^e_{m,1}|D_0 \sim N(a_1, P_1), \tag{A.32}$$

where $a_1 = [E_m \quad E_m]'$ and $P_1 = V_{mm}$. Then note that

$$z_1|D_0 \sim N(f_1, S_1), \tag{A.33}$$

where $f_1 = E_z$ and $S_1 = V_{zz}$. Further, $G_1 = V_{zm}$. Then information about $\tilde{r}^e_{m,1}$ is derived from the observation of z_1 since

$$z_1|\tilde{r}^e_{m,1}, D_0 \sim N(e_1, R_1), \tag{A.34}$$

where $e_1 = f_1 + G_1 P_1^{-1}(\tilde{r}^e_{m,1} - a_1)$ and $R_1 = S_1 - G_1 P_1^{-1} G_1'$. Additional information about $\tilde{R}^e_{m,1}$ comes from the cross section of stock returns through equation (1.10). Combining this information using Bayes' rule,

$$\tilde{r}^e_{m,1}|D_1 \sim N(b_1, Q_1), \tag{A.35}$$

where

$$\begin{aligned}
b_1 =& Q_1 \Bigg(P_1^{-1} a_1 + P_1^{-1} G_1' R_1^{-1} (G_1 P_1^{-1} a_1 + (z_1 - f_1)) \\
&+ \sum_{i=1}^{N_1} \begin{bmatrix} 1 \\ -\phi_m \end{bmatrix} \frac{(\beta_{i,1}^{\overline{m}})^2}{\sigma_{i,1}^2} \Bigg((1-\phi_m) E_m \\
&+ \frac{r_{i,1}^e - \alpha_{i,1} - \beta_{i,1}^m r_{m,1}^e - \beta_{i,1}^r \eta_{r,1}}{\beta_{i,1}^{\overline{m}}} \Bigg) \Bigg),
\end{aligned} \tag{A.36}$$

and

$$Q_1 = \left(P_1^{-1} + P_1^{-1} G_1' R_1^{-1} G_1 P_1^{-1} + \sum_{i=1}^{N_1} \begin{bmatrix} 1 & -\phi_m \\ -\phi_m & \phi_m^2 \end{bmatrix} \frac{(\beta_{i,1}^{\overline{m}})^2}{\sigma_{i,1}^2} \right)^{-1}. \tag{A.37}$$

Continuing for $t^* = 2, ..., T^*$, we have

$$a_{t^*} = (I - \tilde{\phi}_m) E_m \iota + \tilde{\phi}_m b_{t^*-1,1}, \tag{A.38}$$

$$P_{t^*} = \tilde{\phi}_m Q_{t^*-1} \tilde{\phi}_m' + \tilde{\Sigma}_{ww}, \tag{A.39}$$

$$f_{t^*} = \begin{bmatrix} b_{t^*-1,1} \\ (I - \phi_x) E_x + \phi_x x_{t^*-1} \end{bmatrix}, \tag{A.40}$$

$$S_{t^*} = \begin{bmatrix} Q_{t^*-1,1,1} + \tilde{\Sigma}_{uu} & \tilde{\Sigma}_{uv} \\ \tilde{\Sigma}_{vu} & \tilde{\Sigma}_{vv} \end{bmatrix}, \tag{A.41}$$

$$G_{t^*} = \begin{bmatrix} \phi_m Q_{t^*-1,1,1} & Q_{t^*-1,1,1} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \tilde{\Sigma}_{uw} \\ \tilde{\Sigma}_{vw} \end{bmatrix}, \tag{A.42}$$

$$R_{t^*} = S_{t^*} - G_{t^*} P_{t^*}^{-1} G_{t^*}', \tag{A.43}$$

$$
\begin{aligned}
b_{t^*} =& Q_{t^*}\Bigg(P_{t^*}^{-1}a_{t^*} + P_{t^*}^{-1}G'_{t^*}R_{t^*}^{-1}(G_{t^*}P_{t^*}^{-1}a_{t^*} + (z_{t^*} - f_{t^*})) \\
&+ \sum_{i=1}^{N_{t^*}} \begin{bmatrix} 1 \\ -\phi_m \end{bmatrix} \frac{(\beta_{i,t^*}^{\overline{m}})^2}{\sigma_{i,t^*}^2} \Bigg((1-\phi_m)E_m \\
&+ \frac{r_{i,t^*}^e - \alpha_{i,t^*} - \beta_{i,t^*}^m r_{m,t^*}^e - \beta_{i,t^*}^r \eta_{r,t^*}}{\beta_{i,t^*}^{\overline{m}}}\Bigg)\Bigg),
\end{aligned} \tag{A.44}
$$

$$
Q_{t^*} = \left(P_{t^*}^{-1} + P_{t^*}^{-1}G'_{t^*}R_{t^*}^{-1}G_{t^*}P_{t^*}^{-1} + \sum_{i=1}^{N_{t^*}} \begin{bmatrix} 1 & -\phi_m \\ -\phi_m & \phi_m^2 \end{bmatrix} \frac{(\beta_{i,t^*}^{\overline{m}})^2}{\sigma_{i,t^*}^2}\right)^{-1}. \tag{A.45}
$$

If cross-sectional information is ignored while estimating the market risk premium, the summations in N are omitted from the specifications of b_t and Q_t for $t = 1, ..., T$. The sequences of a_{t^*}, b_{t^*}, f_{t^*}, G_{t^*}, P_{t^*}, Q_{t^*}, and S_{t^*} are retained for the backward sampling step.

(b) *Backward Sampling*: Draw $\bar{r}_{m,t^*}^e, \eta_{m,t^*} | a_{t^*}, b_{t^*}, f_{t^*}, G_{t^*}, Q_{t^*}, P_{t^*}, S_{t^*}$ for $t^* = 0, ..., T^*$. First, draw

$$
\tilde{r}_{m,T^*}^e \sim N(b_{T^*}, Q_{T^*}). \tag{A.46}
$$

The draw of $\tilde{r}_{m,T^*}^e$ defines $\bar{r}_{m,T^*}^e$ and $\bar{r}_{m,T^*-1}^e$. Then draw $\tilde{r}_{m,t^*}^e$ for $t^* = T^* - 1, ..., 0$, where $\tilde{r}_{m,t^*}^e$ is the last two elements of the vector

$$
\zeta_{t^*} | \zeta_{t^*+1}, D_{t^*} \sim N(h_{t^*}, H_{t^*}), \tag{A.47}
$$

where

$$h_{t^*} = \begin{bmatrix} r^e_{m,t^*} \\ x_{t^*} \\ b_{t^*} \end{bmatrix} + \tilde{\phi}_x \begin{bmatrix} S_{t^*+1} & G_{t^*+1} \\ G'_{t^*+1} & P_{t^*+1} \end{bmatrix}^{-1} \begin{bmatrix} z_{t^*+1} - f_{t^*+1} \\ \bar{r}^e_{m,t^*+1} - a_{t^*+1} \end{bmatrix}, \quad \text{(A.48)}$$

$$H_{t^*} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Q_{t^*} \end{bmatrix} - \tilde{\phi}_x \begin{bmatrix} S_{t^*+1} & G_{t^*+1} \\ G'_{t^*+1} & P_{t^*+1} \end{bmatrix}^{-1} \tilde{\phi}'_x, \quad \text{(A.49)}$$

$$\quad \text{(A.50)}$$

where

$$\overline{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ Q_{t^*,1,1} & 0 & 0 & \phi_m Q_{t^*,1,1} & Q_{t^*,1,1} \\ Q_{t^*,2,1} & 0 & 0 & \phi_m Q_{t^*,2,1} & Q_{t^*,2,1} \end{bmatrix}, \quad \text{(A.51)}$$

and $Q_{t^*,i,j}$ denotes the (i,j) element of Q_{t^*}. The draw of $\bar{r}^e_{m,t^*-1}$ is the $k+3$ element of ζ_{t^*}. Finally, denote the unanticipated change in $\bar{r}^e_{m,t^*}$ by $\eta_{\overline{m},t^*} = \bar{r}^e_{m,t^*} - ((1-\phi_m)E_m + \phi_m \bar{r}^e_{m,t^*-1})$.

8. Draw $[E'_x \quad E_m]' | \{\bar{r}^e_{m,t^*}\}_{t^*=1}^{T^*}, \phi_m, \phi_x, \Sigma \sim N(\overline{E}_{xm}, \overline{V}_{xm})$. Let $E_{xm} = [E'_x \quad E_m]'$. The prior for E_{xm} is

$$E_{xm} \sim N(E_{xm_0}, V_{xm_0}), \quad \text{(A.52)}$$

where

$$E_{xm_0} = \begin{bmatrix} 0 \\ \bar{r}^e_m \end{bmatrix}, \tag{A.53}$$

$$V_{xm_0} = \begin{bmatrix} \sigma^2_{E_x} I_K & 0 \\ 0 & \sigma^2_{E_m} \end{bmatrix}. \tag{A.54}$$

Information about E_{xm} is obtained from the dynamics of the time series of $[x'_{t^*} \quad \bar{r}^e_{m,t^*}]'$ and the cross section of stock returns through the pricing errors in equation (1.10). Defining

$$L_1 = \begin{bmatrix} \phi_x & 0 \\ 0 & \phi_m \end{bmatrix} \tag{A.55}$$

and $L_2 = I - L_1$, the posterior distribution of E_{xm} is (conditioning is suppressed for ease of notation),

$$E_{xm}|\cdot \sim N(\tilde{E}_{xm}, \tilde{V}_{xm}), \tag{A.56}$$

where

$$\begin{aligned}\tilde{E}_{xm} =& \tilde{V}_{xm} \left(V^{-1}_{xm_0} E_{xm_0} + L'_2 \Sigma^{-1}_{(vw)} \sum_{t^*=1}^{T^*} \left(\begin{bmatrix} x_{t^*} \\ \bar{r}^e_{m,t^*} \end{bmatrix} - L_1 \begin{bmatrix} x_{t^*-1} \\ \bar{r}^e_{m,t^*-1} \end{bmatrix} \right) \right. \\ & + \sum_{t^*=1}^{T^*} \sum_{i=1}^{N_{t^*}} \\ & \left. \begin{bmatrix} 0 \\ \frac{(\beta^{\overline{m}}_{i,t^*}(1-\phi_m))^2}{\sigma^2_{i,t^*}} \frac{\beta^{\overline{m}}_{i,t^*}(\bar{r}^e_{m,t^*} - \phi_m \bar{r}^e_{m,t^*-1}) - (r^e_{i,t^*} - \alpha_{i,t^*} - \beta^m_{i,t^*} r^e_{m,t^*} - \beta^r_{i,t^*} \eta_{r,t^*})}{(1-\phi_m)\beta^{\overline{m}}_{i,t^*}} \end{bmatrix} \right). \end{aligned} \tag{A.57}$$

and

$$\tilde{V}_{xm} = \left(V_{xm_0}^{-1} + T^* L_2' \Sigma_{(vw)}^{-1} L_2 + \begin{bmatrix} 0 & 0 \\ 0 & \sum_{t^*=1}^{T^*} \sum_{i=1}^{N_{t^*}} \frac{(\beta_{i,t^*}^{\overline{m}}(1-\phi_m))^2}{\sigma_{i,t^*}^2} \end{bmatrix} \right)^{-1}. \tag{A.58}$$

If cross-sectional information is ignored while estimating the market risk premium, the summations in N are omitted from the specifications of $\tilde{E}_{xm}$ and $\tilde{V}_{xm}$.

9. Draw $\phi_m, \phi_x | \{\bar{r}_{m,t^*}^e\}_{t^*=1}^{T^*}, E_x, E_m, \Sigma$. Let $x^k \equiv (x_2^k, ..., x_{T^*}^k)'$ be the $(T^* - 1) \times 1$ vector of predictor k in periods $t^* = 2, ..., T^*$. Let $x_{(l)}$ be the $(T^*-1) \times K$ matrix of the K vectors of realizations of the predictors in periods $t^* = 1, ..., T^* - 1$. Also let $\bar{r}_m^e \equiv (\bar{r}_{m,2}^e, ..., \bar{r}_{m,T^*}^e)'$, $\bar{r}_{m,(l)}^e \equiv (\bar{r}_{m,1}^e, ..., \bar{r}_{m,T^*-1}^e)'$, and E_{x^k} be the k-th element of E_x. Define

$$z = \begin{bmatrix} x^1 - \iota_{T^*-1} E_{x^1} \\ \vdots \\ x^K - \iota_{T^*-1} E_{x^K} \\ \bar{r}_m^e - \iota_{T^*-1} E_m \end{bmatrix} \tag{A.59}$$

and

$$Z = \begin{bmatrix} x_{(l)} - \iota_{T^*-1} E_x' & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & x_{(l)} - \iota_{T^*-1} E_x' & 0 \\ 0 & 0 & 0 & \bar{r}_{m,(l)}^e - \iota_{T^*-1} E_m \end{bmatrix}. \tag{A.60}$$

Then z is a $[(T^*-1)(K+1)] \times 1]$ vector and Z is a $[(T^*-1)(K+1)] \times (K^2+1)$

matrix. Further define $b \equiv (\mathrm{vec}(\phi_x')' \quad \phi_m)'$. The prior distribution of b is

$$b \sim N(b_0, V_{b_0}) \times 1_{b \in S}, \tag{A.61}$$

where $1_{b \in S}$ is equal to one if b is in the space S and S is the space such that the eigenvalues of A lie inside the unit circle and $B \in (-1, 1)$. When b is in S, x_{t^*} and $\bar{r}^e_{m,t^*}$ are stationary. Information about b arises from the dynamics of the state variables and market risk premium as well as the cross section of stock returns through the pricing errors in equation (1.10). The posterior distribution of b is (conditioning is suppressed for notational convenience),

$$b|\cdot \sim N(\tilde{b}, \tilde{V}_b), \tag{A.62}$$

where

$$\begin{aligned}\tilde{b} =& \tilde{V}_b \Bigg(V_{b_0}^{-1} b_0 + Z'(\Sigma_{(vw)}^{-1} \otimes I_{T^*-1}) z \\ &+ \sum_{t^*=1}^{T^*} \sum_{i=1}^{N_{t^*}} \\ &\left[\begin{array}{c} 0 \\ \frac{(\beta^{\overline{m}}_{i,t^*}(\bar{r}^e_{m,t^*-1} - E_m))^2}{\sigma^2_{i,t^*}} \frac{\beta^{\overline{m}}_{i,t^*}(\bar{r}^e_{m,t^*} - E_m) - (r^e_{i,t^*} - \alpha_{i,t^*} - \beta^m_{i,t^*} r^e_{m,t^*} - \beta^r_{i,t^*} \eta_{r,t^*})}{\beta^{\overline{m}}_{i,t^*}(\bar{r}^e_{m,t^*-1} - E_m)} \end{array} \right] \Bigg),\end{aligned} \tag{A.63}$$

and

$$\tilde{V}_b = \left(V_{b_0}^{-1} + Z'(\Sigma_{(vw)}^{-1} \otimes I_{T^*-1}) Z + \left[\begin{array}{cc} 0 & 0 \\ 0 & \sum_{t^*=1}^{T^*} \sum_{i=1}^{N_{t^*}} \frac{(\beta^{\overline{m}}_{i,t^*}(\bar{r}^e_{m,t^*-1} - E_m))^2}{\sigma^2_{i,t^*}} \end{array} \right] \right)^{-1}. \tag{A.64}$$

If cross-sectional information is ignored while estimating the market risk premium, the summations in N are omitted from the specifications of $\tilde{b}$ and $\tilde{V}_b$.

Draws of b can be obtained by drawing from $N(\tilde{b}, \tilde{V}_b)$ and accepting the draw if $b \in S$. The parameters ϕ_m and ϕ_x can be found using the definition of b, $b = (\mathrm{vec}(\phi_x')' \quad \phi_m)'$.

10. Draw $\Sigma|\{\bar{r}^e_{m,t^*}\}_{t^*=1}^{T^*}, E_x, E_m, \phi_x, \phi_m, \Sigma^{(p)}, M_{11}, M_{22}$, where $\Sigma^{(p)}$ denotes the previous draw of Σ. Given

$$\Sigma \equiv \begin{bmatrix} \sigma_u^2 & \Sigma_{uv} & \sigma_{uw} \\ \Sigma_{vu} & \Sigma_{vv} & \Sigma_{vw} \\ \sigma_{wu} & \Sigma_{wv} & \sigma_w^2 \end{bmatrix}, \tag{A.65}$$

define

$$\Sigma_{11} \equiv \begin{bmatrix} \sigma_u^2 & \sigma_{uw} \\ \sigma_{wu} & \sigma_w^2 \end{bmatrix}, \tag{A.66}$$

and

$$\hat{\Sigma}_{11,0} \equiv \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix}, \tag{A.67}$$

where M_{11}, M_{12}, and M_{22} are prior parameters and M_{12} is a hyperparameter for Σ. A posterior draw of Σ can be obtained using the following two-step process.

(a) Draw $M_{12}|\Sigma^{(p)}, M_{11}, M_{12} \sim p(M_{12}|\Sigma_{11})$, where

$$p(M_{12}|\Sigma_{11}) = |\hat{\Sigma}_{11,0}|^{\frac{T_0-K}{2}} \exp\left(-\frac{T_0}{2}\mathrm{tr}(\Sigma_{11}^{-1}\hat{\Sigma}_{11,0})\right), \tag{A.68}$$
$$M_{12} \in (\underline{c}\sqrt{M_{11}M_{22}}, \bar{c}\sqrt{M_{11}M_{12}}),$$

and the prior parameter values are $T_0 = T^*/5$, $\underline{c} = -0.90$, and $\bar{c} = -0.35$. See the Technical Appendix of Pástor and Stambaugh (2009) for details on drawing M_{12} from this distribution. The draw of M_{12} defines a new draw of $\hat{\Sigma}_{11,0}$.

(b) Draw $\Sigma|\{\bar{r}^e_{m,t^*}\}_{t^*=1}^{T^*}, E_x, E_m, \phi_x, \phi_m, \hat{\Sigma}_{11,0}$. First compute the time series of residuals $(u_{t^*}, v_{t^*}, w_{t^*})$ for $t^* = 1, ..., T^*$ by

$$\begin{bmatrix} u_{t^*} \\ v_{t^*} \\ w_{t^*} \end{bmatrix} = \begin{bmatrix} r^e_{m,t^*} \\ x_{t^*} \\ \bar{r}^e_{m,t^*} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_x & 0 \\ 0 & 0 & \phi_m \end{bmatrix} \begin{bmatrix} r^e_{m,t^*-1} \\ x_{t^*-1} \\ \bar{r}^e_{m,t^*-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_K - \phi_x & 0 \\ 0 & 0 & 1 - \phi_m \end{bmatrix} \begin{bmatrix} 0 \\ E_x \\ E_m \end{bmatrix}. \tag{A.69}$$

Let X denote the $T^* \times 2$ matrix of $[u_{t^*} \quad w_{t^*}]$ and Y_{2,T^*} be the $T^* \times K$ matrix of v_{t^*}. The posterior distribution of Σ_{11} is inverse Wishart,

$$\Sigma_{11}|\cdot \sim \text{Inverse Wishart}(T_0\hat{\Sigma}_{11,0} + T^*\hat{\Sigma}_{11}, T^* + T_0 - K), \tag{A.70}$$

where $\hat{\Sigma}_{11} = X'X/T^*$. Further, let $\hat{C} = (X'X)^{-1}X'Y_{2,T^*}$, $\hat{\Omega} = (Y_{2,T^*} - X\hat{C})'(Y_{2,T^*} - X\hat{C})/T^*$, $V_C = (X_0'X_0 + X'X)^{-1}$, $\tilde{C} = V_C\left[(X_0'X_0)\hat{C}_0 + (X'X)\hat{C}\right]$, and $D = \hat{C}_0'X_0'X_0\hat{C}_0 + \hat{C}'X'X\hat{C} - \tilde{C}'V_C^{-1}\tilde{C}$. Then the posterior distribution of Ω is

$$\Omega|\cdot \sim \text{Inverse Wishart}(S_0\hat{\Omega}_0 + T^*\hat{\Omega} + D, T^* + S_0). \tag{A.71}$$

Then letting $\tilde{c} = \text{vec}(\tilde{C})$, the posterior distribution of $c = \text{vec}(C)$ is

$$c|\Omega, \cdot \sim N(\tilde{c}, \Omega \otimes V_C). \tag{A.72}$$

Finally, given (Σ_{11}, C, Ω), construct a posterior draw of Σ using $\Sigma_{11} \equiv \begin{bmatrix} \sigma_u^2 & \sigma_{uw} \\ \sigma_{wu} & \sigma_w^2 \end{bmatrix}$, $[\Sigma_{vu} \quad \Sigma_{vw}] = C\Sigma_{11}$, and $\Sigma_{vv} = \Omega + C\Sigma_{11}C'$.

Steps 11 to 14 draw the real interest rate and expected inflation along with related parameters from the system of equations (1.15) in a dynamic linear model framework using FFBS.

11. Draw $V|\{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}$. Let

$$y = [\{r_{n,t^*}\}_{t^*=1}^{T^*} \quad \{\pi_{t^*+1}\}_{t^*=1}^{T^*}]' \tag{A.73}$$

and

$$X = [\{r_{f,t^*}\}_{t^*=1}^{T^*} \quad \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}]'. \tag{A.74}$$

Then let

$$S = \left(y - \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X\right)\left(y - \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X\right)'. \tag{A.75}$$

Then V is drawn from an inverse Wishart distribution,

$$V|\cdot \sim \text{Inverse Wishart}(S + \nu_V S_V, \nu_V + T^*), \tag{A.76}$$

where ν_V and S_V are prior parameters.

12. Draw $W|\{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}, \phi_r, \phi_\pi$. Let $X = [\{r_{f,t^*}\}_{t^*=2}^{T^*} \quad \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=2}^{T^*}]'$ and $X_{-1} = [\{r_{f,t^*}\}_{t^*=1}^{T^*-1} \quad \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*-1}]'$. Then let

$$S = \left(X - \begin{bmatrix} \phi_r & 0 \\ 0 & \phi_\pi \end{bmatrix} X_{-1}\right)\left(X - \begin{bmatrix} \phi_r & 0 \\ 0 & \phi_\pi \end{bmatrix} X_{-1}\right)'. \tag{A.77}$$

Then W is drawn from an inverse Wishart distribution,

$$W|\cdot \sim \text{Inverse Wishart}(S + \nu_W S_W, \nu_W + T^* - 1), \tag{A.78}$$

where ν_W and S_W are prior parameters.

13. Draw $\phi_r, \phi_\pi | \{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}, W$ using a Metropolis–Hastings algorithm. Candidate draws are drawn from a normal distribution with standard deviation of 0.01 centered at the values of $\phi = [\phi_r \quad \phi_\pi]$ from the previous iteration. The new draw ϕ^* is accepted with probability

$$\min\left(1, \frac{p(\{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}; W, \phi^*)}{p(\{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}; W, \phi)}\right), \tag{A.79}$$

where the density is from equations (1.15.3) to (1.15.4). Additionally, if cross-sectional information is being used to provide information about the processes for the market risk premium and real interest rate, the density in equation (A.79) also includes the product of densities from equation (1.10) for firms $i = 1, ..., N_{t^*}$.

14. Draw $\{r_{f,t^*}\}_{t^*=1}^{T^*}, \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*} | \phi_r, \phi_\pi, V, W$ using FFBS.

    (a) *Forward Filtering*: Let G be a diagonal matrix with elements ϕ_r and ϕ_π, $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, and $y = [\{r_{n,t^*}\}_{t^*=1}^{T^*} \quad \{\pi_{t^*+1}\}_{t^*=1}^{T^*}]'$. For $t^* = 1$, we start with prior parameters a_1 and R_1 for a and R rather than the equations below. The following equations are stepped through sequentially for $t^* =$

$1, ..., T^*$.

$$a_{t^*} = Gm_{t^*-1} \tag{A.80}$$

$$R_{t^*} = GC_{t^*-1}G' + W \tag{A.81}$$

$$Q_{t^*} = F'R_{t^*}F + V \tag{A.82}$$

$$A_{t^*} = R_{t^*}FQ_{t^*}^{-1} \tag{A.83}$$

$$f_{t^*} = F'a_{t^*} \tag{A.84}$$

$$e_{t^*} = y_{t^*} - f_{t^*} \tag{A.85}$$

$$m_{t^*} = a_{t^*} + A_{t^*}e_{t^*} \tag{A.86}$$

$$C_{t^*} = R_{t^*} - A_{t^*}Q_{t^*}A'_{t^*} \tag{A.87}$$

The sequences of m and C are kept for the backward sampling step.

(b) *Backward Sampling*: For time T^*, the last element of $\theta = [\{r_{f,t^*}\}_{t^*=1}^{T^*} \quad \{E_{t^*}[\pi_{t^*+1}]\}_{t^*=1}^{T^*}]'$ is drawn from

$$\theta_{T^*} \sim N(m_{T^*}, C_{T^*}). \tag{A.88}$$

Then θ is drawn sequentially for $t^* = T^* - 1, ..., 1$. Let G be a diagonal matrix with elements ϕ_r and ϕ_π. At each time t^*,

$$\theta_{t^*} \sim N(m_{t^*}^*, C_{t^*}^*), \tag{A.89}$$

where

$$C_{t^*}^* = \left(C_{t^*}^{-1} + G'W^{-1}G\right)^{-1} \tag{A.90}$$

and

$$m_{t^*}^* = C_{t^*}^* \left(G'W^{-1}\theta_{t^*+1} + C_{t^*}^{-1}m_{t^*}\right). \tag{A.91}$$

The first element of θ_{t^*} is r_{f,t^*}, so the η_{r,t^*} factor in equation (1.10) is calculated as $\eta_{r,t^*} = r_{f,t^*} - \phi_r r_{f,t^*-1}$.

A.3 Long-Run Risk Model

A.3.1 Aggregate Asset Pricing

Expanding on Proposition 1, we have the following aggregate asset pricing results.

- The value function is given by

$$J(C_t, X_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \exp\left(\frac{\lambda(1-\gamma)}{\kappa+\beta} X_t + \frac{1-\gamma}{\beta} \left[\mu_C - \frac{1}{2}\gamma\sigma_C^2 + \frac{\lambda^2(1-\gamma)\sigma_x^2}{2(\kappa+\beta)^2} \right] \right). \tag{A.92}$$

- The pricing kernel is given by

$$\frac{d\Lambda}{\Lambda} = -r_t^f dt - \gamma\sigma_C dW_C - \frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_X dW_X, \tag{A.93}$$

$$r_t^f = \beta + \mu_C + \lambda X_t - \gamma\sigma_C^2, \tag{A.94}$$

where r_t^f is the risk-free rate.

- The price-dividend ratio is

$$\frac{P_t}{D_t} = G(X_t) = \int_t^\infty S(X_t, \tau) ds \qquad [\tau = s - t], \tag{A.95}$$

where

$$S(X_t, \tau) = \exp(P_1(\tau) X_t + P_2(\tau)), \tag{A.96}$$

with P_1 and P_2 given by

$$P_1(\tau) = \frac{1-\lambda}{\kappa}\left(1 - e^{-\kappa\tau}\right) \tag{A.97}$$

$$P_2(\tau) = a\tau + b(e^{-\kappa\tau} - 1) + c(1 - e^{-2\kappa\tau}) \tag{A.98}$$

$$a = \left(\mu_D - \mu_C + \gamma\sigma_C^2 - \beta + \frac{1-\lambda}{\kappa}\sigma_X^2\left(\frac{1-\lambda}{2\kappa} - \frac{(\gamma-1)\lambda}{\kappa+\beta}\right)\right) \tag{A.99}$$

$$b = \frac{(1-\lambda)\sigma_X^2}{\kappa^2}\left(\frac{1-\lambda}{\kappa} - \frac{(\gamma-1)\lambda}{\kappa+\beta}\right) \tag{A.100}$$

$$c = \frac{(1-\lambda)^2\sigma_X^2}{4\kappa^3}. \tag{A.101}$$

- The instantaneous risk premium and variance of return are given by

$$\mu_t = \frac{\lambda(\gamma-1)\sigma_X^2}{\kappa+\beta}\frac{G_X}{G}, \tag{A.102}$$

$$\sigma_t^2 = \sigma_D^2 + \left(\frac{G_X}{G}\sigma_X\right)^2, \tag{A.103}$$

with μ_t and σ_t^2 increasing in X_t.

Proof of Theorem 1:

Proof. Consider an asset that pays a single cash flow equal to the aggregate dividend at time T, D_T. The value of this asset at time t is given by

$$\begin{aligned} p_t(\tau, D_t, X_t) &= E_t\left[\frac{\Lambda_T}{\Lambda_t}D_T\right] \\ &= \frac{1}{\Lambda_t}E_t\left[\Lambda_T D_T\right] \\ &= D_t S(X_t,\tau) \qquad [\tau = T-t], \end{aligned} \tag{A.104}$$

where the final step includes the equivalence $E_t[\Lambda_T D_T] = \Lambda_t D_t S(X_t,\tau)$ where $S(X_t,\tau)$ is defined above. Then the risk premium on this asset is

$$\begin{aligned} -Cov\left(\frac{d\Lambda_t}{\Lambda_t}, \frac{dp_t}{p_t}\right) &= \frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_X^2\frac{S_X(X_t,\tau)}{S(X_t,\tau)} \\ &= \frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_X^2\frac{S(X_t,\tau)P_1(\tau)}{S(X_t,\tau)} \\ &= \frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_X^2\frac{1-\lambda}{\kappa}(1 - e^{-\kappa\tau}), \end{aligned} \tag{A.105}$$

which is increasing in τ since $\gamma > 1$, $\lambda < 1$, $\kappa > 0$, and $\beta > 0$. Since the asset is paying a single cash flow, notice that its duration is equal to τ. The duration and expected return of a portfolio of cash flows are value-weighted averages of the duration and expected returns of its cash flows, respectively. Therefore, it is straightforward to show that the positive duration–expected return relation generalizes to any portfolio whose payoffs are non-negative fractions of the aggregate dividend. □

A.3.2 Cross Section of Dividends

Proof of Theorem 2:

Proof. The sum of two dividend shares $\theta_t^p = \theta_t^i + \theta_t^j$ follows a WF process. Applying Itô's Lemma with $\rho_t(dW_{\theta^i}, dW_{\theta^j}) = -\sqrt{\frac{\theta_t^i\theta_t^j}{(1-\theta_t^i)(1-\theta_t^j)}}$,

$$\begin{aligned} d\theta_t^p &= d\theta_t^i + d\theta_t^j \\ &= \alpha(\bar{\theta}^i + \bar{\theta}^j - \theta_t^i - \theta_t^j)dt + \delta\sqrt{(1-\theta_t^i)\theta_t^i}dW_{\theta^i} + \delta\sqrt{(1-\theta_t^j)\theta_t^j}dW_{\theta^j} \\ &= \alpha(\bar{\theta}^p - \theta_t^p)dt \\ &\quad +\delta\sqrt{(1-\theta_t^i)\theta_t^i + (1-\theta_t^j)\theta_t^j - 2\sqrt{(1-\theta_t^i)\theta_t^i(1-\theta_t^j)\theta_t^j}\frac{\theta_t^i\theta_t^j}{(1-\theta_t^i)(1-\theta_t^j)}}dW_{\theta^p} \\ &= \alpha(\bar{\theta}^p - \theta_t^p)dt + \delta\sqrt{(1-\theta_t^p)\theta_t^p}dW_{\theta^p}, \end{aligned} \tag{A.106}$$

where $\bar{\theta}^p \equiv \bar{\theta}^i + \bar{\theta}^j$. □

Additional Aggregation Results:

We now show that the portfolio relative share is given by equation (2.11) and that the cross section of firms aggregates properly. From Theorem 2, the dividend share of a portfolio p^* holding the full value of all firms in $\mathcal{P}$ follows the WF process,

$$d\theta_t^{p^*} = \alpha(\bar{\theta}^{p^*} - \theta_t^{p^*})dt + \delta\sqrt{(1-\theta_t^{p^*})\theta_t^{p^*}}dW_{\theta^{p^*}}, \tag{A.107}$$

where $\bar{\theta}^{p^*} = \sum_{i\in\mathcal{P}}\bar{\theta}^i$ and $\theta_t^{p^*} = \sum_{i\in\mathcal{P}}\theta_t^i$. Then consider a portfolio p which holds the proportion of firm value ω of each firm in $\mathcal{P}$. By Itô's lemma, the process for the portfolio's dividend share is

$$d\theta_t^p = \alpha(\bar{\theta}^p - \theta_t^p)dt + \delta\sqrt{(\omega-\theta_t^p)\theta_t^p}dW_{\theta^p}, \tag{A.108}$$

where $\bar{\theta}^p = \sum_{i \in \mathcal{P}} \omega \bar{\theta}^i$ and $\theta_t^p = \sum_{i \in \mathcal{P}} \omega \theta_t^i$. Asset pricing results must be equivalent for $\omega < 1$ and $\omega = 1$ to prevent the existence of arbitrage opportunities. Equation (2.11) is derived by

$$\frac{\bar{\theta}^p}{\theta_t^p} = \frac{\sum_{i \in \mathcal{P}} \omega \bar{\theta}^i}{\sum_{i \in \mathcal{P}} \omega \theta_t^i} = \sum_{i \in \mathcal{P}} \frac{\theta_t^i}{\sum_{i \in \mathcal{P}} \theta_t^i} \frac{\bar{\theta}^i}{\theta_t^i}. \tag{A.109}$$

To establish the aggregation result, let all firms $i = 1, ..., n$ be included in $\mathcal{P}$ and let $\omega = 1$ so all firms are held in their entirety within the portfolio. Then $\sum_{i \in \mathcal{P}} \theta_0^i = 1$ at time 0 and $d(\sum_{i \in \mathcal{P}} \theta_t^i) = 0$ by equation (A.108). Therefore, $1 = \sum_{i \in \mathcal{P}} \theta_t^i = \sum_{i=1}^n \theta_t^i$ for all t, hence $D_t = \sum_{i=1}^n \theta_t^i D_t$.

A.3.3 Cross-Sectional Asset Pricing

Proof of Theorem 3:

Proof. The value of a firm is the expected discounted value of firm dividends,

$$\begin{aligned} P_t^i &= E_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} D_s^i ds \\ &= E_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} D_s \theta_s^i ds. \end{aligned} \tag{A.110}$$

In order to value future dividends, we need to compute $E_t[\theta_s^i] = f(\theta_t^i, \tau)$, where $\tau = s - t$. Applying Itô's Lemma and Feynman–Kac,

$$\alpha(\bar{\theta}^i - \theta_t^i) f_\theta + \frac{1}{2}\delta^2 \theta_t^i (1 - \theta_t^i) f_{\theta\theta} - f_\tau = 0, \tag{A.111}$$

which has a solution

$$f(\theta_t^i, \tau) = E_t[\theta_s^i] = \theta_t^i e^{-\alpha\tau} + \bar{\theta}^i (1 - e^{-\alpha\tau}). \tag{A.112}$$

Notice that this expectation is simply a weighted average of the current dividend share θ_t^i and the long-run dividend share $\bar{\theta}^i$, with $f(\theta_t^i, \tau) = \theta_t^i$ at $\tau = 0$ and $\lim_{\tau \to \infty} f(\theta_t^i, \tau) = \bar{\theta}^i$ as our expectation of θ^i converges to the long-run dividend share as τ grows large.

The price of security i is

$$\begin{aligned} P_t^i &= \frac{1}{\Lambda_t} E_t \int_t^\infty \Lambda_s \theta_s^i D_s ds \\ &= \frac{1}{\Lambda_t} \int_t^\infty E_t[\Lambda_s \theta_s^i D_s] ds \\ &= \frac{1}{\Lambda_t} \int_t^\infty E_t[\Lambda_s D_s] E_t[\theta_s^i] ds \qquad [independence] \\ &= \frac{1}{\Lambda_t} \int_t^\infty \Lambda_t D_t S(X_t, \tau)[\theta_t^i e^{-\alpha\tau} + \bar{\theta}^i(1 - e^{-\alpha\tau})] ds \qquad [\tau = s - t], \end{aligned} \tag{A.113}$$

where $S(X_t, \tau)$ is defined in Proposition 1. The price-dividend ratio of firm i follows by dividing both sides by $D_t^i = \theta_t^i D_t$

$$\frac{P_t^i}{D_t^i} \equiv G^i(X_t, \theta_t^i; \bar{\theta}^i, \alpha) = \int_t^\infty S(X_t, \tau) \left[e^{-\alpha\tau} + \frac{\bar{\theta}^i}{\theta_t^i}(1 - e^{-\alpha\tau}) \right] ds. \tag{A.114}$$

Firm i's price-dividend ratio is increasing in relative share given $\alpha > 0$ since $(1 - e^{-\alpha\tau}) > 0$ and $S(X_t, \tau) > 0$ for all X_t and τ.

Individual firm excess return obeys

$$\begin{aligned} dR_t^i &= \frac{dP_t^i + D_t^i dt}{P_t^i} - r_t^f dt \\ &= \mu_t^i dt + \sigma_D dW_D + \frac{G_X^i}{G^i} \sigma_X dW_X + \delta \sqrt{\frac{1 - \theta_t^i}{\theta_t^i}} \left(1 + \frac{\theta_t^i G_\theta^i}{G^i} \right) dW_{\theta^i}, \end{aligned} \tag{A.115}$$

where μ_t^i is the expected excess return given by

$$\begin{aligned} \mu_t^i &= -Cov\left(\frac{dP^i}{P^i}, \frac{d\Lambda}{\Lambda} \right) \\ &= \frac{\lambda(\gamma - 1)}{\kappa + \beta} \sigma_X^2 \frac{G_X^i}{G^i}. \end{aligned} \tag{A.116}$$

We now establish conditions under which μ_t^i is increasing in relative share. First, define relative share $r_t^i = \frac{\bar{\theta}^i}{\theta_t^i}$. Assuming $\gamma > 1$, μ_t^i is increasing in relative share if $\frac{G_X^i}{G^i}$ is increasing in relative share, where

$$G_X^i = r_t^i \int_t^\infty S(X_t, \tau) P_1(\tau)(1 - e^{-\alpha\tau}) ds + \int_t^\infty S(X_t, \tau) P_1(\tau) e^{-\alpha\tau} ds \tag{A.117}$$

$$G^i = r_t^i \int_t^\infty S(X_t, \tau)(1 - e^{-\alpha\tau}) ds + \int_t^\infty S(X_t, \tau) e^{-\alpha\tau} ds. \tag{A.118}$$

Since G^i and G_X^i are both linear in r_t^i, the sign of $\frac{\partial PV_t}{\partial r_t^i}$ depends on the sign of

$$\begin{aligned} c(r_t^i) = &\int_t^\infty S(X_t, \tau) P_1(\tau) ds \int_t^\infty S(X_t, \tau) e^{-\alpha\tau} ds \\ &- \int_t^\infty S(X_t, \tau) P_1(\tau) e^{-\alpha\tau} ds \int_t^\infty S(X_t, \tau) ds. \end{aligned} \tag{A.119}$$

Substituting in $P_1(\tau) = \frac{1-\lambda}{\kappa}\left(1 - e^{-\kappa\tau}\right)$ the sign of c is determined by[15]

$$\begin{aligned}\mathrm{sgn}(c(r_t^i)) = \mathrm{sgn}\Bigg(&\int_t^\infty S(X_t,\tau)ds \int_t^\infty S(X_t,\tau)e^{-(\alpha+\kappa)\tau}ds \\ &- \int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds \int_t^\infty S(X_t,\tau)e^{-\kappa\tau}ds\Bigg). \end{aligned} \tag{A.120}$$

In order to show $\mathrm{sgn}(c(r_t^i)) > 0$, we need to show that

$$\int_t^\infty S(X_t,\tau)ds \int_t^\infty S(X_t,\tau)e^{-(\alpha+\kappa)\tau}ds > \int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds \int_t^\infty S(X_t,\tau)e^{-\kappa\tau}ds, \tag{A.121}$$

which follows by Chebyshev's algebraic inequality since $e^{-\alpha\tau}$ and $e^{-\kappa\tau}$ are monotonic functions of τ ((Mitrinović, Pečarić, and Fink, 1993, p. 239)).

Firm i's expected return μ_t^i has a beta pricing relation with an economic growth hedge portfolio,

$$\mu_t^i = \beta_t^i \mu_t, \tag{A.122}$$

where $\beta_t^i\left(X_t, \frac{\bar{\theta}^i}{\theta_t^i}\right)$ is firm i's beta on an economic growth hedge portfolio with expected return μ_t. The economic growth hedge portfolio is constructed to have returns that follow

$$dR_t = \mu_t dt + \frac{G_X}{G}\sigma_X dW_X, \tag{A.123}$$

where μ_t is defined in Proposition 1. Notice that

$$\mu_t = -Cov\left(dR_t, \frac{d\Lambda}{\Lambda}\right) = \frac{\lambda(\gamma-1)}{\kappa+\beta}\sigma_X^2 \frac{G_X}{G}, \tag{A.124}$$

so the economic growth hedge portfolio is correctly priced. Firm i's beta is

$$\begin{aligned}\beta_t^i\left(X_t, \frac{\bar{\theta}^i}{\theta_t^i}\right) &= \frac{Cov(dR_t^i, dR_t)}{Var(dR_t)} \\ &= \frac{\sigma_X^2 \frac{G_X^i}{G^i}\frac{G_X}{G}}{\sigma_X^2\left(\frac{G_X}{G}\right)^2} \\ &= \frac{G_X^i}{G^i}\frac{G}{G_X}. \end{aligned} \tag{A.125}$$

Then substituting for $\frac{G_X^i}{G^i}$ in the equation of μ_t^i gives

$$\begin{aligned}\mu_t^i &= \frac{\lambda(\gamma-1)}{\kappa+\beta}\sigma_X^2 \frac{G_X^i}{G^i} \\ &= \beta_t^i \frac{\lambda(\gamma-1)}{\kappa+\beta}\sigma_X^2 \frac{G_X}{G} \\ &= \beta_t^i \mu_t. \end{aligned} \tag{A.126}$$

[15] Hore (2008) and Bansal and Yaron (2004) show that $\lambda < 1$ in the post-war US economy.

Finally, given the process for firm i's return in equation (A.115), the instantaneous variance of the firm return is

$$\sigma_{i,t}^2 = \sigma_D^2 + \left(\frac{G_X^i}{G^i}\sigma_X\right)^2 + \delta^2\frac{1-\theta_t}{\theta_t}\left(1+\frac{\theta_t G_\theta^i}{G^i}\right)^2, \tag{A.127}$$

since all covariance terms equal zero
(i.e. $Cov(dW_X, dW_D) = Cov(dW_{\theta^i}, dW_D) = Cov(dW_{\theta^i}, dW_X) = 0$). □

A.3.4 Dispersion and Idiosyncratic Volatility

Proof of Theorem 4:

Proof. Dispersion is defined by

$$Dispersion_t^i = \left|\frac{\sqrt{\delta^2\left(\frac{r_t^i}{\bar\theta^i}-1\right)+\sigma_D^2}}{\alpha(r_t^i-1)+\mu_D+X_t}\right|, \tag{A.128}$$

where $r_t^i = \frac{\bar\theta^i}{\theta_t^i}$. Assume

$$\alpha(r_t^i-1)+\mu_D+X_t > 0, \tag{A.129}$$

so expected dividend growth for firm i is positive. The direction of the relation between r_t^i and $Dispersion_t^i$ will be the same as between r_t^i and $(Dispersion_t^i)^2$ since expected dividend growth is positive. The partial derivative is

$$\frac{\partial(Dispersion_t^i)^2}{\partial r_t^i} = \frac{\frac{\delta^2}{\bar\theta^i}(-\alpha r_t^i+\alpha(2\bar\theta^i-1)+\mu_D+X_t)-2\alpha\sigma_D^2}{(\alpha(r_t^i-1)+\mu_D+X_t)^3}. \tag{A.130}$$

From equation (A.129), the denominator is positive. Therefore, the sign of $\frac{\partial(Dispersion_t^i)^2}{\partial r_t^i}$ depends on the sign of

$$f(r_t^i) = -\alpha r_t^i+\alpha(2\bar\theta^i-1)+\mu_D+X_t-2\frac{\bar\theta^i}{\delta^2}\alpha\sigma_D^2. \tag{A.131}$$

From equation (A.129), $-\alpha r_t^i < (\mu_D+X_t)-\alpha$, so

$$\begin{aligned} f(r_t^i) &= -\alpha r_t^i+\alpha(2\bar\theta^i-1)+\mu_D+X_t-2\frac{\bar\theta^i}{\delta^2}\alpha\sigma_D^2 && \text{(A.132)}\\ &< 2\left(-\alpha\left(1-\bar\theta^i\left(1-\frac{\sigma_D^2}{\delta^2}\right)\right)+\mu_D+X_t\right) && \text{(A.133)}\end{aligned}$$

provides an upper bound for $f(r_t^i)$. A sufficient (but not necessary) condition for a negative relation between r_t^i and $Dispersion_t^i$ is

$$\alpha\left(1-\bar\theta^i\left(1-\frac{\sigma_D^2}{\delta^2}\right)\right) > \mu_D+X_t, \tag{A.134}$$

which holds quite generally since the term in parentheses is close to one for reasonable parameters and α estimates are much larger than estimates of μ_D (the unconditional mean of X_t is zero). $\square$

When expected dividend growth of firm i is negative, the relation between dispersion and relative share is generally positive. However, during normal economic conditions (i.e. X_t reasonably close to zero), relatively few firms in the cross section will have negative expected dividend growth. Furthermore, dispersion for firms with negative expected dividend growth that is close to zero will have higher dispersion than firms with expected dividend growth equal to that of the aggregate dividend, preserving the negative cross-sectional relation. Simulations show that a negative cross-sectional relation between relative share and dispersion arises quite generally under a variety of economic conditions and specifications of the cross section of firms.

Proof of Theorem 5:

Proof. Idiosyncratic volatility of asset i is given by

$$\begin{aligned}
IV_t^i &= \delta\sqrt{\left(\frac{1}{\theta_t^i}-1\right)}\left(1+\frac{\theta_t^i G_\theta^i}{G^i}\right) \\
&= \delta\sqrt{\left(\frac{1}{\theta_t^i}-1\right)}\left(\frac{\int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds}{G+\int_t^\infty S(X_t,\tau)\left(\frac{\bar{\theta}^i}{\theta_t^i}-1\right)(1-e^{\alpha\tau})ds}\right) \\
&= \delta\left(\int_t^\infty S(X_t,\tau)e^{\alpha\tau}ds\right)\frac{\sqrt{\frac{r_t^i}{\bar{\theta}^i}-1}}{G+(r_t^i-1)\int_t^\infty S(X_t,\tau)(1-e^{-\alpha\tau})ds}. \qquad \text{(A.135)}
\end{aligned}$$

Notice the change of variable $\frac{\bar{\theta}^i}{\theta_t^i}=r_t^i$ which implies that $\frac{1}{\theta_t^i}=\frac{r_t^i}{\bar{\theta}^i}$. By construction, $\frac{r_t^i}{\bar{\theta}^i}>1$.

The sign of the derivative $\frac{\partial IV_t^i}{\partial r_t^i}$ depends on whether the quantity

$$c(r_t^i)=\frac{\sqrt{\frac{r_t^i}{\bar{\theta}^i}-1}}{G+(r_t^i-1)\int_t^\infty S(X_t,\tau)(1-e^{-\alpha\tau})ds} \qquad \text{(A.136)}$$

is increasing or decreasing in r_t^i. By taking the derivative of $c(r_t^i)$ with respect to r_t^i, $\operatorname{sgn}\left(\frac{\partial IV_t^i}{\partial r_t^i}\right)$ is determined by the sign of the quantity

$$\bar{c} = (2\bar{\theta}^i - r_t^i)\int_t^\infty S(X_t,\tau)(1-e^{-\alpha\tau})ds + \int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds. \tag{A.137}$$

Rewrite $\bar{c}$ as

$$\begin{aligned}\bar{c} &= (2\bar{\theta}^i - r_t^i)\int_t^\infty S(X_t,\tau)ds + (r_t^i - 2\bar{\theta}^i + 1)\int_t^\infty S(X_t,\tau)e^{-\alpha\tau}ds &\text{(A.138)}\\ &\leq (2\bar{\theta}^i - r_t^i)|S(X_t,\tau)|_1 + \frac{(r_t^i - 2\bar{\theta}^i + 1)}{\sqrt{2\alpha}}|S(X_t,\tau)|_2, &\text{(A.139)}\end{aligned}$$

where the last inequality is due to the Cauchy–Schwarz inequality and $|\cdot|_i$ is the i-th norm. The success of equation (A.139) is to separate α from inside the integral, thus disentangling any non-linear dependence from other variables inside the integral. Now, we can enforce a condition for the idiosyncratic volatility result in the $(\bar{\theta}^i, \alpha)$-plane that is independent from the other variables in the system. Since, $S(X_t,\tau)$ is positive $\forall\tau$, the first norm is much greater than the second norm because the cross-terms are all positive. Thus, for $\bar{c} < 0$ in (A.139), all we need is to pick α such that

$$(r_t^i - 2\bar{\theta}^i)(\sqrt{2\alpha}M - 1) > 1, \tag{A.140}$$

where $M = \frac{|S(X_t,\tau)|_1}{|S(X_t,\tau)|_2} >> 1$ for reasonable parameter values that we consider here. The first term $r_t^i - 2\bar{\theta}^i > 0$ for $\theta_t^i < 1/2$. In fact, $r_t^i - 2\bar{\theta}^i$ is increasing as θ_t^i decreases. Fix $\theta_t^i < 1/2$ and pick a $\bar{\theta}^i$. Then, one can pick α according to the above condition to ensure that $\bar{c} < 0$.[16] Simulations show that this bound can be easily achieved. □

A.4 Long-Run Risk Model Estimation

A.4.1 Data

Analyst earnings forecast data is from I/B/E/S. Dispersion is calculated following Avramov, Chordia, Jostova, and Philipov (2009) as the standard deviation of analyst fiscal year one earnings estimates divided by the absolute value of the mean estimate. Firms must have at least two analyst forecasts in a month to be included in the sample. Data for calculating IV and credit risk is from Compustat

[16] Notice, this holds true for the upper bound. It is possible that $\bar{c} < 0$ under looser restrictions.

and CRSP. Following Ang, Hodrick, Xing, and Zhang (2006), IV is defined as the standard deviation of pricing errors relative to the Fama–French (1993) three-factor model in a regression using daily returns and a one-month period.[17] Firms with fewer than 18 daily return observations in the month are omitted from the sample.[18] Our proxy for credit risk is calculated following Campbell, Hilscher, and Szilagyi (2008) who model the probability of corporate failure over the next year as a function of accounting and market variables. See Campbell, Hilscher, and Szilagyi (2008) for details on this measure.

Our sample period is July 1981 to June 2008. We choose this period since the credit risk anomaly has been demonstrated for the post-1980 period and the number of firms with valid dispersion data increases sharply in the early 1980s. We form value-weighted decile portfolios based on the dispersion, IV, and credit risk measures. Dispersion and IV portfolios are rebalanced monthly following the related literature, while credit risk portfolios is rebalanced annually at the beginning of July. All firms with a price less than $1 at the time of portfolio formation are excluded from the portfolios. Stock return data is from CRSP, and all returns are adjusted for inflation. After imposing all data requirements, there are an average of 2,761 firms per month that are eligible for inclusion in dispersion portfolios, 5,456 firms in IV portfolios, and 4,355 firms in credit risk portfolios.

Our aggregate dividend and consumption data sample is annual, post-war

---

[17] We thank Kenneth French for making factor returns available through his website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[18] In September 2001, the stock market was closed for four unscheduled days. We therefore require at least 14 daily returns for each firm in September 2001.

US data from 1948 to 2008. Aggregate dividends are calculated for the CRSP value-weighted index following Cochrane (2008a).[19] Aggregate consumption is non-durable goods plus services from the NIPA tables of the Bureau of Economic Analysis. We convert the data to a real, per capita basis by adjusting for CPI inflation and the US population.

A.4.2 Estimation

A.4.2.1 Aggregate Parameter Estimation

We estimate the aggregate parameters of the Euler approximations of equations (2.3), (2.4), and (2.5),

$$\frac{C_{t+1} - C_t}{C_t} = \mu_C + \lambda X_t + \sigma_C \epsilon_C, \tag{A.141}$$

$$\frac{D_{t+1} - D_t}{D_t} = \mu_D + X_t + \sigma_D \epsilon_D, \tag{A.142}$$

$$X_{t+1} = (1 - \kappa) X_t + \sigma_X \epsilon_X, \tag{A.143}$$

using a Bayesian approach. We use a Markov chain Monte Carlo (MCMC) technique to draw from the posterior distribution of the aggregate parameters. To draw from the posterior, we iterate through the following steps with the draws for each parameter being conditioned on the most recent draws of the other parameters.

1. Draw $\{X_t\}_{t=1}^{T} | \mu_D, \mu_C, \kappa, \sigma_D, \sigma_C, \sigma_X$ using a forward filtering–backward sampling algorithm (FFBS).[20] This step produces a draw from the posterior dis-

[19] Cochrane (2008a) points out that CRSP dividends capture all payments to investors including cash mergers, liquidations, and repurchases.

[20] FFBS was developed by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Also see West and Harrison (1997) for details on FFBS.

tribution of the unobservable sequence $\{X_t\}_{t=1}^T$ conditional on the remaining parameters. Define $F = [1 \quad \lambda]$ and $W = \begin{bmatrix} \sigma_D^2 & 0 \\ 0 & \sigma_C^2 \end{bmatrix}$. FFBS is performed in two steps:

(a) *Forward filtering*: Calculate $\{m_t\}_{t=1}^T, \{C_t\}_{t=1}^T | \mu_D, \mu_C, \kappa, \sigma_D, \sigma_C, \sigma_X, a_1, R_1$, where a_1 and R_1 are prior parameters. The time series of m_t and C_t, which are intermediate parameters in the FFBS procedure, are found by iterating through $t = 1$ to T with the following equations (with a_t and R_t being set to the prior parameters a_1 and R_1 when $t = 1$):

$$a_t = (1 - \kappa) m_{t-1}, \tag{A.144}$$

$$R_t = (1 - \kappa)^2 C_{t-1} + \sigma_X^2, \tag{A.145}$$

$$Q = F' R_t F + W, \tag{A.146}$$

$$A = R_t F Q^{-1}, \tag{A.147}$$

$$m_t = a_t + A \left(\begin{bmatrix} \frac{D_{t+1} - D_t}{D_t} - \mu_D, \\ \frac{C_{t+1} - C_t}{C_t} - \mu_C \end{bmatrix} - F' a_t \right), \tag{A.148}$$

$$C_t = R_t - A Q A'. \tag{A.149}$$

(b) *Backward sampling*: Draw the sequence $\{X_t\}_{t=1}^T | \{m_t\}_{t=1}^T, \{C_t\}_{t=1}^T, \kappa, \sigma_X$ by first drawing

$$X_T \sim N(m_T, C_T), \tag{A.150}$$

and then iterating from $t = T-1$ to 1 drawing from

$$X_t \sim N\left(\left((1-\kappa)^2 W^{-1} + C_t^{-1}\right)^{-1}\left((1-\kappa)W^{-1}X_{t+1} + C_t^{-1}m_t\right)\right. \tag{A.151}$$

$$\left., \left((1-\kappa)^2 W^{-1} + C_t^{-1}\right)^{-1}\right), \tag{A.152}$$

to draw a full sequence $\{X_t\}_{t=1}^T$ from the posterior distribution of possible sequences.

2. Draw $\mu_D, \sigma_D | \{X_t\}_{t=1}^T$. Define $y_t = \frac{D_{t+1}-D_t}{D_t} - X_t$ to be the sequence of observed dividend growth minus the current draw of $\{X_t\}_{t=1}^T$. Calculate $\bar{y} = \frac{1}{T}\sum_{t=1}^T y_t$ and $s^2 = \sum_{t=1}^T y_t^2$. Then calculate $\overline{\mu_D} = (\underline{H} + T)^{-1}(\underline{H}\underline{\mu_D} + T\bar{y})$ and $\overline{H} = (\underline{H} + T)$. The parameter σ_D is calculated from the draw

$$\frac{\bar{s}^2}{\sigma_D^2} \sim \chi^2(\bar{\nu}), \tag{A.153}$$

where $\bar{s}^2 = \underline{s}^2 + Q^*$, $\bar{\nu} = \underline{\nu} + T$, and $Q^* = s^2 + T(\overline{\mu_D} - \bar{y})^2 + (\overline{\mu_D} - \underline{\mu_D})'\underline{H}(\overline{\mu_D} - \underline{\mu_D})$. Finally, the draw of μ_D conditional on σ_D^2 is from the univariate normal distribution

$$\mu_D \sim N(\overline{\mu_D}, \sigma_D^2 \overline{H}^{-1}). \tag{A.154}$$

3. Draw $\mu_C, \lambda, \sigma_C | \{X_t\}_{t=1}^T$ using a Bayesian regression.[21] Define $\mathbf{X} = [1 \quad X_t]$ as the $T \times 2$ matrix containing a constant and the sequence $\{X_t\}_{t=1}^T$ and $\mathbf{y} = \left[\frac{C_{t+1}-C_t}{C_t}\right]$ as a $T \times 1$ vector with the sequence of consumption growth. Also define $\beta = [\mu_C \quad \lambda]'$. Calculate $b = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $s^2 = (\mathbf{y} - \mathbf{X}b)'(\mathbf{y} - \mathbf{X}b)$. Then calculate $\bar{\beta} = (\underline{H} + \mathbf{X}'\mathbf{X})^{-1}(\underline{H}\underline{\beta} + \mathbf{X}'\mathbf{X}b)$ and $\overline{H} = (\underline{H} + \mathbf{X}'\mathbf{X})$. The

[21] See Rossi, Allenby, and McCulloch (2005b) and Geweke (2005) for further discussions of Bayesian regressions.

parameter σ_C is calculated from the draw

$$\frac{\bar{s}^2}{\sigma_C^2} \sim \chi^2(\bar{\nu}), \tag{A.155}$$

where $\bar{s}^2 = \underline{s}^2 + Q^*$, $\bar{\nu} = \underline{\nu} + T$, and $Q^* = s^2 + (\bar{\beta} - b)'\mathbf{X}'\mathbf{X}(\bar{\beta} - b) + (\bar{\beta} - \underline{\beta})'\underline{H}(\bar{\beta} - \underline{\beta})$. Finally, the draw of $\beta = [\mu_C \quad \lambda]'$ conditional on σ_C^2 is from the multivariate normal distribution

$$\beta \sim N(\bar{\beta}, \sigma_C^2 \overline{H}^{-1}). \tag{A.156}$$

4. Draw $\kappa, \sigma_X | \{X_t\}_{t=1}^T$ using a Bayesian AR(1) regression. Define $\mathbf{X_T}$ to be a $T-1 \times 1$ vector containing the sequence $\{X_t\}_{t=2}^T$ and $\mathbf{X_{T-1}}$ to be a $T-1 \times 1$ vector containing the sequence $\{X_t\}_{t=1}^{T-1}$. Calculate $b = (\mathbf{X_{T-1}}'\mathbf{X_{T-1}})^{-1}\mathbf{X_{T-1}}'\mathbf{X_T}$ and $s^2 = (\mathbf{X_T} - \mathbf{X_{T-1}}b)'(\mathbf{X_T} - \mathbf{X_{T-1}}b)$. Then calculate $\bar{\beta} = (\underline{H} + \mathbf{X_{T-1}}'\mathbf{X_{T-1}})^{-1}(\underline{H}\underline{\beta} + \mathbf{X_{T-1}}'\mathbf{X_{T-1}}b)$ and $\overline{H} = (\underline{H} + \mathbf{X_{T-1}}'\mathbf{X_{T-1}})$. The parameter σ_X is calculated from the draw

$$\frac{\bar{s}^2}{\sigma_X^2} \sim \chi^2(\bar{\nu}), \tag{A.157}$$

where $\bar{s}^2 = \underline{s}^2 + Q^*$, $\bar{\nu} = \underline{\nu} + T - 1$, and $Q^* = s^2 + (\bar{\beta} - b)'\mathbf{X_{T-1}}'\mathbf{X_{T-1}}(\bar{\beta} - b) + (\bar{\beta} - \underline{\beta})'\underline{H}(\bar{\beta} - \underline{\beta})$. Finally, the draw of κ conditional on σ_X^2 is calculated from the draw of $(1-\kappa)$ from the univariate normal distribution

$$(1 - \kappa) \sim N(\bar{\beta}, \sigma_X^2 \overline{H}^{-1}). \tag{A.158}$$

We draw 6,000 times from the posterior distribution of the parameters and discard the initial 1,000 draws as a burn-in period. The mean draws from the posterior distribution appear in Panel A of Table 1.3.

A.4.2.2 Portfolio Parameter Estimation

We calculate portfolio dividends following Bansal, Dittmar, and Lundblad (2005).[22] Specifically, portfolio value evolves according to

$$H_{t+1}^{i} = H_{t}^{i} R_{t}^{ix}, \tag{A.159}$$

where H_t^i is the value of the portfolio at month t and R_t^{ix} is the ex-dividend portfolio return. Since the relative size of portfolio dividends in the cross section is important to our empirical analysis, we set the initial portfolio values at the beginning of July 1981, H_0^i, to the aggregate market value of the firms in each portfolio. The portfolio dividend, D_t^i, is therefore

$$D_{t}^{i} = H_{t}^{i} (R_{t}^{i} - R_{t}^{ix}), \tag{A.160}$$

where R_t^i is the portfolio return. We sum monthly dividends within each quarter to obtain quarterly dividends. Quarterly dividends continue to exhibit strong seasonality. In response, we use a trailing four-quarter average of the portfolio dividend as the portfolio cash flow following Bansal, Dittmar, and Lundblad (2005). Portfolio cash flows are adjusted for inflation. Given the cash flows for each of the ten portfolios, the portfolio dividend share is

$$\theta_{t}^{i} = \frac{D_{t}^{i}}{\sum_{j=1}^{10} D_{t}^{j}}. \tag{A.161}$$

We therefore have a quarterly time series of dividend share for each portfolio from Q2 1982 to Q2 2008 (after allowing for three lags of quarterly dividends for a four-quarter average) while ensuring that the sum of portfolio dividend shares is always equal to one.

[22]See also Menzly, Santos, and Veronesi (2004) and Hansen, Heaton, and Li (2008).

For each of the three sets of decile portfolios, we estimate the parameter $\bar{\theta}^i$ for each portfolio and the parameters α and δ using a Bayesian MCMC approach. Using the Euler approximation of equation (2.8) along with the instantaneous correlation in equation (2.9), we have a system of ten equations with one equation for each portfolio i of the form,

$$\theta^i_{t+1} = \theta^i_t + \alpha(\bar{\theta}^i - \theta^i_t) + \delta\epsilon^i_{t+1}, \quad i = 1, ..., n, \quad \epsilon_{t+1} \sim N(0, \Sigma_t), \tag{A.162}$$

where

$$\Sigma_t = \begin{bmatrix} \theta^1_t(1-\theta^1_t) & \cdots & -\theta^1_t\theta^i_t & \cdots & -\theta^1_t\theta^n_t \\ \vdots & \ddots & & & \vdots \\ -\theta^i_t\theta^1_t & \cdots & \theta^i_t(1-\theta^i_t) & \cdots & -\theta^i_t\theta^n_t \\ \vdots & & & \ddots & \vdots \\ -\theta^n_t\theta^1_t & \cdots & -\theta^n_t\theta^i_t & \cdots & \theta^n_t(1-\theta^n_t) \end{bmatrix}. \tag{A.163}$$

Further, we have the parameter restrictions that $\bar{\theta}^j > 0$ for all portfolios j, $\alpha > 0$, $\delta > 0$, and $\sum_{j=1}^n \bar{\theta}^j = 1$. Note that the full variance-covariance matrix Σ is not invertible. To see this, consider the variance of θ^n_{t+1} given $\{\theta^i_{t+1}\}_{i=1}^{n-1}$, which is

$$\delta^2\theta^n_t(1-\theta^n_t) - \delta^2\Sigma_{(n,1:n-1)}\Sigma^{-1}_{(1:n-1,1:n-1)}\Sigma_{(1:n-1,n)} = 0, \tag{A.164}$$

while the mean of θ^n_{t+1} is $1 - \sum_{i=1}^{n-1}\theta^i_{t+1}$. Therefore, the variance-covariance matrix Σ automatically enforces the constraint $\sum_{j=1}^n \bar{\theta}^j = 1$. In practice, a combination of the variance-covariance matrix of the first $n-1$ dividend shares, $\Sigma_{(1:n-1,1:n-1)}$, combined with the constraint $\sum_{j=1}^n \bar{\theta}^j = 1$ are used to simulate or estimate the system.

The parameters in equation (A.162) do not have a convenient posterior distribution, so we employ a Metropolis–Hastings algorithm to draw from the posterior distribution. The Metropolis–Hastings algorithm can be used to draw from an unknown posterior distribution by initially drawing a candidate set of parameters from an arbitrary transition distribution and then accepting the candidate draw or rejecting the draw in favor of the previous draw based on the likelihoods of the two draws.

Below, we outline the MCMC chain for estimating parameters for a single set of test assets. Defining the vector of parameters $\Omega = [\bar{\theta}, \alpha, \delta]$, the draw of Ω in iteration m of the MCMC chain is as follows:

1. Draw a vector $\hat{\theta} \sim N_{[0,\infty)}(\bar{\theta}^{(m-1)}, A)$, where A is a diagonal matrix scaled to provide for sufficient mixing and the multivariate normal distribution is truncated at zero in all n dimensions because of the restrictions $\bar{\theta}^j > 0$ for $j = 1, ..., n$. The candidate vector is then normalized to sum to one, $\bar{\theta}^* = \frac{\hat{\theta}}{\hat{\theta}'\iota}$, to satisfy the restriction $\sum_{j=1}^{n} \bar{\theta}^j = 1$. Finally, $\Omega^* = [\bar{\theta}^*, \alpha^{(m-1)}, \delta^{(m-1)}]$ is accepted with probability

$$\min\left\{\frac{p(\Omega^*)/q(\Omega^*|\Omega^{(m-1)})}{p(\Omega^{(m-1)})/q(\Omega^{(m-1)}|\Omega^*)}, 1\right\}, \tag{A.165}$$

where

$$p(\Omega) \propto \prod_{t=1}^{T} \exp\left(-\frac{\left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)' \Sigma_{t,-n}^{-1} \left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)}{2\delta^2}\right), \tag{A.166}$$

is the kernel of a multivariate normal distribution, θ_{t+1}^{-n} and $\bar{\theta}^{-n}$ are vectors of dividend shares and long-run means of dividend shares of the first $n-1$ firms,

Σ_{-n} is the $(n-1 \times n-1)$ block of the variance-covariance matrix defined in equation (A.163) formed by dropping the nth row and column, and $q(\Omega^a|\Omega^b)$ can be replaced by the probability that all elements of a draw $\bar{\theta}^b \sim N(\bar{\theta}^a, A)$ are greater than zero.[23] If the transition density was not truncated, this step would be a random-walk Metropolis–Hastings step and the $q(\Omega^*|\Omega^{(m-1)})$ and $q(\Omega^{(m-1)}|\Omega^*)$ terms would cancel. Scaling the likelihoods of Ω^* and $\Omega^{(m-1)}$ by the probabilities of positive draws adjusts for the differences in transition probabilities arising from the truncation. Transition probabilities are unaffected by the rescaling of $\hat{\theta}$ to $\bar{\theta}^*$ since A is diagonal. If the draw $\bar{\theta}^*$ is accepted then $\bar{\theta}^{(m)} = \bar{\theta}^*$. Otherwise, $\bar{\theta}^{(m)} = \bar{\theta}^{(m-1)}$.

2. Draw a candidate $\alpha^* \sim N_{[0,\infty)}(\alpha^{(m-1)}, \sigma_\alpha)$. The normal distribution is truncated at zero because of the restriction $\alpha > 0$. Then the draw $\Omega^* = [\bar{\theta}^{(m)}, \alpha^*, \delta^{(m-1)}]$ is accepted with probability

$$\min\left\{\frac{p(\Omega^*)/q(\Omega^*|\Omega^{(m-1)})}{p(\Omega^{(m-1)})/q(\Omega^{(m-1)}|\Omega^*)}, 1\right\}, \tag{A.167}$$

where

$$p(\Omega) \propto \prod_{t=1}^{T} \exp\left(-\frac{\left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)' \Sigma_{t,-n}^{-1} \left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)}{2\delta^2}\right), \tag{A.168}$$

and $q(\Omega^a|\Omega^b)$ can be replaced by the probability that a draw $\alpha^b \sim N(\alpha^a, \sigma_\alpha)$

[23] Note that the constants multiplying the kernels to arrive at likelihoods are the same for $p(\Omega^*)$ and $p(\Omega^{(m-1)})$ so they can safely be ignored.

is greater than zero. If the draw α^* is accepted then $\alpha^{(m)} = \alpha^*$, else $\alpha^{(m)} = \alpha^{(m-1)}$.

3. Draw a candidate $\delta^* \sim N_{[0,\infty)}(\delta^{(m-1)}, \sigma_\delta)$. Then the draw $\Omega^* = [\bar{\theta}^{(m)}, \alpha^{(m)}, \delta^*]$ is accepted with probability

$$\min\left\{\frac{p(\Omega^*)/q(\Omega^*|\Omega^{(m-1)})}{p(\Omega^{(m-1)})/q(\Omega^{(m-1)}|\Omega^*)}, 1\right\}, \tag{A.169}$$

where

$$p(\Omega) \propto \frac{1}{\delta^{n-1}} \times \prod_{t=1}^{T} \exp\left(-\frac{\left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)' \Sigma_{t,-n}^{-1} \left(\theta_{t+1}^{-n} - (1-\alpha)\theta_t^{-n} - \alpha\bar{\theta}^{-n}\right)}{2\delta^2}\right), \tag{A.170}$$

and $q(\Omega^a|\Omega^b)$ can be replaced by the probability that a draw $\delta^b \sim N(\delta^a, \sigma_\delta)$ is greater than zero. If the draw δ^* is accepted then $\delta^{(m)} = \delta^*$, else $\delta^{(m)} = \delta^{(m-1)}$.

We draw 250,000 parameter vectors from the posterior distribution and discard the first 50,000 as a burn-in period. Due to the relatively strong autocorrelation in draws that can arise when using a Metropolis–Hastings algorithm, we use every fortieth draw from the posterior as the 5,000 draws that are used in the simulation below. The resulting draws are nearly serially uncorrelated.

A.4.3 Dividend Share Process Tests

We perform two types of specification tests for our dividend share process. First, we follow Da (2009), who performs tests on a similar consumption share process for book-to-market, size, and reversal portfolios, by testing various aspects of the AR(1) process for dividend share implied by our specification. The most important feature of our dividend share process is mean reversion. We test this assumption for portfolios formed on dispersion, IV, and credit risk by implementing an Augmented Dickey–Fuller test. The Dickey–Fuller test examines the stationarity of dividend share (i.e. the existence of mean reversion) versus the hypothesis of a unit root. We assume a non-zero mean and a single lag for the dividend share process. Table A.1 of this document contains the p-values from this test. The hypothesis of a unit root is rejected at the 5% level for 19 of the 30 portfolios. This hypothesis is rejected at the 10% level for three additional portfolios. Overall, the assumption of mean reversion in dividend share appears to be supported among the anomaly portfolios. In addition for testing for mean reversion, we also test the residuals from the AR(1) process as a test of the AR(1) assumption. The Ljung–Box Q test examines whether the AR(1) residuals are white noise. The p-values from this test appear in Table A.1. The AR(1) assumption can not be rejected for 28 of the 30 portfolios, providing further support for our specification as a model for portfolio dividend shares.

Second, we compare our dividend share process to several alternative specifications. We compare models using the AICM and BICM criteria of Raftery, Newton, Satagopan, and Krivitsky (2007). AICM and BICM are posterior simulation-based versions of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) model selection criteria. Hence, the measures select models based on the likelihood of the data given the model, with a penalty for additional parameters. More detail on these model selection criteria is available in Raftery, Newton, Satagopan, and Krivitsky (2007).

For each set of portfolios, we compare our dividend share specification, given by equation (2.8) with correlation described by equation (2.9), to alternative processes using AICM and BICM. We compare our specification to several alternatives. We investigate reasonable alternatives with differing assumptions than our own. Alternative 1 is a Wright–Fisher process analogous to the base specification, but with α and δ allowed to vary across firms,

$$d\theta_t^i = \alpha^i(\bar{\theta}^i - \theta_t^i)dt + \delta^i\sqrt{(1-\theta_t^i)\theta_t^i}dW_{\theta^i}. \tag{A.171}$$

For Alternative 2, we specify a constant, firm-specific diffusion term,

$$d\theta_t^i = \alpha(\bar{\theta}^i - \theta_t^i)dt + \sigma_i dW_{\theta^i}. \tag{A.172}$$

This setup is similar to the share process used by Da (2009) without consumption shocks. Finally, Alternative 3 is a simple specification with firm-specific drift and diffusion terms which remain constant through the sample period,

$$d\theta_t^i = \mu_i dt + \sigma_i dW_{\theta^i}. \tag{A.173}$$

Note that none of the alternatives is able to satisfy the constraint that firm dividends add to the aggregate dividend. However, each alternative alters the assumptions of our dividend share process, so the model comparisons can be viewed as a test of our specification. For each model, we run the MCMC chain for 60,000 iterations and discard the first 10,000 draws as a burn-in period. The AICM and BICM model selection criteria are reported in Table A.2. Our base specification performs well in the model comparisons. For IV and credit risk portfolios, the base specification substantially outperforms all three alternatives whether the model selection criterion is AICM or BICM. For dispersion portfolios, the base specification substantially outperforms Alternatives 2 and 3, while Alternative 1 provides a slightly better fit to the data. In general, our dividend share process is a parsimonious specification which provides a good fit to the data.

A.4.4 Simulation

In each iteration of the simulation, we use a draw from the posterior distributions of the aggregate and portfolio parameters obtained from the estimation procedures in Appendix A.4.2.1 and Appendix A.4.2.2. We have closed-form solutions for several quantities in the unlevered case, such as firm value, expected returns, betas, dispersion, idiosyncratic volatility, and expected dividend growth. For the levered case we use simulations to estimate expected returns, betas, and the probability of bankruptcy. We run the simulation for 5,000 iterations.

We work within the Merton (1974) framework, where firms default if the value of the firm at debt maturity is less than the face value of the debt. In each iteration, we simulate one year of the time series of aggregate dividends D_t, the economic growth variable X_t, the pricing kernel Λ_t, and dividend share θ_t from the Euler approximations of equations (2.3), (2.4), (2.5), and (2.8), respectively, using 250 subperiods during the year while simulating the diffusions. Using these simulated time series, we can calculate the firm value at the end of the year at time T,

$$P_T = D_T \theta_T \int_T^\infty S(X_T, \tau) ds \qquad [\tau = s - t], \tag{A.174}$$

using numerical integration. We can then estimate bankruptcy probabilities by comparing the firm value, P_T, with the face value of debt, B. The proportion of iterations in which $B > P_T$ provides a proxy for distress risk.

Expected returns on levered equity can be estimated by calculating the average return to levered equity across iterations. The (one-year) return from time t to time T in each iteration is

$$R_T = \frac{V_T + \int_t^T D_s \theta_s ds}{V_t}, \tag{A.175}$$

with V_t representing the value of levered equity and $\int_t^T D_s \theta_s ds$ providing the cumulative firm dividend paid over the period. Given the series of D_t, X_t, Λ_t, and θ_t simulated above for this iteration,

$$V_T = \max\{P_T - B, 0\}, \tag{A.176}$$

which can be explicitly calculated and $\int_t^T D_s \theta_s ds$ is estimated using the discrete approximation $\sum_{s=1}^{250} D_{s/250} \theta_{s/250} \frac{1}{250}$. Finally, the value of levered equity at time t, V_t, must be estimated. Within each iteration in the simulation, we run a second loop for 500 iterations to estimate V_t. In each subiteration, we simulate paths for D_t, X_t, Λ_t, and θ_t using the same posterior draws of aggregate and portfolio parameters as in the main iteration. The estimate for V_t is the average across the 500 subiterations of

$$V_t = \int_t^T D_s \theta_s \Lambda_s ds + \Lambda_T \max\{P_T - B, 0\}, \tag{A.177}$$

where the integral is again estimated using a discrete approximation. The expected levered return is estimated using these quantities. In the last step, we estimate the levered equity beta by dividing the estimated expected levered return by the expected return on the economic growth hedge portfolio.

To examine the cross-sectional relation between relative share and dispersion, IV, and credit risk in Figure 2.4, we first develop a cross section of 100 firms. For each firm, we draw a random number from a χ^2 distribution with two degrees of freedom, then scale this number by the sum of the 100 random draws to create the firm's long-run dividend share $\bar{\theta}^i$. This procedure creates a cross section with more small firms than large firms while ensuring proper aggregation of the cross section. Each firm's dividend share at time 0, θ_0^i, is initially set to its long-run dividend share before simulating each dividend share process for a period of 50 years. This procedure creates a cross section of firms which is close to its steady-state distribution in terms of dividend share θ_t^i and relative share $\bar{\theta}^i/\theta_t^i$, with minimal dependence on our initial assumptions about θ_0^i. We then calculate dispersion, IV, and credit risk as discussed above.

A.5 CAPM Model and Estimation

Section A.5.1 provides details about the MCMC estimation algorithm, and Section A.5.2 presents a simulation study that demonstrates the ability of the algorithm to accurately recover parameters.

A.5.1 Estimation Methodology

The model outlined in equations (3.5) to (3.7) can be estimated by repeatedly cycling through steps 1 to 6 below. As discussed in the text, we place a hierarchical structure on alphas, but not on betas. Instead we impose a proper, but diffuse, prior directly on betas in the base specification, $\beta_{i,y} \sim N\left(\mu = 1, \sigma_\beta^2 = 10\right)$. Let $r_{i,t,y}^e$ denote the excess return on stock i in month t of year y and $r_{m,t,y}^e$ the excess return on the market portfolio. Further, let Z denote a matrix in which the first column is a vector of ones and the second column is the excess returns on the market portfolio, and let X denote a matrix of a constant and firm-year characteristics associated with anomalies.

1. Draw $\alpha_{i,y}, \beta_{i,y} | \sigma_{i,y}^2, \delta_y, \sigma_{\alpha,y}^2$ for each stock $i = 1, ..., N$, in each year $y = 1, ..., Y$. We obtain a draw from the marginal posterior distribution of $\alpha_{i,y}$ and $\beta_{i,y}$ as follows:

$$\begin{bmatrix} \alpha_{i,y} \\ \beta_{i,y} \end{bmatrix} \sim N\left(\overline{\lambda}_i, (\sigma_{i,y}^{-2} Z_{i,y}' Z_{i,y} + \mathbf{V}_\lambda^{-1})^{-1}\right), \tag{A.178}$$

where

$$\overline{\lambda}_i = (\sigma_{i,y}^{-2} Z_{i,y}' Z_{i,y} + \mathbf{V}_\lambda^{-1})^{-1} (\sigma_{i,y}^{-2} Z_{i,y}' Z_{i,y} \widehat{\lambda}_i + \mathbf{V}_\lambda^{-1} \overline{\overline{\lambda}}_{i,y}), \tag{A.179}$$

$$\widehat{\lambda}_i = \left(Z_{i,y}' Z_{i,y}\right)^{-1} Z_{i,y}' r_{i,y}^e, \tag{A.180}$$

$$\overline{\overline{\lambda}}_{i,y} = \begin{bmatrix} X_{i,y}\delta_y \\ 1 \end{bmatrix}, \tag{A.181}$$

and

$$\mathbf{V}_\lambda = \begin{bmatrix} \sigma_{\alpha,y}^2 & 0 \\ 0 & 10 \end{bmatrix}. \tag{A.182}$$

2. Draw $\sigma_{i,y}^2 \, | \alpha_{i,y}, \beta_{i,y}$ for each stock $i = 1, ..., N$, in each year $y = 1, ..., Y$. We obtain a draw from the marginal posterior distribution of $\sigma_{i,y}^2$ as follows:

$$\sigma_{i,y}^2 \sim \text{Inverse Gamma}\left(\frac{v_1 s_1^2}{2}, \frac{v_1}{2}\right), \tag{A.183}$$

$$v_1 = v_0 + M, \tag{A.184}$$

and

$$s_1^2 = \frac{v_0 s_0^2 + s^2}{v_0 + M}, \tag{A.185}$$

where s^2 is the sample sum of squared errors and M denotes the number of observations. The priors, v_0 and s_0^2, are determined by the researcher. We set v_0 equal to 3 and s_0^2 equal to the variance of the monthly returns for stock i in year y.

3. Draw $\delta_y | \{\alpha_{i,y}\}, \sigma^2_{\alpha,y}, \overline{\delta}, \mathbf{V}$ for each year $y = 1, ..., Y$. Let α denote a column vector composed of draws of $\alpha_{i,y}$ for all firms i in the dataset in year y. We obtain a draw from the marginal posterior distribution of δ_y as follows:

$$\delta_y \sim N\left(\overline{\overline{\delta}}_y, (\sigma^{-2}_{\alpha,y} X_y' X_y + \mathbf{V}^{-1})^{-1}\right), \tag{A.186}$$

where

$$\overline{\overline{\delta}}_y = (\sigma^{-2}_{\alpha,y} X_y' X_y + \mathbf{V}^{-1})^{-1} (\sigma^{-2}_{\alpha,y} X_y' X_y \widehat{\delta}_y + \mathbf{V}^{-1} \overline{\delta}), \tag{A.187}$$

and

$$\widehat{\delta}_y = \left(X_y' X_y\right)^{-1} X_y' \alpha. \tag{A.188}$$

4. Draw $\sigma^2_{\alpha,y} \mid \{\alpha_{i,y}\}, \delta_y$ for each year $y = 1, ..., Y$. We obtain a draw from the marginal posterior distribution of $\sigma^2_{\alpha,y}$ as follows:

$$\sigma^2_{\alpha,y} \sim \text{Inverse Gamma}\left(\frac{v_1 s_1^2}{2}, \frac{v_1}{2}\right), \tag{A.189}$$

$$v_1 = v_0 + M, \tag{A.190}$$

and

$$s_1^2 = \frac{v_0 s_0^2 + s^2}{v_0 + M}, \tag{A.191}$$

where s^2 is the sample sum of squared errors and M denotes the number of observations. The priors, v_0 and s_0^2, are determined by the researcher. We set v_0 equal to 3. We elicit priors for s_0^2 in the following manner. For each stock in year y we estimate equation (3.5) using OLS and store $\widehat{\alpha}$. We set s_0^2 equal to the variance of $\widehat{\alpha}$ across all firms in year y.

Having drawn the firm- and year-level coefficients we proceed to draw the aggregate-level parameters. Let P denote a $Y \times nvar$ matrix comprised of a draw of $\{\delta_y\}_{y=1}^{Y}$, where $nvar$ denotes the number of columns in X, and let H be a matrix of covariates the researcher believes to be associated with the evolution of the parameter vector δ_y over time. In our specification, H is a column vector of ones, but could easily be extended, for example, to include macroeconomic variables.

5. Draw $\mathbf{V}|\{\delta_y\}$. We obtain a draw from the marginal posterior distribution of $\mathbf{V}$ as follows:

$$\mathbf{V} \sim \text{Inverse Wishart}\left(nvar + Nu + Y, \mathbf{V}_0 + \mathbf{S}\right), \tag{A.192}$$

where

$$\mathbf{S} = \left(P - H\widetilde{\mathbf{\Gamma}}\right)' \left(P - H\widetilde{\mathbf{\Gamma}}\right) + \left(\widetilde{\mathbf{\Gamma}} - \overline{\mathbf{\Gamma}}\right)' \mathbf{A} \left(\widetilde{\mathbf{\Gamma}} - \overline{\mathbf{\Gamma}}\right), \tag{A.193}$$

$$\widetilde{\mathbf{\Gamma}} = \left(\left(H'H + \mathbf{A}\right)^{-1} \left(H'H\widehat{\mathbf{\Gamma}} + \mathbf{A}\overline{\mathbf{\Gamma}}\right)\right), \tag{A.194}$$

and

$$\widehat{\mathbf{\Gamma}} = \left(H'H\right)^{-1} \left(H'P\right). \tag{A.195}$$

$\mathbf{A}$, $\overline{\mathbf{\Gamma}}, Nu$ and $\mathbf{V}_0$ are priors specified by the researcher. We set $\mathbf{A}^{-1} = 100\mathbf{I}$ and define $\overline{\mathbf{\Gamma}}$ to be an $n_H \times nvar$ matrix of zeros, where n_H denotes the number of columns in H. Nu is set to $nvar + 3$, and $\mathbf{V}_0 = Nu\mathbf{I}$. $\mathbf{I}$ denotes an appropriately dimensioned identity matrix.

6. Draw $\gamma | \{\delta_y\}, \mathbf{V}$. We obtain a draw from the marginal posterior distribution of γ as follows:

$$\gamma \sim N\left(\widetilde{\gamma}, \mathbf{V} \otimes \left(H'H + \mathbf{A}\right)^{-1}\right), \tag{A.196}$$

where $\widetilde{\gamma} = vec\left(\widetilde{\mathbf{\Gamma}}\right)$. Given that H is a vector of ones, $\overline{\delta} = \gamma$.

A.5.2 Model Simulation

In this section, we conduct a simulation exercise and show our estimation algorithm successfully recovers the parameters of interest. Data are created for 1,000 firms over a 45-year period. The length of each time period, y, is set to 12 months. We assume there are two firm characteristics associated with firm-level alphas, x_1 and x_2, which are both uniformly distributed over the range -0.5 to +0.5. The parameters in the simulation are set to ensure that the simulated firm-level returns, alphas, betas, and market returns are consistent with the actual values observed using the CRSP return data.

1. Draw $\delta_y \sim \text{MVN}\left(\mu = \overline{\delta}, \sigma^2 = \mathbf{V}\right)$ for each 12-month time period, y. We set

$$\overline{\delta} = \begin{bmatrix} \overline{\delta}_0 = 0 \\ \overline{\delta}_1 = 1 \\ \overline{\delta}_2 = 1 \end{bmatrix}, \text{ and } \mathbf{V} = \begin{bmatrix} 1.5 & 0.5 & 0.5 \\ 0.5 & 1.5 & 0.5 \\ 0.5 & 0.5 & 1.5 \end{bmatrix}.$$

2. Draw $\alpha_y \sim \text{MVN}\left(\mu = \delta_{0,y} + \delta_{1,y}x_1 + \delta_{2,y}x_2, \sigma^2 = \mathbf{\Sigma}_\alpha\right)$, where α_y is a column vectors of firm-specific alphas in time period y. We consider two specifications for the variance-covariance matrix, $\mathbf{\Sigma}_\alpha$, one in which the error terms are independent across firms, and one in which the error terms are correlated across firms. We examine two different levels of correlations, low to medium with correlations ranging from -0.5 to $+0.5$, and medium to high with correlations ranging from -0.9 to $+0.9$. The diagonal elements of $\mathbf{\Sigma}_\alpha$ are set equal to $\sigma_\alpha^2 = 2$.[24]

3. Draw $\beta_{i,y} \sim \text{N}\left(\mu = 1, \sigma^2 = 4\right)$ for each firm i in each time period y.

4. Generate excess monthly returns on the market: $r^e_{m,t,y} \sim \text{N}\left(\mu = 0.5, \sigma^2 = 25\right)$.

5. Generate monthly excess returns for each firm in each month of each time period: $r^e_{t,y} = \alpha_y + \beta_y r^e_{m,t,y} + \epsilon_{t,y}$, where $\epsilon_{t,y} \sim \text{MVN}\left(\mu = 0, \sigma^2 = \mathbf{\Sigma}_{ret}\right)$ and $r^e_{t,y}$ denotes a column vector of excess returns for all firms in month t of time period y. α_y and β_y are column vectors of firm-specific alphas and betas. The specifications for the variance-covariance matrix, $\mathbf{\Sigma}_{ret}$, are constructed in a similar manner to those for $\mathbf{\Sigma}_\alpha$. The only difference is that the diagonal elements of $\mathbf{\Sigma}_{ret}$, σ^2_{ret}, are set equal to 169.

---

[24]We use the following procedure to create a 1,000 $\times$ 1,000 variance-covariance matrix. First, create a column vector, u, with 1,000 draws from the Uniform(-1,1) distribution. Second, calculate $\kappa uu'$ where $\kappa = p\sigma^2_\alpha$. The parameter, p is a scaling factor, between 0 and 1, for the maximum level of correlation in the error terms across firms. If $p = 0$, firm-level alphas are independent. If $p = 1$, $\kappa uu'$ correlations range from -1 to $+1$. For low to medium correlations we set $p = 0.5$, while for medium to high correlations we set $p = 0.9$. Finally, set $\mathbf{\Sigma}_\alpha = \kappa uu'$ and replace the diagonal elements with $\sigma^2_\alpha = 2$.

We examine seven different scenarios to investigate the sensitivity of our model to different correlation structures in the error terms of equations (3.5) and (3.6). The MCMC algorithm is run for 1,000 iterations for each scenario. The algorithm converges quickly. The posterior distributions are characterized using the final 500 iterations. We use the same seed for the random number generator for each scenario. Table A.3 reports the results from the simulation study. Regardless of the correlation structure in the error terms of equations (3.5) and (3.6), the estimation algorithm is able to accurately recover the aggregate-level model parameters, $\overline{\delta}$ and $\mathbf{V}$, indicating that the approach is not sensitive to the possibility of cross-correlations across firms.

A.6 Data Formation

We obtain accounting data from the Compustat Fundamentals Annual files and stock return data from the CRSP monthly return files. Each of the anomaly variables is measured once a year at the end of June in calendar year j. The variables are matched to returns from July of calendar year j to June of calendar year $j+1$. To ensure that the accounting data are known prior to the returns they are used to forecast, we lag all accounting variables by six months. The sample includes all NYSE, Amex, and NASDAQ ordinary common stocks with the data required to compute at least one of the following anomaly variables:

1. $\underline{M \text{ (Size)}}$: The natural log of price per share times the number of shares outstanding at the end of June of year j.

2. $\underline{BM \text{ (Book-to-market)}}$: The natural log of the ratio of book value of equity

to market value of equity. Following Fama and French (2008), we define the book value of equity as total assets (at), minus total liabilities (lt), plus balance sheet deferred taxes and investment tax credits (txditc) if available, minus the book value of preferred stock if available. Depending on availability, we use liquidating value (pstkl), redemption value (pstkrv), or carrying value (upstk) for the the book value of preferred stock. The market value of equity is price per share times the number of shares outstanding at the end of December of year $j-1$.

3. MOM (Momentum): The continuously compounded stock return from January to June of year j. We require a firm to have a price for the end of December of year $j-1$ and a good return for June of year j.

4. REV (Reversal): The continuously compounded stock return from July of year $j-5$ to June of year $j-1$. We require a firm to have a price for the end of June of year $j-5$ and a good return for June of year $j-1$.

5. ROA (Profitability): Income before extraordinary items (ib), minus dividends on preferred (dvp) if available, plus income statement deferred taxes (txdi) if available divided by total assets (at).

6. AG (Asset growth): Total assets (at) at the fiscal year end in year $j-1$, minus total assets at the fiscal year end in year $j-2$ divided by total assets at the fiscal year end in year $j-2$. We also require a firm to have non-zero total assets in both year $j-1$ and $j-2$.

7. NS (Net stock issues): The natural log of the ratio of split-adjusted shares at the fiscal year end in year $j-1$ divided by split-adjusted shares at the fiscal year end in year $j-2$. The number of split-adjusted shares outstanding is common shares outstanding from Compustat (csho) times the cumulative adjustment factor by ex-date (adjex_f).

8. ACC (Accruals): The change in current assets (act) from the fiscal year end in year $j-2$ to $j-1$, minus the change in current liabilities (lct), minus the change in cash and short-term investments (che), plus the change in debt in current liabilities (dlc), minus depreciation (dp) in fiscal year $j-1$ divided by total assets (at) from the fiscal year end in year $j-2$.

9. OS (Financial distress): Ohlson's (1980) O-score:

$$O\text{-score} = \frac{1}{1+\exp(-x)},$$

where

$$\begin{aligned} x = & -1.32 - 0.407\,(SIZE) + 6.03\,(TLTA) - 1.43\,(WCTA) \\ & + 0.076\,(CLCA) - 1.72\,(OENEG) - 2.37\,(NITA) - 1.83\,(FUTL) \\ & + 0.285\,(INTWO) - 0.521\,(CHIN)\,, \end{aligned}$$

where $SIZE$ is the log of the ratio of total assets (at) to the GNP price-level index, $TLTA$ is the ratio of total liabilities (lt) to total assets, $WCTA$ is the ratio of working capital (act – lct) to total assets, $CLCA$ is the ratio of current liabilities (lct) to current assets (act), $OENEG$ is a dummy variable equal to one if total liabilities exceeds total assets and zero otherwise, $NITA$

is the ratio of net income (ni) to total assets, $FUTL$ is the ratio of funds from operations (pi) to total liabilities, $INTWO$ is a dummy variable equal to one if total net income was negative for the past two years and zero otherwise, and $CHIN$ is the change in net income from fiscal year $j-2$ to $j-1$ divided by the sum of the absolute values of net income in fiscal years $j-2$ and $j-1$. Data on the GNP price-level index are from the Federal Reserve Bank of St. Louis website.[25] Following Ohlson (1980), we assign the index a value of 100 in 1968, and the index year is as of the year prior to the year of the balance sheet date.

We exclude financial firms (SIC codes between 6000 and 6999) and firms with negative book equity. The sample period is July 1963 to June 2008. To alleviate the influence of outliers, we winsorize ROA, AG, NS, and ACC at the 1st and 99th percentiles. For cases in which a firm is delisted from an exchange during a given month, we replace any missing returns with the delisting returns provided by CRSP.

[25] http://research.stlouisfed.org/fred2/.

Table A.1: Dividend Share Process Specification Tests

Portfolio	Dickey–Fuller p-value	Ljung–Box p-value
D1	0.90	0.07
D2	0.10	0.96
D3	0.42	0.81
D4	0.00	0.32
D5	0.48	0.18
D6	0.00	0.09
D7	0.30	0.52
D8	0.09	0.14
D9	0.63	0.35
D10	0.01	0.37
I1	0.03	0.22
I2	0.07	0.49
I3	0.01	0.94
I4	0.00	0.27
I5	0.03	0.77
I6	0.00	0.78
I7	0.05	0.36
I8	0.47	0.11
I9	0.01	0.76
I10	0.00	0.99
C1	0.00	0.04
C2	0.05	0.64
C3	0.00	0.05
C4	0.00	0.30
C5	0.01	0.81
C6	0.01	0.48
C7	0.02	0.19
C8	0.36	0.06
C9	0.01	0.63
C10	0.01	0.09

Note: This table reports the results of an Augmented Dickey–Fuller test and a Ljung–Box Q test for each portfolio. Dividend shares are assumed to follow an AR(1) process as implied by our model. The Dickey–Fuller test assumes a constant and a single lag, and tests for the existence of a unit root. The Ljung–Box Q test examines the residuals from the AR(1) process and tests whether these errors are white noise. Portfolios are named using D for dispersion, I for idiosyncratic volatility, and C for credit risk, along with the portfolio decile rank. The table reports p-values for both tests.

Table A.2: Dividend Share Model Comparison

Model	Dispersion	Idiosyncratic Volatility	Default Probability
Panel A: AICM Criterion			
Base Specification	6625.7	7319.9	7198.4
	(0.6)	(1.6)	(1.2)
Alternative 1	6657.8	6362.7	7152.8
	(0.7)	(16.6)	(2.2)
Alternative 2	5239.1	5365.5	5255.2
	(12.9)	(16.0)	(16.8)
Alternative 3	5067.0	5239.0	4887.8
	(15.3)	(17.7)	(22.7)
Panel B: BICM Criterion			
Base Specification	6545.4	7066.6	7020.2
	(1.9)	(5.5)	(4.0)
Alternative 1	6558.6	3659.3	6808.5
	(2.3)	(56.9)	(7.4)
Alternative 2	3139.5	2758.8	2527.1
	(44.2)	(54.9)	(57.4)
Alternative 3	2574.4	2363.5	1198.1
	(52.5)	(60.5)	(77.6)

Note: This table compares the performance of our dividend share process with several alternative specifications. The base specification is given by equations (2.8) and (2.9). Alternatives 1, 2, and 3 are given by equations (A.171), (A.172), and (A.173), respectively. We compare models using AICM and BICM model comparison criteria of Raftery, Newton, Satagopan, and Krivitsky (2007), for which a larger value indicates better model fit. The criteria are based on 50,000 iterations of the MCMC procedure following 10,000 burn-in iterations. Standard errors appear in parentheses.

Table A.3: Model Estimation on Simulated Data

Case	Cross-Correlation (in Equation (1))	Cross-Correlation (in Equation (2))	$\bar{\delta}_0$	$\bar{\delta}_1$	$\bar{\delta}_2$	$\mathbf{V}_{11}$	$\mathbf{V}_{22}$	$\mathbf{V}_{33}$	$\mathbf{V}_{12}$	$\mathbf{V}_{13}$	$\mathbf{V}_{23}$
True Values			0.00	1.00	1.00	1.50	1.50	1.50	0.50	0.50	0.50
			Posterior Means for the Aggregate-Level Parameters, $\bar{\delta}$ and $\mathbf{V}$								
Case 1	None	None	-0.22	0.94	1.07	1.86	1.74	1.67	0.57	0.68	0.75
			(0.20)	(0.20)	(0.21)	(0.39)	(0.38)	(0.40)	(0.29)	(0.31)	(0.28)
Case 2	Low	None	-0.21	0.87	1.06	1.86	1.79	1.65	0.51	0.67	0.78
			(0.20)	(0.20)	(0.21)	(0.39)	(0.39)	(0.40)	(0.29)	(0.31)	(0.29)
Case 3	High	None	-0.20	0.84	1.08	1.87	1.79	1.65	0.49	0.66	0.78
			(0.20)	(0.20)	(0.21)	(0.39)	(0.39)	(0.40)	(0.29)	(0.31)	(0.29)
Case 4	None	Low	-0.22	0.93	1.07	1.86	1.74	1.72	0.56	0.68	0.79
			(0.20)	(0.20)	(0.21)	(0.39)	(0.38)	(0.41)	(0.29)	(0.31)	(0.29)
Case 5	None	High	-0.22	0.93	1.07	1.86	1.74	1.72	0.56	0.67	0.80
			(0.20)	(0.20)	(0.21)	(0.39)	(0.38)	(0.41)	(0.29)	(0.31)	(0.29)
Case 6	Low	Low	-0.21	0.86	1.06	1.86	1.82	1.70	0.51	0.66	0.82
			(0.20)	(0.20)	(0.21)	(0.39)	(0.40)	(0.41)	(0.30)	(0.31)	(0.30)
Case 7	High	High	-0.20	0.83	1.08	1.88	1.82	1.69	0.49	0.64	0.84
			(0.20)	(0.20)	(0.21)	(0.39)	(0.40)	(0.41)	(0.30)	(0.31)	(0.30)

Note: The table presents the results from the estimation of the model described in equations (3.5) to (3.7) for simulated data. We report the posterior mean and standard deviation for the aggregate-level parameters $\bar{\delta}$ and $\mathbf{V}$. We simulate data for 1,000 firms over a 45-year period. We create seven different sets of data using the same seed for the random number generator in each scenario. Each set of data differs only with respect to the assumptions about cross-correlations in the error terms of equation (3.5) (monthly firm returns) and/or equation (3.6) (firm-year alphas). Specifically, we allow cross-correlations in each equation to take on one of three levels: zero, low ($\mp$0.5), or high ($\mp$0.9). We run the Gibbs sampler for 1,000 iterations and discard the first 500 as a burn-in period. An * (**) indicates that the 95% (99%) credible interval of the posterior distribution does not include the true value.

REFERENCES

Abel, A. (1999): "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics*, 43, 3–33.

Ahn, D.-H., J. Conrad, and R. F. Dittmar (2009): "Basis assets," *Review of Financial Studies*, 22, 5133–5174.

Ai, H., and D. Kiku (2010): "Growth to value: Option exercise and the cross section of equity returns," Working paper, Duke University.

Akaike, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19, 716–723.

Altman, E. I. (1968): "Finanical ratios, discriminant analysis and the prediction of corporate bankruptcy," *Journal of Finance*, 23, 589–609.

Ang, A., and G. Bekaert (2007): "Stock return predictability: Is it there?," *Review of Financial Studies*, 20, 651–707.

Ang, A., and J. Chen (2007): "CAPM over the long run: 1926–2001," *Journal of Empirical Finance*, 14, 1–40.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The cross-section of volatility and expected returns," *Journal of Finance*, 61, 259–299.

Ang, A., J. Liu, and K. Schwarz (2010a): "Using stocks or portfolios in tests of factor models," Working paper, Columbia University.

——— (2010b): "Using stocks or portfolios in tests of factor models," Working paper, Columbia University.

Attanasio, O. P., and G. Weber (1989): "Intertemporal substitution, risk aversion and the Euler equation for consumption," *Economic Journal*, 99, 59–73.

Avramov, D., S. Cederburg, and S. Hore (2009): "Cross-sectional asset pricing puzzles: An equilibrium perspective," Working paper, University of Maryland.

Avramov, D., and T. Chordia (2006): "Asset pricing models and financial market anomalies," *Review of Financial Studies*, 19, 1001–1040.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov (2009): "Dispersion in analysts' earnings forecasts and credit rating," *Journal of Financial Economics*, 91, 83–101.

Avramov, D., and S. Hore (2008): "Momentum, information uncertainty, and leverage – An explanation based on recursive preferences," Working paper, University of Maryland.

Bansal, R., R. Dittmar, and D. Kiku (2009): "Cointegration and consumption risks in asset returns," *Review of Financial Studies*, 22, 1343–1375.

Bansal, R., R. F. Dittmar, and C. T. Lundblad (2005): "Consumption, dividends, and the cross section of equity returns," *Journal of Finance*, 60, 1639–1672.

Bansal, R., D. Kiku, and A. Yaron (2007): "Risks for the long run: Estimation and inference," Working paper, Duke University.

——— (2009): "An empirical evaluation of the long-run risks model for asset prices," Working paper, Duke University.

Bansal, R., and A. Yaron (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *Journal of Finance*, 59, 1481–1509.

Banz, R. W. (1981): "The relationship between return and market value of common stocks," *Journal of Financial Economics*, 9, 3–18.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens (1998): "Using analysts' forecasts to measure properties of analysts' information environment," *The Accounting Review*, 73, 421–433.

Barron, O. E., M. H. Stanford, and Y. Yu (2009): "Further evidence on the relation between analysts' forecast dispersion and stock returns," *Contemporary Accounting Research*, 26, 329–357.

Berk, J. B. (1995): "A critique of size-related anomalies," *Review of Financial Studies*, 8, 275–286.

Black, F. (1993): "Estimating Expected Return," *Financial Analysts Journal*, 49, 36–38.

Black, F., M. Jensen, and M. Scholes (1972): "The Capital Asset Pricing Model: Some empirical tests," in *Studies in the Theory of Capital Markets*, ed. by M. C. Jensen. New York: Praeger.

Brennan, M. J., T. Chordia, and A. Subrahmanyam (1998): "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns," *Journal of Financial Economics*, 49, 345–373.

Brennan, M. J., A. W. Wang, and Y. Xia (2001): "Intertemporal capital asset pricing and the Fama–French three-factor model," Working Paper, UCLA.

——— (2004): "Estimation and test of a simple model of intertemporal capital asset pricing," *Journal of Finance*, 59, 1743–1775.

Campbell, J. Y. (1987): "Stock returns and the term structure," *Journal of Financial Economics*, 18, 373–399.

——— (1991): “A variance decomposition for stock returns,” *The Economic Journal*, 101, 157–179.

——— (1993): “Intertemporal asset pricing without consumption data,” *American Economic Review*, 83, 487–512.

——— (1996): “Understanding risk and return,” *Journal of Political Economy*, 104, 298–345.

Campbell, J. Y., and J. Ammer (1993): “What moves the stock and bond markets? A variance decomposition for long-term asset returns,” *Journal of Finance*, 48, 3–37.

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

——— (2000): “Explaining the poor performance of consumption-based asset pricing models,” *Journal of Finance*, 55, 2863–2878.

Campbell, J. Y., J. Hilscher, and J. Szilagyi (2008): “In search of distress risk,” *Journal of Finance*, 63, 2899–2939.

Campbell, J. Y., and H. K. Koo (1997): “A comparison of numerical and analytic approximate solutions to an intertemporal consumption choice problem,” *Journal of Economic Dynamics and Control*, 21, 273–295.

Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001): “Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk,” *Journal of Finance*, 56, 1–43.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1996): *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Campbell, J. Y., and G. Mankiw (1989): “Consumption, income and interest rates: Reinterpreting the time series evidence,” *NBER Macroeconomics Annual*, 4, 185–216.

Campbell, J. Y., and R. J. Shiller (1988): “The dividend-price ratio and expectations of future dividends and discount factors,” *Review of Financial Studies*, 1, 195–228.

Campbell, J. Y., and G. B. Taksler (2003): “Equity volatility and corporate bond yields,” *Journal of Finance*, 58, 2321–2349.

Carter, C. K., and R. Kohn (1994): “On Gibbs sampling for state space models,” *Biometrika*, 81, 541–553.

Cederburg, S., P. Davies, and M. O’Doherty (2010): “Do multifactor models explain asset-pricing anomalies?,” Working paper, University of Iowa.

Chacko, G., and L. M. Viceira (2005): "Dynamic consupmtion and portfolio choice with stochastic volatility in incomplete markets," *Review of Financial Studies*, 18, 1369–1402.

Chan, L. K., Y. Hamao, and J. Lakonishok (1991): "Fundamentals and stock returns in Japan," *Journal of Finance*, 46, 1739–1764.

Chava, S., and A. K. Purnanandam (2010): "Is default risk negatively related to stock returns?," *Review of Financial Studies*, 23, 2523–2559.

Chen, H. (2009): "Macroeconomic conditions and the puzzles of credit spreads and capital structure," Forthcoming in *Journal of Finance*.

Chen, J. (2003): "Intertemporal CAPM and the cross-section of stock returns," Working paper, University of Southern California.

Chen, N.-F., R. Roll, and S. A. Ross (1986): "Economic forces and the stock market," *Journal of Business*, 59, 383–403.

Chopra, N., J. Lakonishok, and J. R. Ritter (1992): "Measuring abnormal performance: Do stocks overreact?," *Journal of Financial Economics*, 31, 235–268.

Cochrane, J. H. (1992): "Explaining the variance of price-dividend ratios," *Review of Financial Studies*, 5, 243–280.

——— (1996): "A cross-sectional test of an investment-based asset pricing model," *Journal of Political Economy*, 104, 572–621.

——— (2005): *Asset Pricing*. Princeton University Press, New Jersey.

——— (2008a): "The dog that did not bark: A defense of return predictability," *Review of Financial Studies*, 21, 1533–1575.

——— (2008b): "Financial Markets and the Real Economy," in *Handbook of the Equity Risk Premium*, ed. by R. Mehra. Elsevier: Amsterdam.

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho (2002): "Who underreacts to cash-flow news? Evidence from trading between individuals and institutions," *Journal of Financial Economics*, 66, 409–462.

Collins, D. W., and P. Hribar (2000): "Earnings-based and accrual-based market anomalies: One effect or two?," *Journal of Accounting and Economics*, 29, 101–123.

Conrad, J., M. Cooper, and G. Kaul (2003): "Value versus glamour," *Journal of Finance*, 58, 1969–1996.

Constantinides, G. M., and A. Ghosh (2008): "Asset pricing tests with long run risks in consumption growth," Working paper, University of Chicago.

Cooper, M. J., H. Gulen, and M. J. Schill (2008): "Asset growth and the cross-section of stock returns," *Journal of Finance*, 63, 1609–1651.

Cosemans, M., R. Frehen, P. C. Schotman, and R. Bauer (2009): "Efficient estimation of firm-specific betas and its benefits for asset pricing tests and portfolio choice," Working Paper, Maastricht University.

Cremers, K. J. M. (2006): "Multifactor efficiency and Bayesian inference," *Journal of Business*, 79, 2951–2998.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2009): "Investor information, long-run risk, and the duration of risky cash flows," Working Paper, UNC Chapel Hill.

Crow, J. F., and M. Kimura (1970): *An introduction to population genetics theory.* Burgess Publishing Co., Minneapolis, MN.

Da, Z. (2009): "Cash flow, consumption risk, and the cross-section of stock returns," *Journal of Finance*, 64, 923–956.

Da, Z., and M. C. Warachka (2009): "Cashflow risk, systematic earnings revisions, and the cross-section of stock returns," *Journal of Financial Economics*, 94, 448–468.

Daniel, K., and S. Titman (2006a): "Market reactions to tangible and intangible information," *Journal of Finance*, 61, 1605–1643.

——— (2006b): "Testing factor-model explanations of market anomalies," Working Paper, Northwestern University.

Davies, P. (2010): "A cross-sectional test of the CAPM at the firm level," Working paper, Rutgers University.

DeBondt, W. F. M., and R. Thaler (1985): "Does the stock market overreact?," *Journal of Finance*, 40, 793–805.

Dichev, I. D. (1998): "Is the risk of bankruptcy a systematic risk?," *Journal of Finance*, 53, 1131–1147.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002): "Differences of opinion and the cross section of stock returns," *Journal of Finance*, 57, 2113–2141.

Dimson, E. (1979): "Risk measurement when shares are subject to infrequent trading," *Journal of Financial Economics*, 7, 197–226.

Duffie, D., and L. G. Epstein (1992): "Asset pricing with stochastic differential utility," *Review of Financial Studies*, 5, 411–436.

Duffie, D., L. Saita, and K. Wang (2007): "Multi-period corporate default prediction with stochastic covariates," *Journal of Financial Economics*, 83, 635–665.

Epstein, L. G., and S. E. Zin (1989): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 57, 937–969.

——— (1991): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis," *Journal of Political Economy*, 99, 263–286.

Fairfield, P. M., J. S. Whisenant, and T. L. Yohn (2003): "Accrued earnings and growth: Implications for future profitability and market mispricing," *The Accounting Review*, 78, 353–371.

Fama, E. F. (1970): "Multiperiod consumption-investment decisions," *American Economic Review*, 60, 163–174.

——— (1990): "Stock returns, expected returns, and real activity," *Journal of Finance*, 45, 1089–1108.

Fama, E. F., and K. R. French (1988): "Dividend yields and expected stock returns," *Journal of Financial Economics*, 22, 3–25.

——— (1989): "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25, 23–49.

——— (1992): "The cross-section of expected stock returns," *Journal of Finance*, 47, 427–465.

——— (1993): "Common risk factors in the returns on bonds and stocks," *Journal of Financial Economics*, 33, 3–53.

——— (1995): "Size and book-to-market factors in earnings and returns," *Journal of Finance*, 50, 131–155.

——— (1996): "Multifactor explanations of asset pricing anomalies," *Journal of Finance*, 51, 55–84.

——— (2002): "The equity premium," *Journal of Finance*, 57, 637–659.

——— (2006): "The value premium and the CAPM," *Journal of Finance*, 61, 2163–2185.

——— (2008): "Dissecting anomalies," *Journal of Finance*, 63, 1653–1678.

Fama, E. F., and J. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

Fama, E. F., and G. W. Schwert (1977): "Asset returns and inflation," *Journal of Financial Economics*, 5, 115–146.

Ferson, W. E., and S. R. Foerster (1994): "Finite sample properties of the generalized method of moments in tests of conditional asset pricing models," *Journal of Financial Economics*, 36, 29–56.

Ferson, W. E., and C. R. Harvey (1999): "Conditioning variables and the cross section of stock returns," *Journal of Finance*, 54, 1325–1360.

Ferson, W. E., S. Sarkissian, and T. Simin (1999): "The alpha factor asset pricing model: A parable," *Journal of Financial Markets*, 2, 49–68.

Fisher, R. (1930): *The genetical theory of natural selection*. Clarendon Press, Oxford.

Frühwirth-Schnatter, S. (1994): "Data augmentation and dynamic linear models," *Journal of Time Series Analysis*, 15, 183–202.

Garlappi, L., T. Shu, and H. Yan (2008): "Default risk, shareholder advantage, and stock returns," *Review of Financial Studies*, 21, 2743–2778.

Gelman, A., and I. Pardoe (2006): "Bayesian measures of explained variance and pooling in multilevel (hierarchical) models," *Technometrics*, 48, 241–251.

George, T. J., and C.-Y. Hwang (2010): "A resolution of the distress risk and leverage puzzles in the cross section of stock returns," *Journal of Financial Economics*, 96, 56–79.

Geweke, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. Wiley, Hoboken, NJ.

Geweke, J., and G. Zhou (1996): "Measuring the pricing error of the arbitrage pricing theory," *Review of Financial Studies*, 9, 557–587.

Gomes, J., L. Kogan, and L. Zhang (2003): "Equilibrium cross section of returns," *Journal of Political Economy*, 111, 693–732.

Guvenen, F. (2006): "Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective," *Journal of Monetary Economics*, 53, 1451–1472.

Hahn, J., and H. Lee (2006): "Yield spreads as alternative risk factors for size and book-to-market," *Journal of Financial and Quantitative Analysis*, 41, 245–269.

Hall, R. E. (1988): "Intertemporal substitution in consumption," *Journal of Political Economy*, 96, 339–357.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption strikes back? Measuring long-run risk," *Journal of Political Economy*, 116, 260–302.

Hansen, L. P., and R. Jagannathan (1991): "Implications of security market data for models of dynamic economies," *Journal of Political Economy*, 91, 225–262.

Hansen, L. P., and T. J. Sargent (2010): "Fragile beliefs and the price of uncertainty," *Quantitative Economics*, 1, 129–164.

Hansen, L. P., and K. J. Singleton (1982): "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, 50, 1269–1286.

Harvey, C. R. (1989): "Time-varying conditional covariances in tests of asset pricing models," *Journal of Financial Economics*, 25, 289–317.

Harvey, C. R., and G. Zhou (1990): "Bayesian inference in asset pricing tests," *Journal of Financial Economics*, 26, 221–254.

Hastings, W. K. (1970): "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, 57, 97–109.

Haugen, R. A., and N. L. Baker (1996): "Commonality in the determinants of expected stock returns," *Journal of Financial Economics*, 41, 401–439.

Hodrick, R. J., D. T.-C. Ng, and P. Sengmueller (1999): "An international dynamic asset pricing model," *International Tax and Public Finance*, 6, 597–620.

Hore, S. (2008): "Equilibrium predictability and other asset pricing characteristics," Working paper, University of Iowa.

Huang, A. G. (2009): "The cross section of cashflow volatility and expected stock returns," *Journal of Empirical Finance*, 16, 409–429.

Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995): "Market underreaction to open market share repurchases," *Journal of Financial Economics*, 39, 181–208.

Jagannathan, R., and Z. Wang (1996): "The conditional CAPM and the cross-section of expected returns," *Journal of Finance*, 51, 3–53.

——— (1998): "An asymptotic theory for estimating beta-pricing models using cross-sectional regression," *Journal of Finance*, 53, 1285–1309.

Jegadeesh, N. (1990): "Evidence of predictable behavior of security returns," *Journal of Finance*, 45, 881–898.

Jegadeesh, N., and S. Titman (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency," *Journal of Finance*, 48, 65–91.

——— (2001): "Profitability of momentum strategies: An evaluation of alternative explanations," *Journal of Finance*, 56, 699–720.

Jensen, M. C. (1978): "Some anomalous evidence regarding market efficiency," *Journal of Financial Economics*, 6, 95–102.

Johnson, T. C. (2004): "Forecast dispersion and the cross section of expected returns," *Journal of Finance*, 59, 1957–1978.

Kaltenbrunner, G., and L. Lochstoer (2008): "Long-run risk through consumption smoothing," Forthcoming in *Review of Financial Studies*.

Kan, R. (2004): "On the explanatory power of asset pricing models across and within portfolios," Working paper, University of Toronto.

Kandel, S., R. McCulloch, and R. F. Stambaugh (1995): "Bayesian inference and portfolio efficiency," *Review of Financial Studies*, 8, 1–53.

Kandel, S., and R. F. Stambaugh (1991): "Asset returns and intertemporal preferences," *Journal of Monetary Economics*, 27, 39–71.

——— (1995): "Portfolio inefficiency and the cross section of expected returns," *Journal of Finance*, 50, 157–184.

Karlin, S., and H. M. Taylor (1981): *A Second Course in Stochastic Processes*. Academic Press, New York.

Karolyi, G. A. (1992): "Predicting risk: Some new generalizations," *Management Science*, 38, 57–74.

Keim, D. B., and R. F. Stambaugh (1986): "Predicting returns in the stock and bond markets," *Journal of Financial Economics*, 17, 357–390.

Kiku, D. (2006): "Is the value premium a puzzle?," Working paper, University of Pennsylvania.

Kothari, S., and J. Shanken (1997): "Book-to-market, dividend yield, and expected market returns: A time series analysis," *Journal of Financial Economics*, 44, 169–203.

Kreps, D. M., and E. L. Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 46, 185–200.

LeRoy, S. F., and R. D. Porter (1981): "The present-value relation: Tests based on implied variance bounds," *Econometrica*, 49, 555–574.

Lettau, M., and S. Ludvigson (2001a): "Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying," *Journal of Political Economy*, 109, 1238–1287.

Lettau, M., and S. C. Ludvigson (2001b): "Consumption, aggregate wealth, and expected stock returns," *Journal of Finance*, 56, 815–849.

Lettau, M., and J. A. Wachter (2007): "Why is long-horizon equity less risky? A duration-based explanation of the value premium," *Journal of Finance*, 62, 55–92.

Lewellen, J., and S. Nagel (2006): "The conditional CAPM does not explain asset-pricing anomalies," *Journal of Financial Economics*, 82, 289–314.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A skeptical appraisal of asset-pricing tests," *Journal of Financial Economics*, 96, 175–194.

Li, Q., M. Vassalou, and Y. Xing (2006): "Sector investment growth rates and the cross section of equity returns," *Journal of Business*, 79, 1637–1665.

Liew, J., and M. Vassalou (2000): "Can book-to-market, size and momentum be risk factors that predict economic growth?," *Journal of Financial Economics*, 57, 221–245.

Lintner, J. (1965): "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Statistics*, 47, 13–37.

Litzenberger, R. H., and K. Ramaswamy (1979): "The effect of personal taxes and dividends on capital asset prices," *Journal of Financial Economics*, 7, 163–196.

Lo, A. W., and A. C. MacKinlay (1990): "Data-snooping biases in tests of financial asset pricing models," *Review of Financial Studies*, 3, 431–468.

Loughran, T. (1997): "Book-to-market across firm size, exchange, and seasonality," *Journal of Financial and Quantitative Analysis*, 32, 249–268.

Loughran, T., and J. R. Ritter (1995): "The new issues puzzle," *Journal of Finance*, 50, 23–51.

Lustig, H. N., and S. G. Van Nieuwerburgh (2005): "Housing collateral, consumption insurance, and risk premia: An empirial perspective," *Journal of Finance*, 60, 1167–1219.

Mayers, D. (1972): "Nonmarketable assets and capital market equilibrium under uncertainty," in *Studies in the Theory of Capital Markets*, ed. by M. C. Jensen. New York: Praeger.

McCulloch, R., and P. E. Rossi (1991): "A Bayesian approach to testing the arbitrage pricing theory," *Journal of Econometrics*, 49, 141–168.

Mehra, R., and E. Prescott (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145–161.

Menzly, L., T. Santos, and P. Veronesi (2004): "Understanding predictability," *Journal of Political Economy*, 112, 1–47.

Merton, R. C. (1969): "Lifetime portfolio selection under uncertainty: The continuous-time case," *Review of Economics and Statistics*, 51, 247–257.

——— (1971): "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

——— (1973): "An intertemporal capital asset pricing model," *Econometrica*, 41, 867–887.

——— (1974): "On the pricing of corporate debt: The risk structure of interest rates," *Journal of Finance*, 29, 449–470.

——— (1987): "A simple model of capital market equilibrium with incomplete information," *Journal of Finance*, 42, 483–510.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953): "Equation of state calculations by fast computing machines," *Journal of Chemical Physics*, 21, 1087–1092.

Mitrinović, D. S., J. E. Pečarić, and A. M. Fink (1993): *Classical and new inequalities in analysis*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Ohlson, J. A. (1980): "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, 18, 109–131.

Pástor, Ľ., M. Sinha, and B. Swaminathan (2008): "Estimating the intertemporal risk-return tradeoff using the implied cost of capital," *Journal of Finance*, 63, 2859–2897.

Pástor, Ľ., and R. F. Stambaugh (2009): "Predictive systems: Living with imperfect predictors," *Journal of Finance*, 64, 1583–1628.

Petkova, R. (2006): "Do the Fama–French factors proxy for innovations in predictive variables?," *Journal of Finance*, 61, 581–612.

Pontiff, J., and A. Woodgate (2008): "Share issuance and cross-sectional returns," *Journal of Finance*, 63, 921–945.

Raftery, A. E., M. A. Newton, J. M. Satagopan, and P. N. Krivitsky (2007): "Estimating the integrated likelihood via posterior simulation using the harmonic mean identity," *Bayesian Statistics*, 8, 1–45.

Roll, R. (1977): "A critique of the asset pricing theory's tests – Part I: On past and potential testability of the theory," *Journal of Financial Economics*, 4, 129–176.

Rosenberg, B., K. Reid, and R. Lanstein (1985): "Persuasive evidence of market inefficiency," *Journal of Portfolio Management*, 11, 9–17.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005a): *Bayesian Statistics and Marketing*. John Wiley & Sons, West Sussex, UK.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005b): *Bayesian Statistics and Marketing*. Wiley, Hoboken, NJ.

Santos, T., and P. Veronesi (2006): "Labor income and predictable stock returns," *Review of Financial Studies*, 19, 1–44.

Schwarz, G. (1978): “Estimating the dimension of a model,” *Annals of Statistics*, 6, 416–464.

Schwert, G. W. (2003): “Anomalies and market efficiency,” in *Handbook of the Economics of Finance*, ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. North Holland.

Shanken, J. (1987): “A Bayesian approach to testing portfolio efficiency,” *Journal of Financial Economics*, 19, 195–215.

——— (1992): “On the estimation of beta-pricing models,” *Review of Financial Studies*, 5, 1–33.

Shanken, J., and G. Zhou (2007): “Estimating and testing beta pricing models: Alternative methods and their performance in simulations,” *Journal of Financial Economics*, 84, 40–86.

Shanken, J. A. (1990): “Intertemporal asset pricing: An empirical investigation,” *Journal of Econometrics*, 45, 99–120.

Sharpe, W. F. (1964): “Capital asset prices: A theory of market equilibrium under conditions of risk,” *Journal of Finance*, 19, 425–442.

Shiller, R. J. (1981): “Do stock prices move too much to be justified by subsequent changes in dividends?,” *American Economic Review*, 71, 421–436.

Sloan, R. G. (1996): “Do stock prices fully reflect information in accruals and cash flows about future earnings?,” *The Accounting Review*, 71, 289–315.

Sohn, B. (2009): “Cross-section of equity returns: Stock market volatility and priced factors,” Working paper, University of North Carolina.

Titman, S., K. C. J. Wei, and F. Xie (2004): “Capital investments and stock returns,” *Journal of Financial and Quantitative Analysis*, 39, 677–700.

van Binsbergen, J. H., and R. S. J. Koijen (2010): “Predictive regressions: A present-value approach,” *Journal of Finance*, 65, 1439–1471.

Vasicek, O. A. (1973): “A note on using cross-sectional information in Bayesian estimation of security betas,” *Journal of Finance*, 28, 1233–1239.

Vassalou, M. (2003): “News related to future GDP growth as a risk factor in equity returns,” *Journal of Financial Economics*, 68, 47–73.

Vissing-Jorgensen, A. (2002): “Limited asset market participation and the elasticity of intertemporal substitution,” *Journal of Political Economy*, 110, 825–853.

Wachter, J. A. (2006): “A consumption-based model of the term structure of interest rates,” *Journal of Financial Economics*, 79, 365–399.

Weil, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401–422.

West, M., and J. Harrison (1997): *Bayesian Forecasting and Dynamic Models.* Springer-Verlag, New York.

Wright, S. (1931): "Evolution in Mendelian populations," *Genetics*, 16, 97–159.

Xie, H. (2001): "The mispricing of abnormal accruals," *The Accounting Review*, 76, 357–373.

Yang, W. (2009): "Asset pricing with left-skewed long-run risk in durable consumption," Working paper, University of Rochester.

Yogo, M. (2004): "Estimating the elasticity of intertemporal substitution when instruments are weak," *Review of Economics and Statistics*, 86, 797–810.

Zhang, X. F. (2006): "Information uncertainty and stock returns," *Journal of Finance*, 61, 105–137.

Zurek, P. (2007): "Momentum and long-run risks," Working paper, University of Pennsylvania.