On the Return Decomposition and the Intertemporal CAPM

by Yuan-Szu Chang

B.S. in Quantitative Finance, June 2001, National Tsing Hua University

M.S. in Finance, August 2004, George Washington University

A Dissertation submitted to

The Faculty of
the School of Business
of The George Washington University
in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy

August 31, 2012

Dissertation directed by

Robert Savickas
Associate Professor of Finance

The School of Business of the George Washington University certifies that Yuan-Szu Chang has passed the Final Examination for the degree of Doctor of Philosophy as of May 15, 2012. This is the final and approved form of the dissertation.

On the Return Decomposition and the Intertemporal CAPM

Yuan-Szu Chang

Dissertation Research Committee:

Robert Savickas, Associate Professor of Finance, Advocate

Refik Soyer, Professor of Decision Sciences, Research Advisory Committee Member

Alexandre M. Baptista, Associate Professor of Finance, Research Advisory Committee Member

© Copyright 2012 by Yuan-Szu Chang
All rights reserved.

Acknowledgments

I am grateful to my advisor Dr. Robert Savickas for his guidance and support throughout the completion of my dissertation. I thank Dr. Refik Soyer and Dr. Alexandre Baptista for their valuable suggestions and comments. I also thank Dr. Min Hwang, Dr. Mark Klock, and Dr. Gergana Jostova for their valuable comments. I am thankful to Dr. Arthur Wilson for presiding at my dissertation defense. I also thank Ms. Elizabeth Huff of the PhD. Program in the School of Business for her help. Last but not least, I am grateful to my family members for their understanding, patience, and support.

Abstract of Dissertation

On the Return Decomposition and the Intertemporal CAPM

In this dissertation, I test the implications of Campbell and Vuolteenaho's (2004) intertemporal capital asset pricing model (ICAPM) based on the return decomposition of Campbell (1991). The return decomposition states that the unexpected market return is composed of cash-flows news and discount-rate news. The previous literature uses VAR models to implement the return decomposition. My dissertation adds to the literature by proposing a state-space model approach. Importantly, the proposed approach has two advantages over the VAR approach. First, the state-space approach directly and simultaneously models both cash-flow news and discount-rate news by expressing the relation between returns and news in terms of linear regressions, whereas the VAR approach models discount-rate news and backs out cash-flow news as the residual between the unexpected market return and discount-rate news. Second, the information that I use to extract cash-flow news and discount-rate news is directly related to the market return, whereas the information used in the VAR approach is not. A Bayesian Markov chain Monte Carlo methodology is developed and used to estimate the state-space model. Based on the state-space model and the data set used in this dissertation, I am not able to explain the value premium because I do not find that value stocks have higher cash-flow betas than do growth stocks. The reason is likely that the cash-flow news and discount-rate news estimated in this dissertation are strongly negatively correlated. With this state-space model, I am now in a position to systematically examine the implications of Campbell and Vuolteenaho's (2004) ICAPM using other data sets and other time periods in future studies.

Table of Contents

Acknowledgments……..iv

Abstract of Dissertation…….....v

List of Figures……..……viii

List of Tables……..ix

Chapter 1: Introduction……...1

Chapter 2: Literature Review……14

2.1 CAPM and some empirical evidence……14

2.2 The intertemporal capital asset pricing model (ICAPM)……..17

2.3 Conditional capital asset pricing model……26

2.4 State-space model and some applications…….27

Chapter 3: Methodology……...31

3.1 The ICAPM framework……..31

3.1.1 Return decomposition……31

3.1.2 Cash-flow and discount-rate betas…….33

3.1.3 An intertemporal asset pricing formula…….35

3.2 State-space model approach to ICAPM……..37

3.2.1 A classical state-space model…….37

3.2.2 State-space model for ICAPM……...40

3.2.3 Bayesian Markov chain Monte Carlo methodology……..44

3.2.4 Simulation study……66

3.2.5 Data description…….70

Chapter 4: Results……..73

4.1 Summary of cash-flow news and discount-rate news……..76

4.2 Cash-flow betas and discount-rate betas……..79

4.3 Cash-flow and discount-rate betas calculated from orthogonalized cash-flow and discount-rate news……..82

4.4 Correlation between news estimated using the state-space model and Campbell and Vuolteenaho's (2004) estimated news…….87

4.5 Correlation between returns and betas…….89

4.6 Sensitivity analysis……...92

4.7 Model checking: Posterior predictive check……96

Chapter 5: Conclusions……..99

References……103

Tables and figures (Tables 1 to 12 and Figures 1 to 3)……110

Appendix A……..137

Appendix B: Additional Tables (Tables 13 to 20)……...138

Appendix C: Preliminary analysis using VAR approach in the early stage of the research……148

List of Figures

Figure 1: Smoothed cash-flow news and discount-rate news for the period 1952:7-2010:12……126

Figure 2: Density plots of cash-flow betas for the 25 Fama-French portfolios during 1963:7-2001:12……127

Figure 3: Density plots of discount-rate betas for the 25 Fama-French portfolios during 1963:7-2001:12……132

List of Tables

Table 1: Posterior means and variances for parameters in the state-space model used to analyze SIMDATA1 (sample size = 500) and SIMDATA2 (sample size = 50) in the simulation study……..... 111

Table 2: Descriptive statistics of the observed variables in the state-space model…. 112

Table 3: Summary of cash-flow news and discount-rate news……... 113

Table 4: Cash-flow betas and discount-rate betas for the 1963:7-2001:12 period…. 115

Table 5: Difference in cash-flow and discount-rate betas between value and growth stocks for 1963:7-2001:12…….. 117

Table 6: Difference in cash-flow and discount-rate betas between large and small stocks for 1963:7-2001:12…….. 118

Table 7: Posterior means and standard deviations of cash-flow and discount-rate betas for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news……... 119

Table 8: Posterior medians and 95% intervals of cash-flow and discount-rate betas for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news..... 120

Table 9: Difference in cash-flow and discount-rate betas between value and growth stocks for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news……... 121

Table 10: Difference in cash-flow and discount-rate betas between large and small stocks for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news……... 122

Table 11: Descriptive statistics for the distribution of correlation between the estimated state variables in this study and Campbell and Vuolteenaho's (2004) estimated conditional expectation of the excess log market return, cash-flow news, and discount-rate news for 1952:7-2001:12…… 123

Table 12: Correlation between average returns and cash-flow betas when there is a positive or negative cash-flow shock and correlations between average returns and discount-rate betas when there is a positive or negative discount-rate shock for 1963:7-2001:12……... 125

Table 13: Descriptive statistics for the posterior distributions of parameters in the state-space model…… 139

Table 14: Cash-flow betas and discount-rate betas for the 1952:7-2010:12 period... 140

Table 15: Difference in cash-flow and discount-rate betas between value and growth stocks for 1952:7-2010:12…….. 142

Table 16: Difference in cash-flow and discount-rate betas between large and small stocks for 1952:7-2010:12…….. 143

Table 17: Posterior means and standard deviations of cash-flow and discount-rate betas for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news……... 144

Table 18: Posterior medians and 95% intervals of cash-flow and discount-rate betas for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news……... 145

Table 19: Difference in cash-flow and discount-rate betas between value and growth stocks for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news……..... 146

Table 20: Difference in cash-flow and discount-rate betas between large and small stocks for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news……... 147

Chapter 1: Introduction

In this dissertation, I test the implications of Campbell and Vuolteenaho's (2004) intertemporal capital asset pricing model (ICAPM) based on the return decomposition of Campbell (1991). The return decomposition states that the unexpected market return is composed of news about future cash flows and news about future discount rates. Previous studies use vector autoregressive (VAR) models with ad-hoc variables to implement the return decomposition but yield conflicting results. To avoid the ad-hoc nature of the empirical specification, I use a state-space model with restrictions derived from the return decomposition instead of a VAR model. My study attempts to resolve the ambiguity in the literature.

The return decomposition of Campbell (1991) is based on the framework of Campbell and Shiller's (1988) dividend-ratio model, which approximates the log dividend-price ratio by the discounted value of all future returns, dividend growth rates, and a constant term. The return decomposition states that the difference between the realized market return and the conditional expectation of the market return (i.e., unexpected market return) consists of two components—change in expectations of all discounted future dividend growth rates and change in expectations of all discounted future market returns. The former component of the unexpected market return is referred to as news about future cash flows, whereas the latter is referred to as news about future discount rates.

Campbell and Vuolteenaho (2004) interpret news about future cash flows and news about future discount rates as permanent and transitory shocks to wealth,

respectively. A decline in the value of the market portfolio could be the result of either a negative shock to future cash flows or a positive shock to future discount rates. A negative shock to future cash flows results in declining wealth and no changes to future investment opportunities because future discount rates are unchanged. However, when there is a positive shock to future discount rates, wealth declines because investors use higher discount rates to discount cash flows, but future investment opportunities improve because of higher expected returns in the future. Cash-flow news is interpreted as a permanent shock because future investment opportunities do not change as a result of the news, and discount-rate news is interpreted as a transitory shock because future investment opportunities change so that declining wealth today is offset by higher expected returns in the future. It is from this perspective that Campbell and Vuolteenaho (2004) believe that conservative investors are more averse to risk brought about by cash-flow news than to risk brought about by discount-rate news. Therefore, intuitively, investors demand a greater compensation for bearing cash-flow risk than for bearing discount-rate risk.

Using the return decomposition, Campbell and Vuolteenaho (2004) rewrite Campbell's (1993) asset pricing formula to obtain a two-beta ICAPM, which states that the risk premium on an asset is expressed as a linear combination of the asset's cash-flow beta and discount-rate beta. They define cash-flow beta as the covariance between the asset's log return and news about future cash flows and discount-rate beta as the covariance between the asset's log return and the negative of news about future discount rates. Both covariances are divided by the variance of the unexpected market return.

The two-beta ICAPM delivers the intuition that conservative investors price cash-flow risk higher than discount-rate risk. In the two-beta ICAPM, the weight (i.e., price of risk) on cash-flow beta is γ (coefficient of relative risk aversion) times the weight of discount-rate beta, and the weight on discount-rate beta is the variance of the unexpected market return. If the investor is sufficiently risk-averse in the sense that $\gamma > 1$, then cash-flow beta has a greater price of risk than discount-rate beta. That is, this investor demands a higher compensation for bearing the risk characterized by cash-flow beta than by discount-rate beta.

An implication of the two-beta ICAPM is that stocks with higher expected returns should have higher cash-flow betas and may or may not have higher discount-rate betas. This is because conservative investors price cash-flow risk more than discount-rate risk. Therefore, the two-beta ICAPM's implication for the value effect (i.e., value stocks have higher returns than growth stocks) is that value stocks should have higher cash-flow betas than do growth stocks, and its implication for the size effect (i.e., small stocks have higher returns than large stocks) is that small stocks should have higher cash-flow betas than do large stocks.

Following Campbell (1991), Campbell and Vuolteenaho (2004) use the VAR model approach to implement the return decomposition. In this approach, discount-rate news are directly modeled whereas cash-flow news are backed out as the residual. Assuming that the data are generated by a first-order VAR model, the formula for discount-rate news is derived using the properties of a first-order VAR model. The formula for cash-flow news is backed out as the residual between the unexpected market return and discount-rate news.

Variables selected in the VAR model should predict market returns. In the VAR model approach, the first variable in the model has to be the market return. Other variables that Campbell and Vuolteenaho (2004) choose to include in the VAR model are the yield spread between long-term and short-term bonds (term yield spread), the 10-year smoothed price-earnings ratio from Shiller (2000), and the small-stock value spread.

Campbell and Vuolteenaho (2004) find that cash-flow news and discount-rate news are only slightly positively correlated and that discount-rate news has a larger variance than cash-flow news during 1929:1-2001:12. They calculate cash-flow betas and discount-rate betas for the 25 portfolios sorted by size and book-to-market equity ratio (i.e., the 25 Fama-French portfolios) and find that value stocks have higher cash-flow betas but lower discount-rate betas than growth stocks for the 1963:7-2001:12 period. Campbell and Vuolteenaho (2004) conclude that this finding provides an explanation for the value effect (i.e., value stocks have higher returns than do growth stocks), which could not be explained by the CAPM for this period.

Chen and Zhao (2009) find that Campbell and Vuolteenaho's (2004) results are sensitive to the choice of variables included in the VAR model. In particular, the results of the patterns in cash-flow betas between growth and value stocks and the relative importance of variances of cash-flow news and discount-rate news are sensitive to close substitutes of the price-earnings ratio or of the value spread. They consider a large pool of variables known to predict stock returns and examine various VAR model specifications based on different combinations of variables. They find that in most of the model specifications, value stocks do not have higher cash-flow betas than do growth stocks. Regarding the relative importance of variances of cash-flow news and discount-

rate news, they are not able to conclude what the common pattern for most model specifications is because the pattern changes frequently. Chen and Zhao (2009) do not draw a conclusion about the correlation between cash-flow news and discount-rate news. However, Table 2 in their paper shows that some model specifications yield negative correlations while others result in positive correlations. It seems that the only results that Chen and Zhao (2009) and Campbell and Vuolteenaho (2004) have in common are that growth stocks have higher discount-rate betas than do value stocks and small stocks have both higher cash-flow betas and higher discount-rate betas than do large stocks.

Even with a VAR model specification that increases the predictive power on market returns in the VAR regression, Chen and Zhao (2009) do not find the same patterns in cash-flow betas as Campbell and Vuolteenaho (2004) do. This alternative VAR model that they consider increases the R-squared of the equity return prediction equation (i.e., the first equation in the VAR model) from 2.57% in Campbell and Vuolteenaho (2004) to 14%. However, the log consumption-wealth ratio *cay* , one of the variables included in this alternative model, seems ad hoc because it has a look-ahead bias. The log consumption-wealth ratio *cay* is from Lettau and Ludvigson (2001a), who use the equilibrium error in the cointegration relation among consumption, asset wealth, and labor income as a proxy for *cay* . Guo (2003) argues that there is a look-ahead bias in *cay* because data on consumption and labor income are subject to substantial revisions to reflect the incorporation of new information and new methodologies and to reduce noise.

The variable *cay* does not seem to be an appropriate variable to use in the VAR model approach to implementing the return decomposition because it has a look-ahead

bias, which means that *cay* may contain information of future cash flows. Recall that the VAR approach directly models discount-rate news. Therefore, discount-rate news obtained from a VAR model with *cay* may be contaminated with cash-flows news. On the other hand, cash-flow news may be contaminated as well because it is backed out as the residual between the unexpected market return and discount-rate news.

Chen and Zhao (2009) argue that the VAR model approach has a limitation in that discount-rate news cannot be accurately estimated because of the small predictive power of the economic variables, and cash-flow news is discount-rate news in disguise because it is backed out as the residual and depends on how well discount-rate news is estimated.

To address the limitations of the VAR model approach, I propose a state-space model approach to model both cash-flow news and discount-rate news directly. In addition, instead of using information contained in variables motivated by economic reasons, I use information directly related to the market return to extract cash-flow news and discount-rate news from the market return.

The main difference between the state-space model approach and Campbell and Vuolteenaho's (2004) VAR model approach is that the former approach is more direct. It is more direct in two aspects. First, the state-space approach directly and simultaneously models both cash-flow news and discount-rate news by expressing the relation between returns and news in terms of linear regressions, whereas the VAR approach models discount-rate news but backs out cash-flow news as the residual. Second, the information that I use to extract cash-flow news and discount-rate news from the market return is directly related to the market return, whereas the information used by Campbell and Vuolteenaho (2004) is only indirectly related to the market return. The data used in the

state-space model approach are returns on the market and returns on portfolios that make up the market, which are directly related to the market return. With the exception of the excess log market return, the other data used in Campbell and Vuolteenaho (2004) (i.e., the term yield spread, the price-earnings ratio, and the value spread) are not that directly related to the market return.

A state-space model consists of a measurement equation and a state equation. The measurement equation relates observed variables to unobserved variables (i.e., state variables), and the state equation describes the development of the unobserved variables over time. The primary interest of an analysis of a state-space model is to analyze the unobserved state variables. However, because the state variables are not observable, information of the state variables must be extracted from the observed variables.

In the state-space model constructed in this dissertation, the state variables are cash-flow news, discount-rate news, and the conditional expectation of the excess log market return, which are all unobserved and are components of the return decomposition. The observed variables include the excess log market return, five log returns of Fama-French portfolios sorted by the ratio of book equity to market equity (BE/ME), and five Fama-Bliss discount bond log returns with maturities from one to five years.

To ensure that the three state variables represent cash-flow news, discount-rate news, and the conditional expectation of the excess log market return, I impose restrictions on the parameters of the state-space model. The restrictions are derived from the return decomposition.

I use a Bayesian Markov chain Monte Carlo (MCMC) methodology to estimate the state-space model in this dissertation. The state-space model is a high-dimensional

model, and MCMC techniques are particularly useful for high-dimensional models; see Casella and George (1992) and Gamerman and Lopes (2006, p. 4). Furthermore, it is well known that MCMC algorithms make the implementation of a Bayesian approach conceptually feasible. Good examples of MCMC algorithms used for Bayesian analysis can be found in, for example, Robert and Casella (2004).

I design an MCMC algorithm that makes use of the techniques of Gibbs sampler, Metropolis-Hastings within Gibbs sampler, and Durbin and Koopman's (2002) simulation smoother to generate samples from the joint posterior distribution. Gibbs sampling and Metropolis-Hastings within Gibbs sampling are two of the most important techniques of MCMC methods. The general MCMC sampling scheme used in this dissertation is most similar to the one studied in Strickland, Turner, Denham, and Mengersen (2009), and it is the preferred scheme among the different schemes compared by Reis, Salazar, and Gamerman (2006).

While MCMC has been widely useful in Bayesian inference, its implementation in state-space models requires substantial work as exemplified in Reis, Salazar, and Gamerman (2006) and Strickland, Turner, Denham, and Mengersen (2009). Compared with Reis, Salazar, and Gamerman (2006) and Strickland, Turner, Denham, and Mengersen (2009), the state-space model in this dissertation contains unknown parameters in the measurement matrix, which introduces an additional level of complexity, and hence, I need to design more elaborate algorithms. In addition, the data and state variables to be analyzed by the state-space model are of high dimension, and therefore, the implementation is computationally demanding. Running the MCMC algorithm for each Bayesian analysis of one real data set presented in this study took

approximately 42 hours of CPU time on a machine that has an AMD Phenom II X6 processor (a six-core processor with 3.2GHz of operating frequency for each processor, six 512 KB private L2 caches, and a shared 6 MB L3 cache) and 16 GB of main memory (DRAM).

Parameters of the prior distributions are obtained by analyzing the data and applying substantive knowledge in finance. This approach of obtaining the prior parameters is in some sense an empirical Bayes approach, see, for example, Gelman, Carlin, Stern, and Rubin (2004, p. 121).

Sensitivity analysis is used to assess the effect of alternative specifications in the prior distributions on the posterior inferences. Qualitatively, my results indicate that all prior specifications considered in this dissertation produced very similar posterior inferences. However, the alternative prior specification yields the unappealing result that market betas for the 25 Fama-French portfolios are all less than one. Therefore, only the posterior inferences of the primary analysis are presented. Results from posterior predictive check do not indicate strong evidence that the model does not fit the data. Therefore, there is no strong evidence against using this model. This gives support for the posterior inferences obtained under the model.

The main results of my dissertation are as follows. First, posterior inferences indicate that growth stocks have higher discount-rate betas than value stocks and that small stocks have both higher cash-flow betas and higher discount-rate betas than large stocks. These findings are consistent with those of Campbell and Vuolteenaho (2004) and Chen and Zhao (2009). The finding that growth stocks have higher discount-rate betas than do value stocks can be explained by the fact that growth stocks have most of

their cash flows in the future, making their durations longer and their sensitivity to discount-rate news higher.

Second, unlike Campbell and Vuolteenaho (2004), this dissertation finds that growth stocks have higher cash-flow betas than value stocks. This finding is consistent with Chen and Zhao (2009).

Third, I also find that cash-flow news and discount-rate news are strongly negatively correlated (-0.94), consistent with Campbell (1991) and Campbell (1996). In contrast, Campbell and Vuolteenaho (2004) find that cash-flow news and discount-rate news are only slightly positively correlated (0.11). Although Chen and Zhao (2009) do not draw a conclusion about the correlation between the two types of news, Table 2 in their paper shows that cash-flow news and discount-rate news are negatively correlated (-0.68) for the VAR model specified in column 7. This model has the best predictive power for market returns among all the VAR model specifications in Table 2 of their paper, including the VAR model of Campbell and Vuolteenaho (2004).

One plausible explanation for the strong negative correlation between cash-flow news and discount-rate news is as follows. When economic conditions deteriorate, investors' expectations of future cash flows deteriorate and become more uncertain, that is, there is a negative shock to expected cash flows. This negative shock to expected cash flows likely increases the uncertainty of future cash flows[1], which leads to increase in risk and, therefore, increase in expected returns (used to discount cash flows), which is the positive shock to discount rates. The opposite occurs in the case of positive shock to

[1] There exist conditions under which one would expect cash-flow news and discount-rate news be strongly negatively correlated. One condition is that $\frac{\partial \sigma_{CF}}{\partial \mu_{CF}} < 0$, where σ_{CF} denotes the standard deviation of future cash flows, and μ_{CF} denotes the mean of future cash flows.

economic outlook. Both cash-flow news and discount-rate news are heavily influenced by the business cycle, hence the strong negative correlation between them.

The strong negative correlation between cash-flow news and discount-rate news helps explain the result that growth stocks have higher cash-flow betas than do value stocks. This strong negative correlation, coupled with the fact that discount-rate betas are based on the negative of discount-rate news, suggests that the pattern of cash-flow betas would approximately mirror that of discount-rate betas. Namely, cash-flow betas, like discount-rate betas, are higher for growth stocks than for value stocks. Also, like discount-rate betas, cash-flow betas are higher for small stocks than for large stocks. These effects are exactly what I find.

Fourth, consistent with the results in Campbell and Vuolteenaho (2004), this dissertation finds that the variance of cash-flow news is smaller than that of discount-rate news, which suggests that discount-rate news is the dominant component of the market return. Chen and Zhao (2009) are unable to find a common pattern in the relative importance of variances of cash-flow news and discount-rate news because the pattern changes frequently depending on the VAR model specifications.

Fifth, posterior inferences indicate that when there is a positive shock to expected cash flows, the correlation between average returns and cash-flow betas of the 25 Fama-French portfolios is positive, and when there is a negative shock to expected cash flows, the correlation is negative. Similarly, when there is a positive shock to the negative of future discount rates, average returns and discount-rate betas are positively correlated, and when the shock to the negative of discount rates is negative, the two are negatively correlated. Using the estimated cash-flow news and discount-rate news from Campbell

and Vuolteenaho (2004), I also find these patterns. In addition, these patterns are also found when I use the cash-flow news and discount-rate news obtained by replicating Chen and Zhao's (2009) results of the VAR model containing the variable *cay* . Therefore, all three studies are consistent with the economic intuition that securities with higher cash-flow betas are likely to have higher contemporaneous returns than are those with lower cash-flow betas when there is a positive shock to expected cash flows, and securities with higher cash-flow betas are likely to have lower contemporaneous returns than are those with lower cash-flow betas when there is a negative shock to expected cash flows. All three studies are also consistent with the similar economic intuition regarding the effect of shocks to discount rates on the relation between contemporaneous returns and discount-rate betas.

Broadly speaking, estimates of discount-rate news in this dissertation are in line with those of Campbell and Vuolteenaho (2004) and Chen and Zhao (2009). Indeed, I find the same patterns in discount-rate betas as in such papers.

Finally, discount-rate news obtained from the state-space model in this dissertation and discount-rate news of Campbell and Vuolteenaho (2004) are strongly positively correlated. Total news (i.e., cash-flow news minus discount-rate news) estimated from the state-space model and total news from Campbell and Vuolteenaho (2004) are also strongly positively correlated. However, the correlation between cash-flow news of this dissertation and cash-flow news of Campbell and Vuolteenaho (2004) is positive but not strong.

In summary, to implement the return decomposition, I propose a state-space model approach that addresses the limitations of the VAR model approach. The data

analyzed by the state-space model do not show that value stocks have higher cash-flow betas than do growth stocks. The reason is likely that the cash-flow news and discount-rate news estimated using the state-space model and the data set in this dissertation are strongly negatively correlated, contrary to the correlation of the two news series estimated by Campbell and Vuolteenaho (2004).

The remainder of this dissertation is organized as follows. Chapter 2 reviews related literature. Chapter 3 first describes the ICAPM framework of Campbell and Vuolteenaho (2004), including the return decomposition, cash-flow beta, and discount-rate beta, then describes the state-space model approach for the implementation of the return decomposition. Details of the state-space model, including restrictions on the parameters, and the Bayesian MCMC methodology used to estimate the state-space model are given in this chapter. Chapter 3 also includes a simulation study and a description of the observed data (real data) analyzed using the state-space model. Chapter 4 presents and discusses the results of this study obtained using the state-space model approach, and Chapter 5 concludes. Some preliminary analysis using the VAR model approach done at the early stage of this research is in the appendix.

Chapter 2: Literature Review

I start by reviewing the static capital asset pricing model (CAPM) and some empirical evidence. Then, I review the intertemporal capital asset pricing model (ICAPM) of Merton (1973) and other intertemporal capital asset pricing models of the sort proposed by Merton. These include work by Campbell (1993), Campbell (1996), Chen (2003), Campbell and Vuolteenaho (2004), and Brennan, Wang, and Xia (2004). I also briefly review conditional capital asset pricing models (CCAPM). Finally, I provide a short review of the literature on state-space models.

2.1 CAPM and some empirical evidence

The CAPM is an equilibrium asset-pricing model where the expected excess return on an asset is a linear function of its beta (i.e., the covariance between its return and the market portfolio's return). The model is built on the portfolio selection theory of Markowitz (1959) and Tobin (1958). The CAPM is first developed by Sharpe (1963, 1964) and Treynor (1961) and then extended by Lintner (1965, 1969), Mossin (1966), and Black (1972); see Copeland and Weston (1992, p. 193). In some of these extensions, the authors modify the model to relax some of the assumptions underlying the CAPM. For example, one of the assumptions under which the CAPM is derived is that investors can borrow and lend at the risk-free rate. Black (1972) considers the case in which a risk-free asset does not exist and derives a model that shows the expected return on any asset to be a linear combination of the expected return on the market portfolio and the expected return on the minimum variance zero-beta portfolio.

The assumptions under which the Sharpe-Lintner-Mossin CAPM is derived are that investors are single-period expected utility of wealth maximizers, have homogenous expectations of the end-of-period joint distributions of returns, can borrow or lend at some given risk-free rate, that there are no restrictions on short sales, and securities markets are frictionless and perfectly competitive. Therefore, any test of the CAPM or market efficiency is a joint test of the CAPM and market efficiency (Copeland and Weston (1992, p. 351).

Early empirical work provides some support for the CAPM. Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), and Blume and Friend (1973) all report results that are consistent with the mean-variance efficiency of the market portfolio (Campbell, Lo, and MacKinlay (1997, p. 211).

However, evidence less favorable for the CAPM begins to emerge since the late 1970s. Empirical results that are inconsistent with the CAPM implications are typically referred to as CAPM anomalies. Examples of such anomalies include the size effect (Banz (1981); Reinganum (1981)), price-earnings effect (Basu (1977, 1983); Ball (1978)), book-to-market equity effect (Fama and French (1992, 1993), among others), leverage effect (Bhandari (1988)), price reversals (DeBondt and Thaler (1985, 1987)), and momentum (Jegadeesh and Titman (1993)). The price-earnings effect, book-to-market equity effect, and leverage effect are sometimes referred to as the value effect. A survey of these anomalies is given by Schwert (2002).

Basu's (1977, 1983) price-earnings-ratio effect is one of the earliest documented anomalies. Basu finds that the market portfolio did not appear to be mean-variance efficient relative to portfolios formed on the basis of the price-earnings-ratios (P/E) of

firms. Namely, firms with low (high) P/E have higher (lower) sample average returns than would be the case if the market portfolio were mean-variance efficient. Ball (1978) argues that P/E may serve as a catch-all proxy for unnamed factors in expected returns.

Stattman (1980) and Rosenberg, Reid, and Lanstein (1985) find that average stock returns are positively related to the ratio of book value of equity to market value of equity (BE/ME). Similarly, Fama and French (1992) identify a value premium for the post-1963 period, that is, stocks with high BE/ME (i.e., value stocks) have higher average returns than do stocks with low BE/ME (i.e., growth stocks). Fama and French (1992, 1993) argue that size (market value of equity) and BE/ME jointly capture the cross-sectional variation in expected stock returns associated with the CAPM beta, P/E, BE/ME, size, and leverage. In addition, they argue that size and BE/ME represent two missing factors from the CAPM. Fama and French (1996) find that except for momentum, most anomalies disappear in the three factor model, which is a model that includes market risk premium, a size factor, and a BE/ME factor as the three factors. Fama and French (1998) also show that the value effect exists in a sample of 13 countries from 1975-1995. Using data from the period 1929-1997, Davis, Fama, and French (2000) document the value premium for both the pre-1963 and post-1963 periods, and also find that the size effect is subsumed by the value effect in the pre-1963 period.

More recently, Fama and French (2006) study the value premium using data from 1926 to 2004 and find that the value premium exists and is similar before and after 1963. However, like Franzoni (2001), they find that the CAPM betas of value stocks decreased dramatically in the second half of the sample period. In the pre-1963 period, value stocks have higher CAPM betas than do growth stocks, and therefore, the CAPM explains the

value premium, which is also shown by Ang and Chen (2007). However, in the post-1963 period, value stocks have lower CAPM betas than do growth stocks, and therefore, the CAPM fails to explain the value premium for that period. The fact that value stocks have higher average returns but lower CAPM betas than do growth stocks is referred to as the value effect.

2.2 The intertemporal capital asset pricing model (ICAPM)

The CAPM is a single-period model because it is derived under the assumption that investors are single-period expected utility of wealth maximizers. Although the CAPM has been generally treated as if it holds intertemporally, Merton (1971) argues that this should not be the case since the portfolio decision making of an intertemporal expected utility maximizer will be very different when he/she faces a changing investment opportunity set rather than a constant one.

Merton (1973) derives an intertemporal capital asset pricing model (ICAPM), which is an equilibrium asset pricing model, under the assumptions that investors are maximizers of the expected utility of lifetime consumption and that trading in assets take place continuously in time, among other assumptions. Under this model, investors' current demands for portfolios are affected by the possibility of uncertain changes in future investment opportunities. Merton (1973) shows that when the investment opportunity set is constant through time, the ICAPM collapses to the classical CAPM. The ICAPM of Merton (1973) states that expected returns are compensation to investors for bearing the market risk and for bearing the risk of unfavorable shifts in the investment

opportunity set. Hence, risky securities with no market risk may have expected returns that differ from the risk-free rate.

It should be noted that the empirical operationalization of Merton's ICAPM is not immediately available. In particular, while the model specifies that shocks to investment opportunities is a source of risk, it is silent on how such shocks could be captured in empirical or applied work. This issue is described by Campbell (1993).

Campbell's (1993) asset-pricing model is an approximate discrete-time version of Merton's (1973) ICAPM. Campbell (1993) derives an intertemporal asset-pricing model by assuming that the variation in the consumption-wealth ratio is small, which makes the intertemporal budget constraint approximately loglinear and the portfolio-choice problem more tractable. The intertemporal asset-pricing model of Campbell (1993) is derived based on the first-order condition of an investor with a recursive preference function proposed by Epstein and Zin (1989, 1991). This intertemporal asset pricing model says that the expected excess log return on an asset, i.e., the risk premium on an asset, is expressed as a linear combination of two covariances. One of them is the covariance of the asset's log return with the unexpected log market return, and the other is the covariance of the asset's log return with news about future returns on the market (i.e., discount rates). The weight (i.e., price of risk) on the first covariance is the coefficient of relative risk aversion, and the weight (i.e., price of risk) on the second covariance is one minus the coefficient of relative risk aversion.

Campbell (1996) expands upon the intertemporal asset pricing model of Campbell (1993). The intertemporal asset pricing model of Campbell (1996) takes human capital into account. Omitting human capital can be serious since human capital accounts for

approximately two-thirds of total wealth (i.e., market portfolio).[2] To introduce human capital into the model, Campbell (1996) starts with the relationship that return on the market portfolio is equal to the weighted average of the return on financial wealth and return on human wealth. The resulting asset-pricing model says that the risk premium on an asset is expressed as the linear combination of three covariances. The first covariance is the covariance of the asset with the financial wealth component, the second covariance is the covariance of the unexpected return on the asset with news about current and future labor income, and the third covariance is the covariance of the unexpected return on the asset with news about future returns on the market. Campbell (1996) uses a VAR approach similar to the one used in Campbell (1991) to estimate news about future stock returns and labor income. Variables included in the VAR model should predict the market return as well as labor income.

The dividend-ratio model of Campbell and Shiller (1988) is a cornerstone for Campbell's (1991) return decomposition equation, which is used by Campbell and Vuolteenaho (2004) to develop an intertemporal capital asset pricing model. To derive the dividend-ratio model, Campbell and Shiller (1988) start with the identity of log return on a dividend-paying stock and take the first-order Taylor approximation of this identity to obtain a linear relationship among log returns, log prices, and log dividends. Then, they rewrite this linear relationship in terms of the log dividend-price ratio and the dividend growth rate and solved the equation forward with the imposition of the terminal condition that the log dividend-price ratio does not explode through time. The result is

[2] This is because labor costs account for approximately two-thirds of gross national product (see Campbell (1996)).

the dividend-ratio model, which approximates the log dividend-price ratio by the discounted value of all future returns and dividend growth rates and a constant term.

Based on the framework of Campbell and Shiller (1988), Campbell (1991) decomposes the unexpected stock return into two components. Specifically, he substitutes the *ex ante* version (conditional expectation version) of the dividend-ratio model of Campbell and Shiller (1988) for the log dividend-price ratio in the first-order Taylor approximation of the log return to derive the return decomposition equation for a stock. The return decomposition equation says that the difference between the realized stock return and the conditional expectation of the stock return (i.e., unexpected stock return) consists of two components—change in expectations of all discounted future dividend growth rates and change in expectations of all discounted future stock returns. The former component is called news about future dividends (i.e., news about future cash flows), and the latter is called news about future returns (i.e., news about future discount rates). A rise in the unexpected returns could be the result of either higher expectations of discounted future dividend growth rates or lower expectations of discounted future returns, or the combined effect of both. Campbell (1991) uses a VAR approach to estimate news about future returns and news about future dividends. Variables selected in the VAR model should have predictability of market returns.

Campbell and Vuolteenaho (2004) extend the intertemporal asset pricing model of Campbell (1993) and develop a two-beta intertemporal asset pricing model (two-beta ICAPM). Specifically, they apply the return decomposition equation of Campbell (1991) to the stock market portfolio, and using the return decomposition, they rewrite Campbell's (1993) asset pricing formula to express the risk premium on any asset as a

linear combination of the covariance of the asset's log return with news about future (market) cash flows and the covariance of the asset's log return with news about future (market) discount rates. Then, multiplying and dividing the revised asset pricing formula by the conditional variance of market return, they obtain the two-beta ICAPM, which states that the risk premium on an asset is expressed as a linear combination of the cash-flow beta and the discount-rate beta. As defined by Campbell and Vuolteenaho (2004), cash-flow beta measures the covariance of the asset's log return with news about future cash flows, and discount-rate beta measures the covariance of the asset's log return with news about future discount rates. According to their definition, the sum of cash-flow and discount-rate betas equals the asset's market beta.

In the two-beta ICAPM, the price of risk of cash-flow beta is γ (coefficient of relative risk aversion) times that of discount-rate beta, and the price of risk of discount-rate beta should equal the variance of the unexpected market return. Therefore, if the investor is sufficiently risk-averse in the sense that $\gamma > 1$, then cash-flow beta has a greater price of risk than does discount-rate beta.

The two-beta ICAPM delivers the intuition that risk-averse investors price the risk characterized by cash-flow beta higher than the risk characterized by discount-rate beta. This intuition comes from the return decomposition, which suggests that a fall in the value of the market portfolio could be the result of investors' downward revision of their expectations about future cash flows, or investors' upward revision of expectations about future discount rates. Although both cases result in the decline of wealth, in the first case, future investment opportunities are unchanged, while in the second case, future investment opportunities improve. This is because higher expectations about future

discount rates indicate improved prospects for future market returns. Thus, cash-flow news can be interpreted as permanent shocks to wealth and discount-rate news as transitory shocks to wealth. Therefore, a risk-averse long-term investor who holds the market portfolio may demand a greater compensation for holding assets that are more sensitive to the market's cash-flow news than for holding assets that are more sensitive to discount-rate news.

The two-beta ICAPM implies that stocks with higher expected returns should exhibit higher cash-flow betas and may or may not have higher discount-rate betas because risk-averse investors price cash-flow risk higher than discount-rate risk. Therefore, implication of the two-beta ICAPM for the value effect is that value stocks should have higher cash-flow betas than do growth stocks, and implication for the size effect is that small stocks should have higher cash-flow betas than do large stocks.

Following Campbell (1991), Campbell and Vuolteenaho (2004) use the VAR approach to implement the return decomposition, that is, to estimate cash-flow news and discount-rate news. Cash-flow news and discount-rate news are used to calculate cash-flow betas and discount-rate betas, respectively. In the VAR approach, discount-rate news is directly modeled using a first-order VAR model, whereas cash-flow news is backed out as the residual between the unexpected market return and discount-rate news. Campbell and Vuolteenaho (2004) specify a vector of variables and assume that the vector follows a first-order VAR model. The first component of this vector is the excess log market return, and the other components are variables that should be able to predict the market return. These variables are the yield spread between long-term and short-term bonds (term yield spread), the 10-year smoothed price-earnings ratio from Shiller (2000),

and the small-stock value spread. The small-stock value spread is defined as the log (BE/ME) of the small-value portfolio minus the log (BE/ME) of the small-growth portfolio.

Using the estimated cash-flow news and discount-rate news, Campbell and Vuolteenaho (2004) calculate cash-flow betas and discount-rate betas of the 25 Fama-French portfolios. The five portfolios with the highest (lowest) BE/ME are referred to as value (growth) portfolios. Campbell and Vuolteenaho (2004) find that for the period of 1963-2001, value portfolios have higher cash-flow betas than do growth portfolios, and growth portfolios have higher discount-rate betas than do value portfolios. This result is consistent with the implication of the two-beta ICAPM for the value effect. The CAPM could not explain the value premium for this time period because value stocks have lower market betas than do growth stocks. Recall that market beta is the sum of cash-flow beta and discount-rate beta. Campbell and Vuolteenaho (2004) point out that value stocks have lower market betas than do growth stocks because value stocks have lower discount-rate betas than do growth stocks, despite the fact that value stocks have higher cash-flow betas than do growth stocks. Moreover, they find that the difference in discount-rate betas between value and growth stocks is greater than the difference in cash-flow betas between value and growth stocks.

The value premium for the period of 1929-1963 could also be explained by the two-beta ICAPM of Campbell and Vuolteenaho (2004), with the exception of the small-value portfolio and the small-growth portfolio[3]. Campbell and Vuolteenaho (2004) find that the small-value portfolio has both lower cash-flow beta and lower discount-rate beta

[3] The small-value portfolio is the portfolio at the intersection of the lowest size quintile and the highest BE/ME quintile. The small-growth portfolio is the portfolio at the intersection of the lowest size quintile and the lowest BE/ME quintile.

than does the small-growth portfolio. For value and growth portfolios in other size categories, they find that value portfolios have both higher cash-flow betas and higher discount-rate betas than do growth portfolios. This result is also consistent with the implication of the CAPM because this evidence indicates that value portfolios have higher market betas than do growth portfolios.

Two recent studies suggest the importance of cash-flow news. In a recent study, Campbell, Polk, and Vuolteenaho (2010) show that the high discount-rate betas of growth stocks and the high cash-flow betas of value stocks are determined by the cash-flow fundamentals of growth and value companies. The study of Campbell, Giglio, and Polk (2010) suggests that negative shocks to future cash flows are much more important in the stock market downturn of 2007-2009 than in the downturn of 2000-2002. The downturn of 2000-2002 is driven primarily by a large increase in discount rates. The study also suggests that compared with the economic downturn of 2000-2002, the downturn of 2007-2009 has more serious implications for long-term investors because there is little increase in expected returns to reassure investors that stock prices are likely to recover over time.

Some studies point out the drawbacks of the VAR model approach in Campbell and Vuolteenaho (2004). Liu and Zhang (2008) cast doubt on the ability of the small-stock value spread to predict the market return. Chen and Zhao (2009) argue that Campbell and Vuolteenaho's (2004) findings regarding cash-flow and discount-rate betas are sensitive to the choice of variables in the VAR model. They consider different VAR model specifications based on different combinations of variables included in a VAR model and find that in most of the model specifications, value stocks do not have higher

cash-flow betas than do growth stocks. In particular, they find that it is necessary to include the small-stock value spread in the VAR model to obtain the patterns in cash-flow betas between value and growth stocks that could explain the value premium. Chen and Zhao (2009) show that even with a VAR model that improves the predictive power of market returns in the VAR regression, they do not find the same patterns in cash-flow betas between value and growth stocks as Campbell and Vuolteenaho (2004) do.

A recent paper by Engsted, Pedersen, and Tanggaard (2012) analyzes the pitfalls involved in VAR-based return decompositions[4]. For example, they argue that it is necessary to include the asset price in a VAR model when implementing return decompositions. In the case of equity return decompositions, this requirement is equivalent to including the dividend-price ratio in a VAR model.

Two papers that also study intertemporal asset-pricing models of the type proposed by Merton (1973) are worth noting. Brennan, Wang, and Xia (2004) develop an ICAPM with an investment opportunity set that is fully described by the real interest rate and the maximum Sharpe ratio (i.e., Sharpe ratio of the market portfolio). They use a Kalman filter to estimate the model parameters and the time series of the real interest rate and the maximum Sharpe ratio from data of U.S. Treasury bond yields and inflation expectations for the period of 1952-2001. Their cross-sectional asset pricing tests show that both the real interest rate and the maximum Sharpe ratio have significant risk premia.

Chen (2003) extends the framework of Campbell (1993) to allow for time-varying market volatilities. He develops an intertemporal asset-pricing model in which changes in the investment opportunity set are captured by both time-varying expected market

[4] Note that my dissertation does not suffer from the drawbacks of the VAR model approach pointed out by Engsted, Pedersen, and Tanggaard (2012).

returns and time-varying market volatilities. However, he finds little evidence that the model explains the value premium.

Recently, Guo, Savickas, Wang, and Yang (2009) provide empirical evidence that suggests that the value premium is a proxy for investment opportunities.

2.3 Conditional capital asset pricing model

Unlike the CAPM and ICAPM, the conditional capital asset pricing model (conditional CAPM) is not derived from a utility-of-consumption maximization problem, but, rather, it is motivated by the empirical observation that covariances of asset returns with risk factors vary over time. The temporal variation serves as an additional risk factor to be hedged.

Jagannathan and Wang (1996) assume that the CAPM holds conditionally, i.e., betas and the market risk premium remain constant in the short periods of time, but vary over long run. Starting with the conditional CAPM, they derive an unconditional asset-pricing model implied by the conditional CAPM, which says that unconditional expected returns are a linear function of the average CAPM beta and a measure of beta instability over time. A special case of this model is the CAPM, which is obtained when betas are constant over time. Jagannathan and Wang (1996) also examine their model empirically, and they include the return on human capital when measuring the return on the market portfolio. They find that their model can explain more variation in the cross-section of average returns when the return on human capital is included than when it is not.

Lettau and Ludvigson (2001b) explore the conditional version of Breeden's (1979) consumption CAPM. Specifically, they express the stochastic discount factor using a

conditional factor model with time-varying parameters. By modeling these parameters as linear functions of a forecasting variable for excess returns (i.e., conditioning variable), they obtain an unconditional multifactor model in which the additional factors are scaled versions of the original factors. Lettau and Ludvigson (2001b) use the log consumption-wealth ratio as the conditioning variable, which is central to their approach. Since the log consumption-wealth ratio is not observable, they use Lettau and Ludvigson's (2001a) *cay* , the equilibrium error in the cointegration relation among consumption, asset wealth, and labor income, as proxy for the log consumption-wealth ratio. Empirically, their conditional model performs better than the unconditional version and approximately as well as the Fama-French three-factor model.

Although several studies suggest that the conditional CAPM holds period-by-period (for example, Jagannathan and Wang (1996)) and that time-varying betas may explain the value premium,[5] Lewellen and Nagel (2006) argue that the conditional CAPM does not explain anomalies such as the value premium and momentum. Their analytical results show that variation in betas and the equity premium would have to be extremely large to explain these anomalies. Moreover, Lewellen and Nagel (2006) provide a new test of the conditional CAPM and find that the conditional CAPM performs almost as poorly as the unconditional CAPM.

2.4 State-space model and some applications

The state-space model, introduced by Kalman (1960) and Kalman and Bucy (1961), is a very general model that subsumes a class of models. For example,

[5] Zhang (2005), Jagannathan and Wang (1996), Lettau and Ludvigson (2001b), Santos and Veronesi (2006), and Lustig and Van Nieuwerburgh (2005) argue that time-varying betas may explain the value premium. See Lewellen and Nagel (2006).

autoregressive-moving average (ARMA) models such as a vector autoregressive (VAR) model can be written in state-space form (see Shumway and Stoffer 2006, p. 329).

The state-space model consists of a measurement equation and a state equation. The state equation describes the development of the state vector, which is comprised of variables that are not observable, over time. The measurement equation relates a vector of observed variables to the state vector through the measurement matrix. The primary interest of an analysis of a state-space model is to analyze the unobserved state variables. However, information of the state variables must be extracted from the observed variables because the state variables are not observable. Techniques for analyzing the state variables include filtering, smoothing, forecasting, and simulation smoothing. Durbin and Koopman (2001) give a good exposition of these techniques.

An approach for estimation of the parameters of a state-space model is maximum likelihood estimation (MLE). One method of obtaining the maximum likelihood estimates is a procedure that usually involves fixing the initial state vector and then developing a set of recursions for the log likelihood function and its first two derivatives (see Gupta and Mehra (1974)). Next, the parameter values are updated successively by a Newton-Raphson algorithm until the log likelihood is maximized.[6] Jones (1980) is an advocate of this approach.

Another method of implementing MLE is to use the Expectation-Maximization (EM) algorithm of Dempster, Laird, and Rubin (1977) to obtain the maximum likelihood estimates of the parameters. In the context of the classical state-space model, in which the measurement matrix is known, Shumway and Stoffer (1982) propose an EM-algorithm-based estimation procedure. A good exposition of this procedure is given by

[6] See Shumway and Stoffer (2006, p. 340).

Shumway and Stoffer (2006). Shumway (2000) extends the work of Shumway and Stoffer (1982) and proposes an EM-algorithm-based procedure to estimate the state-space model in which the measurement matrix is unknown. Stoffer and Wall (1991) propose a bootstrap procedure to estimate the standard errors of the parameters in a state-space model.

Bayesian MCMC is another approach for estimating a state-space model. MCMC techniques are particularly useful for estimation of high-dimensional models, a point discussed in Casella and George (1992). Gamerman and Lopes (2006, p. 4) also point out that since state-space models are high-dimensional models, MCMC techniques are particularly useful. Moreover, it is well known that MCMC algorithms make the implementation of a Bayesian approach conceptually feasible. Good examples of MCMC algorithms used for Bayesian analysis can be found in, for example, Robert and Casella (2004) and Gamerman and Lopes (2006).

Simulation smoothing, the drawing of random samples of the state variables conditional on the observations, is a useful technique when using Bayesian MCMC methods to estimate a state-space model. Fruhwirth-Schnatter (1994) and Carter and Kohn (1994) argue that it is more efficient to sample all the state vectors at once from the full conditional distributions of the state vectors. Both studies independently propose a simulation-smoothing algorithm, widely known as the forward filtering-backward sampling algorithm. de Jong and Shephard (1995) develop this approach further, and Durbin and Koopman (2002) refine the simulation smoother of de Jong and Shephard (1995).

Reis, Salazar, and Gamerman (2006) compare different MCMC sampling schemes for the estimation of a state-space model, and the preferred scheme is similar to the one studied in Strickland, Turner, Denham, and Mengersen (2009), who propose a simulation smoothing algorithm which extends the algorithm of Durbin and Koopman (2002).

The state-space model has a wide range of applications, including aerospace research, economics, finance, medicine, and the soil sciences (Shumway and Stoffer 2006, p.324; Tsay 2005, p.490; Kim and Nelson 1999, p.19). For example, in economics, Engle and Watson (1981) use the state-space model to model the behavior of wage rates. Garbade and Wachtel (1978) and Antoncic (1986) apply the model to modeling the behavior of ex-ante real interest rates. Burmeister and Wall (1982) and Burmeister, Wall, and Hamilton (1986) use the model to estimate expected inflation. Kim and Nelson (1989) use it to model a time-varying monetary reaction function of the Federal Reserve.[7]

The state-space model has not been widely applied to the area of intertemporal asset pricing. Perhaps one of the most well-known papers that used the Kalman filter in the intertemporal asset pricing literature is the one by Brennan, Wang, and Xia (2004). They used the Kalman filter to estimate the time series of the real interest rate and the maximum Sharpe ratio. In a related paper, Brennan and Xia (2006) used the Kalman filter to estimate the time series of the real interest rate, the maximum Sharpe ratio, and the expected rate of inflation.

[7] See Kim and Nelson (1999, p. 19).

Chapter 3: Methodology

3.1 The ICAPM framework

3.1.1 Return decomposition

The return decomposition equation in Campbell and Vuolteenaho (2004) is derived by Campbell (1991) who uses the dividend-ratio model of Campbell and Shiller (1988) as a basis for the derivation. To derive the dividend-ratio model, Campbell and Shiller (1988) take a first-order Taylor approximation of the equation relating the log stock gross return to log stock price and dividends to obtain a linear relationship among log returns, log prices, and log dividends. Then, they rewrite this linear relationship in terms of the log dividend-price ratio and the dividend growth rate and solve the equation forward with the imposition of the terminal condition that the log dividend-price ratio do not explode through time. The result is the dividend-ratio model, which approximates the log dividend-price ratio by the discounted value of all future returns and dividend growth rates and a constant term.

Campbell (1991) substitutes the *ex ante* version (conditional expectation version) of the dividend-ratio model for the log dividend-price ratio in the first-order Taylor approximation of the log return to derive the return decomposition equation,

$$r_{t+1} - E_t r_{t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j}. \qquad (3.1)$$

Here, r_{t+1} is the log return of a stock or portfolio held from the end of period t to the end of period $t+1$, $E_t r_{t+1}$ is the conditional expectation of r_{t+1} given the information available at the end of period t, $\rho < 1$ is a discount rate, d_{t+1} is the log dividend paid during period

$t+1$, and Δd_{t+1} is the dividend growth rate from the end of period t to the end of period $t+1$. Note that $\Delta d_{t+1} \equiv \ln(D_{t+1}) - \ln(D_t)$, where D_{t+1} denotes the dividend paid during period $t+1$. Both terms on the right-hand side of (3.1) are of the form $(E_{t+1} - E_t)X$, which denotes $E_{t+1}X - E_t X$.

The return decomposition equation in (3.1) says that the *unexpected* stock (portfolio) return is comprised of two parts—change in expectations of all discounted future dividend growth rates and change in expectations of all discounted future stock (portfolio) returns. A rise in the unexpected returns could be the result of either: (1) higher expectations of discounted future dividend growth rates, (2) lower expectations of discounted future returns, or (3) the combined effect of both.

Campbell and Vuolteenaho (2004) apply the return decomposition equation to the stock market. Namely,

$$r_{M,t+1} - E_t r_{M,t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{M,t+1+j}$$

$$= N_{CF,t+1} - N_{DR,t+1}, \qquad (3.2)$$

where $r_{M,t+1}$ denotes the log return on the stock market from the end of period t to the end of period $t+1$. The variable N_{CF} denotes news about future cash flows, such as dividends or consumption, and N_{DR} denotes news about future discount rates (i.e., expected returns). The left-hand side of (3.2) is the unexpected market return. The market return decomposition equation in (3.2) says that the unexpected market return can be decomposed into news about future cash flows and news about future discount rates.

When excess log returns are considered, the decomposition equation is of the same form as equation (3.2) if one assumes that the risk-free rate is constant. Let $r^e_{M,t} \equiv r_{M,t} - r_{f,t}$ denote the excess log return on the stock market and $r_{f,t}$ denote the log risk free rate. Campbell (1991) shows that the return decomposition in (3.2) becomes

$$r^e_{M,t+1} - E_t r^e_{M,t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r^e_{M,t+1+j}$$

$$-(E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j r_{f,t+1+j}. \qquad (3.3)$$

If the risk-free rate is constant, then the third term on the right-hand side of equation (3.3) is zero, and equation (3.3) becomes

$$r^e_{M,t+1} - E_t r^e_{M,t+1} = (E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r^e_{M,t+1+j}$$

$$= N_{CF,t+1} - N_{DR,t+1}. \qquad (3.4)$$

In the empirical study of Campbell and Vuolteenaho (2004), they use equation (3.4) rather than equation (3.2). Although they do not mention explicitly, they use $N_{CF,t+1}$ to denote $(E_{t+1} - E_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}$, and $N_{DR,t+1}$ to denote $(E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r^e_{M,t+1+j}$ instead of $(E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r_{M,t+1+j}$. Hereafter, I refer to equation (3.4) as the return decomposition equation.

3.1.2 Cash-flow and discount-rate betas

Campbell and Vuolteenaho (2004) estimate $\{N_{CF,t}\}$ and $\{N_{DR,t}\}$ and find that they are not highly correlated with each other. Thus, they argue that stocks may have different

betas for the two components of the unexpected market return given by the return decomposition in equation (3.4). They define security i's cash-flow and discount-rate betas as

$$\beta_{i,CF} \equiv \frac{Cov(r_{i,t}, N_{CF,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})}, \qquad (3.5)$$

and

$$\beta_{i,DR} \equiv \frac{Cov(r_{i,t}, -N_{DR,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})}, \qquad (3.6)$$

respectively, where $r_{i,t}$ denotes the log return of security i. Note that in definitions (3.5) and (3.6), both cash-flow and discount-rate betas use the variance of the unexpected market returns as the denominator, not the variance of the cash-flow news or discount-rate news. This definition of cash-flow beta and discount-rate beta implies that the sum of the two betas is equal to the market beta,

$$\beta_{i,M} = \beta_{i,CF} + \beta_{i,DR}. \qquad (3.7)$$

This is because by (3.5) and (3.6), $\beta_{i,CF} + \beta_{i,DR} = \frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})}$, and by the return decomposition equation (3.4), $r^e_{M,t+1} - E_t r^e_{M,t+1} = N_{CF,t+1} - N_{DR,t+1}$, one can write

$$\frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})} = \frac{Cov(r_{i,t}, r^e_{M,t} - E_{t-1} r^e_{M,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})}.$$

According to Campbell and Vuolteenaho (2004), the market beta of security i is

$$\beta_{i,M} = \frac{Cov(r_{i,t}, r^e_{M,t} - E_{t-1} r^e_{M,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})}. \qquad (3.8)$$

To obtain empirical estimates of cash-flow beta and discount-rate beta, Campbell and Vuolteenaho (2004) use

$$\hat{\beta}_{i,CF} = \frac{Cov(r_{i,t}, \hat{N}_{CF,t})}{Var(\hat{N}_{CF,t} - \hat{N}_{DR,t})} + \frac{Cov(r_{i,t}, \hat{N}_{CF,t-1})}{Var(\hat{N}_{CF,t} - \hat{N}_{DR,t})} \qquad (3.9)$$

and

$$\hat{\beta}_{i,DR} = \frac{Cov(r_{i,t}, -\hat{N}_{DR,t})}{Var(\hat{N}_{CF,t} - \hat{N}_{DR,t})} + \frac{Cov(r_{i,t}, -\hat{N}_{DR,t-1})}{Var(\hat{N}_{CF,t} - \hat{N}_{DR,t})}, \qquad (3.10)$$

respectively. $\hat{N}_{CF,t}$ and $\hat{N}_{DR,t}$ denote, respectively, the estimated cash-flow news and the estimated discount-rate news. According to Campbell and Vuolteenaho (2004), including one lag of the cash-flow or discount-rate news in, respectively, equations (3.9) and (3.10) is motivated by the possibility that not all stocks in the portfolios were traded frequently and synchronously, especially during the early years of the sample. Infrequent and non-synchronous trading may cause market return, cash-flow news, and discount-rate news to spuriously lead the portfolio returns (Scholes and Williams (1977) and Dimson (1979)). I follow Campbell and Vuolteenaho (2004) and use equations (3.9) and (3.10) to estimate cash-flow and discount-rate betas.

3.1.3 An intertemporal asset pricing formula

Campbell (1993) derives a cross-sectional asset pricing formula that says that the expected excess log return on an asset, i.e., the risk premium on an asset, is determined by its own variance and a linear combination of two covariances. One of them is the covariance of the asset's log return with the unexpected log market return, and the other is the covariance of the asset's log return with news about future discount rates. The weight on the first covariance is the coefficient of relative risk aversion, and the weight

on the second covariance is one minus the coefficient of relative risk aversion. The variance term shows up to adjust for Jensen's Inequality.

Based on equation (3.2), Campbell and Vuolteenaho (2004) rewrite Campbell's (1993) asset pricing formula to express the risk premium on any asset i as a linear combination of the covariance of asset i's log return with N_{CF} and the covariance of its log return with N_{DR},

$$E_t(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma_{i,t}^2}{2} = \gamma Cov_t(r_{i,t+1}, N_{CF,t+1}) + Cov_t(r_{i,t+1}, -N_{DR,t+1}). \qquad (3.11)$$

The term $\frac{\sigma_{i,t}^2}{2}$, one-half the variance of the excess log return, is a Jensen's Inequality effect. The variable γ is the coefficient of relative risk aversion.

Multiplying and dividing equation (3.11) by the conditional variance $\sigma_{M,t}^2 = Var_t(r_{M,t+1}^e - E_t r_{M,t+1}^e)$, Campbell and Vuolteenaho (2004) obtain

$$E_t(r_{i,t+1}) - r_{f,t+1} + \frac{\sigma_{i,t}^2}{2} = \gamma\sigma_{M,t}^2\beta_{i,CF,t} + \sigma_{M,t}^2\beta_{i,DR,t}. \qquad (3.12)$$

They also use simple returns rather than log returns and derive an unconditional version of equation (3.12):

$$E(R_i - R_f) = \gamma\sigma_M^2\beta_{i,CF} + \sigma_M^2\beta_{i,DR}, \qquad (3.13)$$

where R_i and R_f denote, respectively, the return on asset i and the risk-free rate. Equation (3.13) says that the risk price of cash-flow beta is γ times that of discount-rate beta and that the risk price of discount-rate beta should equal the variance of the unexpected market return.

3.2 State-space model approach to ICAPM

3.2.1 A classical state-space model

I propose using the state-space model approach to estimate cash-flow news and discount-rate news in equation (3.4). In this section, I give an introduction of a classical state-space model and the techniques that are typically used to analyze this model. Details of the state-space model designed for the ICAPM are given in section 3.2.2.

The state-space model, which consists of a measurement equation and a state equation, is introduced by Kalman (1960) and Kalman and Bucy (1961). The model has a wide range of applications, including aerospace research, economics, finance, medicine, and the soil sciences (see, for example, Shumway and Stoffer 2006, p. 324; Tsay 2005, p. 490; Kim and Nelson 1999, p. 19). A basic state-space model has the form

$$y_t = Hz_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, R), \tag{3.14}$$

$$z_t = Fz_{t-1} + \eta_t, \quad \eta_t \sim N(0, V), \tag{3.15}$$

$$z_1 \mid \theta_F, V \sim N(a_1, P_1). \tag{3.16}$$

Equation (3.14) is called the measurement equation, and equation (3.15) is called the state equation. In equation (3.14), $t = 1,\ldots, T$, y_t is a $p\times1$ vector of observed data, H is a $p\times m$ measurement or observation matrix, z_t is an $m\times1$ state vector of variables that are not observed, and the observation noise vectors $\{\varepsilon_1,\ldots,\varepsilon_T\}$ are assumed to be serially independent and identically distributed $p\times1$ normal vectors with zero mean and $p\times p$ covariance matrix R. In equation (3.15), $t = 2,\ldots, T$, F is an $m\times m$ matrix, and $\{\eta_2,\ldots,\eta_T\}$ are assumed to be $m\times1$ normal vectors that are serially independent and identically distributed with zero mean and $m\times m$ covariance matrix V. The

measurement matrix H is typically known, and the transition matrix F usually contains both known and unknown parameters. It is assumed that $\{\eta_2,\ldots,\eta_T\}$ and $\{\varepsilon_1,\ldots,\varepsilon_T\}$ are independent of each other at all time points. In expression (3.16), a_1 and P_1 are both assumed to be known conditional on $\{\theta_F, V\}$. Furthermore, the initial state vector z_1 is assumed to be independent of the error terms $\{\eta_2,\ldots,\eta_T\}$ and $\{\varepsilon_1,\ldots,\varepsilon_T\}$.

The primary interest of an analysis of a state-space model would be to analyze the unobserved state vectors z_t given the observations y_t. The idea underlying the state-space model is that the development of the system over time is determined by the state vector z_t according to equation (3.15), but because z_t is not observable, we must base the analysis of z_t on the observations y_t.

Analyzing the state vector z_t involves four techniques – filtering, smoothing, forecasting, and simulation smoothing. Let $I_t = \{y_1,..., y_t\}$ be the information available at time t. Then, calculating $E(z_t \mid I_s)$, $t > s$, is called forecasting, calculating $E(z_t \mid I_t)$ is called filtering, and calculating $E(z_t \mid I_T)$ is called smoothing. Denote $z_{t|s} = E(z_t \mid I_s)$ and $P_{t|s} = E\left[(z_t - z_{t|s})(z_t - z_{t|s})' \mid I_s\right]$. Given the initial values $z_{1|0}$ and $P_{1|0}$, the Kalman filter updates the estimate of the state vector from $z_{t|t-1}$ to $z_{t|t}$ and then to $z_{t+1|t}$ when a new observation y_t becomes available. The Kalman gain determines the contribution of the new prediction error of y_t on the state vector z_t. Upon obtaining $z_{T|T-1}$ and $P_{T|T-1}$ from the Kalman filter, the smoothed states $\{z_{t|T};\ t = 1,\ldots, T\}$ and smoothed state covariance matrices $\{P_{t|T};\ t = 1,\ldots, T\}$ can be obtained via the Kalman smoother.

Simulation smoothing is the drawing of random samples of state or disturbance vectors conditional on the observations. Namely, it is generating random draws from the conditional densities $p(\varepsilon_1,\ldots,\varepsilon_T \mid y_1,\ldots,y_T,\theta)$, $p(\eta_2,\ldots,\eta_T \mid y_1,\ldots,y_T,\theta)$, and $p(z_1,\ldots,z_T \mid y_1,\ldots,y_T,\theta)$ for a given parameter vector θ. Carlin, Polson, and Stoffer (1992) propose to sample individual state vectors z_t from $p(z_t \mid y_1,\ldots,y_T,z_{-t},\theta)$, where z_{-t} denotes all state vectors $\{ z_1,\ldots,z_T \}$ excluding z_t. Fruhwirth-Schnatter (1994) and Carter and Kohn (1994) argue that it is more efficient to sample all the state vectors at once from $p(z_1,\ldots,z_T \mid y_1,\ldots,y_T,\theta)$. de Jong and Shephard (1995) develop this approach further by first sampling the disturbances and then the state vectors. Durbin and Koopman (2002) refine the simulation smoother of de Jong and Shephard (1995).

Simulation smoothing is useful when using Bayesian Markov chain Monte Carlo (MCMC) methods for the estimation of state-space models. In a Bayesian framework, the unknown parameter vector θ of a state-space model is treated as a set of random variables. The aim of a Bayesian analysis of a state-space model is to obtain samples of $\{ z_1,\ldots,z_T \}$ and θ from their joint posterior density $p(z_1,\ldots,z_T,\theta \mid y_1,\ldots,y_T)$. This can be implemented using the MCMC method by repeating the following two steps for a given initialization of the parameter vector $\theta = \theta^0$:

(1) sample $\{ z_1^i,\ldots z_T^i \}$ from $p(z_1,\ldots,z_T \mid y_1,\ldots,y_T,\theta^{i-1})$,

(2) sample θ^i from $p(\theta \mid y_1,\ldots,y_T,z_1^i,\ldots,z_T^i)$,

for iteration i = 1, 2,…, N, where N is some large positive integer. After discarding the first K iterations that are considered "burned-in", one can treat the samples $\{z_1^i,\ldots,z_T^i,\theta^i \mid i = K+1,\ldots,N\}$ as being generated from the joint posterior density

$p(z_1,\ldots,z_T,\theta \mid y_1,\ldots,y_T)$. The first step of the MCMC procedure, sampling from $p(z_1,\ldots,z_T \mid y_1,\ldots,y_T,\theta)$, is done using the simulation smoother. The second step can be done by direct sampling from the density $p(\theta \mid y_1,\ldots,y_T,z_1,\ldots,z_T)$ if an analytical form for the density is available, otherwise, it can be done using the Metropolis-Hastings algorithm. Details of the Bayesian MCMC method for the estimation of a state-space model is given in section 3.2.3.

3.2.2 State-space model for ICAPM

In the ICAPM framework, the main quantities of interest are cash-flow news and discount-rate news, which are used to calculate the cash-flow and discount-rate betas of the 25 Fama-French portfolios. To obtain estimates of cash-flow news and discount-rate news, I construct a state-space model in which the state vector consists of cash-flow news, discount-rate news, and the conditional expectation of the excess log market return. All three components are unobserved, and they are components of the return decomposition equation (3.4). In the state-space model specified in equations (3.14), (3.15), and (3.16), I define the state vector $z_t \equiv \begin{bmatrix} E_{t-1}r^e_{M,t} & N_{CF,t} & N_{DR,t} \end{bmatrix}'$.

To ensure that the three components of the state vector are properly identified as the conditional expectation of the excess log market return, cash-flow news, and discount-rate news, I impose restrictions on the measurement matrix H and the transition matrix F, consistent with the economic intuition and observable data series used. These restrictions are derived from Campbell and Vuolteenaho's (2004) model.

There are three types of restrictions imposed on H. These restrictions are based on the return decomposition equation (3.4), market beta in equation (3.8), and characteristics of zero-coupon bond returns. Recall the return decomposition (3.4),

$$r^e_{M,t} - E_{t-1} r^e_{M,t} = N_{CF,t} - N_{DR,t}.$$

Moving the term $E_{t-1} r^e_{M,t}$ to the right-hand side of the equation, I have

$$r^e_{M,t} = E_{t-1} r^e_{M,t} + N_{CF,t} - N_{DR,t}. \qquad (3.17)$$

Note that the left-hand side of equation (3.17), $r^e_{M,t}$, is the excess log market return and is observed. I use equation (3.17) to place restrictions on H.

The second type of restrictions imposed on H is motivated by the market beta in equation (3.8),

$$\beta_{i,M} = \frac{Cov(r_{i,t}, r^e_{M,t} - E_{t-1} r^e_{M,t})}{Var(r^e_{M,t} - E_{t-1} r^e_{M,t})},$$

which is the sum of cash-flow beta and discount-rate beta defined by Campbell and Vuolteenaho (2004). According to the return decomposition (3.4), the unexpected excess market return, $r^e_{M,t} - E_{t-1} r^e_{M,t}$, is equal to the difference between cash-flow news and discount-rate news, $N_{CF,t} - N_{DR,t}$; therefore, I can re-write $\beta_{i,M}$ in (3.8) as

$$\beta_{i,M} = \frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(N_{CF,t} - N_{DR,t})}. \qquad (3.18)$$

Equation (3.18) suggests that $\beta_{i,M}$ can be thought of as the slope coefficient of the no-intercept classical linear regression

$$r_{i,t} = \beta_{i,M} (N_{CF,t} - N_{DR,t}) + \varepsilon_{i,t}. \qquad (3.19)$$

To see why, consider the OLS estimate of $\beta_{i,M}$ in equation (3.19):

$$\frac{\frac{1}{T}\sum_{t=1}^{T} r_{i,t}(N_{CF,t} - N_{DR,t})}{\frac{1}{T}\sum_{t=1}^{T}(N_{CF,t} - N_{DR,t})^2}, \qquad (3.20)$$

which is approximately equal to

$$\frac{E[r_{i,t}(N_{CF,t} - N_{DR,t})]}{E[(N_{CF,t} - N_{DR,t})^2]}. \qquad (3.21)$$

It is shown in Appendix A that the right-hand side of (3.18) is equal to (3.21); namely,

$$\frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(N_{CF,t} - N_{DR,t})} = \frac{E[r_{i,t}(N_{CF,t} - N_{DR,t})]}{E[(N_{CF,t} - N_{DR,t})^2]}.$$

In regression equation (3.19), $r_{i,t}$ is the log return on stock portfolio i and is observed. I use equation (3.19) to impose restrictions on H.

The third type of restrictions imposed on H is motivated by characteristics of zero-coupon bond returns. The present value of a security is determined by its future cash flows and discount rates. However, zero-coupon bonds do not pay out any coupons; therefore, their prices are not subject to cash-flow news, only to discount-rate news. This implies that zero-coupon bond returns are influenced by discount-rate news, not by cash-flow news. Thus, I use the following regression to impose restrictions on H:

$$r_{i,t}^{b} = \beta_{bi}(-N_{DR,t}) + \varepsilon_{i,t}, \qquad (3.22)$$

where $r_{i,t}^{b}$ is the log return on zero-coupon bond i and is observed.

Restrictions imposed on the transition matrix F are based on the assumption that economic shocks cannot be forecasted. Namely, neither cash-flow news nor discount-rate news can be forecasted. Therefore, I restrict the second and third rows of F to zero

so that both the second and third components of the state vector z_t are noise, which cannot be forecasted. This means that in equation (3.15), $z_{2,t} = \eta_{2,t}$ and $z_{3,t} = \eta_{3,t}$.

The observed vector y_t is an 11-by-1 vector that is composed of the excess log market return, log returns of the five Fama-French portfolios sorted by BE/ME, and log returns of the five Fama-Bliss discount bonds with maturities from one to five years. The observed vector is denoted

$$y_t = \begin{bmatrix} r^e_{M,t} & r_{1,t} & r_{2,t} & \dots & r_{5,t} & r^b_{1,t} & \dots & r^b_{5,t} \end{bmatrix}'.$$

Details of the data including description and summary statistics will be given in section 3.2.5. The observed variables include five stock portfolios and five bonds to ensure that the information used to extract cash-flow news and discount-rate news are balanced between stocks and bonds.

Motivated by equations (3.17), (3.19), and (3.22)[8], I propose to study the state-space model of equations (3.14), (3.15), and (3.16) with H and F of the form

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 0 & \beta_{1,M} & -\beta_{1,M} \\ 0 & \beta_{2,M} & -\beta_{2,M} \\ 0 & \beta_{3,M} & -\beta_{3,M} \\ 0 & \beta_{4,M} & -\beta_{4,M} \\ 0 & \beta_{5,M} & -\beta_{5,M} \\ 0 & 0 & -\beta_{b1} \\ 0 & 0 & -\beta_{b2} \\ 0 & 0 & -\beta_{b3} \\ 0 & 0 & -\beta_{b4} \\ 0 & 0 & -\beta_{b5} \end{bmatrix}, \qquad (3.23)$$

[8] Equations (3.17), (3.19) and (3.22) are used to impose restrictions on the first row of H, the second to sixth rows of H, and the seventh to eleventh rows of H, respectively.

$$F = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad (3.24)$$

3.2.3 Bayesian Markov chain Monte Carlo methodology

I use a Bayesian MCMC methodology to estimate the state-space model of equations (3.14), (3.15), and (3.16) with restrictions imposed on the measurement matrix and transition matrix expressed in (3.23) and (3.24). The aim of Bayesian inference is to summarize the unknown quantities of primary interest through the joint posterior density function of the unknown quantities in the model. I first present the essential part of the joint posterior density, then derive the likelihood, and specify the priors. Then, I give a general Gibbs sampling scheme, within which I make use of Metropolis algorithm, direct sampling from Wishart distribution, and Durbin and Koopman's (2002) simulation smoothing algorithm.

Notation

Next, I introduce some notation. Let θ denote the set of unknown parameters of the state-space model with restrictions. The vectors $\theta = \{ \theta_H, \theta_F, R, V \}$, where θ_H and θ_F are the unknown parameters in expressions (3.23) and (3.24), respectively. Namely,

$\theta_H = \{ \beta_{1,M}, \beta_{2,M}, \beta_{3,M}, \beta_{4,M}, \beta_{5,M}, \beta_{b1}, \beta_{b2}, \beta_{b3}, \beta_{b4}, \beta_{b5} \}$, and

$\theta_F = \{ F_{11}, F_{12}, F_{13} \}$.

Define $y = [y_1', \cdots, y_T']'$ and $z = [z_1', \cdots, z_T']'$.

Joint posterior density

The joint posterior density of the unknown quantities, $\{ z, \theta_H, \theta_F, R, V \}$, denoted by

$$p(z, \theta_H, \theta_F, R, V \mid y),$$

is, by definition, equal to

$$\frac{p(y, z, \theta_H, \theta_F, R, V)}{p(y)}.$$

Because y is observed data, $p(y)$ is fixed. Therefore, the joint posterior density of $\{ z, \theta_H, \theta_F, R, V \}$ is proportional to the joint density of $\{ y, z, \theta_H, \theta_F, R, V \}$:

$$p(z, \theta_H, \theta_F, R, V \mid y) \propto p(y, z, \theta_H, \theta_F, R, V). \qquad (3.25)$$

By Bayes' rule, the right-hand side of (3.25) is equal to

$$p(y \mid z, \theta_H, \theta_F, R, V)\, p(z \mid \theta_H, \theta_F, R, V)\, \pi(\theta_H, \theta_F, R, V), \qquad (3.26)$$

where $p(y \mid z, \theta_H, \theta_F, R, V)$ denotes the joint probability density function of y conditional on $\{ z, \theta_H, \theta_F, R, V \}$, $p(z \mid \theta_H, \theta_F, R, V)$ denotes the joint probability density function of z conditional on $\{ \theta_H, \theta_F, R, V \}$, and $\pi(\theta_H, \theta_F, R, V)$ denotes the joint prior density of $\{ \theta_H, \theta_F, R, V \}$. Throughout the dissertation, $\pi(.)$ denotes prior density. Assuming that all components in θ are *a priori* independent, expression (3.26) can be written as

$$p(y \mid z, \theta_H, \theta_F, R, V)\, p(z \mid \theta_H, \theta_F, R, V)\, \pi(\theta_H)\, \pi(\theta_F)\, \pi(R)\, \pi(V). \qquad (3.27)$$

In (3.27), the conditioning information $\{ z, \theta_H, \theta_F, R, V \}$ in $p(y \mid z, \theta_H, \theta_F, R, V)$ can be reduced to $\{ z, \theta_H, R \}$ because based on equation (3.14), knowing $\{ z, \theta_H, R \}$ is sufficient to obtain the conditional density of y. Similarly, the conditioning information

$\{ \theta_H, \theta_F, R, V \}$ in $p(z \mid \theta_H, \theta_F, R, V)$ can be reduced to $\{ \theta_F, V \}$ because based on equation (3.15), knowing $\{ \theta_F, V \}$ is sufficient to obtain the conditional density of z. Thus, expression (3.27) can be re-written so that the joint posterior density becomes

$$p(z, \theta_H, \theta_F, R, V \mid y) \propto p(y \mid z, \theta_H, R)\, p(z \mid \theta_F, V)\, \pi(\theta_H)\, \pi(\theta_F)\, \pi(R)\, \pi(V). \quad (3.28)$$

The joint conditional densities $p(y \mid z, \theta_H, R)$ and $p(z \mid \theta_F, V)$ at the right-hand side of (3.28) are specified below.

Likelihood

The likelihood is $p(y, z \mid \theta_H, \theta_F, R, V)$, and by Bayes' rule, it can be written as

$$p(y \mid z, \theta_H, \theta_F, R, V)\, p(z \mid \theta_H, \theta_F, R, V). \quad (3.29)$$

In the discussion following expression (3.27), I have shown that the conditioning information $\{ z, \theta_H, \theta_F, R, V \}$ in $p(y \mid z, \theta_H, \theta_F, R, V)$ can be reduced to $\{ z, \theta_H, R \}$, and the information set $\{ \theta_H, \theta_F, R, V \}$ in $p(z \mid \theta_H, \theta_F, R, V)$ can be reduced to $\{ \theta_F, V \}$. Therefore, the likelihood in (3.29) becomes

$$p(y \mid z, \theta_H, R)\, p(z \mid \theta_F, V), \quad (3.30)$$

which appears as the first two terms of the right-hand side of (3.28), the expression describing the joint posterior density. The joint conditional densities $p(y \mid z, \theta_H, R)$ and $p(z \mid \theta_F, V)$ are given in (3.33) and (3.40), respectively.

The joint conditional density of $\{ y_1, \cdots, y_T \}$ given $\{ z, \theta_H, R \}$, denoted $p(y \mid z, \theta_H, R)$ or $p(y_1, \cdots, y_T \mid z, \theta_H, R)$, is equal to

$$\prod_{t=1}^{T} p(y_t \mid z, \theta_H, R), \quad (3.31)$$

because by equation (3.14), given $\{ z, \theta_H \}$, $\{ y_1, \cdots, y_T \}$ are serially independent. Also by equation (3.14), $\varepsilon_t \sim N(0, R)$, and $\varepsilon_t = y_t - Hz_t$; thus,

$$p(y_t \mid z, \theta_H, R) = \frac{1}{(2\pi)^{\frac{p}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} (y_t - Hz_t)' R^{-1} (y_t - Hz_t) \right). \qquad (3.32)$$

Using (3.31) and (3.32), the joint conditional density of y given $\{ z, \theta_H, R \}$ is obtained:

$$p(y \mid z, \theta_H, R) = \frac{1}{(2\pi)^{\frac{Tp}{2}} |R|^{\frac{T}{2}}} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} (y_t - Hz_t)' R^{-1} (y_t - Hz_t) \right). \qquad (3.33)$$

The joint conditional density of $\{ z_1, \cdots, z_T \}$ given $\{ \theta_F, V \}$, denoted $p(z \mid \theta_F, V)$ or $p(z_1, \cdots, z_T \mid \theta_F, V)$, is equal to

$$p(z_T \mid z_1, \cdots, z_{T-1}, \theta_F, V) p(z_{T-1} \mid z_1, \cdots, z_{T-2}, \theta_F, V) \cdots p(z_2 \mid z_1, \theta_F, V) p(z_1 \mid \theta_F, V), \qquad (3.34)$$

using Bayes' rule.

Equation (3.15) indicates that when (z_t, θ_F, V) are given, the density of z_{t+1} is completely specified. Because one does not need to know $(z_1, \cdots, z_{t-1})$ to get the density of z_{t+1}, the conditioning information $(z_1, \cdots, z_t, \theta_F, V)$ in $p(z_{t+1} \mid z_1, \cdots, z_t, \theta_F, V)$ can be reduced to (z_t, θ_F, V). Applying this trick to each term except the last term in expression (3.34), expression (3.34) becomes

$$p(z_T \mid z_{T-1}, \theta_F, V) p(z_{T-1} \mid z_{T-2}, \theta_F, V) \cdots p(z_2 \mid z_1, \theta_F, V) p(z_1 \mid \theta_F, V), \qquad (3.35)$$

which can be written more succinctly as

$$p(z_1 \mid \theta_F, V) \prod_{t=1}^{T-1} p(z_{t+1} \mid z_t, \theta_F, V). \qquad (3.36)$$

By equation (3.15), $\eta_t \sim N(0, V)$ and $\eta_{t+1} = z_{t+1} - Fz_t$; therefore,

$$p(z_{t+1} \mid z_t, \theta_F, V) = \frac{1}{(2\pi)^{\frac{m}{2}} |V|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(z_{t+1} - Fz_t)' V^{-1} (z_{t+1} - Fz_t) \right). \qquad (3.37)$$

Using equation (3.37),

$$\prod_{t=1}^{T-1} p(z_{t+1} \mid z_t, \theta_F, V) = \frac{1}{(2\pi)^{\frac{(T-1)m}{2}} |V|^{\frac{T-1}{2}}} \exp\left(-\frac{1}{2}\sum_{t=1}^{T-1}(z_{t+1} - Fz_t)' V^{-1} (z_{t+1} - Fz_t) \right). \qquad (3.38)$$

By (3.16), $z_1 \mid \theta_F, V \sim N(a_1, P_1)$, so

$$p(z_1 \mid \theta_F, V) = \frac{1}{(2\pi)^{\frac{m}{2}} |P_1|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(z_1 - a_1)' P_1^{-1} (z_1 - a_1) \right). \qquad (3.39)$$

Substituting the right-hand sides of (3.38) and (3.39) into expression (3.36), $p(z \mid \theta_F, V)$, the joint conditional density of z given $\{ \theta_F, V \}$, becomes

$$\frac{1}{(2\pi)^{\frac{Tm}{2}} |P_1|^{\frac{1}{2}} |V|^{\frac{T-1}{2}}} \exp\left(-\frac{1}{2}\left((z_1 - a_1)' P_1^{-1} (z_1 - a_1) + \sum_{t=1}^{T-1}(z_{t+1} - Fz_t)' V^{-1} (z_{t+1} - Fz_t) \right) \right). \qquad (3.40)$$

Mean vector and covariance matrix of the initial state vector

Recall that in expression (3.16), the mean vector (a_1) and variance-covariance matrix (P_1) of the initial state vector z_1 are both assumed to be known, conditional on $\{ \theta_F, V \}$. In this study, a_1 is set equal to $\begin{bmatrix} 0.0105 & 0.0006 & 0.0038 \end{bmatrix}'$, and P_1 is set equal to

$$\begin{bmatrix} 0.0001 & 0 & 0 \\ 0 & 4.0584E-04 & 0 \\ 0 & 0 & 1.9875E-03 \end{bmatrix}.$$

Prior knowledge was used to obtain some of these numbers. The first, second, and third components of a_1 are the averages of Campbell and Vuolteenaho's (2004)

estimates of the conditional expectation of the excess log market return, cash-flow news, and discount-rate news, respectively, during 1952:2-1953:1. The averages are based on 1952:2-1953:1 because the data analyzed using the state-space model span the period 1952:7-2010:12, and a_1 is the mean of the initial state vector. The second and third elements on the diagonal of P_1 are the variances of Campbell and Vuolteenaho's (2004) estimates of cash-flow news and discount-rate news, respectively, during 1952:7-2001:12. The variances are based on 1952:7-2001:12 rather than 1952:7-2010:12 because I do not have Campbell and Vuolteenaho's (2004) estimates of the two types of news since 2001:12.

Priors

The priors on θ_H, θ_F, R, and V are independently assigned. In addition, priors on the components of $\theta_H = \{ \beta_{1,M}, \beta_{2,M}, \beta_{3,M}, \beta_{4,M}, \beta_{5,M}, \beta_{b1}, \beta_{b2}, \beta_{b3}, \beta_{b4}, \beta_{b5} \}$ and priors on the components of $\theta_F = \{ F_{11}, F_{12}, F_{13} \}$ are independently assigned.

Parameters of the prior distributions are obtained by analyzing the data and applying substantive knowledge in finance. This approach is in some sense an empirical Bayes approach, see, for example, Gelman, Carlin, Stern, and Rubin (2004, p. 121).

To assess the effect of prior specifications on the posterior inferences, a sensitivity analysis is performed by considering alternative prior specifications. The alternative prior specification and the results of the sensitivity analysis are discussed in Chapter 4. The results indicate that, qualitatively, the primary analysis and the analysis based on the alternative prior specification produce very similar posterior inferences. Below is a description of the prior specification for the primary analysis.

Priors on $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, ***and*** $\beta_{5,M}$

In this section, I first give the support and prior distribution, and then specify the parameters of the prior distribution.

Components in θ_H represent the stocks' and bonds' loadings on shocks. Specifically, $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$ are the five Fama-French BE/ME portfolios' loadings on the difference between cash-flow shock and discount-rate shock, given in equation 3.19; β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} are the five Fama-Bliss discount bonds' loadings on the discount-rate shock (see equation 3.22). Securities' loadings on economic shocks should be positive because shocks represent risk, and investors are assumed to be risk-averse, demanding compensation of higher returns when faced with higher risk. According to Campbell and Vuolteenaho (2004), the sum of the estimates of cash-flow and discount-rate betas of each of the 25 Fama-French portfolios is between zero and two, indicating that market betas for stock portfolios are probably less than two. Therefore, I assume the prior on each of $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$ has support [0, 3]. Note that by considering a support of [0, 3] rather than [0, 2], the priors are less informative.

A simple and natural way of introducing a prior on a finite interval like [0, 3] is considering a multiple of a Beta random variable. Therefore, I assign the priors on $\{ \beta_{i,M};\ i = 1,\ldots,5 \}$ as Beta distribution multiplied by three so that the support of each component of $\{ \beta_{i,M};\ i = 1,\ldots,5 \}$ is [0, 3].

The two shape parameters of Beta distribution are determined by first setting the prior mean and prior variance, and then using the formulae of the mean and variance of a

Beta random variable to derive the values of the two shape parameters. Suppose a random variable B has a Beta distribution with shape parameters a and b, denoted $B \sim Beta(a,b)$, $a>0$ and $b>0$. Then, the mean and variance of B are

$$E(B)=\frac{a}{a+b}, \text{ and}$$

$$\text{var}(B)=\frac{ab}{(a+b)^2(a+b+1)}.$$

Solving for a and b from the above mean and variance formulae for B, I get

$$a=\frac{E(B)\left(E(B)-(E(B))^2-\text{var}(B)\right)}{\text{var}(B)}, \text{ and} \qquad (3.41)$$

$$b=\frac{\left(E(B)-(E(B))^2-\text{var}(B)\right)\left(1-E(B)\right)}{\text{var}(B)}. \qquad (3.42)$$

Prior knowledge of $\{\beta_{i,M};\ i=1,\ldots,5\}$ can be obtained by considering regression (3.19) and Campbell and Vuolteenaho's (2004) estimates of cash-flow news and discount-rate news. In regression (3.19),

$$r_{i,t}=\beta_{i,M}(N_{CF,t}-N_{DR,t})+\varepsilon_{i,t},$$

where $i=1,\ldots,5$, I use Campbell and Vuolteenaho's (2004) estimates of cash-flow news and discount-rate news to obtain the covariate $N_{CF,t}-N_{DR,t}$. The dependent variable $r_{i,t}$ is the log return of each of the 5 Fama-French BE/ME portfolios. The time period used in estimating regression (3.19) is 1952:7-2001:12. Ordinary Least Squares (OLS) estimates of the slope coefficients $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$ in regression (3.19) are 1.0606, 0.9659, 0.8706, 0.8423, and 0.9014, respectively. Thus, I set the prior means for $\frac{1}{3}\beta_{1,M}$, $\frac{1}{3}\beta_{2,M}$, $\frac{1}{3}\beta_{3,M}$, $\frac{1}{3}\beta_{4,M}$, and $\frac{1}{3}\beta_{5,M}$ as 0.3535, 0.322, 0.2902, 0.2808, and

0.3005, respectively. Note that I report the prior means for $\{ \frac{1}{3}\beta_{i,M};\ i=1,\ldots,5 \}$ because the prior distribution of $\frac{1}{3}\beta_{i,M}$ is Beta distribution, and to obtain its shape parameters using equations (3.41) and (3.42), I need the prior mean and variance of $\frac{1}{3}\beta_{i,M}$.

Prior variances of $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$ are made larger so that the priors are less informative. The upper bound of a Beta random variable's prior variance can be calculated when the prior mean is given and under the condition that the shape parameter $a>1$. Note that the condition that $a>1$ and $b>1$ is necessary for a Beta distribution to have only one mode. The prior variances for all five parameters are 80% of their respective upper bounds. The prior variances for $\frac{1}{3}\beta_{1,M}$, $\frac{1}{3}\beta_{2,M}$, $\frac{1}{3}\beta_{3,M}$, $\frac{1}{3}\beta_{4,M}$, and $\frac{1}{3}\beta_{5,M}$ are 0.0478, 0.0425, 0.0371, 0.0354, and 0.0389, respectively.

With the prior means and prior variances of $\frac{1}{3}\beta_{1,M}$, $\frac{1}{3}\beta_{2,M}$, $\frac{1}{3}\beta_{3,M}$, $\frac{1}{3}\beta_{4,M}$, and $\frac{1}{3}\beta_{5,M}$ noted above, I use equations (3.41) and (3.42) to obtain the shape parameters of the Beta distributions. The prior distributions for $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$ are

$$\beta_{1,M} \sim 3 * Beta(1.3384,\ 2.4474),$$

$$\beta_{2,M} \sim 3 * Beta(1.3305,\ 2.8019),$$

$$\beta_{3,M} \sim 3 * Beta(1.3226,\ 3.2348),$$

$$\beta_{4,M} \sim 3 * Beta(1.3202,\ 3.3819),$$

$$\beta_{5,M} \sim 3 * Beta(1.3251,\ 3.0851).$$

Priors on β_{b1}, β_{b2}, β_{b3}, β_{b4}, ***and*** β_{b5}

There is not sufficient knowledge in the literature about discount bonds' loadings on discount-rate shock. Chen and Zhao (2009) decompose the returns of a bond market portfolio into cash-flow and discount-rate news and used these to obtain cash-flow and discount-rate betas of the Fama-Bliss discount bonds for maturities ranging from one to five years. They obtain positive cash-flow betas and negative discount-rate betas. However, since Chen and Zhao's (2009) discount-rate betas are calculated using news decomposed from a bond market portfolio, they are different from the loadings of discount bonds on discount-rate shock decomposed from an equity market portfolio as in equation (3.22). Since the knowledge about the loadings of discount bonds on discount-rate shock is not sufficient, the priors on β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} are defined by assuming that they have the same support as $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$.

Prior knowledge of $\{\beta_{bi};\ i=1,\ldots,5\}$ can be obtained by considering regression (3.22) and Campbell and Vuolteenaho's (2004) estimate of discount-rate news. In regression (3.22),

$$r_{i,t}^{b} = \beta_{bi}(-N_{DR,t}) + \varepsilon_{i,t},$$

where $i=1,\ldots,5$, I use Campbell and Vuolteenaho's (2004) estimates of discount-rate news, $-1*N_{DR,t}$, to obtain the covariate. The dependent variable $r_{i,t}^{b}$ is the log return of each of the five Fama-Bliss discount bonds. The time period used in estimating regression (3.22) is 1952:7-2001:12. Ordinary Least Squares (OLS) estimates of the slope coefficients β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} in regression (3.22) are 0.0136, 0.0292, 0.038, 0.0558, and 0.0597, respectively. As in the case of $\{\beta_{i,M};\ i=1,\ldots,5\}$, I use prior

means of $\{ \frac{1}{3}\beta_{bi};\ i=1,\ldots,5 \}$ rather than prior means of $\{ \beta_{bi};\ i=1,\ldots,5 \}$ to find the shape parameters of the Beta distributions. Prior means of $\frac{1}{3}\beta_{b1}$, $\frac{1}{3}\beta_{b2}$, $\frac{1}{3}\beta_{b3}$, $\frac{1}{3}\beta_{b4}$, and $\frac{1}{3}\beta_{b5}$ are set as 0.0045, 0.0097, 0.0127, 0.0186, and 0.0199, respectively.

The way in which prior variances of β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} are selected is similar to the case for $\beta_{1,M}$, $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, and $\beta_{5,M}$. Upper bounds of the prior variances for Beta random variables $\frac{1}{3}\beta_{b1}$, $\frac{1}{3}\beta_{b2}$, $\frac{1}{3}\beta_{b3}$, $\frac{1}{3}\beta_{b4}$, and $\frac{1}{3}\beta_{b5}$ are calculated, and 80% of the upper bounds are taken to be the prior variances of these Beta random variables. The prior variances of $\frac{1}{3}\beta_{b1}$, $\frac{1}{3}\beta_{b2}$, $\frac{1}{3}\beta_{b3}$, $\frac{1}{3}\beta_{b4}$, and $\frac{1}{3}\beta_{b5}$ are 1.6292E-5, 7.4329E-5, 1.2514E-4, 2.6666E-4, and 3.0445E-4, respectively.

Given the prior means and prior variances of $\frac{1}{3}\beta_{b1}$, $\frac{1}{3}\beta_{b2}$, $\frac{1}{3}\beta_{b3}$, $\frac{1}{3}\beta_{b4}$, and $\frac{1}{3}\beta_{b5}$, I use equations (3.41) and (3.42) to obtain the shape parameters of the Beta distributions. The prior distributions for β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} are

$$\beta_{b1} \sim 3 * Beta(1.2511,\ 274.7342),$$

$$\beta_{b2} \sim 3 * Beta(1.2524,\ 127.4222),$$

$$\beta_{b3} \sim 3 * Beta(1.2532,\ 97.6810),$$

$$\beta_{b4} \sim 3 * Beta(1.2547,\ 66.1997),$$

$$\beta_{b5} \sim 3 * Beta(1.2550,\ 61.8091).$$

Priors on F_{11}, F_{12}, ***and*** F_{13}

The parameters F_{11}, F_{12}, and F_{13} determine the development of $z_{1,t}$ over time according to equations (3.15) and (3.24). Namely,

$$z_{1,t} = F_{11} z_{1,t-1} + F_{12} z_{2,t-1} + F_{13} z_{3,t-1} + \eta_{1,t}, \qquad (3.43)$$

which can be written as

$$E_{t-1} r^e_{M,t} = F_{11} E_{t-2} r^e_{M,t-1} + F_{12} N_{CF,t-1} + F_{13} N_{DR,t-1} + \eta_{1,t}, \qquad (3.44)$$

since the state vector z_t is composed of $E_{t-1} r^e_{M,t}$, $N_{CF,t}$, and $N_{DR,t}$. Knowledge of F_{11}, F_{12}, and F_{13} can be derived from Campbell and Vuolteenaho's (2004) estimates of cash-flow news, discount-rate news, and conditional expectation of the excess log market return by estimating regression (3.44). The time period used in estimating regression (3.44) is 1952:7-2001:11. OLS estimates of F_{11}, F_{12}, and F_{13} in (3.44) are 0.8279, 0.1088, and -0.0853, respectively, and these are used as their respective prior means.

F_{11} determines the relation between $E_{t-1} r^e_{M,t}$ and $E_{t-2} r^e_{M,t-1}$, and it is reasonable to assume that F_{11} is a number close to but less than one. I define the support of F_{11} as [0.2, 1.2] and use Beta distribution as its prior distribution. Note that I define the upper bound of F_{11}'s support to be 1.2 instead of one so that the prior is less informative, and if the data were to suggest that F_{11}'s value should be greater than one, it could have a chance to reflect this. The prior distribution of F_{11} is $0.2 + Beta(7.5116,\ 4.4515)$.

F_{12} reflects the relation between $E_{t-1} r^e_{M,t}$ and $N_{CF,t-1}$, and F_{13} reflects the relation between $E_{t-1} r^e_{M,t}$ and $N_{DR,t-1}$. The prior for F_{12} is $Normal(0.1088,\ 0.01)$, where 0.01 is the prior variance. The prior for F_{13} is $Normal(-0.0853,\ 0.01)$. I choose Normal

distributions with standard deviations equal to 0.1 as their priors because I wish to have more vague priors.

Prior on R

Prior for R, the covariance matrix of ε_t, is inverse Wishart distribution. Inverse Wishart is often used as a prior for covariance matrices of multivariate Normal random vectors, see, for example, Gamerman (1998), Gelman, Carlin, Stern, and Rubin (2004, p. 87), Gamerman and Lopes (2006, pp. 67, 173), Reis, Salazar, and Gamerman (2006), and Strickland, Turner, Denham, Mengersen (2009). If a $d \times d$ symmetric and positive-definite matrix X has Wishart distribution with degrees of freedom parameter φ and scale matrix parameter B, where B is a $d \times d$ symmetric and positive-definite matrix, denoted $X \sim W_d(\varphi, B)$, then X^{-1} has inverse Wishart distribution with the same parameters as X, denoted $X^{-1} \sim IW_d(\varphi, B)$. A random matrix with $IW_d(\varphi, B)$ distribution is obtained by drawing a Wishart matrix with $W_d(\varphi, B)$ distribution and inverting it (Gamerman and Lopes 2006, p. 24). Therefore, covariance matrix R is obtained by drawing R^{-1}, which is Wishart, and inverting it.

Prior for R^{-1} is Wishart distribution, denoted $R^{-1} \sim W_p(\varphi_R, B_R)$, where $p = 11$ and $\varphi_R = 10001$. B_R is a diagonal matrix of the form

$$B_R = 1000 * \begin{bmatrix} (\varphi_R \operatorname{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\varphi_R \operatorname{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\varphi_R \operatorname{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix},$$

where the diagonal is the inverse of φ_R multiplied by the inverse of error term variances in the measurement equation (3.14). The variances of the error terms are obtained by estimating regressions (3.19) and (3.22), where the covariates are calculated using Campbell and Vuolteenaho's (2004) estimates of cash-flow news and discount-rate news. The OLS estimates of the error term variances from $\text{var}(\varepsilon_{1,t})$ to $\text{var}(\varepsilon_{11,t})$ are 2.2590E-06, 3.1720E-04, 3.1934E-04, 3.9006E-04, 5.1802E-04, 8.0259E-04, 2.5190E-05, 7.5186E-05, 1.4304E-04¸ 2.3624E-04, and 3.1934E-04, respectively.

Prior on V

Like prior for R, prior for V, the covariance matrix of η_t, is also inverse Wishart, denoted $V \sim IW_m(\varphi_V, B_V)$. Covariance matrix V is obtained by drawing V^{-1}, which is Wishart, and inverting it.

Prior for V^{-1} is Wishart distribution, denoted $V^{-1} \sim W_m(\varphi_V, B_V)$, where $m = 3$ and $\varphi_V = 10001$. The scale matrix parameter B_V has the same form as B_R:

$$B_V = 1000 * \begin{bmatrix} (\varphi_V \text{ var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\varphi_V \text{ var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\varphi_V \text{ var}(\eta_{3,t}))^{-1} \end{bmatrix},$$

where the diagonal is the inverse of φ_V multiplied by the inverse of error term variances in the state equation (3.15). The variance of $\eta_{1,t}$ is obtained by estimating regression (3.44), where the dependent variable and covariates are obtained using Campbell and Vuolteenaho's (2004) estimates of cash-flow news, discount-rate news, and conditional expectation of the excess log market return. The OLS estimate of the variance of $\eta_{1,t}$ is

1.7497E-05. According to restrictions imposed on the transition matrix F in (3.24), variances of $\eta_{2,t}$ and $\eta_{3,t}$ are equal to variances of $z_{2,t}$ and $z_{3,t}$, respectively. Furthermore, since $z_{2,t}$ is defined as $N_{CF,t}$, and $z_{3,t}$ is defined as $N_{DR,t}$, variances of $\eta_{2,t}$ and $\eta_{3,t}$ are equal to variances of $N_{CF,t}$ and $N_{DR,t}$, respectively. Estimated variances of $\eta_{2,t}$ and $\eta_{3,t}$ are obtained by calculating the variances of Campbell and Vuolteenaho's (2004) estimates of cash-flow news and discount-rate news, respectively. The estimated variance of $\eta_{2,t}$ is 4.0584E-04, and the estimated variance of $\eta_{3,t}$ is 1.9875E-03.

Gibbs sampling

Gibbs sampling is a technique for generating random variables from a distribution indirectly, without the need to calculate the density. Suppose there is a vector of random variables $\psi = [\psi_1,\ldots,\psi_d]'$ with a joint density $p(\psi_1,\ldots,\psi_d)$, and we are interested in obtaining random draws from $p(\psi_1,\ldots,\psi_d)$. In many cases, $p(\psi_1,\ldots,\psi_d)$ is difficult or even impossible to calculate, but the full conditional densities $p(\psi_j \mid \psi_{-j})$, $j = 1,\ldots,d$, where ψ_{-j} represents all the components of ψ except ψ_j, can be calculated and sampled from (see, for example, Casella and George 1992). The Gibbs sampler generates a sample from $p(\psi_1,\ldots,\psi_d)$ by sampling from the full conditional densities $p(\psi_j \mid \psi_{-j})$ successively, rather than sampling from $p(\psi_1,\ldots,\psi_d)$ directly. This is done by first specifying initial values for the vector ψ as ψ^0 and initializing the iteration counter $i = 1$, and then iteratively sampling from $p(\psi_j \mid \psi_{-j})$, $j = 1,\ldots,d$. At iteration i,

sample ψ_1^i from $p(\psi_1 \mid \psi_2^{i-1},\ldots,\psi_d^{i-1})$,

sample ψ_2^i from $p(\psi_2 \mid \psi_1^i, \psi_3^{i-1}, \ldots, \psi_d^{i-1})$,

$\vdots$

sample ψ_d^i from $p(\psi_d \mid \psi_1^i, \ldots, \psi_{d-1}^i)$;

set $i = i+1$ and repeat the above steps until convergence is reached.

Suppose the above steps are repeated for $i = N$ iterations and convergence is reached at iteration $i = K$, where $K < N$, then $\{\psi^i \mid i = K+1, \ldots, N\}$ is effectively a sample from $p(\psi_1, \ldots, \psi_d)$. The first K iterations are considered the burn-in period and are discarded.

Durbin and Koopman (2001, p. 159) note that it is not straightforward to develop appropriate convergence diagnostics. One way is to compute the potential scale reduction (i.e., Gelman-Rubin statistics) for all scalar quantities of interest (Gelman, Carlin, Stern, and Rubin 2004, pp. 294-297).

A general Gibbs sampling scheme

I design an MCMC algorithm that makes use of the techniques of Gibbs sampler, Metropolis-Hastings within Gibbs sampler, and Durbin and Koopman's (2002) simulation smoother to generate samples from the joint posterior density $p(z, \theta_H, \theta_F, R, V \mid y)$. The following Gibbs sampling scheme will be used:

1. Specify initial values for the set of unknown parameters θ as $\theta = \theta^0$ and initialize the iteration counter $i = 1$.
2. Sample z^i from $p(z \mid y, \theta_H^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$.
3. Sample $\beta_{1,M}^i$ from $p(\beta_{1,M} \mid y, z^i, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$.
4. Sample $\beta_{2,M}^i$ from $p(\beta_{2,M} \mid y, z^i, \theta_{H-\beta_{2,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$.

5. Sample $\beta_{3,M}^{i}$ from $p(\beta_{3,M} \mid y, z^{i}, \theta_{H-\beta_{3,M}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

6. Sample $\beta_{4,M}^{i}$ from $p(\beta_{4,M} \mid y, z^{i}, \theta_{H-\beta_{4,M}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

7. Sample $\beta_{5,M}^{i}$ from $p(\beta_{5,M} \mid y, z^{i}, \theta_{H-\beta_{5,M}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

8. Sample β_{b1}^{i} from $p(\beta_{b1} \mid y, z^{i}, \theta_{H-\beta_{b1}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

9. Sample β_{b2}^{i} from $p(\beta_{b2} \mid y, z^{i}, \theta_{H-\beta_{b2}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

10. Sample β_{b3}^{i} from $p(\beta_{b3} \mid y, z^{i}, \theta_{H-\beta_{b3}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

11. Sample β_{b4}^{i} from $p(\beta_{b4} \mid y, z^{i}, \theta_{H-\beta_{b4}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

12. Sample β_{b5}^{i} from $p(\beta_{b5} \mid y, z^{i}, \theta_{H-\beta_{b5}}^{i-1}, \theta_{F}^{i-1}, R^{i-1}, V^{i-1})$.

13. Sample F_{11}^{i} from $p(F_{11} \mid y, z^{i}, \theta_{H}^{i}, F_{12}^{i-1}, F_{13}^{i-1}, R^{i-1}, V^{i-1})$.

14. Sample F_{12}^{i} from $p(F_{12} \mid y, z^{i}, \theta_{H}^{i}, F_{11}^{i}, F_{13}^{i-1}, R^{i-1}, V^{i-1})$.

15. Sample F_{13}^{i} from $p(F_{13} \mid y, z^{i}, \theta_{H}^{i}, F_{11}^{i}, F_{12}^{i}, R^{i-1}, V^{i-1})$.

16. Sample $(R^{-1})^{i}$ from $p(R^{-1} \mid y, z^{i}, \theta_{H}^{i}, \theta_{F}^{i}, V^{i-1})$.

17. Invert $(R^{-1})^{i}$ to get R^{i}.

18. Sample $(V^{-1})^{i}$ from $p(V^{-1} \mid y, z^{i}, \theta_{H}^{i}, \theta_{F}^{i}, R^{i})$.

19. Invert $(V^{-1})^{i}$ to get V^{i}.

20. Set $i = i+1$.

Note that $\theta_{H-\beta_*}^{i-1}$ represents all the components of θ_H except β_*, at their current values. Namely,

$$\theta_{H-\beta_{1,M}}^{i-1} = \{\beta_{2,M}^{i-1}, \beta_{3,M}^{i-1}, \beta_{4,M}^{i-1}, \beta_{5,M}^{i-1}, \beta_{b1}^{i-1}, \beta_{b2}^{i-1}, \beta_{b3}^{i-1}, \beta_{b4}^{i-1}, \beta_{b5}^{i-1}\},$$

$$\theta_{H-\beta_{2,M}}^{i-1} = \{\beta_{1,M}^{i}, \beta_{3,M}^{i-1}, \beta_{4,M}^{i-1}, \beta_{5,M}^{i-1}, \beta_{b1}^{i-1}, \beta_{b2}^{i-1}, \beta_{b3}^{i-1}, \beta_{b4}^{i-1}, \beta_{b5}^{i-1}\},$$

$$\theta^{i-1}_{H-\beta_{3,M}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i-1}_{4,M}, \beta^{i-1}_{5,M}, \beta^{i-1}_{b1}, \beta^{i-1}_{b2}, \beta^{i-1}_{b3}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{4,M}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i-1}_{5,M}, \beta^{i-1}_{b1}, \beta^{i-1}_{b2}, \beta^{i-1}_{b3}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{5,M}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i-1}_{b1}, \beta^{i-1}_{b2}, \beta^{i-1}_{b3}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{b1}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i}_{5,M}, \beta^{i-1}_{b2}, \beta^{i-1}_{b3}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{b2}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i}_{5,M}, \beta^{i}_{b1}, \beta^{i-1}_{b3}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{b3}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i}_{5,M}, \beta^{i}_{b1}, \beta^{i}_{b2}, \beta^{i-1}_{b4}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{b4}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i}_{5,M}, \beta^{i}_{b1}, \beta^{i}_{b2}, \beta^{i}_{b3}, \beta^{i-1}_{b5}\},$$

$$\theta^{i-1}_{H-\beta_{b5}} = \{\beta^{i}_{1,M}, \beta^{i}_{2,M}, \beta^{i}_{3,M}, \beta^{i}_{4,M}, \beta^{i}_{5,M}, \beta^{i}_{b1}, \beta^{i}_{b2}, \beta^{i}_{b3}, \beta^{i}_{b4}\}.$$

I have written a code in S+ to implement the MCMC algorithm based on the above Gibbs sampling scheme. In the above sampling scheme, Step 2, sampling the state vectors z, is carried out using the simulation smoothing algorithm of Durbin and Koopman (2002), who have written a software, SsfPack, to do this. Steps 3 to 15, sampling the components of θ_H and θ_F, the unknown parameters in the measurement and transition matrices, are each carried out using Metropolis-Hastings within Gibbs sampler. Finally, sampling the covariance matrices R and V are carried out by first sampling R^{-1} and V^{-1} from their respective full conditional posterior densities which are Wishart distribution and then inverting them to get R and V, respectively.

Full conditional posterior densities for components in θ_H and θ_F

The full conditional posterior density for $\beta_{1,M}$, denoted $p(\beta_{1,M} \mid y, z, \theta_{H-\beta_{1,M}}, \theta_F, R, V)$, is, by definition, equal to

$$\frac{p(y,z,\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)}{p(y,z,\theta_{H-\beta_{1,M}},\theta_F,R,V)},$$

which is proportional to

$$p(y,z,\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)$$

because $p(y,z,\theta_{H-\beta_{1,M}},\theta_F,R,V)$ is the joint density of the conditioning information set and is treated as fixed. By Bayes' rule, $p(y,z,\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)$ is equal to

$$p(y \mid z,\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)\,p(z \mid \beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)\pi(\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V),$$

which can be written as

$$p(y \mid z,\theta_H,R)\,p(z \mid \theta_F,V)\pi(\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)$$

because equation (3.14) indicates that when $\{ z,\theta_H,R \}$ are given, the density of y is completely specified, and equation (3.15) indicates that when $\{ \theta_F,V \}$ are given, the density of z is completely specified. Since all components of θ are assumed to be *a priori* independent, $p(y \mid z,\theta_H,R)\,p(z \mid \theta_F,V)\pi(\beta_{1,M},\theta_{H-\beta_{1,M}},\theta_F,R,V)$ is equal to

$$p(y \mid z,\theta_H,R)\,p(z \mid \theta_F,V)\pi(\beta_{1,M})\pi(\theta_{H-\beta_{1,M}})\pi(\theta_F)\pi(R)\pi(V),$$

which is proportional to

$$p(y \mid z,\theta_H,R)\pi(\beta_{1,M})$$

because $\{ z,\theta_{H-\beta_{1,M}},\theta_F,R,V \}$ is the conditioning information set, and therefore, $p(z \mid \theta_F,V)$, $\pi(\theta_{H-\beta_{1,M}})$, $\pi(\theta_F)$, $\pi(R)$, and $\pi(V)$ are treated as fixed. Thus, the full conditional posterior density of $\beta_{1,M}$

$$p(\beta_{1,M} \mid y,z,\theta_{H-\beta_{1,M}},\theta_F,R,V) \propto p(y \mid z,\theta_H,R)\pi(\beta_{1,M}), \qquad (3.45)$$

where $p(y \mid z,\theta_H,R)$ is given in equation (3.33).

The full conditional posterior densities of $\beta_{2,M}$, $\beta_{3,M}$, $\beta_{4,M}$, $\beta_{5,M}$, β_{b1}, β_{b2}, β_{b3}, β_{b4}, and β_{b5} can be obtained in a similar way as that of $\beta_{1,M}$, and they also have forms similar to (3.45). The following are the full conditional posterior densities for these parameters:

$$p(\beta_{2,M} \mid y, z, \theta_{H-\beta_{2,M}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{2,M}), \qquad (3.46)$$

$$p(\beta_{3,M} \mid y, z, \theta_{H-\beta_{3,M}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{3,M}), \qquad (3.47)$$

$$p(\beta_{4,M} \mid y, z, \theta_{H-\beta_{4,M}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{4,M}), \qquad (3.48)$$

$$p(\beta_{5,M} \mid y, z, \theta_{H-\beta_{5,M}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{5,M}), \qquad (3.49)$$

$$p(\beta_{b1} \mid y, z, \theta_{H-\beta_{b1}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{b1}), \qquad (3.50)$$

$$p(\beta_{b2} \mid y, z, \theta_{H-\beta_{b2}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{b2}), \qquad (3.51)$$

$$p(\beta_{b3} \mid y, z, \theta_{H-\beta_{b3}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{b3}), \qquad (3.52)$$

$$p(\beta_{b4} \mid y, z, \theta_{H-\beta_{b4}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{b4}), \qquad (3.53)$$

$$p(\beta_{b5} \mid y, z, \theta_{H-\beta_{b5}}, \theta_F, R, V) \propto p(y \mid z, \theta_H, R)\pi(\beta_{b5}). \qquad (3.54)$$

The full conditional posterior density for F_{11}, denoted $p(F_{11} \mid y, z, \theta_H, F_{12}, F_{13}, R, V)$, is, by definition, equal to

$$\frac{p(y, z, \theta_H, F_{11}, F_{12}, F_{13}, R, V)}{p(y, z, \theta_H, F_{12}, F_{13}, R, V)},$$

which is proportional to

$$p(y, z, \theta_H, F_{11}, F_{12}, F_{13}, R, V)$$

because $\{ y, z, \theta_H, F_{12}, F_{13}, R, V \}$ is the conditioning information set, and thus, $p(y, z, \theta_H, F_{12}, F_{13}, R, V)$ is treated as fixed. By Bayes' rule and equations (3.14) and (3.15), $p(y, z, \theta_H, F_{11}, F_{12}, F_{13}, R, V)$ is equal to

$$p(y \mid z, \theta_H, R)\, p(z \mid \theta_F, V) \pi(\theta_H, F_{11}, F_{12}, F_{13}, R, V),$$

which is equal to

$$p(y \mid z, \theta_H, R)\, p(z \mid \theta_F, V) \pi(\theta_H) \pi(F_{11}) \pi(F_{12}) \pi(F_{13}) \pi(R) \pi(V),$$

since all components of θ are assumed to be *a priori* independent. Because $\{ y, z, \theta_H, F_{12}, F_{13}, R, V \}$ is the conditioning information set and y is observed, $p(y \mid z, \theta_H, R)$, $\pi(\theta_H)$, $\pi(F_{12})$, $\pi(F_{13})$, $\pi(R)$, and $\pi(V)$ are treated as fixed. Therefore, the full conditional posterior density of F_{11}

$$p(F_{11} \mid y, z, \theta_H, F_{12}, F_{13}, R, V) \propto p(z \mid \theta_F, V) \pi(F_{11}), \qquad (3.55)$$

where $p(z \mid \theta_F, V)$ was given in equation (3.40).

Similarly, the full conditional posterior density of F_{12} is

$$p(F_{12} \mid y, z, \theta_H, F_{11}, F_{13}, R, V) \propto p(z \mid \theta_F, V) \pi(F_{12}), \qquad (3.56)$$

and the full conditional posterior density of F_{13} is

$$p(F_{13} \mid y, z, \theta_H, F_{11}, F_{12}, R, V) \propto p(z \mid \theta_F, V) \pi(F_{13}). \qquad (3.57)$$

Metropolis-Hastings within Gibbs sampling for components of θ_H and θ_F

To illustrate Metropolis-Hastings within Gibbs sampler for components of θ_H and θ_F, I use Step 3 in the Gibbs sampling scheme as an example.

STEP 3 Sample $\beta_{1,M}^i$ from $p(\beta_{1,M} \mid y, z^i, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$.

Suppose that $\{ z^i, \beta_{1,M}^{i-1}, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1} \}$ has been generated, and the task is to generate $\beta_{1,M}^i$. The target density is $p(\beta_{1,M} \mid y, z, \theta_{H-\beta_{1,M}}, \theta_F, R, V)$ and is defined in (3.45). The proposal density, denoted by $q_{\beta_{1,M}}(\beta_{1,M}^* \mid z^i, \beta_{1,M}^{i-1}, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$, is the prior density of $\beta_{1,M}$. The following gives the algorithm for Metropolis-Hastings within Gibbs sampler for $\beta_{1,M}^i$.

(a) Draw $\beta_{1,M}^*$ from $q_{\beta_{1,M}}(\beta_{1,M}^* \mid z^i, \beta_{1,M}^{i-1}, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})$.

(b) Calculate the acceptance probability $\alpha(\beta_{1,M}^{i-1}, \beta_{1,M}^*)$, defined as

$$\min\left(1, \frac{p(\beta_{1,M}^* \mid y, z^i, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1}) q_{\beta_{1,M}}(\beta_{1,M}^{i-1} \mid z^i, \beta_{1,M}^*, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})}{p(\beta_{1,M}^{i-1} \mid y, z^i, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1}) q_{\beta_{1,M}}(\beta_{1,M}^* \mid z^i, \beta_{1,M}^{i-1}, \theta_{H-\beta_{1,M}}^{i-1}, \theta_F^{i-1}, R^{i-1}, V^{i-1})}\right).$$

(c) Draw a random number u from $U(0,1)$.

(d) If $\alpha(\beta_{1,M}^{i-1}, \beta_{1,M}^*) \geq u$, then set $\beta_{1,M}^i = \beta_{1,M}^*$, else set $\beta_{1,M}^i = \beta_{1,M}^{i-1}$.

Full conditional posterior distribution for R^{-1}

The full conditional posterior distribution for R^{-1} is Wishart distribution, denoted $(R^{-1} \mid y, z, \theta_H, \theta_F, V) \sim W_p(T + \varphi_R, (S_R + B_R^{-1})^{-1})$, where $S_R = \sum_{t=1}^{T} (y_t - Hz_t)(y_t - Hz_t)'$, and p, φ_R, and B_R are defined in the prior specification for R^{-1}.

Full conditional posterior distribution for V^{-1}

The full conditional posterior distribution for V^{-1} is Wishart distribution, denoted $(V^{-1} \mid y, z, \theta_H, \theta_F, R) \sim W_m(T + \varphi_V - 1, (S_V + B_V^{-1})^{-1})$, where $S_V = \sum_{t=1}^{T-1} (z_{t+1} - Fz_t)(z_{t+1} - Fz_t)'$, and m, φ_V, and B_V are defined in the prior specification for V^{-1}.

3.2.4 Simulation study

A simulation study is performed to check whether the S+ program written to implement the MCMC algorithm for the estimation of the state-space model is correct. The simulation study was done based on a state-space model of equations (3.14), (3.15), and (3.16) with H of the form

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 0 & H_{22} & -H_{22} \\ 0 & H_{32} & -H_{32} \\ 0 & H_{42} & -H_{42} \\ 0 & H_{52} & -H_{52} \\ 0 & H_{62} & -H_{62} \\ 0 & H_{72} & -H_{72} \\ 0 & 0 & -H_{83} \\ 0 & 0 & -H_{93} \\ 0 & 0 & -H_{10,3} \\ 0 & 0 & -H_{11,3} \end{bmatrix},$$

and F of the same form as equation (3.24), namely,

$$F = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that the form of H in the simulation study and the form of H in (3.23) are similar but with a slight difference, and the difference is in the seventh row of H. Although the form of H in the simulation study is slightly different from equation (3.23), this difference should not affect the value of the simulation study because the aim is to check whether the code is correct, and the simulation study serves this purpose.

I use the following criteria to check whether the program is correct: first, posterior means of the parameters should be close to the respective parameter values used to generate the data; second, when two data sets of different sample sizes are generated, analysis of the data set with the larger sample size is expected to result in better parameter estimates.

Two data sets of different samples sizes (i.e., T), one with a sample size of 500 and the other with a sample size of 50, are generated according to the state-space model of equations (3.14), (3.15), and (3.16) with the following parameters:

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1.33 & -1.33 \\ 0 & 1.01 & -1.01 \\ 0 & 0.97 & -0.97 \\ 0 & 1.01 & -1.01 \\ 0 & 0.85 & -0.85 \\ 0 & 0.81 & -0.81 \\ 0 & 0 & -0.2234 \\ 0 & 0 & -0.6281 \\ 0 & 0 & -1.0031 \\ 0 & 0 & -1.8321 \end{bmatrix},$$

$$F = \begin{bmatrix} 0.84 & 0.11 & -0.09 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$R = 0.001 * I_{11}$, and $V = 0.0001 * I_3$, where I_k denotes a $k \times k$ identity matrix. The mean vector and variance-covariance matrix of the initial state vector z_1 in (3.16) are assumed known and defined respectively as

$$a_1 = \begin{bmatrix} 0.5 \\ 0.01 \\ 0.01 \end{bmatrix},$$

and

$$P_1 = \begin{bmatrix} 0.0001 & 0 & 0 \\ 0 & 0.0006 & 0 \\ 0 & 0 & 0.0027 \end{bmatrix}.$$

The generated data set with a sample size of 500 is called SIMDATA1, and the generated data set with a sample size of 50 is called SIMDATA2. Both data sets are analyzed with the state-space model using the MCMC algorithm.

The prior specifications for the analyses of both data sets are the same. The prior distributions for the unknown parameters in H are $H_{22} \sim Beta(66.65, 76.17)$, $H_{32} \sim Beta(75.43, 132.18)$, $H_{42} \sim Beta(52.26, 112.78)$, $H_{52} \sim Beta(230.85, 461.7)$, $H_{62} \sim Beta(132.57, 324.56)$, $H_{72} \sim Beta(70.74, 194.54)$, $H_{83} \sim Beta(1.95, 22.39)$, $H_{93} \sim Beta(1.6, 6.39)$, $H_{10,3} \sim Beta(1.6, 3.16)$, $H_{11,3} \sim Beta(2.01, 1.36)$. The priors on F_{11}, F_{12}, and F_{13} are $F_{11} \sim U(0.7, 1)$, $F_{12} \sim N(0.1, 0.02^2)$, and $F_{13} \sim N(-0.08, 0.01^2)$, respectively.

The prior on R^{-1} is Wishart distribution, denoted $R^{-1} \sim W_p(\varphi_R, B_R)$, where $p = 11$ and $\varphi_R = 10000$. B_R is an 11×11 diagonal matrix of the form

$$B_R = \begin{bmatrix} (0.001*\varphi_R)^{-1} & 0 & \cdots & 0 \\ 0 & (0.001*\varphi_R)^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (0.001*\varphi_R)^{-1} \end{bmatrix}.$$

The prior on V^{-1} is Wishart distribution, denoted $V^{-1} \sim W_m(\varphi_V, B_V)$, where $m = 3$ and $\varphi_V = 100000$. The scale matrix B_V is of the form

$$B_V = \begin{bmatrix} (0.0001*\varphi_V)^{-1} & 0 & 0 \\ 0 & (0.0001*\varphi_V)^{-1} & 0 \\ 0 & 0 & (0.0001*\varphi_V)^{-1} \end{bmatrix}.$$

For each simulated data set, I run one chain with 100,000 iterations and use the latter 50,000 iterations to calculate the posterior distributions of all the parameters in the simulation study. Table 1 shows the descriptive statistics for the posterior distributions of the parameters in the state-space model used to analyze SIMDATA1 and SIMDATA2 in the simulation study. Posterior inferences from analyzing both SIMDATA1 and SIMDATA2 show that the posterior means of the parameters are close to the respective parameter values used to generate the data sets. For example, the posterior means of the diagonal elements of the covariance matrix V are very close to 0.0001, and the posterior means of the off-diagonal elements of V are much smaller than the diagonal elements and are close to 0. Compared with SIMDATA2 (sample size = 50), analysis of SIMDATA1 (sample size = 500) seems to result in better parameter estimates since the average mean square error is smaller[9]. Therefore, the results indicate that the S+ program

[9] Mean square error for a parameter is defined as the posterior variance + (posterior mean - true value)^2, where "true value" is the parameter value used to generate the data. The average mean square error is the average of mean square errors of the parameters.

written to implement the MCMC algorithm for the estimation of the state-space model is correct.

3.2.5 Data description

The data to be analyzed using the state-space model contain 11 observed variables, which are the excess log market return, log returns of the five Fama-French BE/ME portfolios, and log returns of the five Fama-Bliss discount bonds with maturities from one to five years. These are the information used to extract cash-flow news and discount-rate news from the unexpected market return.

Table 2 shows descriptive statistics for the observed variables that span the period 1952:7-2010:12. The observed variables are defined as follows. First, the excess log market return (r_M^e) is the difference between the log return on a value-weighted stock market index and the log risk-free rate. The data of the value-weighted stock market index are constructed by the Center for Research in Securities Prices (CRSP) and are the return on the value-weighted stock index (including dividends) composed of stocks listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the NASDAQ. The risk-free rate data are constructed by CRSP and are the yield of Treasury bills with approximately three months to maturity[10].

Second, the data on the returns of the five Fama-French BE/ME portfolios are downloaded from Professor Kenneth French's website[11]. The portfolios are formed on the ratio of book equity to market equity (BE/ME) at the end of each June using NYSE

[10] Data for the risk-free rate and data for the returns on the CRSP value-weighted stock index (NYSE/AMEX/NASDAQ) including distributions were downloaded from the Wharton Research Data Services (WRDS) website, https://wrds-web.wharton.upenn.edu/wrds/.
[11] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

BE/ME breakpoints. The BE/ME breakpoints are the 20th, 40th, 60th, and 80th NYSE percentiles (i.e., the breakpoints are quintiles). The book equity (BE) used in June of year t is the book equity for the last fiscal year end in year $t-1$. Market equity (ME) is price times shares outstanding at the end of December of year $t-1$. The portfolios are formed using all NYSE, AMEX, and NASDAQ stocks for which market equity for December of year $t-1$ and June of year t and BE for year $t-1$ is available. The five portfolio returns are log transformed and denoted r_1, r_2, r_3, r_4, and r_5 from the portfolio pertaining to the lowest BE/ME quintile to the one belonging to the highest BE/ME quintile.

Third, the data on the prices of the five Fama-Bliss discount bonds with maturities from one to five years are constructed by CRSP and are used to create the log returns data for discount bonds[12]. The log returns for discount bonds with maturities from one to five years are denoted r_1^b, r_2^b, r_3^b, r_4^b, and r_5^b, respectively.

Log returns of the 25 Fama-French portfolios are used to calculate cash-flow betas and discount-rate betas for each portfolio for the time period 1963:7-2001:12. The data for the returns of the 25 Fama-French portfolios are downloaded from Professor Kenneth French's website[13]. The portfolios, which are constructed at the end of each June, are the intersections of five portfolios formed on size (market equity, denoted ME) and five portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoints for year t are the NYSE market equity quintiles at the end of June of year t. BE/ME for June of year t is the book equity for the last fiscal year end in year

[12] Data for the prices of the five Fama-Bliss discount bonds are downloaded from the WRDS website, https://wrds-web.wharton.upenn.edu/wrds/.
[13] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

$t-1$ divided by the market equity for December of year $t-1$. The BE/ME breakpoints are NYSE quintiles. The portfolios for July of year t to June of year $t+1$ include all NYSE, AMEX, and NASDAQ stocks for which market equity data for December of year $t-1$ and June of year t and (positive) book equity data for year $t-1$ are available.

Prior knowledge based on Campbell and Vuolteenaho's (2004) estimates of cash-flow news, discount-rate news, and the conditional expectation of excess log market return is used to determine some of the prior distributions. Campbell and Vuolteenaho (2004) have kindly provided their estimates of the unobserved components of the return decomposition equation, which spans the period 1929:1-2001:12.

Chapter 4: Results

The posterior distributions are obtained using the MCMC algorithm, written based on the Gibbs sampling scheme, described in section 3.2.3 of Chapter 3. In a preliminary study, I calculated the Gelman-Rubin statistic[14] $\hat{R}$, following the suggestion in Gelman and Rubin (1992) and Gelman, Carlin, Stern, and Rubin (2004, pp. 294-297), for the important estimands based on five chains with different initial values. I found that all the $\hat{R}$s were less than 1.1 using the iterations between 150,000 and 300,000. Based on this experience and the fact that the study in this dissertation involves similar computations, I run one chain with 300,000 iterations and use the latter 150,000 iterations to calculate the posterior distributions of all the parameters[15] and state variables in this study. These posterior distributions form the basis for the inference in this dissertation.

For each iteration of the posterior distributions of cash-flow news and discount-rate news, I calculate cash-flow betas and discount-rate betas for the 25 Fama-French portfolios based on equations (3.9) and (3.10) for the period 1963:7-2001:12. For each portfolio, the 150,000 updates of cash-flow beta are considered a sample from the posterior distribution of cash-flow beta, and similarly, the 150,000 updates of discount-rate beta are treated as a sample from its posterior distribution. These posterior samples of cash-flow betas and discount-rate betas form the basis for inference. I find that growth stocks have larger cash-flow betas than do value stocks, which is consistent with the

[14] The purpose of the Gelman-Rubin statistic is to determine whether the chain has converged. It is generally considered that the values below 1.1 signify convergence.
[15] Descriptive statistics for the posterior distributions of parameters in the state-space model are in Table 13 in Appendix B.

findings of Chen and Zhao (2009), but not with those of Campbell and Vuolteenaho (2004).

This study also documents that growth stocks have larger discount-rate betas than do value stocks, and small stocks have both higher cash-flow betas and higher discount-rate betas than do large stocks, which are all consistent with the findings of both Campbell and Vuolteenaho (2004) and Chen and Zhao (2009). The finding that growth stocks have higher discount-rate betas than do value stocks can be explained by the fact that growth stocks have most of their cash flows in the future, which makes their durations longer and their sensitivity to discount-rate news higher. Similar conclusions can be drawn when I calculate cash-flow and discount-rate betas for the 1952:7-2010:12 period.

The discount-rate news estimated in this study and Campbell and Vuolteenaho's (2004) estimated discount-rate news are highly correlated. In addition, Campbell and Vuolteenaho (2004), Chen and Zhao (2009), and I all find the same patterns in discount-rate betas when comparing growth and value stocks or small and large stocks. Therefore, we all seem to estimate discount-rate news relatively precisely.

I find that cash-flow news and discount-rate news are strongly negatively correlated during 1952:7-2010:12, consistent with Campbell (1991) and Campbell (1996). This helps explain my result that growth stocks have larger cash-flow betas than do value stocks. Since discount-rate betas are based on the negative of discount-rate news (i.e., $(-1)*N_{DR}$), it can be expected that the pattern of cash-flow betas would approximately mirror that of discount-rate betas, and hence, growth stocks have higher cash-flow betas than do value stocks, which is exactly what I find, and so do Chen and Zhao (2009).

It is possible that when cash-flow news and discount-rate news are less correlated with each other, as in Campbell and Vuolteenaho (2004), growth stocks will have smaller cash-flow betas than will value stocks, and the value premium could be explained. To test this, I use orthogonalized cash-flow news and discount-rate news to calculate cash-flow betas and discount-rate betas for the 25 Fama-French portfolios. The resulting cash-flow betas from orthogonalized cash-flow news are all tightly distributed around zero, suggesting that the orthogonalized cash-flow news is noise and do not contribute any information.

A frequently asked question in any Bayesian analysis is how much do posterior inferences change when other reasonable models differing in the specification of the prior distribution are used in place of the present model. Sensitivity analysis can be used to assess the effect of alternative prior specifications on the posterior inferences. Throughout, I use the term "the original model" to refer to the model defined by the likelihood and prior distributions given in Chapter 3. In addition to the original model, I consider an alternative model with different specifications of the prior distributions so that the model is less informative than the original model.

Both the original model and the alternative model produce qualitatively very similar posterior inferences. Posterior inferences obtained from both models show that cash-flow news and discount-rate news are strongly negatively correlated, growth stocks have higher cash-flow betas and higher discount-rate betas than do value stocks, and small stocks have higher cash-flow betas and higher discount-rate betas than do large stocks. However, posterior inferences show that market betas for the 25 Fama-French portfolios are all less than one under the alternative model, which is inconsistent with the

basic knowledge in finance. On the other hand, market betas under the original model are larger than those obtained under the alternative model and are larger than one for some of the portfolios. The posterior inferences from the original model do not violate our knowledge about betas, but the posterior estimates of market betas obtained under the alternative model contradict our knowledge about betas. Therefore, I choose to present the original model and posterior inferences obtained under it.

To see whether the model is consistent with the data, I use posterior predictive check, following the suggestion in Gelman, Carlin, Stern, and Rubin (2004, pp. 159-177). Results from posterior predictive check do not indicate strong evidence that the model does not fit the data. This gives support for the posterior inferences obtained under the model.

In the following sections, I provide a summary of the behavior of cash-flow news and discount-rate news, estimated cash-flow betas and discount-rate betas, cash-flow and discount-rate betas calculated from orthogonalized cash-flow news and discount-rate news, the correlation between news estimated in this study and Campbell and Vuolteenaho's (2004) estimated news, the correlation between average returns and betas, sensitivity analysis, and some details of posterior predictive checking.

4.1 Summary of cash-flow news and discount-rate news

Table 3 summarizes the behavior of cash-flow news and discount-rate news for the period 1952:7-2010:12. The top section of Table 3 shows the variance-covariance matrix of cash-flow news and discount-rate news with variances on the diagonal and

covariance at the off-diagonal entries, and the middle section shows the correlation matrix with standard deviations on the diagonal and correlation at the off-diagonal entries.

Based on the restrictions on the transition matrix F in equation (3.24), cash-flow news and discount-rate news are error terms in the state equation, i.e., in equation (3.15), $z_{2,t}$ becomes $\eta_{2,t}$ and $z_{3,t}$ becomes $\eta_{3,t}$. Thus, by equations (3.15) and (3.24), the variance of cash-flow news is the variance of $\eta_{2,t}$, which is V_{22}, the variance of discount-rate news is the variance of $\eta_{3,t}$, which is V_{33}, and the covariance between cash-flow news and discount-rate news is the covariance between $\eta_{2,t}$ and $\eta_{3,t}$, which is V_{23}, where V_{ij} is the entry at the intersection of the i^{th} row and the j^{th} column of V, the variance-covariance matrix of the state equation error terms η_t.

The posterior distributions of V are obtained using the MCMC algorithm based on the Gibbs sampling scheme in section 3.2.3. For each iteration of the posterior distributions of V, I calculate the correlation between cash-flow news and discount-rate news using V_{22}, V_{33}, and V_{23}, namely, $cor(z_{2,t}, z_{3,t}) = V_{23} \Big/ \left(\sqrt{V_{22}} * \sqrt{V_{33}}\right)$. The 150,000 iterations of correlations between cash-flow news and discount-rate news are considered a sample from its posterior distribution. The variance-covariance matrix and correlation matrix reported in this table are based on posterior means. Results based on posterior medians are very similar to those based on posterior means, so they are not reported.

The bottom section of Table 3[16] shows that the correlation between cash-flow news and discount-rate news is tightly distributed around -0.94, consistent with Campbell

[16] Table 3 summarizes the behavior between cash-flow news and discount-rate news for the period 1952:7-2010:12. For the period 1963:7-2001:12, the correlation between cash-flow news and discount-rate news is distributed closely around -0.92, since the posterior minimum, first quartile, median, third quartile, and

(1991) and Campbell (1996) that these two types of news are negatively correlated. Campbell and Vuolteenaho (2004), on the other hand, find that cash-flow news and discount-rate news are slightly positively correlated with each other. Figure 1 illustrates the posterior means of cash-flow news and discount-rate news from the state space model for the period of 1952:7-2010:12.

One plausible explanation for the strong negative correlation between cash-flow news and discount-rate news is as follows. When economic conditions deteriorate, the expectations of future cash flows deteriorate and become more uncertain, i.e., there is a negative shock to expected cash flows, which likely increases the uncertainty of future cash flows and, therefore, leads to increase in risk and, thus, increase in expected returns (used to discount cash flows), which is the positive shock to the discount rates. The opposite occurs in the case of a positive shock to economic outlook. So, both cash-flow news and discount-rate news are heavily influenced by the business cycle, hence the strong negative correlation.

The middle section of Table 3 shows that discount-rate news has a standard deviation of 0.0096, larger than the 0.0021 standard deviation of cash-flow news. The finding that discount-rate news has a larger standard deviation than does cash-flow news is consistent with Campbell (1991) and Campbell and Vuolteenaho (2004), and it indicates that discount-rate news is the dominant component of the market return.

---

maximum of the correlation are -0.9631, -0.9221, -0.9201, -0.9180, and -0.9039, respectively. During 1963:7-2001:12, the posterior mean of the correlation between cash-flow news and discount-rate news is -0.9200.

4.2 Cash-flow betas and discount-rate betas

Table 4 summarizes the posterior distributions of cash-flow betas and discount-rate betas for the 25 Fama-French portfolios during 1963:7-2001:12. Panel A of Table 4 shows the posterior means and standard deviations of the cash-flow and discount-rate betas, Panel B shows the 95% posterior intervals, and Panel C reports the posterior medians. In each panel, the top section reports cash-flow betas and the bottom section reports discount-rate betas. The portfolios are organized in a five-by-five matrix with growth stocks at the left, value stocks at the right, small stocks at the top, and large stocks at the bottom.

The posterior means of cash-flow betas and discount-rate betas are all positive. The magnitudes of discount-rate betas are larger than those of cash-flow betas, which is probably because discount-rate news is the dominant component of the market return. The table also shows that growth stocks have higher cash-flow betas and higher discount-rate betas than do value stocks, and small stocks have higher cash-flow betas and higher discount-rate betas than do large stocks. The posterior distributions of the differences in betas between value and growth stocks and between large and small stocks are summarized in Tables 5 and 6, respectively.

The posterior standard deviations of cash-flow and discount-rate betas, reported in Panel A of Table 4, are quite small compared to the posterior means. For example, the standard deviation of cash-flow beta for the portfolio corresponding to the intersection of the extreme growth and extreme small portfolios is 0.008, which is 2.7% of its mean of 0.288. Panel C of Table 4 shows that the posterior medians of cash-flow betas and discount-rate betas are very similar to the posterior means reported in Panel A.

Panel B of Table 4 shows the 95% posterior intervals of cash-flow betas and discount-rate betas for the 25 Fama-French portfolios. The 95% posterior intervals of all cash-flow betas and discount-rate betas for all 25 portfolios are in the positive region and are quite narrow. For example, the 95% posterior interval of cash-flow beta of the portfolio corresponding to the intersection of the extreme growth and extreme small portfolios is (0.273, 0.304). Figures 2 and 3 illustrate the density plots of cash-flow betas and discount-rate betas, respectively, for the 25 Fama-French portfolios during 1963:7-2001:12.

Summaries of the posterior distributions of the differences in betas between the value and growth portfolios in each size category are reported in Table 5. For each iteration of cash-flow betas, I subtract cash-flow beta of a growth portfolio from that of a value portfolio and obtain 150,000 iterations of differences, which are treated as the posterior distribution of the difference in cash-flow beta between a value and a growth portfolio. This is done for each size category. The posterior distributions of the difference in discount-rate betas between value and growth portfolios are obtained similarly. The top section of the table shows the summaries of the posterior distribution for the difference in cash-flow betas between value and growth stocks, and the bottom section displays the summaries of the posterior distribution for the difference in discount-rate betas. The posterior distributions of the differences between value and growth stocks are summarized by the means, medians, standard deviations, 95% posterior intervals (determined by the 2.5% and 97.5% percentiles), and 50% posterior intervals (determined by the 25% and 75% quartiles).

The posterior means and medians of the difference in cash-flow betas between value and growth stocks are negative for all five size categories as shown in the top half of Table 5. The 95% posterior intervals, shown in the columns labeled "2.5%" and "97.5%", are in the negative region across all size categories. Variation of the differences between value and growth stocks seems to be small as indicated by the small standard deviations and narrow posterior intervals. The bottom half of Table 5 shows that the posterior means and medians of the difference in discount-rate betas between value and growth stocks are negative for all size categories. Furthermore, the 95% posterior intervals are in the negative region for all size groups. Like the differences in cash-flow betas, the differences in discount-rate betas seem to have little variation, as indicated by the small standard deviations and narrow posterior intervals.

In summary, in the 1963:7-2001:12 period, growth stocks have both higher cash-flow betas and higher discount-rate betas than do value stocks. Growth stocks have most of their cash-flows in the future, making their durations longer and their sensitivity to discount-rate news higher, and thus, their discount-rate betas are higher than those of value stocks. The strong negative correlation between cash-flow news and discount-rate news helps explain the observation that growth stocks have higher cash-flow betas than do value stocks. Since discount-rate betas are based on $(-1)*N_{DR}$, it can be expected that the pattern of cash-flow betas would approximately mirror that of discount-rate betas. Therefore, growth stocks also have higher cash-flow betas than do value stocks.

Table 6 summarizes the posterior distributions of the differences in betas between large and small portfolios in each size category. For each iteration of cash-flow betas, the cash-flow beta of a small portfolio is subtracted from that of a large portfolio to obtain

150,000 iterations of differences, which are treated as the posterior distribution of the difference in cash-flow beta between a large and a small portfolio. This is done for each BE/ME category. The posterior distributions of the difference in discount-rate betas between large and small portfolios are obtained similarly.

The top half of Table 6 shows that the posterior means and medians of the difference in cash-flow betas between large and small stocks are negative for all five BE/ME categories. The 95% posterior intervals are in the negative region and are narrow for all BE/ME groups. Standard deviations are small compared to the means, indicating that variation is small. The differences in discount-rate betas between large and small stocks are negative for all BE/ME groups as indicated by the means, medians, and 95% posterior intervals in the bottom half of Table 6. In summary, small stocks have higher cash-flow betas and higher discount-rate betas than do large stocks in the 1963:7-2001:12 period.

Similar conclusions can be drawn when I calculate cash-flow betas and discount-rate betas for the period of 1952:7-2010:12, and the results are shown in Tables 14 to 16 in Appendix B.

4.3 Cash-flow and discount-rate betas calculated from orthogonalized cash-flow and discount-rate news

Growth stocks have higher discount-rate betas than do value stocks, an observation that is generally accepted. Discount-rate betas are based on $(-1)*N_{DR}$, and therefore, if cash-flow news and discount-rate news are negatively correlated, then growth stocks will have higher cash-flow betas than do value stocks, which is what Chen

and Zhao (2009) and this study find. On the other hand, if cash-flow news and discount-rate news are positively correlated, then growth stocks will have lower cash-flow betas than do value stocks, which is what Campbell and Vuolteenaho (2004) find. In fact, using the estimated cash-flow news and discount-rate news from Campbell and Vuolteenaho (2004), I find that the correlation between Campbell and Vuolteenaho's (2004) estimated cash-flow news and discount-rate news is 0.29 in the 1963:7-2001:12 period. The results in all three studies seem to be affected by the correlation between cash-flow news and discount-rate news. Depending on whether the correlation is negative or positive, the pattern of cash-flow betas could mirror or be the opposite of the pattern of discount-rate betas. This motivates me to remove the correlation by orthogonalizing cash-flow news by discount-rate news and to look at the new betas calculated from orthogonalized cash-flow and discount-rate news.

To orthogonalize cash-flow news and discount-rate news, I run a regression with cash-flow news as the dependent variable and discount-rate news as the covariate for each month of cash-flow and discount-rate news. Namely, for a fixed t, I run the regression

$$N_{CF,t,i} = b_1 + b_2 N_{DR,t,i} + e_{t,i}, \qquad (4.1)$$

where $i = 1,\ldots,150000$ and represents the iterations of the posterior distributions of cash-flow news and discount-rate news. Regression (4.1) is run for every t, where $t = 1,\ldots,702$ and represents the 702 months in the 1952:7-2010:12 time period. The residuals $\{e_{t,i}; t = 1,\ldots,702 \text{ and } i = 1,\ldots,150000\}$ estimated from the regressions should be uncorrelated with discount-rate news and are treated as the posterior distribution of orthogonalized cash-flow news. Posterior distribution of discount-rate news is the same

as before, namely, it is obtained from the MCMC algorithm that is based on the Gibbs sampling scheme in section 3.2.3.

I substitute orthogonalized cash-flow news for cash-flow news and follow equations (3.9) and (3.10) to calculate cash-flow betas and discount-rate betas for the 25 Fama-French portfolios for the 1963:7-2001:12 period. Note that discount-rate betas are different from those obtained earlier because the denominator in (3.10) now becomes the variance of the difference between orthogonalized cash-flow news and discount-rate news. Let $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ denote cash-flow betas and discount-rate betas obtained from orthogonalized cash-flow news and discount-rate news, respectively. The 150,000 iterations of cash-flow betas obtained from orthogonalized cash-flow news for portfolio i, denoted $\hat{\beta}_{i,OCF}$, is treated as the posterior distribution of $\hat{\beta}_{i,OCF}$, and the 150,000 iterations of new discount-rate betas for portfolio i, denoted $\hat{\beta}_{i,ODR}$, is considered the posterior distribution of $\hat{\beta}_{i,ODR}$.

Tables 7 to 10 summarize the posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ for the 25 Fama-French portfolios during the 1963:7-2001:12 period. Similar as before, the portfolios are organized in a five-by-five matrix with growth stocks at the left, value stocks at the right, small stocks at the top, and large stocks at the bottom.

The top half of Table 7 shows that the posterior means of $\hat{\beta}_{OCF}$ for all 25 portfolios are negative and close to zero. Posterior medians of $\hat{\beta}_{OCF}$ are displayed in the top half of Table 8 and are also negative and very small, but a bit larger in absolute values than the posterior means. For example, the posterior mean of $\hat{\beta}_{OCF}$ for the portfolio corresponding to the intersection of the extreme growth and extreme small

portfolios is -2.99E-05 , and the posterior median for this portfolio is -2.10E-04 . Posterior standard deviations of $\hat{\beta}_{OCF}$ are shown in Table 7 and are large compared to the posterior means. For instance, the posterior standard deviation of $\hat{\beta}_{OCF}$ for the portfolio corresponding to the intersection of the extreme growth and extreme small portfolios is 0.0089, which is 298 times the absolute value of the posterior mean.

The bottom half of Table 8 shows that the 95% posterior intervals of $\hat{\beta}_{OCF}$ for all 25 Fama-French portfolios cover both negative and positive values and are all within the $(-0.02, 0.02)$ range. For example, the 95% posterior interval of $\hat{\beta}_{OCF}$ for the portfolio corresponding to the intersection of the extreme growth and extreme small portfolios is $(-0.017, 0.018)$. To sum up, the posterior distributions of $\hat{\beta}_{OCF}$ cover both negative and positive values, seem to be almost symmetric around zero, and have extremely small absolute values, suggesting that the orthogonalized cash-flow news is noise which does not contribute any information.

The top half of Table 9 summarizes the posterior distribution of the difference in $\hat{\beta}_{OCF}$ between value and growth portfolios in each size category. The posterior distribution is obtained by calculating the difference in $\hat{\beta}_{OCF}$ between value and growth portfolios in a size group for each iteration, and the resulting 150,000 iterations of the differences are considered the posterior distribution. The table shows that the posterior means and medians of the differences in $\hat{\beta}_{OCF}$ between value and growth stocks are positive and very small across all five size groups, and the 95% posterior intervals are very narrow and range from negative to positive values. This is because the posterior

distributions of $\hat{\beta}_{OCF}$ are very small and range from negative to positive values, as discussed previously.

The top half of Table 10 summarizes the posterior distribution of the difference in $\hat{\beta}_{OCF}$ between large and small portfolios in each BE/ME category. For each iteration, I calculate the difference in $\hat{\beta}_{OCF}$ between large and small portfolios in each BE/ME category, and the resulting 150,000 iterations of differences are treated as the posterior distribution. The posterior means and medians are positive but small, and the 95% posterior intervals are narrow and cover both negative and positive values. This is because the posterior distributions of $\hat{\beta}_{OCF}$ are very small and range from negative to positive values.

The posterior distributions of $\hat{\beta}_{ODR}$ are summarized in Tables 7 to 10. These tables show that $\hat{\beta}_{ODR}$ and discount-rate betas share similar patterns. Namely, Tables 7 and 8 show that growth stocks have higher $\hat{\beta}_{ODR}$ than do value stocks and small stocks have higher $\hat{\beta}_{ODR}$ than do large stocks. Posterior means and medians of $\hat{\beta}_{ODR}$ are about the same but are larger than those of discount-rate betas. For example, the posterior mean of $\hat{\beta}_{ODR}$ for the portfolio corresponding to the intersection of the extreme growth and extreme small portfolios is 2.021, while the posterior mean of discount-rate beta for this portfolio is 1.384 (see Table 4). Furthermore, Table 8 shows that the 95% posterior intervals of $\hat{\beta}_{ODR}$ are all in the positive region, and Table 7 shows that posterior standard deviations of $\hat{\beta}_{ODR}$ are small compared to the posterior means. It can be expected that

$\hat{\beta}_{ODR}$ and discount-rate betas have similar patterns because they only differ in the denominator in equation (3.10), and therefore, they are proportional to each other.

The bottom half of Table 9 summarizes the posterior distributions of the difference in $\hat{\beta}_{ODR}$ between value and growth stocks. The posterior means and medians are negative, and the 95% posterior intervals are in the negative region. The posterior distributions of the difference in $\hat{\beta}_{ODR}$ between large and small stocks summarized in the bottom half of Table 10 show that the posterior means and medians are negative and that the 95% posterior intervals are in the negative region. In summary, growth stocks have higher $\hat{\beta}_{ODR}$ than do value stocks, and small stocks have higher $\hat{\beta}_{ODR}$ than do large stocks, similar to the patterns of discount-rate betas.

Similar conclusions can be drawn when I calculate cash-flow betas and discount-rate betas using orthogonalized cash-flow news and discount-rate news for the period of 1952:7-2010:12, and the results are shown in Tables 17 to 20 in Appendix B.

4.4 Correlation between news estimated using the state-space model and Campbell and Vuolteenaho's (2004) estimated news

Discount-rate betas calculated using news estimated from the state-space model and those calculated with news estimated from Campbell and Vuolteenaho's (2004) VAR model share the same patterns, i.e., growth stocks have higher discount-rate betas than do value stocks, and small stocks have higher discount-rate betas than do large stocks. It would be interesting to see how the news estimated from these two models correlate with each other.

As mentioned in the beginning of this chapter, I implemented the MCMC algorithm to obtain 300,000 iterations of cash-flow news and discount-rate news for each month in the 1952:7-2010:12 period and treat the latter 150,000 iterations as posterior distributions of cash-flow news and discount-rate news, respectively. Since Campbell and Vuolteenaho (2004) do not use a Bayesian model, they do not produce posterior distributions of cash-flow news and discount-rate news, and their estimated news are point estimates for every month during their sample time period of 1929:1-2001:12.

Therefore, I obtain the distribution of empirical correlation between news estimated from the Bayesian state-space model and Campbell and Vuolteenaho's (2004) estimated news for the 1952:7-2001:12 period (a total of 594 months). That is, for each iteration, I calculate the correlation between cash-flow news estimated from the state-space model and cash-flow news from Campbell and Vuolteenaho (2004) using the 594 pairs of cash-flow news and get 150,000 iterations of correlations. The 150,000 iterations of correlations are considered the posterior distribution of the correlation between the estimated cash-flow news in this study and the estimated cash-flow news in Campbell and Vuolteenaho (2004). The distribution of correlation between the estimated discount-rate news in this study and the estimated discount-rate news in Campbell and Vuolteenaho (2004) is obtained similarly.

The posterior distributions of the correlations between the estimated news in this dissertation and the estimated news in Campbell and Vuolteenaho (2004) are summarized in Table 11. The correlation between the estimated discount-rate news in this study and the estimated discount-rate news in Campbell and Vuolteenaho (2004) has a posterior

mean of 0.82 and has a dense distribution, according to the small range and narrow interquartile range.

The table also shows that the correlation between the estimated total news (i.e., $N_{CF} - N_{DR}$) in this study and the estimated total news in Campbell and Vuolteenaho (2004) is tightly distributed around 0.93. The posterior mean of the correlation between the estimated cash-flow news in this dissertation and the estimated cash-flow news in Campbell and Vuolteenaho (2004) is 0.18. The posterior mean of the correlation between the estimated conditional expectation of excess log market return (i.e., $E_{t-1} r^e_{M,t}$) in this study and that in Campbell and Vuolteenaho (2004) is 0.1.

To sum up, there is strong positive correlation between the estimated discount-rate news in this study and that in Campbell and Vuolteenaho (2004) and strong positive correlation between the estimated total news in this study and that in Campbell and Vuolteenaho (2004). The correlation between the estimated cash-flow news in this study and that in Campbell and Vuolteenaho (2004) is also positive but weaker.

4.5 Correlation between returns and betas

The purpose of the analysis in this section is to show that the posterior inferences in this study do not violate the economic intuition described below. Positive shocks and negative shocks to expected cash flows should have different effect on the correlation between contemporaneous returns and cash-flow betas of securities. When there is a positive shock to expected cash flows, investors' expectations of future cash flows improve and contemporaneous returns of securities are generally higher and likely to be positive. By the definition of cash-flow betas in equation (3.5), securities that have

higher cash-flow betas are more sensitive to cash-flow news. Therefore, when there is a positive shock to expected cash flows, securities with higher cash-flow betas are likely to have higher contemporaneous returns than are those with lower cash-flow betas, which suggests that contemporaneous returns and cash-flow betas are likely to be positively correlated.

When there is a negative shock to expected cash flows, investors' expectations of future cash flows deteriorate and contemporaneous returns of securities are generally lower and likely to be negative. Thus, securities with higher cash-flow betas are likely to have lower contemporaneous returns than are those with lower cash-flow betas, which suggests that contemporaneous returns and cash-flow betas are likely to be negatively correlated.

Similarly, positive shocks and negative shocks to discount rates should have different effect on the correlation between contemporaneous returns and discount-rate betas of securities. When there is a positive shock to the negative of future discount rates (i.e., $(-1)*N_{DR}$), contemporaneous returns of securities are generally higher and likely to be positive. By the definition of discount-rate betas in equation (3.6), securities that have higher discount-rate betas are more sensitive to discount-rate news. Thus, when there is a positive shock to the negative of future discount rates, securities with higher discount-rate betas are likely to have higher contemporaneous returns than do those with lower discount-rate betas, suggesting that contemporaneous returns and discount-rate betas are likely to be positively correlated.

On the other hand, when there is a negative shock to the negative of future discount rates, contemporaneous returns of securities are generally lower and likely to be

negative. Therefore, securities with higher discount-rate betas are likely to have lower contemporaneous returns than do those with lower discount-rate betas, suggesting that contemporaneous returns and discount-rate betas are likely to be negatively correlated.

The top half of Table 12 shows the correlations between average returns and cash-flow betas of the 25 Fama-French portfolios for the period of 1963:7-2001:12. For each portfolio, I calculate the average return based on the months in which cash-flow news is positive and the average return based on the months in which cash-flow news is negative. The estimated cash-flow news series is based on the posterior means of cash-flow news for each month. Posterior means of cash-flow betas are used to calculate the correlation between average returns and cash-flow betas. The table shows that when there is a positive shock to future cash flows, the correlation between average returns and cash-flow betas of the 25 Fama-French portfolios is 0.89, and when there is a negative shock to future cash flows, the correlation is -0.86.

The bottom half of Table 12 shows the correlations between average returns and discount-rate betas of the 25 Fama-French portfolios for the period of 1963:7-2001:12. For each portfolio, the average return based on the months in which $(-1)*N_{DR}$ is positive and the average return based on the months in which $(-1)*N_{DR}$ is negative are calculated. The estimated discount-rate news series is based on the posterior means of discount-rate news for each month. Posterior means of discount-rate betas are used to calculate the correlation between average returns and discount-rate betas. The table shows that the correlation between average returns and discount-rate betas of the 25 Fama-French portfolios is 0.87 when there is a positive shock to $(-1)*N_{DR}$ and -0.85 when there is a negative shock to $(-1)*N_{DR}$.

4.6 Sensitivity analysis

To assess the effect of alternative prior specifications on the posterior inferences, a sensitivity analysis, which compares an alternative model with the original model, is performed. The alternative model is less informative than the original model.

The difference between the alternative model and the original model is in the prior specification of R^{-1} and V^{-1}, which are the inverses of the covariance matrices of the error terms ε_t in equation (3.14) and η_t in equation (3.15), respectively. In the alternative model, priors for R^{-1} and V^{-1} are specified such that the mean of R, i.e., $E(R)$, and the mean of V, i.e., $E(V)$, become larger than they are in the original model. By making both $E(R)$ and $E(V)$ larger, the alternative model is less informative than the original model.

Recall that in the original model, the prior distribution for R^{-1} is a Wishart distribution, denoted $R^{-1} \sim W_p(\varphi_R, B_R)$, where $p = 11$, $\varphi_R = 10001$, and B_R is a diagonal matrix defined as

$$B_R = 1000 * \begin{bmatrix} (\varphi_R \operatorname{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\varphi_R \operatorname{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\varphi_R \operatorname{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix},$$

which is approximately equal to

$$\frac{1}{10} * \begin{bmatrix} (\operatorname{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\operatorname{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\operatorname{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix}. \qquad (4.2)$$

The variances of the error terms $\{\varepsilon_{1,t},\ldots,\varepsilon_{11,t}\}$ are obtained by estimating regressions (3.19) and (3.22), and the estimates of the variances are given in Chapter 3. Because $R^{-1} \sim W_p(\varphi_R, B_R)$, the mean of R^{-1}, denoted $E(R^{-1})$, is $\varphi_R B_R$. Therefore, using $\varphi_R = 10001$ and equation (4.2), in the original model, $E(R^{-1})$ is approximately

$$1000 * \begin{bmatrix} (\text{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\text{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\text{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix}. \qquad (4.3)$$

In the alternative model, I assume a less informative Wishart prior distribution for R^{-1} by setting $\varphi_R = 701$ and

$$B_R = 100 * \begin{bmatrix} (\varphi_R \text{ var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\varphi_R \text{ var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\varphi_R \text{ var}(\varepsilon_{11,t}))^{-1} \end{bmatrix},$$

which is approximately equal to

$$\frac{1}{7} * \begin{bmatrix} (\text{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\text{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\text{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix}. \qquad (4.4)$$

In the alternative model, using $\varphi_R = 701$ and equation (4.4), $E(R^{-1})$ approximately equals

$$100 * \begin{bmatrix} (\text{var}(\varepsilon_{1,t}))^{-1} & 0 & \cdots & 0 \\ 0 & (\text{var}(\varepsilon_{2,t}))^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\text{var}(\varepsilon_{11,t}))^{-1} \end{bmatrix}. \qquad (4.5)$$

A Comparison of equations (4.3) and (4.5) shows that $E(R^{-1})$ is smaller in the alternative model than it is in the original model, which suggests that, generally speaking, $E(R)$ is larger in the alternative model than it is in the original model.

Similarly, for the Wishart prior of V^{-1} in the alternative model, the degrees of freedom and the scale matrix are set such that the alternative model is less informative than the original model. Recall that in the original model, the prior distribution for V^{-1} is a Wishart distribution, denoted $V^{-1} \sim W_m(\varphi_V, B_V)$, where $m = 3$, $\varphi_V = 10001$, and

$$B_V = 1000 * \begin{bmatrix} (\varphi_V \text{ var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\varphi_V \text{ var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\varphi_V \text{ var}(\eta_{3,t}))^{-1} \end{bmatrix},$$

which is approximately

$$\frac{1}{10} * \begin{bmatrix} (\text{var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\text{var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\text{var}(\eta_{3,t}))^{-1} \end{bmatrix}. \qquad (4.6)$$

The estimated variances of $\{\eta_{1,t}, \eta_{2,t}, \eta_{3,t}\}$ are given in Chapter 3. Since $V^{-1} \sim W_m(\varphi_V, B_V)$, the mean of V^{-1}, denoted $E(V^{-1})$, is $\varphi_V B_V$. Thus, using $\varphi_V = 10001$ and equation (4.6), $E(V^{-1})$ in the original model is approximately equal to

$$1000 * \begin{bmatrix} (\text{var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\text{var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\text{var}(\eta_{3,t}))^{-1} \end{bmatrix}. \qquad (4.7)$$

In the alternative model, a less informative Wishart prior distribution for V^{-1} is assumed by setting $\varphi_V = 701$ and

$$B_V = 100 * \begin{bmatrix} (\varphi_V \text{ var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\varphi_V \text{ var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\varphi_V \text{ var}(\eta_{3,t}))^{-1} \end{bmatrix},$$

which approximately equals

$$\frac{1}{7} * \begin{bmatrix} (\text{var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\text{var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\text{var}(\eta_{3,t}))^{-1} \end{bmatrix}. \qquad (4.8)$$

Using $\varphi_V = 701$ and equation (4.8), $E(V^{-1})$ in the alternative model approximately equals

$$100 * \begin{bmatrix} (\text{var}(\eta_{1,t}))^{-1} & 0 & 0 \\ 0 & (\text{var}(\eta_{2,t}))^{-1} & 0 \\ 0 & 0 & (\text{var}(\eta_{3,t}))^{-1} \end{bmatrix}. \qquad (4.9)$$

A comparison of equations (4.7) and (4.9) shows that $E(V^{-1})$ in the alternative model is smaller than $E(V^{-1})$ in the original model. Therefore, generally speaking, $E(V)$ is larger in the alternative model than it is in the original model.

Both $E(R)$ and $E(V)$ are larger in the alternative model than they are in the original model. Therefore, the alternative model is less informative than the original model.

Qualitatively, both the original model and the alternative model produce very similar posterior inferences[17]. Posterior inferences obtained from both models show that cash-flow news and discount-rate news are strongly negatively correlated, growth stocks have higher cash-flow betas and higher discount-rate betas than do value stocks, and

[17] In fact, I considered other prior specifications in addition to the alternative prior specification presented in this dissertation, and the posterior inferences are all qualitatively similar.

small stocks have higher cash-flow betas and higher discount-rate betas than do large stocks.

Between the original model and the alternative model, the only difference in posterior inferences is that the market betas for the 25 Fama-French portfolios are all less than one under the alternative model, whereas market betas under the original model are all larger and for some of the portfolios larger than one. The result that market betas for some portfolios are larger than one is consistent with the finance literature. It is a reasonable result because the market beta of the stock market is one and some portfolios are riskier than the market, so the market betas for these portfolios should be greater than one. The posterior inferences from the original model do not violate our knowledge about betas, but the posterior estimates of market betas obtained under the alternative model contradict our knowledge about betas. Therefore, I choose to present the original model and posterior inferences obtained under it.

4.7 Model checking: Posterior predictive check

Does the model fit the data? If not, then one may have cause to cast doubt on the inferences obtained from the model. Following the suggestion in Gelman, Carlin, Stern, and Rubin (2004, pp. 159-177), I use posterior predictive check to see whether the model is consistent with data.

The posterior predictive check is the comparison between the realized test quantities and the predictive test quantities. One way to compare the two is to calculate the posterior predictive p-value, defined as the probability that the replicated data could be more extreme than the observed data, as measured by the test quantity. A test quantity

is a scalar summary of parameters and data that is used as a standard when comparing data to simulations from the posterior predictive distribution. Examples of test quantities include sample mean and sample median, among others.

The posterior predictive distribution is the distribution of replicated data, data that could have been observed if the experiment that produced the observed data were replicated with the same model and the same value of parameters that produced the observed data, given the observed data. A realized test quantity is a test quantity calculated from observed data, and a predictive test quantity is a test quantity calculated from simulations from the posterior predictive distribution.

To obtain simulations from the posterior predictive distribution, I draw 1000 iterations of parameters and state variables $\{\theta, z\}$ from the joint posterior distribution of 150,000 iterations and calculate one set of replicated data from each of the 1000 draws, where a set consists of replicated data for the 11 observed variables in the state-space model. I calculate a test quantity for each of the 11 observed variables and each of the 11 replicated variables. The test quantity that I use is the sample mean of each variable taken over the 702 months for the period 1952:7-2010:12. The estimated posterior predictive p-value is the proportion of the 1000 simulations for which the predictive test quantity equals or exceeds the realized test quantity. A posterior predictive p-value is obtained for every observable variable, so there are a total of 11 p-values.

Eight of the eleven posterior predictive p-values for the original model fall within the reasonable range of 0.025 to 0.975. Therefore, there is no strong evidence that the model does not fit the data. For the alternative model, all of the eleven posterior

predictive p-values lie within the range of 0.025 to 0.975. Results of posterior predictive check suggest that both models seem to be consistent with the data.

Chapter 5: Conclusions

Based on the return decomposition, Campbell and Vuolteenaho (2004) propose the two-beta intertemporal capital asset pricing model (ICAPM) that captures a stock's risk with cash-flow news and discount-rate news separately. This model captures the intuition that risk-averse investors price cash-flow risk higher than discount-rate risk and implies that higher average stock returns are explained by their higher cash-flow betas. The model implies that the anomaly that value stocks have higher average returns than growth stocks can be explained by the higher cash-flow betas of value stocks. Previous studies use vector autoregressive (VAR) models with ad-hoc variables to implement the return decomposition and yield conflicting results regarding the pattern of cash-flow betas from growth to value stocks.

My dissertation adds to the literature by proposing a state-space model approach to implement the return decomposition. Importantly, the state-space model approach has two advantages over the VAR approach. First, the state-space approach directly and simultaneously models both cash-flow news and discount-rate news by expressing the relation between returns and news in terms of linear regressions, whereas the VAR approach models discount-rate news and backs out cash-flow news as the residual between the unexpected market return and discount-rate news. Second, the information that I use to extract cash-flow news and discount-rate news is directly related to the market return, whereas the information used in the VAR approach is not.

Restrictions derived from the return decomposition are imposed on the parameters of the state-space model to ensure identification of the three state variables: cash-flow news, discount-rate news, and the conditional expectation of the excess log market return.

Since state-space models are high-dimensional, a Bayesian Markov chain Monte Carlo (MCMC) methodology is used to estimate the state-space model in this dissertation. I design an MCMC algorithm that uses the techniques of Gibbs sampler, Metropolis-Hastings within Gibbs sampler, and Durbin and Koopman's (2002) simulation smoother to generate samples from the joint posterior distribution.

Posterior inferences obtained under the state-space model indicate that patterns of discount-rate betas from growth to value stocks as well as patterns of cash-flow betas and discount-rate betas from small to large stocks are consistent with the findings of both Campbell and Vuolteenaho (2004) and Chen and Zhao (2009). However, unlike Campbell and Vuolteenaho (2004), I do not find that value stocks have higher cash-flow betas than do growth stocks. This finding is consistent with Chen and Zhao (2009).

While I document that cash-flow news and discount-rate news are strongly negatively correlated, Campbell and Vuolteenaho (2004) find that the two components of unexpected market return are slightly positively correlated. However, the negative correlation between cash-flow news and discount-rate news is also found in Campbell (1991), Campbell (1996), and Chen and Zhao (2009).

The economic intuition behind the strong negative correlation between cash-flow news and discount-rate news is as follows. When economic conditions deteriorate, investors' expectations of future cash flows deteriorate, leading to a negative shock to expected cash flows, which likely increases the uncertainty of future cash flows. This

increase in the uncertainty of future cash flows results in an increase in risk and, consequently, increase in expected returns (used to discount cash flows), which is a positive shock to discount rates. The opposite occurs in the case of improving economic conditions. Since both cash-flow news and discount-rate news are heavily influenced by the business cycle, the correlation is strongly negative.

The strong negative correlation between cash-flow news and discount-rate news, coupled with the fact that discount-rate betas are based on the negative of discount-rate news, suggests that the pattern of cash-flow betas would approximately mirror that of discount-rate betas. Therefore, the strong negative correlation between cash-flow news and discount-rate news helps to explain the finding that growth stocks have higher cash-flow betas than do value stocks.

Campbell and Vuolteenaho's (2004) motivation for measuring cash-flow risk and discount-rate risk separately by defining cash-flow beta and discount-rate beta is implied in their paper (pp. 1256-1257), in which they find that cash-flow news and discount-rate news are only slightly positively correlated with each other. They say that the weak correlation "raises the possibility that different types of stocks may have different betas with the two components of the market". To test this, I use orthogonalized cash-flow news and discount-rate news to calculate cash-flow betas and discount-rate betas for the 25 Fama-French portfolios for the 1963:7-2001:12 period. I find that cash-flow betas from orthogonalized cash-flow news for all of the 25 portfolios are tightly distributed around zero, suggesting that the orthogonalized cash-flow news is noise and does not contribute any information.

There are several directions for improving this study. First, more restrictions can be added to the state-space model to improve identification of cash-flow news. Second, it would be useful to assess the robustness of my findings by using the 25 Fama-French portfolios instead of the five Fama-French BE/ME portfolios. Third, since the mean vector and variance-covariance matrix of the initial state vector are unknown, assuming prior distributions for them would likely improve the model. Finally, it would be beneficial to fine-tune the current prior specifications of the state-space model so that it becomes a hierarchical model, as discussed in Gelman, Carlin, Stern, and Rubin (2004, Chapter 5), in which parameters in the measurement matrix pertaining to stocks will be given a probabilistic specification in terms of further parameters (i.e., hyperparameters), and parameters pertaining to bonds will also be given a probabilistic specification in terms of hyperparameters.

References

Ang, A., and J. Chen, 2007, CAPM over the long run: 1926-2001, *Journal of Empirical Finance*, 14, 1, 1-40.

Antoncic, M., 1986, High and volatile real interest rates: Where does the Fed fit in?, *Journal of Money, Credit, and Banking*, 18, 18-27.

Ball, R., 1978, Anomalies in Relationships between Securities' Yields and Yield-Surrogates, *Journal of Financial Economics*, 6, 103-126.

Banz, R., 1981, The relation between return and market value of common stocks, *Journal of Financial Economics*, 9, 3-18.

Basu, S., 1977, The investment performance of common stocks in relation to their price to earnings ratios: A test of the efficient market hypothesis, *Journal of Finance*, 32, 663-682.

Basu, S., 1983, The Relationship between Earnings Yield, Market Value, and Return for NYSE Common Stocks: Further Evidence, *Journal of Financial Economics*, 12, 129-156.

Bhandari, L. C., 1988, Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence, *Journal of Finance*, 43, 507-528.

Black, F., 1972, Capital market equilibrium with restricted borrowing, *Journal of Business*, 45, July, 444-454.

Black, F., M. Jensen, and M. Scholes, 1972, The Capital Asset Pricing Model: Some empirical tests, in Jensen, M. (ed.), *Studies in the Theory of Capital Markets*, Praeger, New York.

Blume, M., and I. Friend, 1973, A new look at the Capital Asset Pricing Model, *Journal of Finance*, 28, 19-33.

Brennan, M. J., A.W. Wang, and Y. Xia, 2004, Estimation and test of a simple model of intertemporal capital asset pricing, *Journal of Finance*, 59, 1743-1775.

Brennan, M. J, and Y. Xia, 2006, Risk and valuation under an intertemporal capital asset pricing model, *Journal of Business*, 79, 1-35.

Breeden, D. T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics*, 7, 265-296.

Burmeister, E., and K. D. Wall, 1982, Kalman filtering estimation of unobserved rational expectations with an application to German hyperinflation, *Journal of Econometrics*, 20, 255-284.

Burmeister, E., K. D. Wall, and J. Hamilton, 1986, Estimation of unobserved expected monthly inflation using Kalman filtering, *Journal of Business and Economic Statistics*, 4, 147-160.

Campbell, J. Y., 1991, A variance decomposition for stock returns, *The Economic Journal*, 101, 157-179.

Campbell, J. Y., 1993, Intertemporal asset pricing without consumption data, *American Economic Review*, 83, 487-512.

Campbell, J. Y., 1996, Understanding risk and return, *Journal of Political Economy*, 104, 298-345.

Campbell, J. Y., S. Giglio, and C. Polk, 2010, Hard times, NBER working paper no. w16222.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Campbell, J. Y., C. Polk, and T. Vuolteenaho, 2010, Growth or glamour? Fundamentals and systematic risk in stock returns, *Review of Financial Studies*, 23, 305-344.

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 T. Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review*, 94, 1249-1275.

Carlin, B. P., N. G. Polson, and D. S. Stoffer, 1992, A Monte Carlo approach to non-normal and nonlinear state-space modeling, *Journal of the American Statistical Association*, 87, 493-500.

Carter, C. K. and R. Kohn, 1994, On Gibbs sampling for state-space models, *Biometrika*, 81, 541-553.

Casella, G. and E. I. George, 1992, Explaining the Gibbs sampler, *The American Statistician*, 46, 3, 167-174.

Chen, J., 2003, Intertemporal CAPM and the cross-section of stock returns, Working paper.

Chen, L. and X. Zhao, 2009, Return decomposition, *Review of Financial Studies*, forthcoming.

Copeland, T. E., and J. F. Weston, 1992, *Financial Theory and Corporate Policy*, third edition, Addison-Wesley, Canada.

Davis, J. L., Fama, E., and K. French, 2000, Characteristics, Covariances, and Average Returns, *Journal of Finance*, 55, 389-406.

DeBondt, W., and R. Thaler, 1985, Does the stock market overreact?, *Journal of Finance*, 40, 793-805.

DeBondt, W., and R. Thaler, 1987, Further Evidence on Investor Overreaction and Stock Market Seasonality?, *Journal of Finance*, 42, 557-581.

de Jong, P. and N. Shephard, 1995, The simulation smoother for time series models, *Biometrika*, 82, 339-350.

Dempster, A.P., N.M. Laird, and D.B. Rubin, 1977, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. B*, 39, 1-38.

Durbin, J. and S. J. Koopman, 2001, Time series analysis by state space methods.

Durbin, J. and S. J. Koopman, 2002, A simple and efficient simulation smoother for state space time series analysis, *Biometrika*, 89, 603-615.

Engle, R. F, and M. Watson, 1981, A one-factor multivariate time series model of metropolitan wage rates, *Journal of the American Statistical Association*, 76, Applications section, 774-781.

Engsted, T., T. Q. Pedersen, C. Tanggaard, 2012, Pitfalls in VAR based return decompositions: A clarification, *Journal of Banking and Finance*, 36, 1255-1265.

Epstein, L., and S. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica*, 57, 937-969.

Epstein, L., and S. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical investigation, *Journal of Political Economy*, 99, 263-286.

Fama, E., and J. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy*, 71, 607-636.

Fama, E., and K. French, 1992, The cross-section of expected stock returns, *Journal of Finance*, 47, 427-465.

Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, 33, 3-56.

Fama, E., and K. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance*, 51, 55-84.

Fama, E., and K. French, 1998, Value versus Growth: the International Evidence, *Journal of Finance*, 53, 1975-1999.

Fama, E., and K. French, 2006, The value premium and the CAPM, *Journal of Finance*, 61, 2163-2185.

Franzoni, F., 2001, Where is beta going? The riskiness of value and small stocks, PhD dissertation, MIT.

Fruhwirth-Schnatter, S., 1994, Data augmentation and dynamic linear models, Journal of Time Series Analysis, 15, 183-202.

Gamerman, D., 1998, Markov chain Monte Carlo for dynamic generalised linear models, *Biometrika*, 85, 1, 215-227.

Gamerman, D. and H. F. Lopes, 2006, Markov chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd edition.

Garbade, K. D., and P. Wachtel, 1978, Time variation in the relationship between inflation and interest rates, *Journal of Monetary Economics*, 4, 765-775.

Gelman, A. and D. B. Rubin, 1992, Inference from iterative simulation using multiple sequences (with discussion), *Statistical Science*, 7, 457-511.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004, Bayesian data analysis, 2nd edition.

Guo, H., 2003, On the real time forecasting ability of the consumption-wealth ratio, Working Paper 2003-007B, http://research.stlouisfed.org/wp/2003/2003-007.pdf, The Federal Reserve Bank of St. Louis.

Guo, H., R. Savickas, Z. Wang, and J. Yang, 2009, Is value premium a proxy for time-varying investment opportunities: Some time series evidence, *Journal of Financial and Quantitative Analysis*, 44, 1, 133-154.

Gupta, N. K., and R. K. Mehra, 1974, Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations, *IEEE Trans. Automat. Contr.*, AC-19, 774-783.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance*, 51, 3-53.

Jegadeesh, N., and S. Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance*, 48, 65-91.

Jones, R. H., 1980, Maximum likelihood fitting of ARMA models to time series with missing observations, Technometrics, 22, 389-395.

Kalman, R.E., 1960, A new approach to linear filtering and prediction problems, *Trans ASME J. Basic Eng.*, 82, 35-45.

Kalman, R.E. and R.S. Bucy, 1961, New results in filtering and prediction theory, *Trans ASME J. Basic Eng.*, 83, 95-108.

Kim, C., and C. R. Nelson, 1989, The time-varying-parameter model for modeling changing conditional variance: The case of the Lucas Hypothesis, *Journal of Business and Economic Statistics*, 7, 433-440.

Kim, C., and C. R. Nelson, 1999, *State-space models with regime switching*, Massachusetts Institute of Technology.

Lettau, M. and S. C. Ludvigson, 2001a, Consumption, aggregate wealth and expected stock returns, *Journal of Finance*, 56, 3, 815-849.

Lettau, M. and S. C. Ludvigson, 2001b, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy*, 109, 1238-1287.

Lewellen, J., and S. Nagel, 2006, The conditional CAPM does not explain asset-pricing anomalies, *Journal of Financial Economics*, 82, 289-314.

Lintner, J., 1965a, Security prices, risk, and maximal gains from diversification, *Journal of Finance*, 20, 587-615.

Lintner, J., 1965b, The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 47, 13-37.

Lintner, J., 1969, The aggregation of investor's diverse judgments and preferences in purely competitive security markets, *Journal of Financial and Quantitative Analysis*, 4, 347-400.

Liu, N., and L. Zhang, 2008, Is the value spread a useful predictor of returns?, *Journal of Financial Markets*, 11, 199-227.

Ludvigson, S. C. and S. Ng, 2007, The empirical risk-return relation: a factor analysis approach, *Journal of Financial Economics,* 83, 171-222.

Lustig, H., and S. Van Nieuwerburgh, 2005, Housing collateral, consumption insurance, and risk premia: An empirical perspective, *Journal of Finance*, 60, 1167-1219.

Lynch, S. M., 2007, Introduction to applied Bayesian statistics and estimation for social sciences.

Markowitz, H., 1959, *Portfolio selection: efficient diversification of investments*, John Wiley, New York.

Merton, R. C., 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory*, 3, 373-413.

Merton, R. C., 1973, An intertemporal capital asset pricing model, *Econometrica*, 41, 867-887.

Mossin, J., 1966, Equilibrium in a capital asset market, *Econometrica*, 35, 768-783.

Reinganum, M. R., 1981, A New Empirical Perspective on the CAPM, *Journal of Financial and Quantitative Analysis*, 16, 439-462.

Reis, E. A., E. Salazar, and D. Gamerman, 2006, Comparison of sampling schemes for dynamic linear models, *International Statistical Review*, 74, 2, 203-214.

Robert, C. P. and G. Casella, 2004, Monte Carlo statistical methods, 2nd edition.

Rosenberg, B., K. Reid, and R. Lanstein, 1985, Persuasive Evidence of Market Inefficiency, *Journal of Portfolio Management*, 11, 9-17.

Santos, T., and P. Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies*, 19, 1-44.

Schwert, G. W., 2002, Anomalies and Market Efficiency, University of Rochester William E. Simon Graduate School of Business Administration Working Paper No. FR 02-13, 1-52.

Sharpe, W. F., 1963, A simplified model for portfolio analysis, *Management Science*, 2, 277-293.

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., 2000, *Irrational Exuberance*, Princeton University Press, Princeton, N.J.

Shumway, R.H., 2000, Dynamic mixed models for irregularly observed time series, *Resenhas-Reviews of the Institute of Mathematics and Statistics*, 4, 433-456, University of Sao Paulo, Brazil: USP Press.

Shumway, R.H. and D.S. Stoffer, 1982, An approach to time series smoothing and forecasting using the EM algorithm, *J. Time Series Anal.*, 3, 253-264.

Shumway, R.H. and D.S. Stoffer, 2006, *Time series analysis and its applications with R examples,* second edition, Springer Science+Business Media, LLC.

Stattman, D. 1980, Book Values and Stock Returns, *The Chicago MBA: A Journal of Selected Papers*, 4, 25-45.

Stoffer, D. S., and K. D. Wall, 1991, Bootstrapping state space models: Gaussian maximum likelihood estimation and the Kalman filter, *Journal of the American Statistical Association*, 86, 1024-1033.

Strickland, C. M., I. W. Turner, R. Denham, and K. L. Mengersen, 2009, Efficient Bayesian estimation of multivariate state-space models, *Computational Statistics and Data Analysis*, 53, 12, 4116-4125.

Tobin, J., 1958, Liquidity preference as behavior towards risk, *Review of Economic Studies*, 25, 65-86.

Tsay, R.S, 2005, *Analysis of financial time series*, second edition, John Wiley & Sons, Inc.

Treynor, J., 1961, Toward a theory of the market value of risky assets, unpublished manuscript.

Zhang, L., 2005, The value premium, *Journal of Finance*, 60, 67-103.

Tables and Figures

(Tables 1 to 12 and Figures 1 to 3)

Table 1: Posterior means and variances for parameters in the state-space model used to analyze SIMDATA1 (sample size = 500) and SIMDATA2 (sample size = 50) in the simulation study

The table shows the posterior means, posterior variances, and mean-squared errors for parameters in the state-space model used to analyze the simulated data sets SIMDATA1 and SIMDATA2 in the simulation study in Section 3.2.3. The posterior distributions of the parameters are obtained by running one chain of 100,000 iterations and using the latter 50,000 iterations. Column 2 displays the "true values" of the parameters, which are the parameter values used to generate SIMDATA1 and SIMDATA2. The posterior means, posterior variances, and mean-squared errors of the parameters from analyzing SIMDATA1 are displayed in Columns 3, 4, and 5, respectively. The posterior means, posterior variances, and mean-squared errors of the parameters from analyzing SIMDATA2 are displayed in Columns 6, 7, and 8, respectively. Mean-squared error is defined as posterior variance + (posterior mean - true value)^2.

		SIMDATA1			SIMDATA2		
Parameter	True value	Mean	Variance	MSE	Mean	Variance	MSE
H_{22}	1.3300	1.1915694	0.0061236	0.0252866	1.3881713	0.0135350	0.0169189
H_{32}	1.0100	1.0227305	0.0055147	0.0056767	1.0305263	0.0087210	0.0091423
H_{42}	0.9700	0.9165992	0.0061420	0.0089937	0.9384042	0.0108365	0.0118348
H_{52}	1.0100	0.9551632	0.0021347	0.0051418	0.9934910	0.0028236	0.0030962
H_{62}	0.8500	0.8197627	0.0028605	0.0037748	0.8546709	0.0038364	0.0038582
H_{72}	0.8100	0.7646611	0.0043045	0.0063601	0.7824611	0.0061085	0.0068669
H_{83}	-0.2234	-0.1383204	0.0078473	0.0150858	-0.1994266	0.0175687	0.0181434
H_{93}	-0.6281	-0.4685565	0.0368321	0.0622863	-0.5264917	0.1092380	0.1195623
$H_{10,3}$	-1.0031	-1.1027505	0.0392773	0.0492075	-1.2917615	0.2094478	0.2927732
$H_{11,3}$	-1.8321	-1.9508878	0.0386533	0.0527638	-1.9695979	0.2202123	0.2391180
F_{11}	0.8400	0.8825640	0.0001743	0.0019860	0.8344512	0.0002268	0.0002576
F_{12}	0.1100	0.1006527	0.0004219	0.0005093	0.0997153	0.0004000	0.0005058
F_{13}	-0.0900	-0.0794212	0.0001011	0.0002131	-0.0798360	0.0001003	0.0002036
V_{11}	0.0001	0.0000997	0.0000000	0.0000000	0.0001000	0.0000000	0.0000000
V_{12}	0.0000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
V_{13}	0.0000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
V_{22}	0.0001	0.0000998	0.0000000	0.0000000	0.0001000	0.0000000	0.0000000
V_{23}	0.0000	-0.0000002	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
V_{33}	0.0001	0.0001000	0.0000000	0.0000000	0.0001000	0.0000000	0.0000000
Average				0.0124887			0.0380148

Table 2: Descriptive statistics of the observed variables in the state-space model

The table shows the descriptive statistics of the observed variables in the state-space model estimated from the full sample period 1952:7-2010:12, which consists of 702 monthly data points. The variable r_M^e is the excess log return on the CRSP value-weighted stock index. The variables r_1, r_2, r_3, r_4, and r_5 are the log returns of the five Fama-French BE/ME portfolios from the portfolio of the lowest BE/ME quintile to the one of the highest BE/ME quintile, respectively. The variables r_1^b, r_2^b, r_3^b, r_4^b, and r_5^b are the log returns of the five Fama-Bliss discount bonds with maturities from one to five years, respectively. In the table, "Stdev." denotes standard deviation, "Min" denotes minimum, and "Max" denotes maximum.

Variable	Mean	Median	Stdev.	Min	Max
r_M^e	0.0044	0.0089	0.0440	-0.2597	0.1469
r_1	0.0076	0.0102	0.0480	-0.2710	0.1988
r_2	0.0086	0.0119	0.0445	-0.2825	0.1482
r_3	0.0095	0.0126	0.0426	-0.2638	0.1566
r_4	0.0101	0.0133	0.0432	-0.2169	0.2020
r_5	0.0111	0.0147	0.0493	-0.2375	0.2237
r_1^b	0.0000	-0.0001	0.0047	-0.0273	0.0451
r_2^b	0.0000	-0.0002	0.0084	-0.0458	0.0694
r_3^b	0.0000	-0.0003	0.0117	-0.0738	0.0853
r_4^b	0.0000	-0.0004	0.0152	-0.0793	0.0903
r_5^b	0.0000	-0.0006	0.0178	-0.0801	0.1050

Table 3: Summary of cash-flow news and discount-rate news

The table summarizes the behavior of cash-flow news and discount-rate news obtained using the state-space model for the period 1952:7-2010:12. The table is organized into three sections: the top section shows the variance-covariance matrix of cash-flow news and discount-rate news with variances on the diagonal and covariance at the off-diagonal entries, the middle section shows the standard deviation and correlation matrix of cash-flow news and discount-rate news with standard deviations on the diagonal and correlation at the off-diagonal entries, and the bottom section displays summary statistics of the distribution of correlation between cash-flow news and discount-rate news. The posterior distributions of cash-flow news and discount-rate news are obtained based on running one chain of 300,000 iterations and using the latter 150,000 iterations.

The variance of cash-flow news reported in the top section of the table is the mean of the 150,000 iterations of V_{22}. Similarly, the variance of discount-rate news reported is the mean of the 150,000 iterations of V_{33}, and the covariance between cash-flow news and discount-rate news reported is the mean of the 150,000 iterations of V_{23}. In the middle section of the table, the standard deviation of cash-flow news reported is the mean of the 150,000 iterations of $\sqrt{V_{22}}$, the standard deviation of discount-rate news reported is the mean of the 150,000 iterations of $\sqrt{V_{33}}$, and the correlation between cash-flow news and discount-rate news reported is the mean of the 150,000 iterations of $V_{23} \Big/ \left(\sqrt{V_{22}} * \sqrt{V_{33}} \right)$. The bottom section of the table reports descriptive statistics of the posterior distribution of the correlation between cash-flow news and discount-rate news based on the 150,000 iterations of $V_{23} \Big/ \left(\sqrt{V_{22}} * \sqrt{V_{33}} \right)$.

Table 3 (continued from previous page)

Variance-covariance matrix		
	N_{CF}	N_{DR}
N_{CF}	0.0000044	-0.0000192
N_{DR}	-0.0000192	0.0000938

Standard deviation and correlation matrix		
	N_{CF}	N_{DR}
N_{CF}	0.0021	-0.9447
N_{DR}	-0.9447	0.0096

Distribution of correlation	
	$cor(N_{CF,}N_{DR})$
Mean	-0.9447
Min	-0.9749
First quartile	-0.9461
Median	-0.9448
Third quartile	-0.9434
Max	-0.9348

Table 4: Cash-flow betas and discount-rate betas for the 1963:7-2001:12 period

Panel A of the table shows the posterior means and standard deviations [in brackets] of cash-flow and discount-rate betas for the 25 Fama-French portfolios. Panel B and C show the 95% posterior intervals and posterior medians, respectively, of cash-flow and discount-rate betas for the 25 Fama-French portfolios. "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of cash-flow beta and discount-rate beta of 150,000 iterations are obtained and are used to calculate the means, medians, standard deviations, and 95% intervals. The column names "2.5%" and "97.5%" are the 2.5% and 97.5% percentiles of the posterior distribution, respectively. The posterior distributions of cash-flow and discount-rate betas are for the 1963:7-2001:12 period.

Panel A: Posterior Means and Standard Deviations

$\hat{\beta}_{CF}$	Growth		2		3		4		Value	
Small	0.288	[0.008]	0.247	[0.007]	0.221	[0.006]	0.213	[0.006]	0.225	[0.006]
2	0.268	[0.007]	0.228	[0.006]	0.208	[0.005]	0.194	[0.005]	0.209	[0.005]
3	0.244	[0.006]	0.213	[0.005]	0.190	[0.005]	0.174	[0.004]	0.193	[0.005]
4	0.213	[0.006]	0.202	[0.005]	0.180	[0.005]	0.162	[0.004]	0.179	[0.005]
Large	0.165	[0.004]	0.168	[0.004]	0.146	[0.004]	0.134	[0.004]	0.152	[0.004]
$\hat{\beta}_{DR}$	Growth		2		3		4		Value	
Small	1.384	[0.030]	1.185	[0.025]	1.064	[0.022]	1.024	[0.021]	1.080	[0.022]
2	1.286	[0.027]	1.098	[0.022]	1.000	[0.020]	0.933	[0.018]	1.005	[0.020]
3	1.172	[0.025]	1.022	[0.020]	0.914	[0.018]	0.837	[0.017]	0.930	[0.019]
4	1.024	[0.021]	0.970	[0.018]	0.868	[0.017]	0.780	[0.016]	0.861	[0.018]
Large	0.796	[0.017]	0.807	[0.015]	0.702	[0.014]	0.645	[0.014]	0.731	[0.016]

Panel B: 95% Posterior Intervals

$\hat{\beta}_{CF}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	0.273	0.304	0.234	0.260	0.210	0.233	0.203	0.224	0.214	0.236
2	0.254	0.282	0.217	0.240	0.198	0.218	0.184	0.204	0.199	0.220
3	0.232	0.257	0.202	0.223	0.181	0.200	0.165	0.183	0.184	0.203
4	0.202	0.224	0.192	0.212	0.172	0.190	0.154	0.170	0.170	0.188
Large	0.157	0.174	0.160	0.176	0.139	0.153	0.127	0.141	0.144	0.160
$\hat{\beta}_{DR}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	1.325	1.443	1.136	1.235	1.021	1.107	0.984	1.065	1.037	1.123
2	1.233	1.339	1.055	1.141	0.961	1.038	0.897	0.970	0.965	1.045
3	1.125	1.221	0.983	1.061	0.880	0.949	0.805	0.871	0.893	0.967
4	0.982	1.066	0.934	1.006	0.835	0.901	0.749	0.811	0.826	0.896
Large	0.764	0.829	0.777	0.837	0.674	0.730	0.619	0.673	0.701	0.761

Table 4 (continued from previous page)

Panel C: Posterior Medians					
$\hat{\beta}_{CF}$	Growth	2	3	4	Value
Small	0.2883	0.2468	0.2214	0.2131	0.2248
2	0.2677	0.2284	0.2078	0.1940	0.2090
3	0.2438	0.2125	0.1900	0.1740	0.1933
4	0.2129	0.2016	0.1804	0.1620	0.1788
Large	0.1654	0.1677	0.1458	0.1341	0.1519

$\hat{\beta}_{DR}$	Growth	2	3	4	Value
Small	1.3837	1.1849	1.0639	1.0238	1.0797
2	1.2858	1.0977	0.9994	0.9334	1.0046
3	1.1722	1.0216	0.9140	0.8374	0.9297
4	1.0234	0.9697	0.8679	0.7795	0.8607
Large	0.7958	0.8066	0.7017	0.6453	0.7306

Table 5: Difference in cash-flow and discount-rate betas between value and growth stocks for 1963:7-2001:12

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between value and growth stocks for the period 1963:7-2001:12. For each of the five size categories, the posterior distribution of the difference in cash-flow beta between value and growth stocks (i.e., $\hat{\beta}_{value,CF} - \hat{\beta}_{growth,CF}$) of 150,000 iterations is obtained using the posterior distributions of cash-flow betas. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between value and growth stocks for the period 1963:7-2001:12. For each of the five size categories, the posterior distribution of the difference in discount-rate beta between value and growth stocks (i.e., $\hat{\beta}_{value,DR} - \hat{\beta}_{growth,DR}$) of 150,000 iterations is obtained using the posterior distributions of discount-rate betas. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{value,CF} - \hat{\beta}_{growth,CF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.0636	-0.0636	0.0032	-0.0700	-0.0573	-0.0658	-0.0614
2	-0.0587	-0.0587	0.0033	-0.0651	-0.0524	-0.0609	-0.0565
3	-0.0505	-0.0505	0.0034	-0.0572	-0.0439	-0.0527	-0.0482
4	-0.0341	-0.0341	0.0032	-0.0403	-0.0279	-0.0362	-0.0319
Large	-0.0135	-0.0135	0.0028	-0.0191	-0.0080	-0.0154	-0.0116

$\hat{\beta}_{value,DR} - \hat{\beta}_{growth,DR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.3040	-0.3039	0.0148	-0.3331	-0.2750	-0.3139	-0.2940
2	-0.2812	-0.2812	0.0150	-0.3106	-0.2520	-0.2913	-0.2711
3	-0.2425	-0.2425	0.0158	-0.2736	-0.2117	-0.2532	-0.2319
4	-0.1627	-0.1627	0.0150	-0.1923	-0.1334	-0.1728	-0.1527
Large	-0.0652	-0.0652	0.0136	-0.0919	-0.0386	-0.0744	-0.0561

Table 6: Difference in cash-flow and discount-rate betas between large and small stocks for 1963:7-2001:12

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between large and small stocks for the period 1963:7-2001:12. For each of the five BE/ME categories, the posterior distribution of the difference in cash-flow beta between large and small stocks (i.e., $\hat{\beta}_{large,CF} - \hat{\beta}_{small,CF}$) of 150,000 iterations is obtained using the posterior distributions of cash-flow betas. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between large and small stocks for the period 1963:7-2001:12. For each of the five BE/ME categories, the posterior distribution of the difference in discount-rate beta between large and small stocks (i.e., $\hat{\beta}_{large,DR} - \hat{\beta}_{small,DR}$) of 150,000 iterations is obtained using the posterior distributions of discount-rate betas. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{large,CF} - \hat{\beta}_{small,CF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.1230	-0.1230	0.0048	-0.1327	-0.1136	-0.1263	-0.1197
2	-0.0791	-0.0791	0.0041	-0.0872	-0.0713	-0.0818	-0.0764
3	-0.0756	-0.0755	0.0036	-0.0826	-0.0687	-0.0780	-0.0732
4	-0.0791	-0.0790	0.0034	-0.0858	-0.0725	-0.0814	-0.0767
Value	-0.0729	-0.0729	0.0033	-0.0794	-0.0666	-0.0751	-0.0707

$\hat{\beta}_{large,DR} - \hat{\beta}_{small,DR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.5878	-0.5878	0.0211	-0.6293	-0.5465	-0.6020	-0.5736
2	-0.3782	-0.3782	0.0183	-0.4143	-0.3423	-0.3906	-0.3658
3	-0.3622	-0.3621	0.0160	-0.3935	-0.3309	-0.3729	-0.3514
4	-0.3784	-0.3784	0.0151	-0.4081	-0.3488	-0.3886	-0.3683
Value	-0.3491	-0.3490	0.0146	-0.3776	-0.3206	-0.3589	-0.3392

Table 7: Posterior means and standard deviations of cash-flow and discount-rate betas for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news

The table shows the posterior means and standard deviations [in brackets] of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$) calculated using orthogonalized cash-flow news and discount-rate news for the 25 Fama-French portfolios for the period 1963:7-2001:12. "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ of 150,000 iterations are obtained and used to calculate posterior means and standard deviations. The top half of the table reports the mean and standard deviation of the 150,000 iterations of $\hat{\beta}_{OCF}$ for each portfolio, and the bottom half of the table reports the mean and standard deviation of the 150,000 iterations of $\hat{\beta}_{ODR}$ for each portfolio.

$\hat{\beta}_{OCF}$	Growth	2	3	4	Value
Small	-0.0000299	-0.0000244	-0.0000204	-0.0000197	-0.0000203
	[0.0089]	[0.0076]	[0.0068]	[0.0065]	[0.0069]
2	-0.0000261	-0.0000208	-0.0000185	-0.0000165	-0.0000189
	[0.0083]	[0.0070]	[0.0064]	[0.0059]	[0.0064]
3	-0.0000233	-0.0000194	-0.0000167	-0.0000152	-0.0000173
	[0.0075]	[0.0065]	[0.0058]	[0.0053]	[0.0059]
4	-0.0000213	-0.0000177	-0.0000163	-0.0000151	-0.0000153
	[0.0066]	[0.0062]	[0.0055]	[0.0050]	[0.0055]
Large	-0.0000158	-0.0000146	-0.0000126	-0.0000112	-0.0000120
	[0.0051]	[0.0051]	[0.0045]	[0.0041]	[0.0046]

$\hat{\beta}_{ODR}$	Growth	2	3	4	Value
Small	2.021	1.730	1.554	1.495	1.577
	[0.045]	[0.038]	[0.033]	[0.031]	[0.032]
2	1.878	1.603	1.460	1.363	1.467
	[0.040]	[0.033]	[0.029]	[0.027]	[0.030]
3	1.712	1.492	1.335	1.223	1.358
	[0.036]	[0.030]	[0.026]	[0.025]	[0.028]
4	1.495	1.416	1.268	1.139	1.257
	[0.032]	[0.028]	[0.025]	[0.024]	[0.027]
Large	1.162	1.178	1.025	0.943	1.067
	[0.025]	[0.023]	[0.021]	[0.020]	[0.023]

Table 8: Posterior medians and 95% intervals of cash-flow and discount-rate betas for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows the posterior medians of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$) calculated using orthogonalized cash-flow news and discount-rate news for the 25 Fama-French portfolios. The bottom half of the table shows the 95% posterior intervals of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$). "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ of 150,000 iterations are obtained and used to calculate posterior medians and 95% intervals. The posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ are for the 1963:7-2001:12 period. The column names "2.5%" and "97.5%" are the 2.5% and 97.5% percentiles of the posterior distribution, respectively.

Posterior Medians

$\hat{\beta}_{OCF}$	Growth	2	3	4	Value
Small	-0.00021	-0.00017	-0.00013	-0.00013	-0.00013
2	-0.00018	-0.00013	-0.00011	-0.00009	-0.00012
3	-0.00015	-0.00012	-0.00010	-0.00009	-0.00010
4	-0.00014	-0.00011	-0.00010	-0.00009	-0.00009
Large	-0.00010	-0.00009	-0.00007	-0.00006	-0.00006
$\hat{\beta}_{ODR}$	Growth	2	3	4	Value
Small	2.021	1.730	1.554	1.495	1.577
2	1.878	1.603	1.460	1.363	1.467
3	1.712	1.492	1.335	1.223	1.358
4	1.495	1.416	1.268	1.138	1.257
Large	1.162	1.178	1.025	0.942	1.067

95% Posterior Intervals

$\hat{\beta}_{OCF}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	-0.017	0.018	-0.014	0.015	-0.013	0.014	-0.012	0.013	-0.013	0.014
2	-0.015	0.016	-0.013	0.014	-0.012	0.013	-0.011	0.012	-0.012	0.013
3	-0.014	0.015	-0.012	0.013	-0.011	0.012	-0.010	0.011	-0.011	0.012
4	-0.012	0.013	-0.012	0.012	-0.010	0.011	-0.009	0.010	-0.010	0.011
Large	-0.010	0.010	-0.010	0.010	-0.008	0.009	-0.008	0.008	-0.009	0.009
$\hat{\beta}_{ODR}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	1.934	2.109	1.657	1.804	1.491	1.618	1.435	1.556	1.514	1.641
2	1.800	1.957	1.540	1.667	1.403	1.517	1.310	1.418	1.409	1.527
3	1.641	1.784	1.435	1.550	1.284	1.387	1.175	1.272	1.303	1.414
4	1.433	1.558	1.363	1.471	1.219	1.317	1.093	1.185	1.206	1.310
Large	1.114	1.211	1.133	1.224	0.984	1.066	0.903	0.983	1.022	1.113

Table 9: Difference in cash-flow and discount-rate betas between value and growth stocks for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between value and growth stocks for the period 1963:7-2001:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five size categories, the posterior distribution of the difference in cash-flow beta between value and growth stocks (i.e., $\hat{\beta}_{value,OCF} - \hat{\beta}_{growth,OCF}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{OCF}$. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between value and growth stocks for the period 1963:7-2001:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five size categories, the posterior distribution of the difference in discount-rate beta between value and growth stocks (i.e., $\hat{\beta}_{value,ODR} - \hat{\beta}_{growth,ODR}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{ODR}$. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{value,OCF} - \hat{\beta}_{growth,OCF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	0.0000095	0.0000842	0.0021	-0.0042	0.0038	-0.0014	0.0015
2	0.0000071	0.0000605	0.0019	-0.0038	0.0035	-0.0013	0.0013
3	0.0000060	0.0000515	0.0016	-0.0033	0.0030	-0.0011	0.0011
4	0.0000059	0.0000590	0.0012	-0.0024	0.0021	-0.0008	0.0008
Large	0.0000038	0.0000281	0.0006	-0.0011	0.0010	-0.0004	0.0004

$\hat{\beta}_{value,ODR} - \hat{\beta}_{growth,ODR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.4439	-0.4438	0.0217	-0.4866	-0.4015	-0.4584	-0.4293
2	-0.4107	-0.4107	0.0219	-0.4538	-0.3678	-0.4255	-0.3959
3	-0.3542	-0.3542	0.0231	-0.3997	-0.3090	-0.3697	-0.3386
4	-0.2376	-0.2376	0.0219	-0.2809	-0.1948	-0.2523	-0.2229
Large	-0.0952	-0.0952	0.0199	-0.1342	-0.0564	-0.1086	-0.0819

Table 10: Difference in cash-flow and discount-rate betas between large and small stocks for 1963:7-2001:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between large and small stocks for the period 1963:7-2001:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five BE/ME categories, the posterior distribution of the difference in cash-flow beta between large and small stocks (i.e., $\hat{\beta}_{large,OCF} - \hat{\beta}_{small,OCF}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{OCF}$. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between large and small stocks for the period 1963:7-2001:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five BE/ME categories, the posterior distribution of the difference in discount-rate beta between large and small stocks (i.e., $\hat{\beta}_{large,ODR} - \hat{\beta}_{small,ODR}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{ODR}$. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{large,OCF} - \hat{\beta}_{small,OCF}$

	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	0.0000141	0.0001072	0.0038	-0.0078	0.0071	-0.0026	0.0027
2	0.0000098	0.0000804	0.0025	-0.0051	0.0046	-0.0017	0.0018
3	0.0000078	0.0000578	0.0023	-0.0047	0.0044	-0.0016	0.0017
4	0.0000086	0.0000658	0.0025	-0.0050	0.0046	-0.0017	0.0017
Value	0.0000083	0.0000642	0.0023	-0.0046	0.0042	-0.0015	0.0016

$\hat{\beta}_{large,ODR} - \hat{\beta}_{small,ODR}$

	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.8585	-0.8584	0.0311	-0.9194	-0.7978	-0.8794	-0.8374
2	-0.5523	-0.5523	0.0269	-0.6052	-0.4997	-0.5704	-0.5342
3	-0.5289	-0.5288	0.0234	-0.5749	-0.4831	-0.5447	-0.5131
4	-0.5527	-0.5526	0.0222	-0.5962	-0.5091	-0.5676	-0.5377
Value	-0.5098	-0.5097	0.0214	-0.5517	-0.4681	-0.5243	-0.4953

Table 11: Descriptive statistics for the distribution of correlation between the estimated state variables in this study and Campbell and Vuolteenaho's (2004) estimated conditional expectation of the excess log market return, cash-flow news, and discount-rate news for 1952:7-2001:12

Panel A shows the descriptive statistics for the posterior distribution of correlation between the estimated conditional expectation of the excess log market return ($E_{t-1}r^e_{M,t}$) in this study and that in Campbell and Vuolteenaho (2004). The posterior distribution of correlation is calculated based on Campbell and Vuolteenaho's (2004) estimated $E_{t-1}r^e_{M,t}$ and the posterior distribution of 150,000 iterations of $E_{t-1}r^e_{M,t}$ obtained from the state-space model.

Panel B shows the descriptive statistics for the posterior distribution of correlation between the estimated cash-flow news (N_{CF}) in this study and that in Campbell and Vuolteenaho (2004). The posterior distribution of correlation is calculated based on Campbell and Vuolteenaho's (2004) estimated N_{CF} and the posterior distribution of 150,000 iterations of N_{CF} obtained from the state-space model.

Panel C shows the descriptive statistics for the posterior distribution of correlation between the estimated discount-rate news (N_{DR}) in this study and that in Campbell and Vuolteenaho (2004). The posterior distribution of correlation is calculated based on Campbell and Vuolteenaho's (2004) estimated N_{DR} and the posterior distribution of 150,000 iterations of N_{DR} obtained from the state-space model.

Panel D shows the descriptive statistics for the posterior distribution of correlation between the estimated total news ($N_{CF} - N_{DR}$) in this study and that in Campbell and Vuolteenaho (2004). The posterior distribution of correlation is calculated based on Campbell and Vuolteenaho's (2004) estimated $N_{CF} - N_{DR}$ and the posterior distribution of 150,000 iterations of $N_{CF} - N_{DR}$ obtained from the state-space model.

Table 11 (continued from previous page)

Panel A: $cor(E_{t-1}r^e_{M,t}.CV, E_{t-1}r^e_{M,t}.SSM)$	
Mean	0.0951
Min	-0.3241
First quartile	-0.0067
Median	0.0850
Third quartile	0.2008
Max	0.3287

Panel B: $cor(N_{CF}.CV, N_{CF}.SSM)$	
Mean	0.1752
Min	0.1246
First quartile	0.1676
Median	0.1752
Third quartile	0.1827
Max	0.2257

Panel C: $cor(N_{DR}.CV, N_{DR}.SSM)$	
Mean	0.8159
Min	0.7821
First quartile	0.8116
Median	0.8160
Third quartile	0.8203
Max	0.8426

Panel D: $cor\left((N_{CF}-N_{DR}).CV, (N_{CF}-N_{DR}).SSM\right)$	
Mean	0.9291
Min	0.9112
First quartile	0.9265
Median	0.9291
Third quartile	0.9317
Max	0.9467

Table 12: Correlation between average returns and cash-flow betas when there is a positive or negative cash-flow shock and correlations between average returns and discount-rate betas when there is a positive or negative discount-rate shock for 1963:7-2001:12

The top half of the table shows the correlation between average returns and cash-flow betas of the 25 Fama-French portfolios for the period of 1963:7-2001:12 when there is a positive or a negative shock to future cash flows. The bottom half of the table shows the correlation between average returns and discount-rate betas of the 25 Fama-French portfolios for the period of 1963:7-2001:12 when there is a positive or a negative shock to the negative of future discount-rates. In the table, "$N_{CF} > 0$" denotes a positive shock to future cash flows, "$N_{CF} < 0$" denotes a negative shock to future cash flows, "$-1 * N_{DR} > 0$" denotes a positive shock to the negative of future discount-rates, and "$-1 * N_{DR} < 0$"denotes a negative shock to the negative of future discount-rates. The column name "$cor(r, \beta_{CF})$" denotes the correlation between average returns and cash-flow betas of the 25 Fama-French portfolios, and "$cor(r, \beta_{DR})$"denotes the correlation between average returns and discount-rate betas of the 25 Fama-French portfolios.

	$cor(r, \beta_{CF})$
$N_{CF} > 0$	0.8919
$N_{CF} < 0$	-0.8601

	$cor(r, \beta_{DR})$
$-1 * N_{DR} > 0$	0.8658
$-1 * N_{DR} < 0$	-0.8476

Figure 1: This figure plots smoothed cash-flow news and discount-rate news for the period 1952:7-2010:12. The smoothed cash-flow (discount-rate) news series are 10-month moving averages of the posterior means of cash-flow (discount-rate) news for each month. The posterior mean of cash-flow (discount-rate) news for a month is calculated based on the 150,000 iterations of the posterior distribution of cash-flow (discount-rate) news for that month.

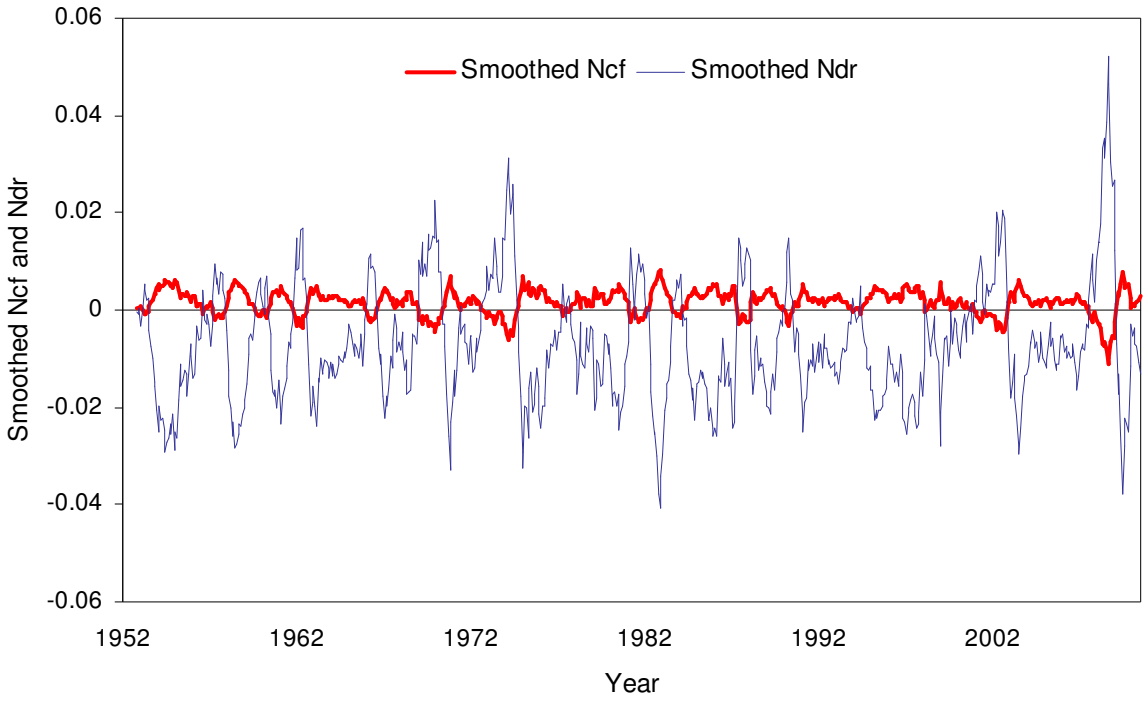


Figure 2: Density plots of cash-flow betas for the 25 Fama-French portfolios during 1963:7-2001:12

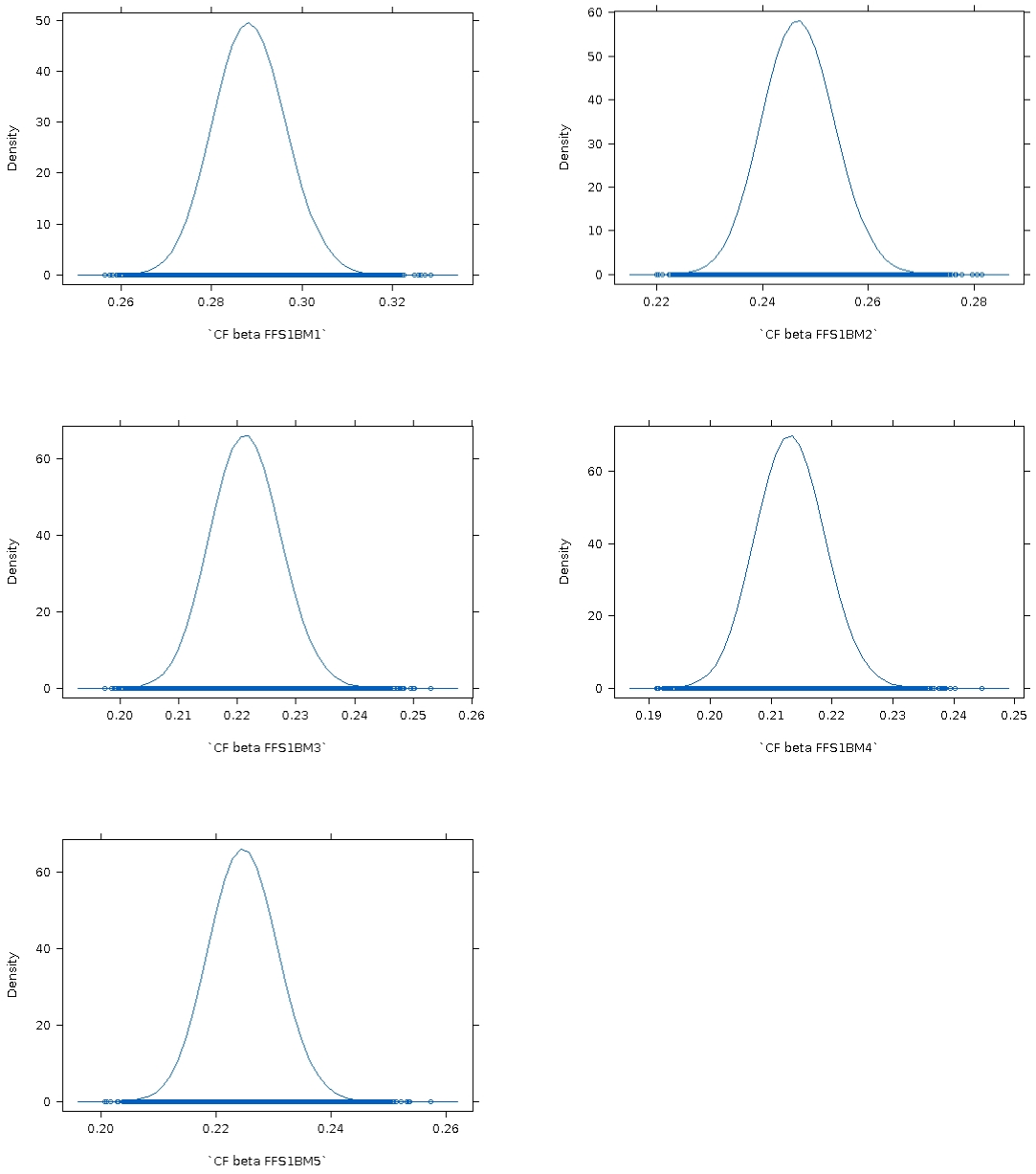
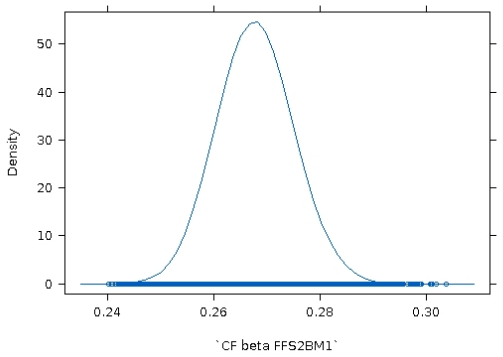
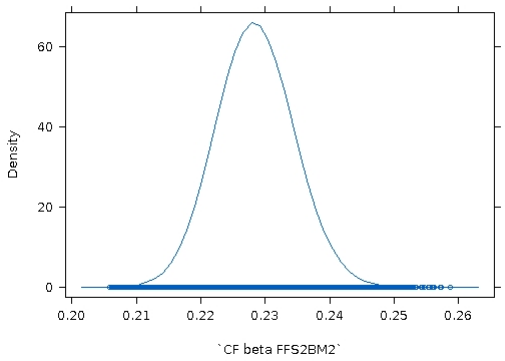
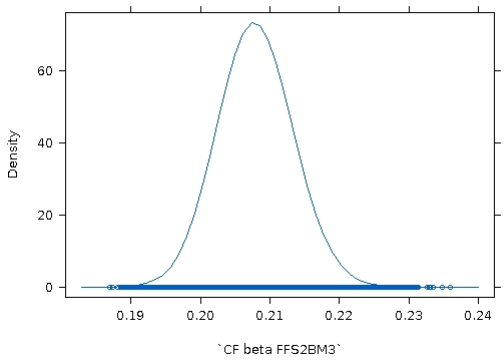
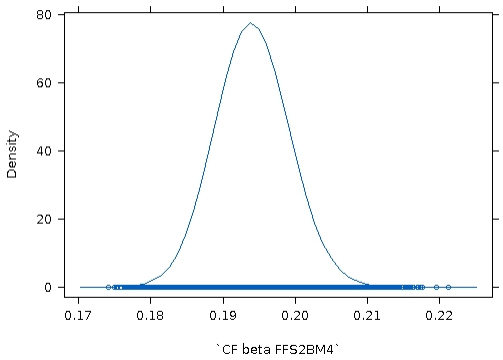


Figure 2 (continued)









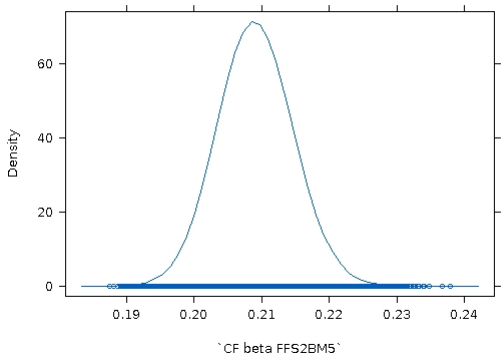
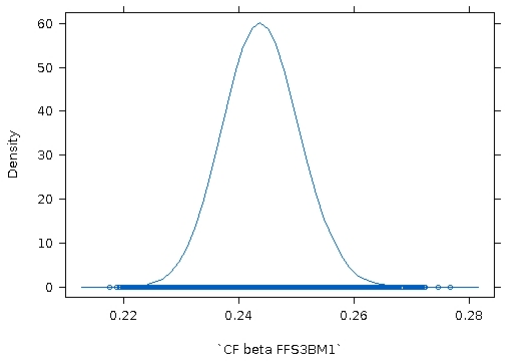
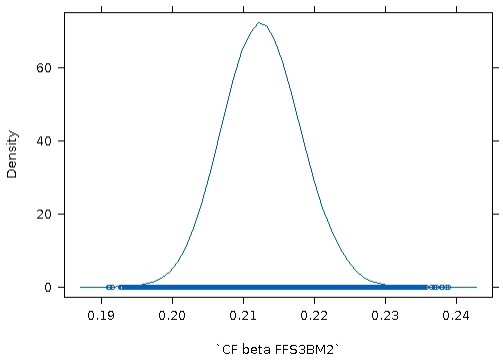
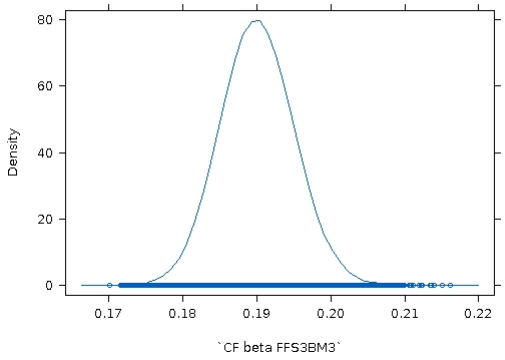
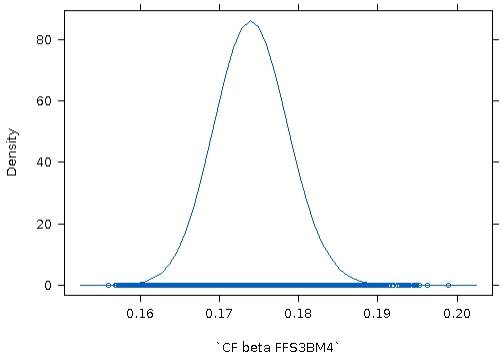


Figure 2 (continued)









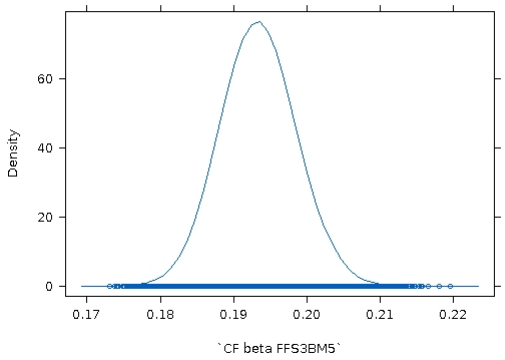
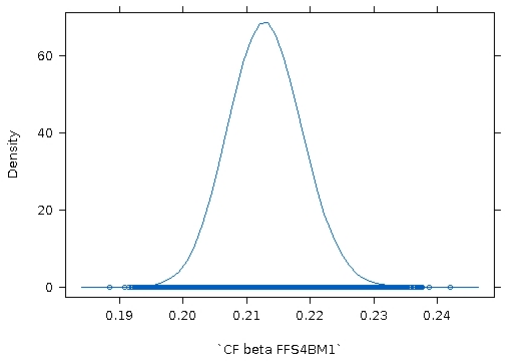
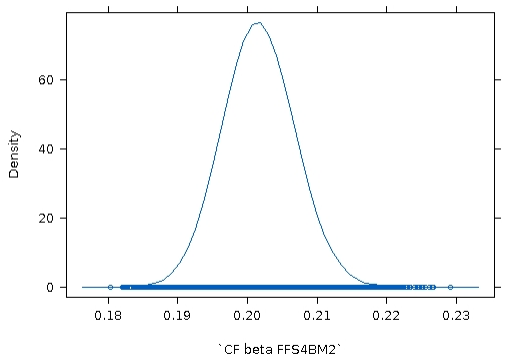
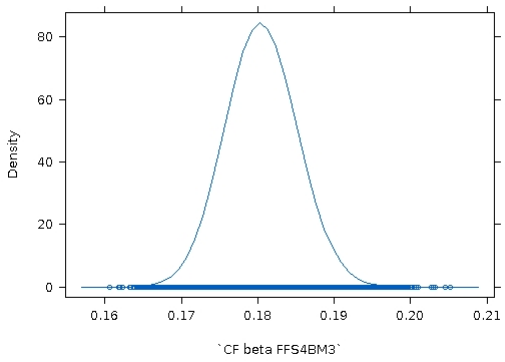
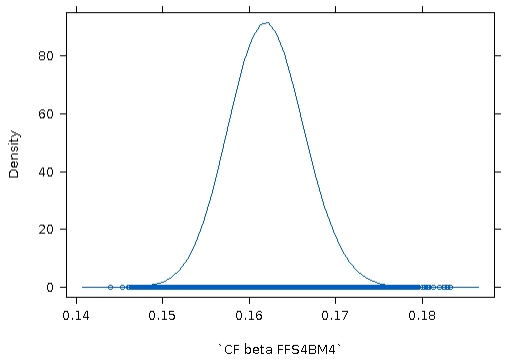


Figure 2 (continued)









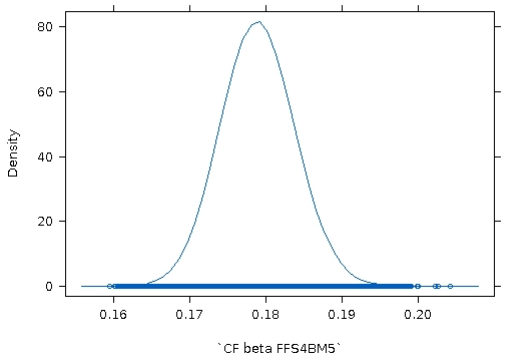
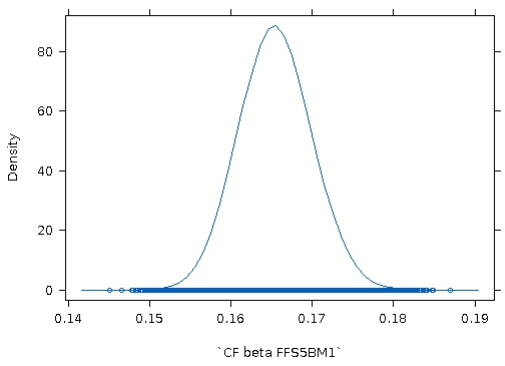
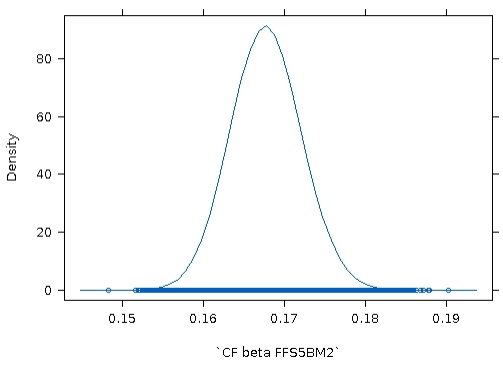
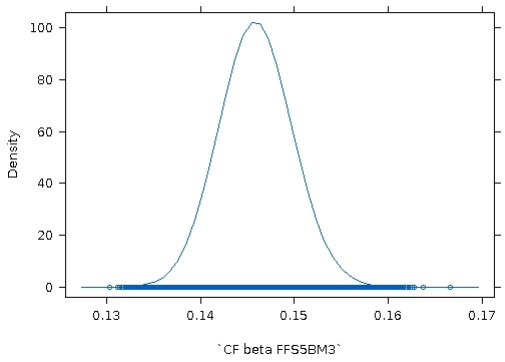
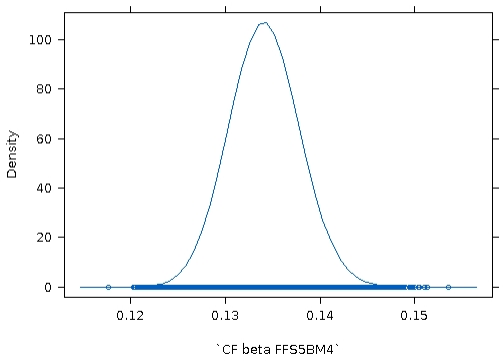


Figure 2 (continued)









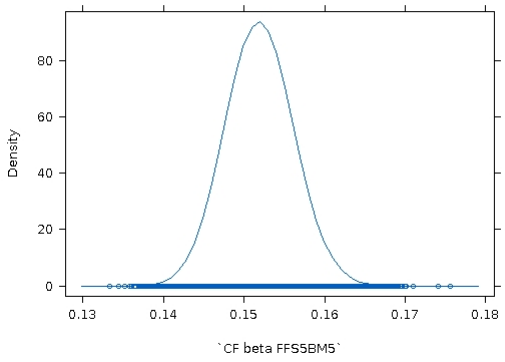
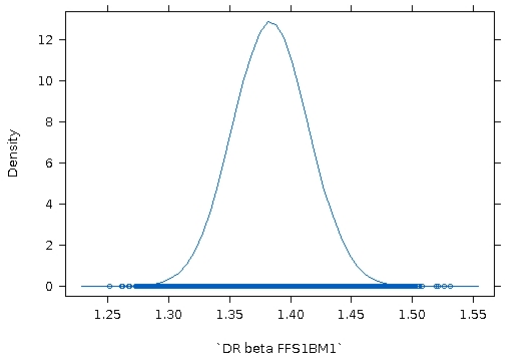
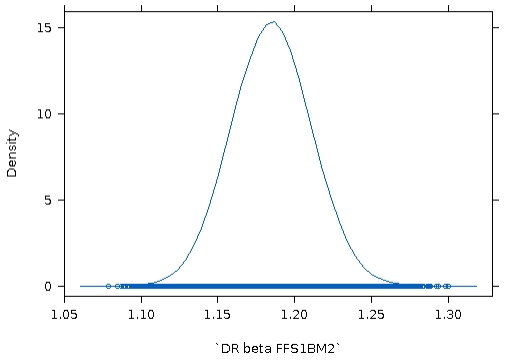
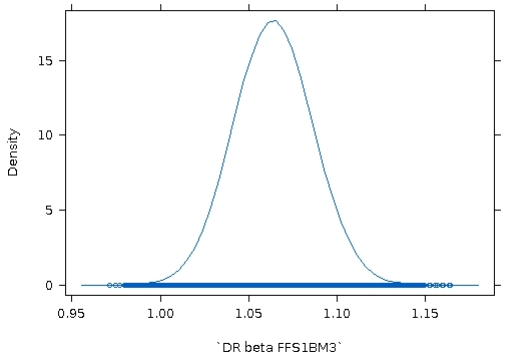
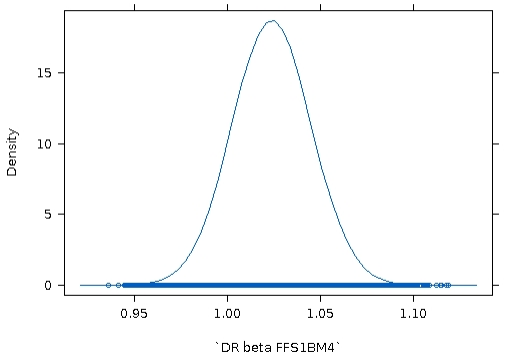


Figure 3: Density plots of discount-rate betas for the 25 Fama-French portfolios during 1963:7-2001:12









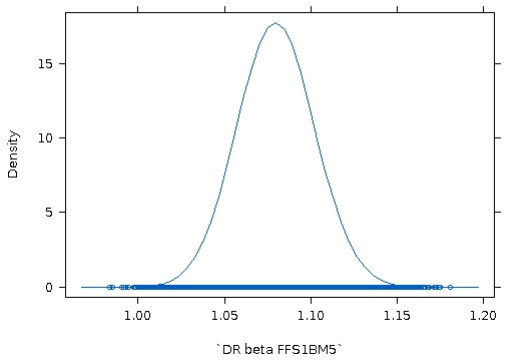
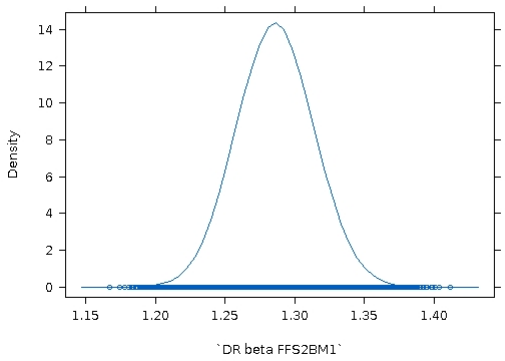
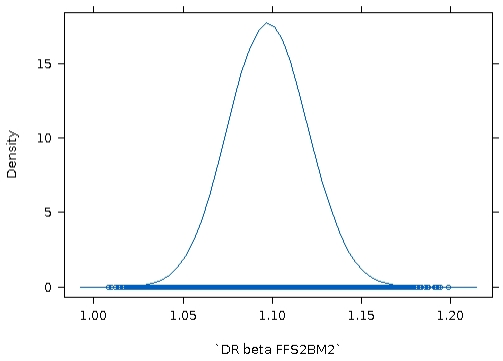
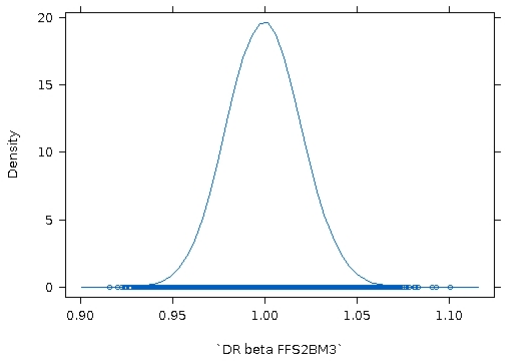
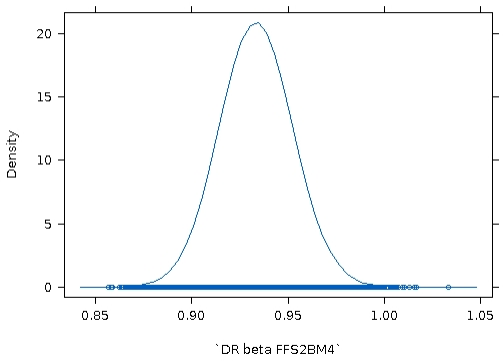


Figure 3 (continued)









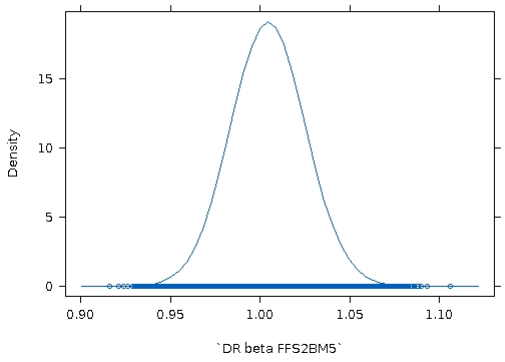
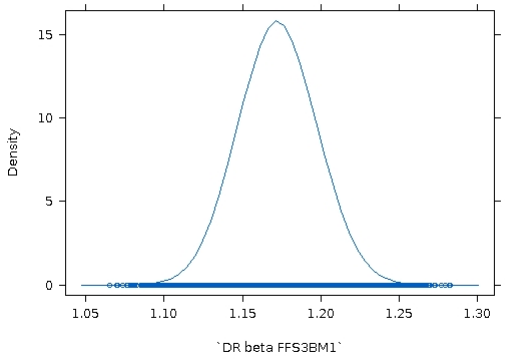
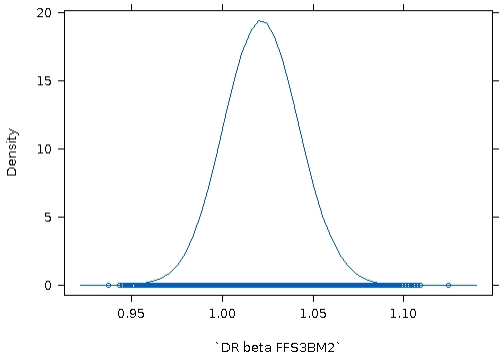
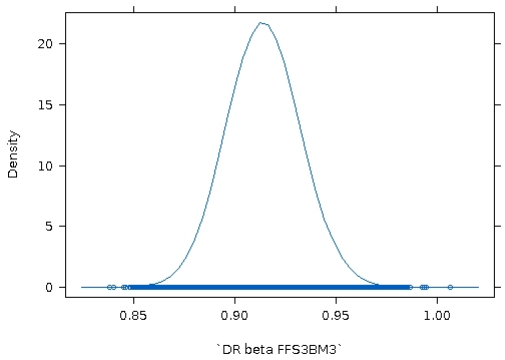
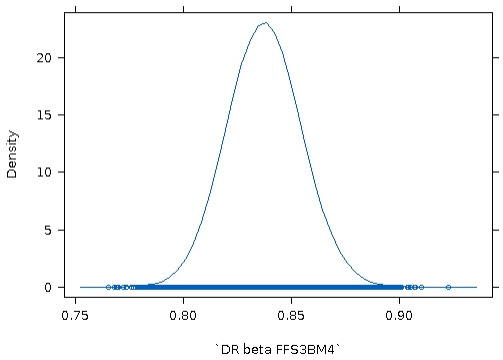


Figure 3 (continued)









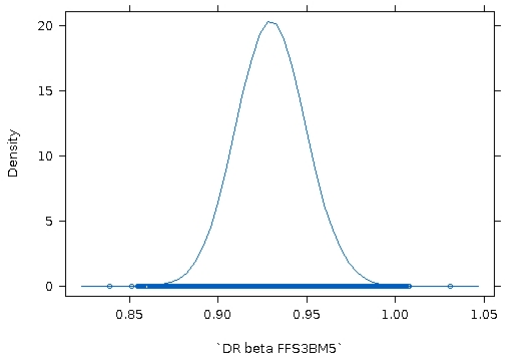
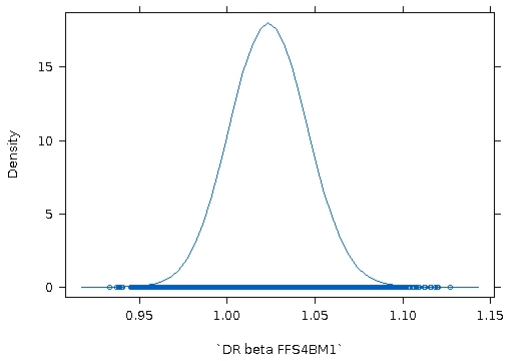
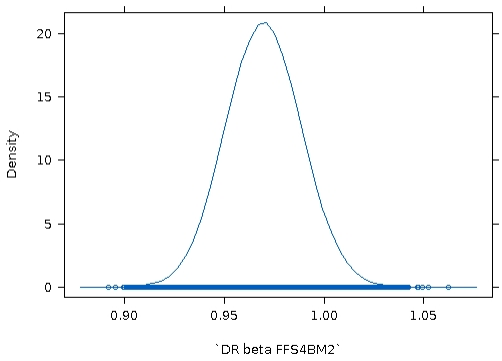
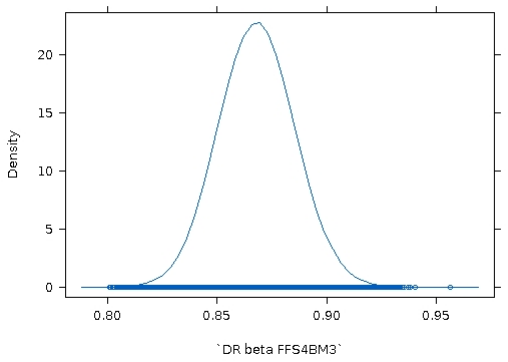
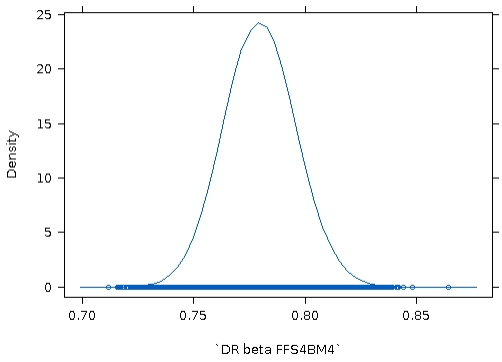


Figure 3 (continued)









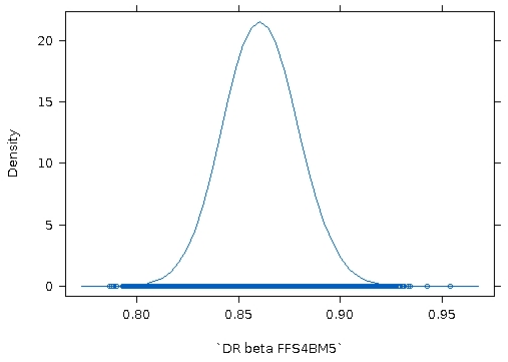
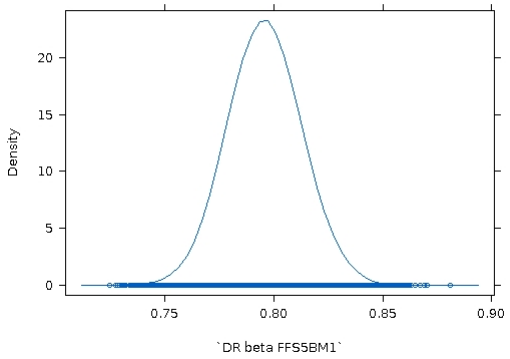
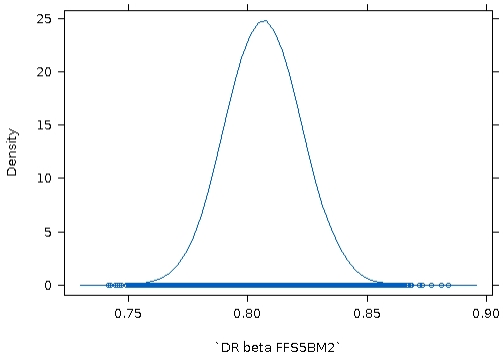
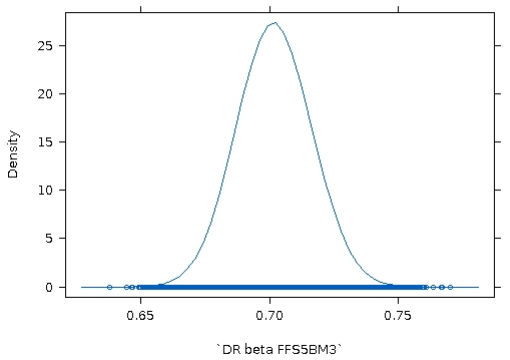
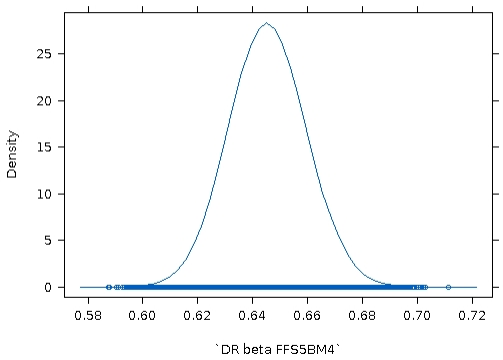


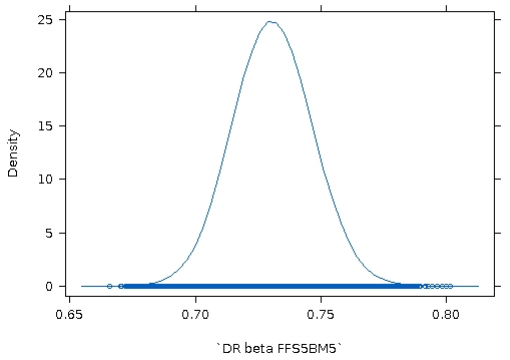
Figure 3 (continued)











Appendix A

I show that $\frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(N_{CF,t} - N_{DR,t})} = \frac{E[r_{i,t}(N_{CF,t} - N_{DR,t})]}{E[(N_{CF,t} - N_{DR,t})^2]}$.

We know that $Var(N_{CF,t} - N_{DR,t}) = E[(N_{CF,t} - N_{DR,t})^2] - [E(N_{CF,t} - N_{DR,t})]^2$. (A.1)

Based on equation (3.4), the return decomposition equation $r^e_{M,t} - E_{t-1}r^e_{M,t} = N_{CF,t} - N_{DR,t}$, I have

$$\begin{aligned} E(N_{CF,t} - N_{DR,t}) &= E(r^e_{M,t} - E_{t-1}r^e_{M,t}) \\ &= E(r^e_{M,t}) - E(E_{t-1}r^e_{M,t}) \\ &= 0, \end{aligned} \quad \text{(A.2)}$$

since $E(E_{t-1}r^e_{M,t}) = E(r^e_{M,t})$ by the Law of Iterated Expectations.

According to (A.1) and (A.2),

$$Var(N_{CF,t} - N_{DR,t}) = E[(N_{CF,t} - N_{DR,t})^2]. \quad \text{(A.3)}$$

By the definition of covariance, I have

$$Cov(r_{i,t}, N_{CF,t} - N_{DR,t}) = E[r_{i,t}(N_{CF,t} - N_{DR,t})] - E(r_{i,t})E(N_{CF,t} - N_{DR,t}). \quad \text{(A.4)}$$

I have already shown that $E(N_{CF,t} - N_{DR,t}) = 0$ in (A.2). Therefore,

$$Cov(r_{i,t}, N_{CF,t} - N_{DR,t}) = E[r_{i,t}(N_{CF,t} - N_{DR,t})]. \quad \text{(A.5)}$$

Based on equations (A.3) and (A.5), I have shown show that

$$\frac{Cov(r_{i,t}, N_{CF,t} - N_{DR,t})}{Var(N_{CF,t} - N_{DR,t})} = \frac{E[r_{i,t}(N_{CF,t} - N_{DR,t})]}{E[(N_{CF,t} - N_{DR,t})^2]}.$$

Appendix B

Additional tables (Tables 13 to 20)

Table 13: Descriptive statistics for the posterior distributions of parameters in the state-space model

The table shows the means, variances, 2.5% percentiles, and 97.5% percentiles of the posterior distributions (150,000 iterations) of parameters in the state-space model.

	Mean	Variance	2.50%	97.50%
$\beta_{1,M}$	1.0372782	0.0001237	1.0156170	1.0545990
$\beta_{2,M}$	0.9854012	0.0000666	0.9704689	1.0006130
$\beta_{3,M}$	0.9275894	0.0000747	0.9089295	0.9426939
$\beta_{4,M}$	0.9143618	0.0000884	0.8958690	0.9337263
$\beta_{5,M}$	1.0092124	0.0001441	0.9842414	1.0337780
$-\beta_{b1}$	-0.0139696	0.0000040	-0.0178977	-0.0099368
$-\beta_{b2}$	-0.0233103	0.0000118	-0.0300623	-0.0160738
$-\beta_{b3}$	-0.0288435	0.0000230	-0.0383427	-0.0187383
$-\beta_{b4}$	-0.0361210	0.0000386	-0.0482521	-0.0234466
$-\beta_{b5}$	-0.0407857	0.0000537	-0.0554230	-0.0259904
F_{11}	0.7357286	0.0858727	0.3702284	0.9958299
F_{12}	-0.0934053	0.0001188	-0.1062451	-0.0753246
F_{13}	-0.0202012	0.0000038	-0.0228379	-0.0165799
R_{11}	0.0000130	0.0000000	0.0000113	0.0000147
R_{22}	0.0000330	0.0000000	0.0000302	0.0000359
R_{33}	0.0000151	0.0000000	0.0000136	0.0000167
R_{44}	0.0000163	0.0000000	0.0000148	0.0000179
R_{55}	0.0000218	0.0000000	0.0000199	0.0000238
R_{66}	0.0000373	0.0000000	0.0000344	0.0000403
R_{77}	0.0000015	0.0000000	0.0000014	0.0000015
R_{88}	0.0000046	0.0000000	0.0000045	0.0000047
R_{99}	0.0000090	0.0000000	0.0000088	0.0000092
$R_{10,10}$	0.0000152	0.0000000	0.0000148	0.0000156
$R_{11,11}$	0.0000208	0.0000000	0.0000202	0.0000213
V_{11}	0.0000000	0.0000000	0.0000000	0.0000000
V_{12}	0.0000000	0.0000000	-0.0000001	0.0000000
V_{13}	0.0000001	0.0000000	0.0000000	0.0000004
V_{22}	0.0000044	0.0000000	0.0000041	0.0000048
V_{23}	-0.0000192	0.0000000	-0.0000205	-0.0000180
V_{33}	0.0000938	0.0000000	0.0000888	0.0000990

Table 14: Cash-flow betas and discount-rate betas for the 1952:7-2010:12 period

Panel A of the table shows the posterior means and standard deviations [in brackets] of cash-flow and discount-rate betas for the 25 Fama-French portfolios. Panel B and C show the 95% posterior intervals and posterior medians, respectively, of cash-flow and discount-rate betas for the 25 Fama-French portfolios. "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of cash-flow beta and discount-rate beta of 150,000 iterations are obtained and used to calculate the means, standard deviations, and 95% intervals. The column names "2.5%" and "97.5%" are the 2.5% and 97.5% percentiles of the posterior distribution, respectively. The posterior distributions of cash-flow and discount-rate betas are for the 1952:7-2010:12 period.

Panel A: Posterior Means and Standard Deviations

$\hat{\beta}_{CF}$	Growth		2		3		4		Value	
Small	0.281	[0.007]	0.245	[0.006]	0.220	[0.005]	0.213	[0.005]	0.239	[0.006]
2	0.257	[0.006]	0.226	[0.005]	0.208	[0.005]	0.204	[0.005]	0.230	[0.005]
3	0.234	[0.006]	0.210	[0.005]	0.195	[0.004]	0.191	[0.004]	0.204	[0.005]
4	0.208	[0.005]	0.202	[0.005]	0.195	[0.004]	0.183	[0.004]	0.206	[0.005]
Large	0.167	[0.004]	0.169	[0.004]	0.156	[0.004]	0.157	[0.004]	0.177	[0.004]

$\hat{\beta}_{DR}$	Growth		2		3		4		Value	
Small	1.347	[0.024]	1.175	[0.020]	1.053	[0.018]	1.019	[0.017]	1.144	[0.019]
2	1.229	[0.021]	1.085	[0.018]	0.995	[0.016]	0.976	[0.016]	1.104	[0.018]
3	1.123	[0.020]	1.005	[0.016]	0.933	[0.015]	0.914	[0.015]	0.977	[0.016]
4	0.998	[0.017]	0.968	[0.015]	0.934	[0.015]	0.879	[0.014]	0.989	[0.016]
Large	0.800	[0.014]	0.809	[0.013]	0.749	[0.013]	0.752	[0.013]	0.849	[0.015]

Panel B: 95% Posterior Intervals

$\hat{\beta}_{CF}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	0.268	0.295	0.234	0.257	0.210	0.230	0.203	0.222	0.228	0.250
2	0.245	0.269	0.216	0.237	0.198	0.217	0.195	0.213	0.220	0.241
3	0.224	0.245	0.200	0.219	0.186	0.203	0.182	0.199	0.195	0.213
4	0.199	0.218	0.193	0.211	0.186	0.204	0.175	0.192	0.197	0.216
Large	0.159	0.175	0.161	0.176	0.149	0.164	0.150	0.164	0.169	0.186

$\hat{\beta}_{DR}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	1.301	1.395	1.135	1.214	1.018	1.087	0.986	1.052	1.108	1.181
2	1.188	1.271	1.050	1.120	0.964	1.026	0.946	1.007	1.069	1.139
3	1.085	1.161	0.973	1.037	0.904	0.963	0.885	0.943	0.945	1.008
4	0.964	1.032	0.938	0.998	0.904	0.963	0.851	0.907	0.958	1.022
Large	0.772	0.827	0.784	0.835	0.725	0.774	0.727	0.777	0.821	0.878

Table 14 (continued from previous page)

Panel C: Posterior Medians					
$\hat{\beta}_{CF}$	Growth	2	3	4	Value
Small	0.2813	0.2452	0.2196	0.2126	0.2388
2	0.2565	0.2263	0.2075	0.2036	0.2304
3	0.2342	0.2096	0.1946	0.1906	0.2036
4	0.2081	0.2018	0.1948	0.1833	0.2064
Large	0.1667	0.1688	0.1563	0.1570	0.1773

$\hat{\beta}_{DR}$	Growth	2	3	4	Value
Small	1.3474	1.1746	1.0526	1.0187	1.1441
2	1.2293	1.0849	0.9948	0.9762	1.1041
3	1.1227	1.0048	0.9331	0.9138	0.9765
4	0.9977	0.9678	0.9336	0.8785	0.9894
Large	0.7994	0.8093	0.7492	0.7520	0.8494

Table 15: Difference in cash-flow and discount-rate betas between value and growth stocks for 1952:7-2010:12

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between value and growth stocks for the period 1952:7-2010:12. For each of the five size categories, the posterior distribution of the difference in cash-flow beta between value and growth stocks (i.e., $\hat{\beta}_{value,CF} - \hat{\beta}_{growth,CF}$) of 150,000 iterations is obtained using the posterior distributions of cash-flow betas. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between value and growth stocks for the period 1952:7-2010:12. For each of the five size categories, the posterior distribution of the difference in discount-rate beta between value and growth stocks (i.e., $\hat{\beta}_{value,DR} - \hat{\beta}_{growth,DR}$) of 150,000 iterations is obtained using the posterior distributions of discount-rate betas. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{value,CF} - \hat{\beta}_{growth,CF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.0425	-0.0425	0.0026	-0.0477	-0.0375	-0.0443	-0.0407
2	-0.0261	-0.0261	0.0025	-0.0310	-0.0212	-0.0278	-0.0244
3	-0.0306	-0.0306	0.0026	-0.0356	-0.0255	-0.0323	-0.0288
4	-0.0017	-0.0017	0.0025	-0.0065	0.0032	-0.0033	0.0000
Large	0.0106	0.0106	0.0023	0.0062	0.0151	0.0091	0.0121

$\hat{\beta}_{value,DR} - \hat{\beta}_{growth,DR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.2034	-0.2033	0.0121	-0.2273	-0.1797	-0.2116	-0.1952
2	-0.1253	-0.1253	0.0119	-0.1486	-0.1021	-0.1333	-0.1174
3	-0.1462	-0.1462	0.0123	-0.1704	-0.1221	-0.1545	-0.1379
4	-0.0085	-0.0085	0.0118	-0.0317	0.0147	-0.0164	-0.0005
Large	0.0500	0.0500	0.0108	0.0288	0.0712	0.0427	0.0573

Table 16: Difference in cash-flow and discount-rate betas between large and small stocks for 1952:7-2010:12

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between large and small stocks for the period 1952:7-2010:12. For each of the five BE/ME categories, the posterior distribution of the difference in cash-flow beta between large and small stocks (i.e., $\hat{\beta}_{large,CF} - \hat{\beta}_{small,CF}$) of 150,000 iterations is obtained using the posterior distributions of cash-flow betas. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between large and small stocks for the period 1952:7-2010:12. For each of the five BE/ME categories, the posterior distribution of the difference in discount-rate beta between large and small stocks (i.e., $\hat{\beta}_{large,DR} - \hat{\beta}_{small,DR}$) of 150,000 iterations is obtained using the posterior distributions of discount-rate betas. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{large,CF} - \hat{\beta}_{small,CF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.1146	-0.1146	0.0039	-0.1224	-0.1071	-0.1172	-0.1120
2	-0.0765	-0.0764	0.0031	-0.0827	-0.0704	-0.0786	-0.0744
3	-0.0634	-0.0633	0.0027	-0.0687	-0.0582	-0.0651	-0.0615
4	-0.0556	-0.0556	0.0025	-0.0606	-0.0508	-0.0573	-0.0539
Value	-0.0615	-0.0615	0.0026	-0.0667	-0.0564	-0.0633	-0.0597

$\hat{\beta}_{large,DR} - \hat{\beta}_{small,DR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.5480	-0.5479	0.0164	-0.5803	-0.5161	-0.5590	-0.5369
2	-0.3653	-0.3653	0.0137	-0.3921	-0.3384	-0.3745	-0.3560
3	-0.3034	-0.3034	0.0118	-0.3266	-0.2803	-0.3113	-0.2955
4	-0.2666	-0.2666	0.0111	-0.2886	-0.2448	-0.2741	-0.2591
Value	-0.2946	-0.2946	0.0117	-0.3176	-0.2718	-0.3025	-0.2867

Table 17: Posterior means and standard deviations of cash-flow and discount-rate betas for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news

The table shows the posterior means and standard deviations [in brackets] of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$) calculated using orthogonalized cash-flow news and discount-rate news for the 25 Fama-French portfolios for the period 1952:7-2010:12. "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ of 150,000 iterations are obtained and used to calculate posterior means and standard deviations. The top half of the table reports the mean and standard deviation of the 150,000 iterations of $\hat{\beta}_{OCF}$ for each portfolio, and the bottom half of the table reports the mean and standard deviation of the 150,000 iterations of $\hat{\beta}_{ODR}$ for each portfolio.

$\hat{\beta}_{OCF}$	Growth	2	3	4	Value
Small	-0.0000285	-0.0000243	-0.0000206	-0.0000197	-0.0000219
	[0.0087]	[0.0076]	[0.0068]	[0.0066]	[0.0074]
2	-0.0000253	-0.0000213	-0.0000195	-0.0000183	-0.0000215
	[0.0079]	[0.0070]	[0.0064]	[0.0063]	[0.0071]
3	-0.0000237	-0.0000201	-0.0000181	-0.0000176	-0.0000182
	[0.0073]	[0.0065]	[0.0060]	[0.0059]	[0.0063]
4	-0.0000213	-0.0000189	-0.0000192	-0.0000177	-0.0000191
	[0.0065]	[0.0062]	[0.0060]	[0.0057]	[0.0064]
Large	-0.0000173	-0.0000158	-0.0000152	-0.0000151	-0.0000163
	[0.0052]	[0.0052]	[0.0048]	[0.0048]	[0.0055]

$\hat{\beta}_{ODR}$	Growth	2	3	4	Value
Small	1.969	1.717	1.538	1.489	1.672
	[0.036]	[0.030]	[0.026]	[0.025]	[0.028]
2	1.797	1.585	1.454	1.427	1.613
	[0.032]	[0.027]	[0.024]	[0.024]	[0.027]
3	1.641	1.468	1.364	1.335	1.427
	[0.029]	[0.024]	[0.023]	[0.022]	[0.024]
4	1.458	1.414	1.364	1.284	1.446
	[0.026]	[0.023]	[0.023]	[0.021]	[0.025]
Large	1.168	1.183	1.095	1.099	1.241
	[0.021]	[0.020]	[0.019]	[0.019]	[0.022]

Table 18: Posterior medians and 95% intervals of cash-flow and discount-rate betas for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows the posterior medians of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$) calculated using orthogonalized cash-flow news and discount-rate news for the 25 Fama-French portfolios. The bottom half of the table shows the 95% posterior intervals of cash-flow betas ($\hat{\beta}_{OCF}$) and discount-rate betas ($\hat{\beta}_{ODR}$). "Growth" denotes stocks with the lowest BE/ME, "Value" denotes stocks with the highest BE/ME, "Small" denotes stocks with the lowest market equity, and "Large" denotes stocks with the highest market equity. For each portfolio, posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ of 150,000 iterations are obtained and used to calculate posterior medians and 95% intervals. The posterior distributions of $\hat{\beta}_{OCF}$ and $\hat{\beta}_{ODR}$ are for the 1952:7-2010:12 period. The column names "2.5%" and "97.5%" are the 2.5% and 97.5% percentiles of the posterior distribution, respectively.

Posterior Medians

$\hat{\beta}_{OCF}$	Growth	2	3	4	Value
Small	-0.00021	-0.00018	-0.00015	-0.00014	-0.00015
2	-0.00018	-0.00015	-0.00014	-0.00013	-0.00015
3	-0.00018	-0.00014	-0.00013	-0.00012	-0.00012
4	-0.00016	-0.00013	-0.00013	-0.00013	-0.00013
Large	-0.00013	-0.00011	-0.00011	-0.00011	-0.00012
$\hat{\beta}_{ODR}$	Growth	2	3	4	Value
Small	1.969	1.717	1.538	1.489	1.672
2	1.797	1.585	1.454	1.427	1.613
3	1.641	1.468	1.364	1.335	1.427
4	1.458	1.414	1.364	1.284	1.446
Large	1.168	1.183	1.095	1.099	1.241

95% Posterior Intervals

$\hat{\beta}_{OCF}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	-0.016	0.018	-0.014	0.015	-0.013	0.014	-0.012	0.013	-0.014	0.015
2	-0.015	0.016	-0.013	0.014	-0.012	0.013	-0.012	0.012	-0.013	0.014
3	-0.014	0.015	-0.012	0.013	-0.011	0.012	-0.011	0.012	-0.012	0.012
4	-0.012	0.013	-0.012	0.012	-0.011	0.012	-0.011	0.011	-0.012	0.013
Large	-0.010	0.010	-0.010	0.010	-0.009	0.010	-0.009	0.010	-0.010	0.011
$\hat{\beta}_{ODR}$	Growth		2		3		4		Value	
	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%	2.5%	97.5%
Small	1.899	2.040	1.657	1.776	1.487	1.590	1.439	1.538	1.617	1.727
2	1.734	1.859	1.533	1.638	1.407	1.501	1.381	1.473	1.561	1.667
3	1.583	1.699	1.421	1.516	1.320	1.408	1.292	1.379	1.380	1.475
4	1.407	1.509	1.369	1.460	1.320	1.409	1.242	1.326	1.398	1.494
Large	1.127	1.210	1.145	1.221	1.058	1.132	1.062	1.137	1.199	1.285

Table 19: Difference in cash-flow and discount-rate betas between value and growth stocks for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between value and growth stocks for the period 1952:7-2010:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five size categories, the posterior distribution of the difference in cash-flow beta between value and growth stocks (i.e., $\hat{\beta}_{value,OCF} - \hat{\beta}_{growth,OCF}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{OCF}$. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between value and growth stocks for the period 1952:7-2010:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five size categories, the posterior distribution of the difference in discount-rate beta between value and growth stocks (i.e., $\hat{\beta}_{value,ODR} - \hat{\beta}_{growth,ODR}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{ODR}$. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{value,OCF} - \hat{\beta}_{growth,OCF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	0.0000067	0.0000599	0.0014	-0.0029	0.0026	-0.0009	0.0010
2	0.0000038	0.0000341	0.0009	-0.0017	0.0016	-0.0006	0.0006
3	0.0000055	0.0000570	0.0011	-0.0022	0.0019	-0.0007	0.0008
4	0.0000022	-0.0000630	0.0003	-0.0005	0.0005	-0.0002	0.0003
Large	0.0000010	0.0000125	0.0004	-0.0008	0.0007	-0.0003	0.0003

$\hat{\beta}_{value,ODR} - \hat{\beta}_{growth,ODR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Small	-0.2972	-0.2971	0.0178	-0.3322	-0.2625	-0.3092	-0.2852
2	-0.1832	-0.1832	0.0174	-0.2172	-0.1492	-0.1949	-0.1715
3	-0.2137	-0.2137	0.0180	-0.2491	-0.1784	-0.2258	-0.2016
4	-0.0123	-0.0124	0.0173	-0.0463	0.0215	-0.0240	-0.0007
Large	0.0730	0.0730	0.0158	0.0421	0.1040	0.0624	0.0837

Table 20: Difference in cash-flow and discount-rate betas between large and small stocks for 1952:7-2010:12 calculated using orthogonalized cash-flow news and discount-rate news

The top half of the table shows descriptive statistics of the posterior distribution of the difference in cash-flow beta between large and small stocks for the period 1952:7-2010:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five BE/ME categories, the posterior distribution of the difference in cash-flow beta between large and small stocks (i.e., $\hat{\beta}_{large,OCF} - \hat{\beta}_{small,OCF}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{OCF}$. The bottom half of the table shows descriptive statistics of the posterior distribution of the difference in discount-rate beta between large and small stocks for the period 1952:7-2010:12 obtained using orthogonalized cash-flow news and discount-rate news. For each of the five BE/ME categories, the posterior distribution of the difference in discount-rate beta between large and small stocks (i.e., $\hat{\beta}_{large,ODR} - \hat{\beta}_{small,ODR}$) of 150,000 iterations is obtained using the posterior distributions of $\hat{\beta}_{ODR}$. The column name "Stdev." denotes standard deviation. The column names "2.5%", "97.5%", "25%", and "75%" denote the 2.5%, 97.5%, 25%, and 75% percentiles, respectively.

$\hat{\beta}_{large,OCF} - \hat{\beta}_{small,OCF}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	0.0000112	0.0000845	0.0036	-0.0071	0.0066	-0.0024	0.0025
2	0.0000084	0.0000697	0.0024	-0.0048	0.0045	-0.0016	0.0017
3	0.0000054	0.0000368	0.0020	-0.0039	0.0037	-0.0013	0.0014
4	0.0000046	0.0000321	0.0017	-0.0034	0.0032	-0.0012	0.0012
Value	0.0000056	0.0000391	0.0019	-0.0038	0.0036	-0.0013	0.0013

$\hat{\beta}_{large,ODR} - \hat{\beta}_{small,ODR}$							
	Mean	Median	Stdev.	2.5%	97.5%	25%	75%
Growth	-0.8008	-0.8007	0.0242	-0.8484	-0.7538	-0.8171	-0.7845
2	-0.5338	-0.5338	0.0201	-0.5733	-0.4943	-0.5474	-0.5202
3	-0.4434	-0.4434	0.0173	-0.4775	-0.4095	-0.4550	-0.4317
4	-0.3896	-0.3896	0.0164	-0.4218	-0.3576	-0.4007	-0.3786
Value	-0.4305	-0.4305	0.0171	-0.4642	-0.3970	-0.4421	-0.4190

Appendix C

Preliminary analysis using the VAR model approach in the early stage of the research

1. Introduction

In the early stage of the research, preliminary analysis using the VAR model approach is carried out to examine the applicability of this approach to implementing Campbell's (1991) return decomposition, that is, extracting cash-flow and discount-rate news from the unexpected market return. The findings suggest that the VAR approach does not have sufficient power to extract the cash-flow and discount-rate news.

Campbell and Vuolteenaho (2004) develop the two-beta ICAPM by utilizing the return decomposition to rewrite Campbell's (1993) asset pricing formula. The two-beta ICAPM states that the risk premium on any asset is a linear combination of the asset's cash-flow and discount-rate betas. Furthermore, the model captures the intuition that investors price cash-flow risk higher than discount-rate risk. Following Campbell (1991), Campbell and Vuolteenaho (2004) use the VAR approach to implement the return decomposition for the period of 1929:1-2001:12. Variables included in the VAR model should predict market returns, and the variables that they select are the excess log market return, the yield spread between long-term and short-term bonds, the 10-year smoothed price-earnings ratio from Shiller (2000), and the small-stock value spread. Campbell and Vuolteenaho (2004) use the estimated cash-flow and discount-rate news to calculate the cash-flow and discount-rate betas of the 25 Fama-French portfolios for both the 1929:1-1963:6 and 1963:7-2001:12 periods. They find that value stocks have higher cash-flow

betas than do growth stocks, which indicates that the value effect is consistent with the implication of the two-beta ICAPM.

Although the two-beta ICAPM is theoretically sound, its empirical performance has been challenged by, for example, Liu and Zhang (2008) and Chen and Zhao (2009). Liu and Zhang (2008) cast doubt on the ability of the small-stock value spread to predict the market return, and therefore, the small-stock value spread may not be an appropriate variable to be used in the VAR model. Chen and Zhao (2009) find that Campbell and Vuolteenaho's (2004) results of the cash-flow and discount-rate betas are sensitive to the choice of variables selected in the VAR model. In particular, they find that it is necessary to include the small-stock value spread in a VAR model to obtain the patterns in cash-flow betas that could explain the value effect. Chen and Zhao (2009) also find that Campbell and Vuolteenaho's (2004) results are sensitive to the time period used to estimate cash-flow and discount-rate news. Using the same variables in a VAR model as do Campbell and Vuolteenaho (2004) but a different time period of 1952-2001 to estimate the two news components, Chen and Zhao (2009) do not find the same patterns in cash-flow betas as Campbell and Vuolteenaho (2004) do.

Among the different VAR model specifications that Chen and Zhao (2009) consider, one VAR model increases the predictive power on market returns in the VAR regression but does not yield the patterns in cash-flow betas that could explain the value premium. Specifically, they find that the difference in cash-flow beta between value and growth stocks becomes less positive and even negative for three of the five size categories of the 25 Fama-French portfolios. However, this VAR model includes the variable *cay* , the log consumption-wealth ratio from Lettau and Ludvigson (2001a), and

Guo (2003) argues that there is a look-ahead bias in *cay* because data on consumption and labor income (data used to construct *cay*) are subject to substantial revisions to reflect the incorporation of new information and new methodologies and to reduce noise. Therefore, it may not be appropriate to use *cay* in the VAR model approach to implement the return decomposition, because it is likely that a variable with look-ahead bias may cause distortion to the estimated discount-rate news and to the estimated cash-flow news as well, since cash-flow news is backed out as the residual between the unexpected market return and discount-rate news.

While I use the VAR approach following Campbell and Vuolteenaho (2004), I utilize a different set of variables motivated by the identity of returns on a dividend-paying asset to estimate cash-flow news and discount-rate news, with which I test the implications of the two-beta ICAPM, particularly the ability of the model to explain the value premium. Starting with this identity, I present a simple derivation which shows that the log return on a dividend-paying asset is equal to the sum of: (1) the differenced log price-earnings ratio, (2) log real earnings growth rate, (3) log inflation, and (4) log dividend-price ratio. Accordingly, the log market return is equal to the sum of the differenced log price-earnings ratio of the market, log real earnings growth rate of the market, log inflation, and log dividend-price ratio of the market. Since these four variables are motivated by the identity of asset returns, there is an economic reason that they should predict the market return, and I use them in the VAR model approach to estimate the cash-flow and discount-rate news. Using the cash-flow and discount-rate news estimated from this VAR model, I calculate cash-flow and discount-rate betas for the 25 Fama-French portfolios for the periods 1929:1-1963:6 and 1963:7-2001:12. I find

that value stocks have both higher cash-flow betas and higher discount-rate betas than do growth stocks during 1929:1-1963:6, and, therefore, the value effect is consistent with the implications of the two-beta ICAPM. However, during 1963:7-2001:12, value stocks have both lower cash-flow betas and lower discount-rate betas than do growth stocks; thus, the value premium during this period could not be explained.

To show the effect of using look-ahead biased variables in the VAR model approach on the estimated cash-flow and discount-rate news, I create a synthetic look-ahead biased variable and include it in a VAR model that also includes the four variables that Campbell and Vuolteenaho (2004) use in their VAR model to obtain estimates of cash-flow and discount-rate news. I compare the cash-flow and discount-rate betas calculated using the news estimated from this VAR model and those calculated using the news estimated from the VAR model that includes only the four variables that Campbell and Vuolteenaho (2004) use. The results show that the difference in cash-flow beta between value and growth stocks becomes less positive when a synthetic look-ahead biased variable is included in a VAR model than when it is not. In addition, the difference in discount-rate beta between value and growth stocks becomes less negative when a synthetic look-ahead biased variable is included in a VAR model than when it is not. These results suggest the possibility that using look-ahead biased variables, such as *cay*, may weaken the patterns in cash-flow and discount-rate betas found in Campbell and Vuolteenaho (2004).

2. Campbell and Vuolteenaho's (2004) VAR implementation

Following Campbell (1991), Campbell and Vuolteenaho (2004) use a vector autoregressive (VAR) model to estimate the discount-rate news series, $\{N_{DR,t}\}$, and the cash-flow news series, $\{N_{CF,t}\}$.

Campbell and Vuolteenaho (2004) define a $m\times 1$ vector z_{t+1} with the excess log market return $r^e_{M,t+1}$ as the first component, and the other components of z_{t+1} are variables known to investors by the end of period $t+1$. These variables are the term yield spread, the 10-year smoothed price-earnings ratio, and the small-stock value spread. They assume that z_{t+1} follows a first-order VAR model:

$$z_{t+1} = a + Gz_t + u_{t+1}, \qquad \text{(A.6)}$$

where a is a $m\times 1$ vector of constants, G is a $m\times m$ matrix of parameters, and u_{t+1} is an i.i.d. $m\times 1$ vector of white noise. Define a $m\times 1$ vector $e1 \equiv \begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}'$ so that $e1'z_{t+1} = r^e_{M,t+1}$. Since higher-order VAR models can be represented in VAR(1) form, the assumption that the VAR is first order is not restrictive, and their procedure, described below, can be implemented as long as z_{t+1} follows a VAR model.

Based on (A.6), Campbell and Vuolteenaho (2004) present formulae for $N_{DR,t+1}$ and $N_{CF,t+1}$, which are originally derived by Campbell (1991). From (A.6), one can obtain

$$E_{t+1}z_{t+1+j} = a + Ga + G^2a + \ldots + G^{j-1}a + G^j z_{t+1} = \left(\sum_{i=0}^{j-1} G^i a\right) + G^j z_{t+1}, \qquad \text{(A.7)}$$

$$E_t z_{t+1+j} = a + Ga + G^2 a + ... + G^j a + G^{j+1} z_t = \left(\sum_{i=0}^{j} G^i a \right) + G^{j+1} z_t . \qquad \text{(A.8)}$$

Using (A.7) and (A.8), one obtains

$$\begin{aligned} E_{t+1} z_{t+1+j} - E_t z_{t+1+j} &= G^j z_{t+1} - G^j a - G^{j+1} z_t \\ &= G^j (z_{t+1} - a - G z_t) \\ &= G^j u_{t+1}, \qquad \text{(A.9)} \end{aligned}$$

where the last equality follows from (A.6).

$$\begin{aligned} (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j z_{t+1+j} &= \sum_{j=1}^{\infty} \rho^j [E_{t+1}(z_{t+1+j}) - E_t(z_{t+1+j})] \\ &= \sum_{j=1}^{\infty} \rho^j G^j u_{t+1} \\ &= \rho G (I - \rho G)^{-1} u_{t+1} \qquad \text{(A.10)} \end{aligned}$$

where the second equality in (A.10) follows from (A.9). ρ is a scalar with a value smaller than one.

Pre-multiplying both sides of (A.10) by $e1'$, one obtains

$$(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j e1' z_{t+1+j} = e1' \rho G (I - \rho G)^{-1} u_{t+1} . \qquad \text{(A.11)}$$

Since $e1' z_{t+1} = r^e_{M,t+1}$, the left hand side of (A.11) is $(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r^e_{M,t+1+j}$, which is denoted by $N_{DR,t+1}$. Campbell and Vuolteenaho (2004) define $\lambda \equiv \rho G (I - \rho G)^{-1}$, so the formula for the discount rate news is

$$N_{DR,t+1} = e1' \lambda u_{t+1} . \qquad \text{(A.12)}$$

Note that $E_t r^e_{M,t+1} = E_t e1' z_{t+1} = e1' (a + G Z_t)$. (A.13)

Rearranging the terms in (3.4), one obtains

$$N_{CF,t+1} = r^e_{M,t+1} - E_t r^e_{M,t+1} + N_{DR,t+1}.$$

Substituting (A.12), (A.13), and $r^e_{M,t+1} = e1'z_{t+1}$ into the above equation yields

$$\begin{aligned} N_{CF,t+1} &= e1'z_{t+1} - e1'(a + GZ_t) + e1'\lambda u_{t+1} \\ &= e1'[z_{t+1} - (a + GZ_t)] + e1'\lambda u_{t+1} \\ &= e1'u_{t+1} + e1'\lambda u_{t+1} \\ &= (e1' + e1'\lambda)u_{t+1}. \end{aligned} \qquad \text{(A.14)}$$

To sum up, the procedure in Campbell and Vuolteenaho (2004) is to first estimate the parameters and the noise in (A.6) and then to use the formulae (A.12) and (A.14) to obtain estimates of the discount-rate news series and the cash-flow news series. The estimated series $\{N_{CF,t+1}\}$ and $\{N_{DR,t+1}\}$ are used to calculate the cash-flow and discount-rate betas of stock portfolios. Finally, the empirical performance of the asset pricing model in (3.13) can be evaluated by running a cross-sectional regression of test asset returns on their cash-flow betas and discount-rate betas to see how well the model explains the cross section of average returns.

3. My implementation using VAR approach

Choice of variables

Following Campbell and Vuolteenaho (2004), I use the VAR methodology to estimate $\{N_{CF,t+1}\}$ and $\{N_{DR,t+1}\}$. However, I utilize a different set of variables included in the vector z_{t+1} of (A.6). These variables are the log price-earnings ratio, log real

earnings growth rate, log inflation, and log dividend yield. The choice of variables to include in z_{t+1} is motivated by the definition of the return on a dividend-paying asset,

$$R_t \equiv \frac{P_t + D_t - P_{t-1}}{P_{t-1}}, \qquad \text{(A.15)}$$

where P_t denotes the asset's price at the end of period t, and D_t denotes the dividend paid during period t.

I start with the definition of the return on an asset, and then decompose it to several components including price-earnings ratio (P/E), dividend yield (D/P), inflation, and real earnings growth rate. Namely, using elementary algebra, I have

$$\begin{aligned} R_t &\equiv \frac{P_t + D_t - P_{t-1}}{P_{t-1}} \\ &= \frac{P_t}{P_{t-1}} + \frac{D_t}{P_{t-1}} - 1 \\ &= \frac{P_t}{P_{t-1}}\left(1 + \frac{D_t}{P_t}\right) - 1 \\ &= \left(\frac{E_{t-1}}{E_{t-1}}\right)\left(\frac{E_t}{E_t}\right)\left(\frac{P_t}{P_{t-1}}\right)\left(1 + \frac{D_t}{P_t}\right) - 1 \\ &= \left(\frac{P_t}{E_t}\right)\left(\frac{1}{P_{t-1}/E_{t-1}}\right)\left(\frac{E_t}{E_{t-1}}\right)\left(1 + \frac{D_t}{P_t}\right) - 1, \qquad \text{(A.16)} \end{aligned}$$

where E_t denotes nominal earnings by the end of period t. Define the price-earnings ratio $\psi_t \equiv P_t / E_t$ and the dividend yield $\delta_t \equiv D_t / P_t$.

Let $g_t \equiv \frac{E_t}{E_{t-1}} - 1$ denote the nominal earnings growth rate, g_t^r the real earnings growth rate, and π_t the inflation in period t. Then,

$$\frac{E_t}{E_{t-1}} = 1 + g_t = (1 + g_t^r)(1 + \pi_t). \qquad \text{(A.17)}$$

Substituting (A.17) into (A.16), I get

$$R_t = \frac{\psi_t}{\psi_{t-1}}(1 + g_t^r)(1 + \pi_t)\left(1 + \delta_t\right) - 1. \qquad \text{(A.18)}$$

Moving the 1 from the right hand side to the left hand side of (A.18) and then taking the natural logarithm of both sides of the equation yields

$$\log(1 + R_t) = \log\psi_t - \log\psi_{t-1} + \log(1 + g_t^r) + \log(1 + \pi_t) + \log\left(1 + \delta_t\right). \qquad \text{(A.19)}$$

According to equation (A.19), the log market return is the sum of the: (1) differenced log price-earnings ratio of the market, (2) log real earnings growth rate of the market, (3) log inflation, and (4) log dividend-price ratio of the market. Therefore, these four variables should be able to predict the market return. Accordingly, I use them in the VAR approach to estimate the cash-flow and discount-rate news.

Implementation using a VAR model

The formulae for $N_{DR,\,t+1}$ and $N_{CF,\,t+1}$ is of the same form as in Campbell and Vuolteenaho (2004). I define the first-order difference of the log price-earnings ratio $\Delta\log\psi_t \equiv \log\psi_t - \log\psi_{t-1}$. Let $z_{t+1} = \left[\Delta\log\psi_t \quad \log\left(1 + \delta_t\right) \quad \log(1 + \pi_t) \quad \log(1 + g_t^r)\right]'$ and assume that z_{t+1} follows the first-order VAR model in (A.6). Define a vector of ones $c = [1 \quad 1 \quad 1 \quad 1]'$, and let $r_{M,t+1}$ denote $\log(1 + R_{M,t+1})$. By (A.19), $r_{M,t+1} = c'z_{t+1}$. By definition, the discount-rate news is

$$N_{DR,\,t+1} \equiv (E_{t+1} - E_t)\sum_{j=1}^{\infty}\rho^j r_{M,t+1+j}$$

$$= (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j c' z_{t+1+j}$$

$$= c' \lambda u_{t+1}, \qquad \text{(A.20)}$$

where the second equality follows from (A.19) and the third equality follows from (A.10) and $\lambda \equiv \rho G(I - \rho G)^{-1}$. Note that

$$E_t r_{M,t+1} = E_t c' z_{t+1} = c'(a + GZ_t). \qquad \text{(A.21)}$$

By (3.2), cash-flow news is

$$N_{CF,t+1} = r_{M,t+1} - E_t r_{M,t+1} + N_{DR,t+1}.$$

Substituting (A.20) and (A.21) into the above equation, I obtain

$$N_{CF,t+1} = \mathrm{c}' z_{t+1} - c'(a + GZ_t) + c' \lambda u_{t+1}$$

$$= c'[z_{t+1} - (a + GZ_t)] + c' \lambda u_{t+1}$$

$$= c' u_{t+1} + c' \lambda u_{t+1}$$

$$= (c' + c' \lambda) u_{t+1}. \qquad \text{(A.22)}$$

I use (A.20) and (A.22) to estimate the time series of the discount-rate news and cash-flow news implemented with a VAR model that includes the following four variables: (1) differenced log price-earnings ratio of the market, (2) log real earnings growth rate of the market, (3) log inflation, and (4) log dividend-price ratio of the market.

4. Data description

As noted earlier, I use the differenced log price-earnings ratio of the market (ΔlPE), log real earnings growth rate of the market (lEG), log inflation ($l\pi$), and log dividend-price ratio of the market (lDP) in the VAR model approach to estimate the time

series of the discount-rate and cash-flow news. The time period for the data is 1928:12-2001:12, the same time period that Campbell and Vuolteenaho (2004) use to estimate the two news series. The data frequency for the sample is monthly.

Data for ΔlPE, $l\pi$, and lDP are constructed using data of S&P 500 price level, S&P 500 earnings, consumer price index (CPI), and dividend of S&P 500 firms from Shiller (2000).[18] The price-earnings ratio is constructed by dividing the S&P 500 price level by the earnings of S&P 500. I take the natural logarithm of the price-earnings ratio and then take its difference to obtain ΔlPE. Note that I do not use the 10-year smoothed price-earnings ratio provided by Shiller. Inflation, π_t, is calculated as $(CPI_t - CPI_{t-1})/CPI_{t-1}$. I take the log of $1+\pi_t$ to get $l\pi$. The dividend-price ratio is constructed by dividing the S&P 500 dividend by the S&P 500 price level. Since the data for dividends provided by Shiller (2000) are on an annual basis, I divide the dividends by 12 to adjust for monthly data. I take the natural log of one plus the dividend-price ratio to obtain lDP. After obtaining data for ΔlPE, $l\pi$, and lDP, I use equation (A.19) and the log return of a value weighted equity index[19] to construct data for the log real earnings growth rate of the market (lEG).

I also use the same variables that Campbell and Vuolteenaho (2004) use in their VAR model approach, but with different time periods. These variables are the excess log market return (R_Me),[20] the yield spread between long-term bonds and short-term bonds (TY), the market's 10-year smoothed price-earnings ratio (PE) from Shiller (2000), and

[18] I thank Professor Shiller for making the data available on his website: http://www.econ.yale.edu/~shiller/data.htm.

[19] This is the value-weighted equity index of NYSE, AMEX, and Nasdaq stocks including dividends from the Center for Research in Security Prices (CRSP). The data for this equity index are obtained from Wharton Research Data Services (WRDS), http://wrds.wharton.upenn.edu/.

[20] The excess log market return, R_Me, is the difference between the log return on the CRSP value-weighted stock index and the log risk-free rate.

the small-stock value spread (VS).[21] I thank Professor Campbell for providing the data for these variables. For a detailed description of these variables, see Campbell and Vuolteenaho (2004).

I use log returns on the 25 Fama-French portfolios sorted by size and book-to-market equity to calculate the cash-flow betas and the discount-rate betas of these portfolios. The data of the returns are made available by Professor French on his website[22] and are also provided by Professor Campbell.

5. Results

5.1 Using data from Campbell and Vuolteenaho (2004) but different time periods

Chen and Zhao (2009) show that Campbell and Vuolteenaho's (2004) results of cash-flow and discount-rate betas for the 25 Fama-French portfolios during the 1963:7-2001:12 period are sensitive to the time period used to estimate the cash-flow and discount-rate news. Campbell and Vuolteenaho (2004) use data of the period 1928:12-2001:12 to estimate the cash-flow news and discount-rate news series ranging from 1929:1-2001:12[23], and they use the two estimated news series to compute the cash-flow and discount-rate betas of the 25 Fama-French portfolios for both the 1929:1-1963:6 and 1963:7-2001:12 periods. Chen and Zhao (2009) use post-1952 data to estimate the two news series and calculate cash-flow and discount-rate betas for 1963:7-2001:12, but they do not find the same patterns in cash-flow betas as do Campbell and Vuolteenaho (2004).

[21] The small-stock value spread, VS, is the difference between the log (BE/ME) of the small value portfolio and the log (BE/ME) of the small growth portfolio.
[22] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[23] The cash-flow and discount-rate news are estimated using the noise term of the VAR model (see equations A.12 and A.14). Since I do not have the noise term for the first observation, i.e., 1928:12, the range of the two estimated news series starts from 1929:1.

Campbell and Vuolteenaho (2004) provide empirical support for their theory that investors demand a greater reward for bearing the risk determined by cash-flow betas than by discount-rate betas. In particular, these findings explain why value portfolios have higher average returns than growth portfolios, a phenomenon that cannot be justified by the CAPM market beta for data in the post 1963 period. They find that during 1929:1-1963:6, value portfolios have higher cash-flow and discount-rate betas than do growth portfolios, suggesting that the overall risk, as determined by the market beta, of value portfolios is greater than that of growth portfolios, and therefore, value portfolios have higher average returns. During 1963:7-2001:12, although value portfolios have lower discount-rate betas than do growth portfolios, they have higher-cash flow betas. The fact that value portfolios have higher cash-flow betas than do growth portfolios explains why value portfolios have higher average returns despite their having lower market betas than do growth portfolios. Thus, the two-beta ICAPM explains the value effect for both 1929:1-1963:6 and 1963:7-2001:12.

Using the same variables in a VAR model as Campbell and Vuolteenaho (2004), Chen and Zhao (2009) find that when post-1952 data are used to estimate the cash-flow news series and the discount-rate news series, value portfolios have lower cash-flow betas than do growth portfolios for the 1963:7-2001:12 period. They selected 1952 as the separating year because Campbell (1991) discovered a shift in variance from cash-flow news to discount-rate news since 1952.

I use the same variables that Campbell and Vuolteenaho (2004) use (R_Me, TY, PE, and VS), but various time periods to estimate the cash-flow and discount-rate news. Then, I use the two types of news to calculate cash-flow and discount-rate betas for the

25 Fama-French portfolios for the period 1963:7-2001:12. Consistent with Chen and Zhao (2009), I find that whether value portfolios have larger cash-flow betas than growth portfolios during 1963:7-2001:12 seems to depend on the time period used to estimate the cash-flow and discount-rate news.

The break down year appears to be 1936. When data prior to 1936 are included in the sample to estimate the cash-flow and discount-rate news, the cash-flow beta of value portfolios is mostly higher than that of growth portfolios for 1963:7-2001:12. In particular, I use data for the different time periods of 1928:12-2001:12 (Table C1), 1929:12-2001:12 (Table C2), 1932:12-2001:12 (Table C3), 1933:12-2001:12 (Table C4), and 1934:12-2001:12 (Table C5) to estimate the discount-rate news and the cash-flow news for each time period. Then, I use each pair of estimated news series to calculate the cash-flow and discount-rate betas of the 25 Fama-French portfolios for 1963:7-2001:12. Table C1 replicates Campbell and Vuolteenaho's (2004) results. Tables C2 to C5 show that among the five size categories, most of the value portfolios have cash-flow betas that are either higher than or close to those of growth portfolios.

Table C6 shows that when data from 1935:12-2001:12 are used to estimate the discount-rate and cash-flow news series, cash-flow betas of value portfolios are mostly lower than or approximately equal to those of growth portfolios. I also use data for the periods of 1939:12-2001:12 (Table C7), 1949:12-2001:12 (Table C8), and 1963:6-2001:12 (Table C9) to estimate the cash-flow and discount-rate news for each time period, and I find that the cash-flow betas of value portfolios are predominantly smaller than those of growth portfolios during 1963:7-2001:12.

Comparing the difference between the cash-flow betas of value and growth portfolios for 1963:7-2001:12, there seems to be a decreasing trend in the *number* of positive differences as the length of the time period used to estimate the cash-flow and discount-rate news series shortens, fixing the ending year of the sample period at 2001. Note that the difference is calculated as a value portfolio's cash-flow beta minus a growth portfolio's cash-flow beta. Since there are five value portfolios and five growth portfolios, there are at most five positive differences in cash-flow beta. For example, Table C1 reports five positive differences, and Table C9 reports zero positive differences. Figure C1 gives a good idea of the decreasing trend.

5.2 Implementation using a VAR model with alternative variables

Chen and Zhao (2009) point out that the cash-flow and discount-rate beta results for the period of 1963:7-2001:12 are sensitive to the choice of variables in a VAR model used to estimate the cash-flow and discount-rate news series. Their empirical findings suggest that it is necessary to use the four variables used by Campbell and Vuolteenaho (2004) to get the same patterns in cash-flow betas from growth to value stocks. Omitting any of the four variables as well as using additional variables does not produce results that support the theory.

In particular, Chen and Zhao (2009) present a VAR model with the variables R_Me, TY, PE, VS, *cay* , volfac, and rpfac to estimate the cash-flow and discount-rate news, which are used to calculate the cash-flow and discount-rate betas of the 25 Fama-French portfolios. The first four variables are used by Campbell and Vuolteenaho (2004). The variable *cay* is the log consumption-wealth ratio from Lettau and Ludvigson

(2001a). The variables volfac and rpfac are the volatility factor and the risk premium factor, both from Ludvigson and Ng (2007). Chen and Zhao (2009) find that the R^2 of the excess market return forecasting equation in this VAR model is about 14%, and thus higher than the R^2 in Campbell and Vuolteenaho (2004), 2.57%, suggesting that this seven-variable VAR model does a better job at predicting the market return. Despite the higher R^2, Chen and Zhao (2009) do not find evidence that the cash-flow and discount-rate betas for 1963:7-2001:12 explain the value effect, since the cash-flow betas of value portfolios are not significantly different from those of growth portfolios, and the discount-rate betas of value portfolios are lower than those of growth portfolios.

Two variables in the VAR model of Chen and Zhao (2009) are difficult to interpret. Specifically, the volatility factor and the risk premium factor are the principal components from a dataset of 172 financial indicators (see Ludvigson and Ng (2007)), and there is no clear economic interpretation attached to these two factors. Furthermore, some researchers argue that *cay* incorporates a look-ahead bias.[24]

In this preliminary analysis, I use variables motivated by economic reasons in a VAR model to implement the return decomposition. Motivated by the identity of return on a dividend-paying asset, I include the first difference in the market's log price-earnings ratio (ΔlPE), the market's log dividend yield (lDP), log inflation ($l\pi$), and the market's log real earnings growth rate (lEG) in a VAR model to estimate the cash-flow news series and the discount-rate news series. The reason for choosing these variables is that starting from the definition of asset return, by derivation, they are the components of an asset's log return.

[24] See, e.g., Guo (2003).

I use data from 1928:12-2001:12, the same time period as Campbell and Vuolteenaho (2004), to estimate a first-order VAR model. Upon obtaining the parameter estimates and the estimated residuals of the VAR model, I use the formulae for the discount-rate and cash-flow news in equations (A.20) and (A.22) to compute the estimated discount-rate news series and cash-flow news series. Next, I use the two estimated news series to calculate the cash-flow and discount-rate betas of the 25 Fama-French portfolios for both 1929:1-1963:6 and 1963:7-2001:12. The cash-flow and discount-rate betas are calculated according to equations (3.9) and (3.10).

The results of the cash-flow and discount-rate betas for the period of 1929:1-1963:6 are reported in Table C10. The value portfolios have higher cash-flow betas than do growth portfolios, except the value portfolio in the first (smallest) size quintile. Value portfolios also have higher discount-rate betas than do growth portfolios. Therefore, this explains the higher average returns of value portfolios than growth portfolios during 1929:1-1963:6, except perhaps the value and growth portfolios in the smallest size quintile. Furthermore, the higher average returns of the small portfolios than of the large portfolios are explained by the fact that the small portfolios have both higher cash-flow betas and higher discount-rate betas than do large portfolios.

Table C11 shows the cash-flow and discount-rate betas of the 25 Fama-French portfolios for 1963:7-2001:12. During 1963:7-2001:12, the value effect could not be explained by the cash-flow and discount-rate betas obtained using this VAR model. Both cash-flow and discount-rate betas of value portfolios are lower than those of growth portfolios. On the other hand, the cash-flow and discount-rate betas of the small

portfolios are higher than those of the large portfolios, so the higher average returns of small portfolios are explained.

5.3 Synthetic look-ahead biased variables

To illustrate the point that the results of the cash-flow and discount-rate betas for the period of 1963:7-2001:12 are sensitive to the choice of variables in the VAR methodology, Chen and Zhao (2009) use a VAR(1) model with the variables R_Me, TY, PE, VS, *cay* , volfac, and rpfac to estimate the cash-flow and discount-rate news. As mentioned in section 5.2 of this appendix, they find that although the R^2 of the excess market return forecasting equation in this VAR model is higher than that of Campbell and Vuolteenaho (2004), the cash-flow and discount-rate betas computed from the two estimated news series do not exhibit the patterns that explain the value effect for 1963:7-2001:12.

One problem with the seven-variable VAR model lies in *cay* , which is argued to have look-ahead bias (e.g., Guo (2003)). Lettau and Ludvigson (2001a) estimate the log consumption-wealth ratio, *cay* , as the equilibrium error of the cointegration relation among consumption, net worth, and labor income. Guo (2003) notes that data on consumption and labor income are subject to substantial revisions to reflect the incorporation of new information and new methodologies and to reduce noise. Therefore, *cay* may have look-ahead bias and may contain information of future cash-flow news. As a result, using look-ahead biased variables in the VAR methodology to estimate discount-rate news may cause distortion in the estimated discount-rate news in the sense that it contains some cash-flow news. Consequently, there may also be distortion in the

estimated cash-flow news because it is backed out as a residual between the unexpected market return and the estimated discount-rate news.

As a result, the estimated discount-rate betas and cash-flow betas may be biased, and their patterns from growth to value stocks may be weaker than those in Campbell and Vuolteenaho (2004). Recall that Campbell and Vuolteenaho (2004) report results that show the value portfolios have higher cash-flow betas but lower discount-rate betas than do the growth portfolios during 1963:7-2001:12. Since the usage of look-ahead biased variables in the estimation of the two news series may cause the estimated discount-rate news to be contaminated with cash-flow news, the difference in discount-rate betas between value portfolios and growth portfolios ($\hat{\beta}_{Value,DR} - \hat{\beta}_{Growth,DR}$) may become less negative, compared with Campbell and Vuolteenaho (2004). Moreover, since the estimated cash-flow news becomes distorted as a result of distortion in the estimated discount-rate news, the difference in cash-flow betas between value and growth portfolios ($\hat{\beta}_{Value,CF} - \hat{\beta}_{Growth,CF}$) may be less positive or even negative, compared with Campbell and Vuolteenaho (2004).

To illustrate the effect of look-ahead biased variables on the estimated cash-flow and discount-rate betas, I use synthetic look-ahead biased variables in the VAR approach of Campbell and Vuolteenaho (2004) to estimate the cash-flow and discount-rate news. Specifically, I create a synthetic look-ahead biased variable and use it together with the four variables (R_Me, TY, PE, VS) that Campbell and Vuolteenaho (2004) use in their VAR(1) model to obtain estimates of the cash-flow and discount-rate news.

The look-ahead biased variable is created using the future values of the log dividend-price ratio plus some noise:

$$LABias_t = \log(1+DP_{t+k}) + 0.01 * v_t, \qquad \text{(A.23)}$$

where DP_{t+k}, the dividend-price ratio at $t+k$, is constructed by dividing the S&P 500 dividend at $t+k$ by the S&P 500 price level at $t+k$. The data of the S&P 500 dividends and price levels are both from Shiller (2000) and are available from his webpage[25]. Here, $k = 72$, that is, the synthetic look-ahead biased variable is created using the 72-months lead of the log dividend-price ratio. The noise v_t is a random number drawn from a normal distribution with mean and variance equal to the sample mean and sample variance of $\log(1+DP_{t+k})$ for the sample period[26] (i.e., $v_t \sim N(0.0032, 0.0013)$).

Table C12 shows the cash-flow and discount-rate betas of the 25 Fama-French portfolios for 1963:7-2001:12 calculated using the two estimated news series from the VAR approach with a synthetic look-ahead biased variable and the four variables used by Campbell and Vuolteenaho (2004) (R_Me, TY, PE, and VS). The news series are estimated using data from the period 1928:12-2001:12. For comparison, I use the cash-flow and discount-rate betas calculated from the estimated cash-flow and discount-rate news with the variables R_Me, TY, PE, and VS, which I refer to as the benchmark case shown in Table C1.

A comparison of Table C1 and Table C12 shows that for all five size categories, the difference in discount-rate betas between value and growth stocks ($\hat{\beta}_{Value,DR} - \hat{\beta}_{Growth,DR}$) becomes less negative when a synthetic look-ahead biased variable is included in a VAR model to estimate cash-flow and discount-rate news. In addition, for all five size categories, the difference in cash-flow betas between value and growth

[25] http://www.econ.yale.edu/~shiller/data.htm.
[26] I use 1928:12-2001:12 as the sample period, since this is what Campbell and Vuolteenaho (2004) use.

stocks ($\hat{\beta}_{Value,CF} - \hat{\beta}_{Growth,CF}$) becomes less positive when a synthetic look-ahead biased variable is included in a VAR model to extract cash-flow and discount-rate news. These results suggest the possibility that using look-ahead biased variables may weaken the patterns in cash-flow and discount-rate betas found in Campbell and Vuolteenaho (2004).

(Appendix C continued)

Figure and tables from the preliminary analysis using the VAR model approach

Figure C1: Trend in the cash-flow beta differences between value portfolios and growth portfolios

This figure displays the number of positive differences in cash flow beta between value portfolios and growth portfolios when the cash flow news and discount rate news are estimated from data of different time periods. The horizontal axis represents the time period by which the cash flow news series and the discount rate news series are estimated. The Fama and French 25 portfolios are sorted by five size and five B/M quintiles. Within each size quintile, the portfolio with the largest B/M is the value portfolio, and the one with the smallest B/M is the growth portfolio. Using each pair of estimated news series, I compute the cash flow beta of Fama and French 25 portfolios for the period 1963:7-2001:12. These results are in Table C1 to Table C9. I also calculate the difference in cash flow beta between value portfolios and growth portfolios (value portfolio's β_{CF} minus growth portfolio's β_{CF}). The vertical axis represents the number of differences that are positive.

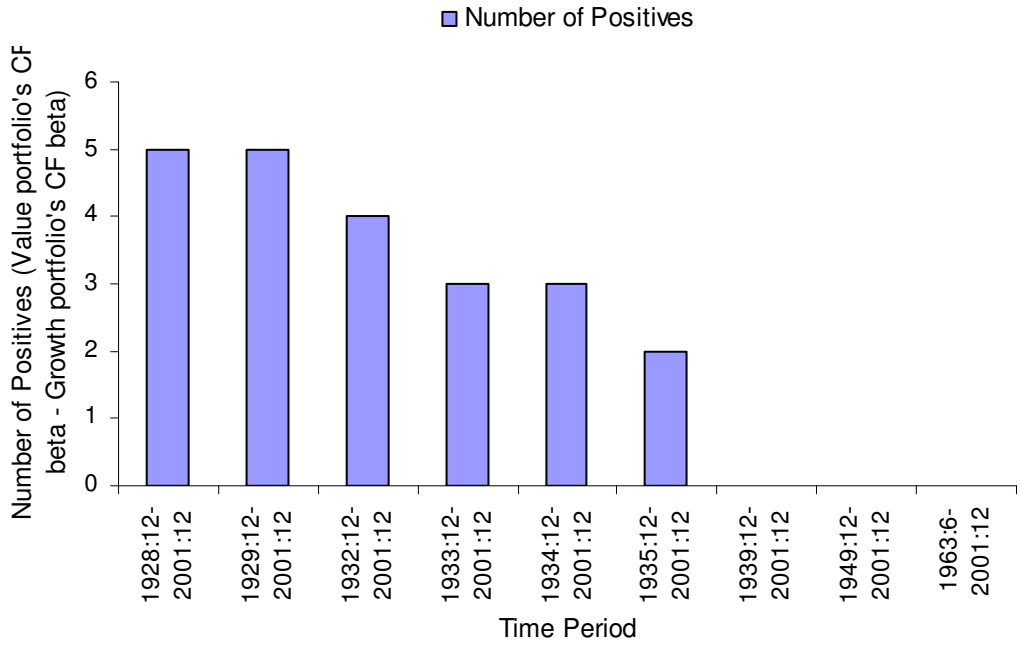


Table C1: Replication of Campbell and Vuolteenaho's (2004) Table 5

This table shows a replication of Campbell and Vuolteenaho's (2004) Table 5. I use the same variables (R_Me, TY, PE, VS) and the same time period (1928:12-2001:12) as Campbell and Vuolteenaho (2004) to estimate $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$. The two estimated news series are used to calculate the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period of 1963:7-2001:12.[27] The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.06	0.07	0.09	0.09	0.13	0.07
2	0.04	0.08	0.10	0.11	0.12	0.09
3	0.03	0.09	0.11	0.12	0.13	0.09
4	0.03	0.10	0.11	0.11	0.13	0.10
Large	0.03	0.08	0.08	0.11	0.11	0.09
Diff.	-0.03	0.02	-0.01	0.02	-0.01	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.66	1.37	1.18	1.12	1.12	-0.54
2	1.54	1.22	1.06	0.95	1.02	-0.52
3	1.41	1.10	0.95	0.82	0.94	-0.47
4	1.27	1.04	0.89	0.78	0.86	-0.41
Large	1.00	0.87	0.74	0.63	0.68	-0.33
Diff.	-0.66	-0.50	-0.44	-0.49	-0.44	

[27] I follow Campbell and Vuolteenaho (2004) and use the full sample, 1928:12-2001:12, to estimate the cash flow and discount rate news. The news series are used to compute the cash flow and discount rate betas of the 25 Fama-French portfolios during 1963:7-2001:12, which are shown in Table C1.

Table C2: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1929:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1929:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.10	0.11	0.12	0.12	0.16	0.05
2	0.08	0.11	0.13	0.13	0.15	0.07
3	0.08	0.12	0.13	0.14	0.15	0.07
4	0.06	0.13	0.14	0.13	0.15	0.08
Large	0.06	0.11	0.10	0.13	0.13	0.08
Diff.	-0.05	0.00	-0.02	0.00	-0.02	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.62	1.33	1.15	1.09	1.10	-0.52
2	1.50	1.19	1.04	0.93	1.00	-0.50
3	1.38	1.08	0.93	0.80	0.92	-0.46
4	1.24	1.02	0.86	0.77	0.85	-0.39
Large	0.98	0.85	0.72	0.62	0.66	-0.31
Diff.	-0.64	-0.49	-0.43	-0.47	-0.43	

Table C3: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1932:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1932:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.18	0.16	0.15	0.15	0.17	-0.01
2	0.15	0.15	0.15	0.14	0.15	0.00
3	0.13	0.15	0.14	0.13	0.15	0.01
4	0.12	0.14	0.14	0.13	0.14	0.03
Large	0.09	0.12	0.11	0.12	0.12	0.03
Diff.	-0.10	-0.04	-0.05	-0.04	-0.05	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.57	1.31	1.14	1.09	1.11	-0.47
2	1.46	1.18	1.04	0.94	1.01	-0.45
3	1.34	1.07	0.93	0.82	0.93	-0.42
4	1.20	1.01	0.87	0.77	0.86	-0.34
Large	0.95	0.84	0.71	0.63	0.68	-0.27
Diff.	-0.62	-0.47	-0.43	-0.46	-0.43	

Table C4: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1933:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1933:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.24	0.21	0.20	0.19	0.21	-0.03
2	0.21	0.19	0.19	0.18	0.19	-0.01
3	0.18	0.19	0.18	0.17	0.18	0.00
4	0.16	0.18	0.17	0.16	0.18	0.01
Large	0.12	0.15	0.14	0.14	0.15	0.02
Diff.	-0.12	-0.06	-0.06	-0.05	-0.07	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.53	1.27	1.10	1.06	1.07	-0.45
2	1.41	1.14	1.01	0.91	0.98	-0.43
3	1.30	1.04	0.90	0.79	0.90	-0.40
4	1.16	0.98	0.84	0.74	0.83	-0.33
Large	0.92	0.81	0.69	0.61	0.66	-0.26
Diff.	-0.61	-0.46	-0.41	-0.45	-0.41	

Table C5: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1934:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1934:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.24	0.20	0.19	0.18	0.21	-0.03
2	0.20	0.19	0.18	0.17	0.18	-0.02
3	0.18	0.18	0.17	0.16	0.18	0.00
4	0.16	0.17	0.17	0.15	0.17	0.01
Large	0.12	0.14	0.13	0.13	0.14	0.02
Diff.	-0.12	-0.06	-0.06	-0.05	-0.07	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.53	1.28	1.11	1.06	1.08	-0.45
2	1.42	1.15	1.01	0.92	0.99	-0.43
3	1.30	1.04	0.91	0.80	0.91	-0.40
4	1.16	0.98	0.85	0.75	0.84	-0.32
Large	0.93	0.82	0.70	0.62	0.67	-0.26
Diff.	-0.61	-0.46	-0.41	-0.45	-0.41	

Table C6: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1935:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1935:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.23	0.19	0.18	0.17	0.19	-0.04
2	0.19	0.17	0.17	0.15	0.17	-0.03
3	0.17	0.17	0.16	0.14	0.16	-0.01
4	0.15	0.16	0.15	0.14	0.15	0.00
Large	0.11	0.13	0.12	0.12	0.12	0.02
Diff.	-0.12	-0.06	-0.06	-0.05	-0.07	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.55	1.29	1.13	1.08	1.10	-0.44
2	1.43	1.16	1.03	0.93	1.01	-0.42
3	1.31	1.06	0.92	0.82	0.92	-0.39
4	1.17	1.00	0.86	0.77	0.86	-0.31
Large	0.94	0.83	0.71	0.63	0.68	-0.25
Diff.	-0.61	-0.46	-0.42	-0.45	-0.42	

Table C7: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1939:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1939:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.35	0.27	0.23	0.21	0.22	-0.13
2	0.30	0.23	0.20	0.17	0.18	-0.12
3	0.26	0.21	0.18	0.14	0.17	-0.09
4	0.24	0.20	0.17	0.14	0.15	-0.08
Large	0.16	0.15	0.13	0.11	0.11	-0.04
Diff.	-0.19	-0.11	-0.10	-0.10	-0.11	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.42	1.21	1.07	1.03	1.07	-0.35
2	1.32	1.10	1.00	0.92	1.00	-0.32
3	1.23	1.02	0.90	0.82	0.92	-0.31
4	1.09	0.97	0.85	0.77	0.86	-0.23
Large	0.89	0.82	0.70	0.64	0.70	-0.19
Diff.	-0.53	-0.39	-0.37	-0.39	-0.37	

Table C8: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1949:12-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1949:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.47	0.39	0.36	0.34	0.37	-0.10
2	0.42	0.36	0.33	0.31	0.34	-0.09
3	0.38	0.33	0.31	0.29	0.33	-0.05
4	0.35	0.32	0.30	0.27	0.31	-0.04
Large	0.27	0.27	0.25	0.24	0.26	-0.01
Diff.	-0.21	-0.12	-0.11	-0.10	-0.11	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	1.21	1.08	0.97	0.91	0.93	-0.28
2	1.21	1.03	0.93	0.89	0.95	-0.26
3	1.17	0.97	0.89	0.84	0.89	-0.28
4	1.09	0.95	0.88	0.84	0.88	-0.21
Large	0.94	0.86	0.79	0.75	0.76	-0.18
Diff.	-0.27	-0.22	-0.18	-0.16	-0.17	

Table C9: Cash-flow betas and discount-rate betas calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated from 1963:6-2001:12

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow betas and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the same variables as Campbell and Vuolteenaho (2004), but a different time period of 1963:6-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.86	0.73	0.66	0.63	0.67	-0.19
2	0.79	0.68	0.62	0.57	0.62	-0.17
3	0.73	0.63	0.57	0.52	0.58	-0.15
4	0.64	0.60	0.54	0.49	0.54	-0.10
Large	0.50	0.50	0.43	0.41	0.44	-0.06
Diff.	-0.36	-0.23	-0.23	-0.22	-0.23	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	0.92	0.76	0.65	0.62	0.63	-0.29
2	0.84	0.67	0.59	0.53	0.57	-0.27
3	0.77	0.61	0.53	0.46	0.53	-0.25
4	0.69	0.57	0.49	0.44	0.49	-0.21
Large	0.55	0.48	0.40	0.35	0.39	-0.17
Diff.	-0.36	-0.28	-0.25	-0.27	-0.24	

Table C10: Cash-flow and discount-rate betas for 1929:1-1963:6 calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated using the VAR approach with the variables ΔlPE, lDP, $l\pi$, and lEG

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1929:1-1963:6. The cash-flow and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the VAR approach with the variables ΔlPE, lDP, $l\pi$, and lEG with a time period from 1928:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	1.20	1.05	0.99	1.01	1.07	-0.13
2	0.81	0.89	0.86	0.91	1.02	0.21
3	0.84	0.78	0.81	0.85	1.04	0.20
4	0.64	0.71	0.75	0.83	1.09	0.45
Large	0.63	0.59	0.68	0.81	0.90	0.28
Diff.	-0.57	-0.46	-0.31	-0.20	-0.17	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	0.65	0.86	0.73	0.68	0.67	0.02
2	0.53	0.58	0.58	0.61	0.67	0.14
3	0.59	0.51	0.56	0.55	0.69	0.10
4	0.43	0.52	0.52	0.57	0.75	0.32
Large	0.45	0.43	0.48	0.58	0.65	0.20
Diff.	-0.20	-0.43	-0.25	-0.10	-0.02	

Table C11: Cash-flow and discount-rate betas for 1963:7-2001:12 calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated using the VAR approach with the variables ΔlPE, lDP, $l\pi$, and lEG

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the VAR approach with the variables ΔlPE, lDP, $l\pi$, and lEG with a time period from 1928:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	1.34	1.11	0.96	0.92	0.96	-0.38
2	1.21	0.99	0.88	0.81	0.87	-0.34
3	1.11	0.91	0.79	0.70	0.79	-0.32
4	0.97	0.85	0.74	0.65	0.73	-0.24
Large	0.75	0.70	0.60	0.54	0.59	-0.16
Diff.	-0.59	-0.40	-0.36	-0.38	-0.37	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small	0.44	0.38	0.34	0.32	0.33	-0.11
2	0.43	0.36	0.33	0.30	0.32	-0.11
3	0.41	0.34	0.30	0.27	0.31	-0.10
4	0.37	0.33	0.29	0.26	0.30	-0.08
Large	0.31	0.28	0.24	0.22	0.23	-0.07
Diff.	-0.13	-0.09	-0.10	-0.11	-0.10	

Table C12: Cash-flow and discount-rate betas for 1963:7-2001:12 calculated using $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$ that are estimated using the VAR approach with the variables R_Me, TY, PE, VS, and a synthetic look-ahead biased variable

This table shows the cash-flow betas and discount-rate betas of the 25 Fama-French portfolios for the period 1963:7-2001:12. The cash-flow and discount-rate betas are calculated from $\{ N_{CF,t} \}$ and $\{ N_{DR,t} \}$, which are estimated using the VAR approach with the variables R_Me, TY, PE, VS, and *LABias* (created according to equation A.23) with a time period from 1928:12-2001:12. The column titled "Diff." is the difference in cash-flow (discount-rate) beta between a value portfolio and a growth portfolio. The row titled "Diff." is the difference in cash-flow (discount-rate) beta between a large portfolio and a small portfolio.

$\hat{\beta}_{CF}$	Growth	2	3	4	Value	Diff.
Small	0.14	0.14	0.15	0.15	0.18	0.04
2	0.11	0.14	0.15	0.15	0.17	0.06
3	0.10	0.15	0.15	0.16	0.17	0.07
4	0.09	0.15	0.15	0.15	0.17	0.08
Large	0.07	0.12	0.12	0.14	0.15	0.07
Diff.	-0.06	-0.01	-0.03	-0.01	-0.04	

$\hat{\beta}_{DR}$	Growth	2	3	4	Value	Diff.
Small.	1.58	1.29	1.11	1.06	1.06	-0.52
2	1.46	1.15	1.01	0.90	0.97	-0.49
3	1.34	1.05	0.90	0.78	0.89	-0.45
4	1.20	0.99	0.85	0.74	0.82	-0.38
Large	0.95	0.82	0.70	0.60	0.64	-0.31
Diff.	-0.63	-0.47	-0.41	-0.46	-0.42	