Regime-Switching Model and Stock Return Predictability

Dr. Zhijun Yang (Attender/Presenter)

Associate Professor of Finance

Metropolitan State University

1501 Hennepin Avenue

Minneapolis, MN 55403

USA

Email: Zhijun.yang@metrostate.edu

EFM Classification Codes:

310 Asset Pricing Models and Tests

760 Methodological Issues

Abstract

This paper investigates the statistical and economic importance of modeling regime switches in the context of stock return predictability with model uncertainty. Several interesting results are presented. First, when placing the linear and regime-switching models in juxtaposition, posterior odds ratios and model probabilities unanimously favor regime-switching models. Second, the best predictors picked by linear models differ from those chosen by nonlinear models, which indicates that focusing on linear models alone could lead to potentially misleading inferences. Third, the support for regime switches is economically significant as measured by the utility loss perceived by investors who are forced to use linear forecasting models. These results are robust to the use of different priors.

I. Introduction

Are stock returns predictable? If so, which models and predictive variables are the best performers? Answers to these intriguing questions are of great importance to financial economists as well as investors. During the past two decades, an increasing number of academic studies have identified some market, macroeconomic, and calendar variables that appear to do a good job in terms of predicting stock returns. See, for example, Fama and French (1989), Pontiff and Schall (1998), Cooper, McConnell, and Ovtchinnikov (2006), among many others.

Arguments against the findings of return predictability are often based on some statistical considerations. For instance, Lo and MacKinlay (1990) point out that because of the non-experimental nature of stock market data, researchers should guard against data- snooping biases when searching for the best model. To address concerns of data-snooping biases, Bossaerts and Hillion (1999) (BH) implement several statistical model selection criteria that are designed to choose a model with the best out-of-sample validity. Although their results confirm in-sample return predictability, they find no evidence of out-of-sample forecasting power. A potential explanation offered by BH is that stock return predictability is subject to model nonstationarity, i.e. model parameters tend to change over time.

Due to the lack of theoretical guidance, a researcher often arbitrarily chooses a set of predictors that either are convenient, intuitive, or seem to work from past experience. Hence studies in search of the best prediction model using the classical approach can be misleading because they fail to recognize the presence of model uncertainty that is intrinsic to such an exercise. Perhaps the best way to account for model uncertainty in the context of predicting stock returns is the Bayesian model selection/averaging approach advocated in two articles. Cremers (2002) simultaneously compares the prediction performance of all linear models that are spanned by a set of 14 commonly used predictors and shows that the posterior probabilities are supportive of predictability. Avramov (2002) demonstrates that in terms of forecasting power the Bayesian model averaging approach is superior to traditional model selection criteria as implemented in BH.

The Bayesian approach adopted in Cremers and Avramov is undoubtedly a significant step forward. Unlike the frequentist approach, which draws inferences regarding the significance of a predictive variable in the context of one individual model, the Bayesian paradigm weighs the empirical evidence across all the models. This is achieved by calculating posterior model probabilities via Bayes' rule. Therefore the Bayesian methodology solves the model uncertainty problem and is robust to model misspecification, at least within the general class of models under consideration.

It is important to realize that, similar to many other papers in the literature, both Avramov and Cremers study exclusively linear normal prediction models. A priori, however, there is no compelling reason why we should confine our focus within the class of linear models. In fact, the choice is likely to be driven by analytical tractability more than anything else. In the Bayesian framework, exclusive focus on linear models is equivalent to placing a dogmatic prior belief that assigns zero probability to nonlinear models. Thus if the true data-generating process (DGP) is nonlinear, the choice of linear models could result in erroneous inferences.

Tu and Zhou (2004) relax the normality assumption and investigate the effects of DGP uncertainty on investors' portfolio decisions. They introduce the student t distribution with varying degrees of freedom to replace the normal distribution assumption. They find that accounting for fat tails leads to nontrivial changes in parameter estimates and optimal portfolio weight. But the economic losses associated with the data-generating uncertainty are found to be relatively small. It should be pointed out, however, that Tu and Zhou's approach still maintains the i.i.d. assumption and therefore, by design, they rule out predictability in stock returns. Empirical evidence presented in BH and Pesaran and Timmermann (1995) (PT) appear to support the notion that a nonlinear model that can account for model nonstationarity is needed. PT finds that the predictive power of various predictors changes through time. For example, they show that stock return predictability seemed quite low during the relatively calm market in the 1960s, but increased substantially in the volatile markets of the 1970s. Pontiff and Schall (1998) also document a structural difference in the predictive power of the book-to-market ratio in pre- and post-1960 samples of Dow Jones Industrial Average index.

Tu (2010) studies the importance of modeling regime switching when an investor faces mispricing uncertainty and parameter uncertainty. Tu shows that ignoring regime switching can result in significant certainty-equivalent losses that can exceed 10% per year during market downturns. Tu's article focuses on investors' beliefs on different asset pricing models and the associated asset allocation problem. Thus his paper can be viewed as a natural extension of Pastor and Stambaugh (2000)'s framework. In contrast, our paper emphasizes the role of regime switching in the context of predicting stock market returns.

Lettau and Nieuwerburgh (2008) present interesting empirical evidence that adjusting financial ratios for shifts in the steady state mean of the economy can explain the seemingly incompatible in-sample and out-of-sample results regarding stock return predictability. While Lettau and Nieuwerburgh focus on the shifts in the mean, our approach is to model the regime shifts in the regression coefficients. In addition, our approach can account for model and parameter uncertainty. Ang and Timmermann (2011) show that regime switching model matches many properties of asset returns, in particular skewness and fat tails, downside risk properties, and time-varying correlations. Dangl and Halling (2012) find that predictive models with constant coefficients are dominated by models with time-varying coefficients.

We contribute to the literature by extending the Bayesian analysis of stock return predictability to nonlinear models as well as linear ones. More specifically, we consider the case where the underlying DGP follows the regime-switching models of Hamilton (1989). Analytical solutions are unavailable once we go beyond linear normal models with conjugate priors. Hence we follow the Gibbs sampling approach of Albert and Chib (1993) for the Bayesian analysis of regime-switching models.

We find several significant results. First, when placing the linear and regime-switching models in juxtaposition, the posterior odds ratios unanimously favor regime-switching models over their linear counterparts. Thus our results highlight the importance of considering non-linear models. Second, the best predictors picked by linear models can differ from those chosen by nonlinear models. This confirms our intuition that confining our focus exclusively to linear models could result in potentially misleading inferences. Third, conditional on the set of predictors included in this paper, we find substantial evidence in favor of a regime-switching model that uses lagged stock returns and T-bill rates as its predictive variables. Last but not the least, the support for regime switches is economically meaningful as measured by the utility loss perceived by investors who are forced to rely only on linear forecasting models even after taking model uncertainty into account.

This paper is organized as follows. Section II provides further justifications for regime-switching models and describes the empirical approach. Section III presents the main results.

Section IV evaluates the economic significance of incorporating regime switches when predicting stock returns. Section V offers some concluding remarks.

II. Methodology

A. The Regime-Switching Prediction Model

We start with the following linear predictive regression model:

$$r_t = x_{t-1} * \beta + \varepsilon_t, \;\; \varepsilon_t \sim i.i.d. N\,(0, \sigma^2) \tag{1}$$

where r_t denotes the return on a stock market index at time t in excess of the risk-free rate. $x_{t-1} \equiv [1, x_{t-1}]$ is a $1 \times K$ vector consisting of a constant and the values of $K - 1$ lagged predictive variables (x_{t-1}). Only lagged predictors are used to ensure they are present in an investor's information set at the time of prediction. β is the $K \times 1$ vector of slope coefficients. This is the standard linear predictive model considered by the majority of articles in the extant literature on stock return predictability. The differences among these papers are mainly in terms of which predictors are included or excluded in equation (1). The assumption that the return innovation ε_t is i.i.d. normal is also a standard setup although it is well-known that stock returns do exhibit heteroscedasticity. In the Bayesian setting, the normality assumption is particularly convenient as it allows for analytical derivations of the posterior and predictive densities as long as the priors are chosen appropriately (Zellner, 1971).

Using a data set of international stock market returns, BH notice that the in-sample and out-of-sample forecasting performance of linear prediction models are very inconsistent. For example, they discover that even the best prediction models have no out-of-sample predictive power. They conclude that "the poor external validity of the prediction models that the formal model selection criteria chose indicates model nonstationarity: the parameters of the 'best' prediction models change over time."

PT also presents convincing empirical evidence in support of model nonstationarity. Utilizing a recursive model selection and estimation strategy, they find it important to allow for "changes in the underlying process of excess returns in the U.S. stock market". In particular, PT show that the behavior of two predictors, the inflation rate and interest rate, is closely related to economic "regime switches". PT concludes that "in analyzing stock return predictability it is advisable to use forecasting procedures that allow for possible regime changes."

Given the interesting empirical evidence shown in PT and BH, we propose to incorporate regime switches in the slope coefficients of equation (1) to allow for the effects of structural breaks in the linear prediction model. To be precise, we consider the following two-state regime-switching model:

$$r_t = x_{t-1} * \beta_{S_t} + \varepsilon_t, \ \varepsilon_t \sim N\,(0, \sigma^2), t = 1,2, \dots, T \qquad (2)$$

$$\beta_{S_t} = \beta_0 * (1 - S_t) + \beta_1 * S_t \qquad (3)$$

where S_t = 0, 1 denotes the unobserved regime indicator. In other words, at any time t, the slope coefficient can switch between two values β_0 and β_1, contingent upon the realizations of the regime indicator S_t. A nontrivial question is how to model the time-series behavior of S_t.

Following the path-breaking work of Hamilton (1989), we assume that S_t follows a two-state, first-order Markov process with the following transition probability matrix

$$P = \begin{pmatrix} q & 1-p \\ 1-q & p \end{pmatrix} \quad (4)$$

where q = prob(S_t = 0|S_{t-1} = 0) and p = prob(S_t= 1|S_{t-1} = 1). In this model, the transition probabilities are constant and the regime at time t depends only on the regime at t − 1, which makes it a tractable and flexible model.

Due to their flexibility, regime-switching models have been successfully used to model random structural breaks in financial data, e.g. Gray's (1996) study on regime changes in interest rates. In the case of stock market returns, a two-state regime-switching model appears to be consistent with the common practice by investors and the financial press alike to categorize the market into a bull and a bear market. From a technical perspective, a regime-switching model is particularly useful to characterize stock returns because it can capture negative skewness and leptokurtosis in stock returns by modeling the stock return distribution as a mixture of normal distributions. For example, Veronesi (1999) presents a model of stock price where dividends follow a continuous-time Markov switching model. Compared with conventional models with no regime shifts, Veronesi's model can capture many salient features of the stock returns, such as volatility clustering, leverage effects, and time-varying expected returns. Other examples include Hamilton and Susmel (1994) who study regime shifts in stock return volatilities. Turner, Startz, and Nelson (1989) use a regime-switching model to characterize stock returns and volatilities. Maheu and McCurdy (2000) identify a high-return stable bull market regime and a low-return volatile bear market regime with a duration-dependent regime-switching model. Ang and Bekaert (2002) study the implication of regime shifts in stock returns for the asset allocation problem in international equity markets. Given the empirical evidence documented in BH and PT, we expect the regime-switching model consisting of equations (2) and (3) to capture parameter changes that the linear predictive regression equation (1) is unable to model.

B. Gibbs Sampler for Regime-Switching Models

To account for model uncertainty, we follow the Bayesian approach advocated by Avramov (2002) and Cremers (2002) to analyze stock return predictability. However, unlike the case of linear prediction models studied by Avramov and Cremers where the posterior densities are known in closed-forms, Bayesian analysis of regime-switching models does not have analytical solutions. Hence we rely on the Markov Chain Monte Carlo (MCMC) approach for posterior analysis. To be precise, we use the Gibbs sampling technique introduced by Albert and Chib (1993) to the class of regime-switching model.

A key feature of the Bayesian approach is that unknown model parameters are treated as random variables. Bayes' rule dictates that

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{5}$$

where $p(\theta)$ denotes the prior density for model parameters, $p(y|\theta)$ the likelihood function, and $p(\theta|y)$ the posterior density. In the context of the regime-switching model, our goal is to derive the following joint posterior density:

$$p(\tilde{S}_{T,}\beta_0, \beta_1, \sigma^2, q, p|\tilde{y}_T) = h(\beta_0, \beta_1, \sigma^2|\tilde{S}_{T,}\tilde{y}_T)h(q, p|\tilde{S}_{T,})h(\tilde{S}_{T,}|\tilde{y}_T) \tag{6}$$

where $\tilde{S}_T = [S_{1,}S_{2,} \dots , S_{T,}]^t$ denotes the vector of latent regimes and $\tilde{y}_T$ denotes the data. Equation (6) uses the fact that conditional on probabilities, p and q, are independent of other model parameters. In addition, conditional on $\tilde{S}_T$ equation (2) and (3) are simply regression models with a known dummy variable S_t. These conditioning features of the model allow us to use the Gibbs sampling technique for Bayesian inference. When the joint posterior density is of an unknown form, the Gibbs sampler allows us to split up the parameter space into blocks of parameters for which it is possible to specify the full conditionals. The idea is to recursively draw from the conditional densities for the various blocks of parameters until convergence is achieved.

The Gibbs sampler for regime-switching models consists of the following steps:

1. Generate a whole block of $\tilde{S}_T$ from its conditional distribution

   $h(\tilde{S}_{T,}|\beta_0, \beta_1, \sigma^2, q, p, \tilde{y}_T)$

2. Generate the transition probabilities, p and q from $h(q, p|\tilde{S}_T)$

3. Generate $\beta_0, \beta_1, \sigma^2$ from $h(\beta_0, \beta_1, \sigma^2|\tilde{S}_{T,}\tilde{y}_T)$

Note that, in step 1, Albert and Chib (1993) use a single-move strategy to draw $\tilde{S}_T$, we

Instead use a computationally more efficient multi-move algorithm proposed by Chib (1996) and Kim and Nelson (1998). Readers are referred to above-mentioned articles for details about the multi-move algorithm. The following subsection provides further details on the second and third steps. We cycle through the Gibbs sampling procedure for 20,000 iterations. To alleviate the transient effects induced by initial values, we discard the first 10,000 iterations, which are treated as the “burn-in” period.

C. Prior Specifications

To implement the second step and draw q and p, we use the beta distributions as conjugate priors for the transition probabilities.

$$q \sim B(q|u_{00}, u_{01}) \tag{7}$$

$$p \sim B(p|u_{11}, u_{10}) \tag{8}$$

where u_{ij} , i, j = 0, 1 are known hyperparameters of the priors. B(z|α, δ) denotes a beta distribution with the expected value $\frac{\alpha}{\alpha+\delta}$ and the variance $\frac{\alpha\delta}{(\alpha+\delta)^2(\alpha+\delta+1)}$. To ensure that our base prior for q and p is uninformative, we set α = δ = 1. In this case, the beta distribution reduces to the uniform distribution. Assuming independence, the joint prior for q and p is given by

$$B(p,q) \propto p^{u_{11}-1}(1-p)^{u_{10}-1}q^{u_{00}-1}(1-q)^{u_{01}-1} \tag{9}$$

The likelihood function in this case is given by

$$L = p^{n_{11}}(1-p)^{n_{10}}q^{n_{00}}(1-q)^{n_{01}} \tag{10}$$

where n_{ij} refers to the transitions from regime i to j, which can be easily counted given the regime realizations of $\tilde{S}_T$ generated from step 1. With Bayes' rule in equation (5), we can combine the prior distribution and the likelihood function to obtain the following joint posterior distribution p and q:

$$h(q,p|\tilde{S}_T) \propto p^{u_{11}+n_{11}-1}(1-p)^{u_{10}+n_{10}-1}q^{u_{00}+n_{00}-1}(1-q)^{u_{01}+n_{01}-1} \tag{11}$$

This suggests that the posterior distribution is given by two independent beta distributions:

$$q \sim B(q|u_{00}+n_{00}, u_{01}+n_{01}) \tag{12}$$

$$p \sim B(p|u_{11}+n_{11}, u_{10}+n_{10}) \tag{13}$$

from which p and q are drawn.

To implement step 3, we follow standard practice in Bayesian analysis and specify conjugate priors for β and σ^2. The choice of conjugate priors is indispensable for the Gibbs sampler as it allows us to obtain the posterior in closed-forms. Specifically, we assume a normal prior for $\tilde{\beta} \equiv [\beta_0, \beta_1]$,

$$\tilde{\beta} \sim N(b_0, V_0) \tag{14}$$

where b_0 and V_0 are known hyperparameters. By the conjugacy property, the posterior distribution is given by

$$\tilde{\beta} \big| \sigma^2, \tilde{S}_{T,} \tilde{y}_T \sim N\ (b_1, V_1) \tag{15}$$

where

$$V_1 = (V_0^{-1} + \sigma^{-2} X'X)^{-1} \tag{16}$$

$$b_1 = V_1 (V_0^{-1} b_0 + \sigma^{-2} X'Y) \tag{17}$$

X and Y are the matrix notation for the regressors and dependent variable respectively.

We specify an inverted Gamma distribution as a conjugate prior for σ^2,

$$\sigma^2 \sim IG\ (\frac{c_0}{2}, \frac{d_0}{2}) \tag{18}$$

where c_0 and d_0 are the known hyperparameters. The posterior distribution is given by

$$\sigma^2 | \tilde{\beta}, \tilde{S}_{T,} \tilde{y}_T \sim IG\ (\frac{c_1}{2}, \frac{d_1}{2}) \tag{19}$$

where $c_1 = c_0 + T$ and $d_1 = d_0 + (Y - X\tilde{\beta})'(Y - X\tilde{\beta})$.

Our objective is to minimize the impact of the priors on the posterior estimates and, especially, on the model comparisons. Accordingly, we employ relatively uninformative (i.e. imprecise) priors. As a result, the posterior distributions of the model parameters are driven primarily by the sample data. Our uninformative or "base" priors are chosen as follows. First, we specify that u_{00} = u_{01} = 1 and u_{10} = u_{11} = 1. In this case, the priors for q and p become uniform distributions. Second, we set the normal prior for the slope coefficients to be N (0, 100). Third, the prior parameters for σ^2 are chosen so that the prior contains only about as much information as a sample of 4 observations and its mean match the sample variance.

We also performance a prior sensitivity analysis to ensure our results are robust to variations in the prior. Specifically, we also consider the following "no-predictability" prior, which specifies that the normal prior for the slope coefficients for all the predictive variables to have a mean of zero (the prior mean for the constant is set to the sample mean) and a variance of 1. Other aspect of the no-predictability prior is the same as our base prior. Hence investors who hold this no-predictability prior have a strong view that stock returns are unpredictable using any of the predictive variables.

To further disturb the prior specifications, we replace the uninformed base prior with priors "informed" by analysis of pre-sample data (i.e. training samples). We employ two sets of training samples. The first training sample spans a 10-year period from January 1934 to December 1943. The second training sample covers a 20-year period from January 1934 to December 1953. We report the results of this sensitivity analysis in Section III.

D. Bayesian Model Comparison

One of the objectives of this article is to compare the performance of linear models versus regime-switching models. In the Bayesian approach, model comparison is conceptually simple but computationally burdensome due to the need to calculate marginal likelihood. To set notation, assume that we have m different models. We rewrite the Bayes' rule as follows:

$$p(\theta^i|y, M_i) = \frac{p(y|\theta^i, M_i)p(\theta^i|M_i)}{p(y|M_i)}, \quad i = 1, \dots, m, \tag{20}$$

where M_i stands for the ith model that depends on the parameter vector θ^i. $p(y|M_i)$ is called the marginal likelihood (ML) and is given by

$$p(y|M_i) = \int p(y|\theta^i, M_i)\, p(\theta^i|M_i) d\theta^i. \tag{21}$$

The integration in equation (21) is often difficult to compute directly, especially for high-dimensional problems. A popular and very general approach to calculate the marginal likelihood was proposed by Gelfand and Dey (1994) (GD):

$$\widehat{ML}^{-1} = E\left[\frac{g(\theta)}{p(y|\theta^i, M_i)p(\theta^i|M_i)}\middle| y, M_i\right] \tag{22}$$

where the denominator on the right hand side of equation (22) is simply the product of the model-specific likelihood and prior. $g(\theta)$ is a tuning function and has to be chosen carefully for GD's method to work. We follow Geweke (1999) and use a truncated multivariate normal density. Let $\hat{\theta}$ and $\hat{\Sigma}$ be estimates of $E(\theta|y, Mi)$ and $var(\theta|y, Mi)$, which can be easily obtained from the posterior simulator. Moreover, for some probability p ∈ (0, 1), let $\widehat{\Theta}$ denote the support of $g(\theta)$, which is defined by

$$\widehat{\Theta} = \left\{\theta: (\hat{\theta} - \theta)' \hat{\Sigma}^{-1}(\hat{\theta} - \theta) \le \chi^2_{1-p}(k)\right\} \tag{23}$$

where $\chi^2_{1-p}(k)$ is the $(1-\text{p})$th percentile of the Chi-squared distribution with k degrees of freedom and k is the dimension of θ. Geweke (1999) suggests that we choose the multivariate Normal density truncated to the region $\widehat{\Theta}$,

$$g(\theta) = \frac{1}{p(2\pi)^{\frac{k}{2}}} \left|\hat{\Sigma}\right|^{-\frac{1}{2}} exp\left[-\frac{1}{2}\left(\hat{\theta}-\theta\right)' \hat{\Sigma}^{-1}(\hat{\theta}-\theta)\right] 1(\theta \in \widehat{\Theta}) \tag{24}$$

where $1(\theta \in \widehat{\Theta})$ is the indicator function. As pointed out by Geweke, the computational costs of experimenting with several different values of p is quite low. In our empirical implementation, we find our results are insensitive to these choices.

Once the marginal likelihood is calculated for each model, model comparison is relatively straightforward in the Bayesian approach. For example, to compare two competing models, we can calculate the posterior odds ratio (POR), which is defined as the the ratio of two posterior model probabilities:

$$POR_{ij} = \frac{p(M_i|y)}{p(M_j|y)} \tag{25}$$

The posterior model probabilities are easily derived from Bayes' rule

$$p(M_i|y) = \frac{p(y|M_i)p(M_i)}{p(y)} = \frac{p(y|M_i)p(M_i)}{\sum_{i=1}^{M} p(y|M_i)p(M_i)} \;,\; i = 1, \dots, M, \tag{26}$$

where $p(M_i)$ is the prior model probability for the ith model. Substituting equation (26) into equation (25), we have

$$POR_{ij} = \frac{p(y|M_i)p(M_i)}{p(y|M_j)p(M_j)} \tag{27}$$

Where $\frac{p(M_i)}{p(M_j)}$ is known as the prior odds ratio. A natural choice to assume that the prior model probabilities are the same across all models. In this case, the POR is also known as the Bayes factor, defined as:

$$BF_{ij} = \frac{p(y|M_i)}{p(y|M_j)} \tag{28}$$

which is simply the ratio of marginal likelihood values for the two models that we want to compare.

III. Empirical Evidence

A. Data

In this study we focus on the predictability of the excess monthly returns of the value-weighted index obtained from the Center for Research in Security Prices (CRSP) database. As a proxy for the risk-free rate, we use the one-month U.S. Treasury bill yield from the CRSP files. The main sample period is from January 1954 to December 2005, a total of 624 observations. The choice of this sample period is similar to previous studies in the literature on stock return predictability. To formulate the training sample priors, we also collect the pre-sample data spanning a 20-year period from January 1934 to December 1953.

At the beginning of each month we allow the investor to choose from a base set of five predictors, including one market variable and four business cycle-related variables. As pointed out by Honda and Tiwari (2006), the business cycle-related variables are motivated by Fama and French (1989) and have attracted a lot of attention in the literature. They include excess CRSP value-weighted index monthly return lagged by one month, lagged one-month Treasury bill yield, lagged credit spread calculated by taking the difference between monthly Moody's BAA bond return and AAA bond return, lagged term premium constructed by taking the difference between 10-year Treasury constant maturity rate and 3-month Treasury bill rate, and the dividend yield. As suggested by Campbell (1991) and Hodrick (1992), the one-month T-bill yield is stochastically detrended by subtracting its 12-month backward moving average. The dividend yield is constructed by dividing the cumulative dividends over the previous 12 months by the current index level. All interest rate-related data are obtained from St. Louis Federal Reserve Bank's web site. The long- term interest rate data used in our training sample are obtained from Robert Shiller's web site. It should be noted that the number of predictors that we use are less than those chosen by Avramov (2002) and Cremers (2002). This is due to the fact that the MCMC approach used in this article is significantly more demanding in terms of computational costs than Avramov and Cremers, who focus exclusively on linear models with known analytical solutions.

Table I provides some descriptive statistics for all the variables. The numbers are expressed in percentages. We report the mean, standard deviation, skewness, excess kurtosis, as well as the first order autocorrelation. Consistent with previous findings, we find the four business cycle-related variables, Treasury bill rate, credit spread, term premium, and dividend yield, show very high persistence with their first-order autocorrelation all above 0.8. In contrast the lagged monthly value-weighted CRSP index returns exhibits almost no autocorrelation. In addition, the interest rate variable has a very large excess kurtosis, indicating that it is unlikely to be normally distributed. The stock market return variable also appears more volatile than other predictive variables.

B. Pairwise Model Comparison

To sharpen the contrast between linear and the regime-switching forecasting models, we only consider linear models spanned by the five predictors and their regime-switching competitors. In other words, we exclude models where some of the predictors follow the regime-switching process and others do not. This choice allows us to make a pairwise comparison of linear versus regime-switching models. All models include the constant term and we also consider the null model, where the constant is the only regressor. This leaves us with a total of 64 models evenly divided between linear and regime-switching models.

The pairwise model comparison is made by calculating Bayes factors for the 32 pairs of models. We first calculate the marginal likelihood values of all the linear and regime- switching models based on the GD method using outputs from the posterior simulators. Then the Bayes factors are computed using equation (28) (ML of regime-switching model over ML of linear model). Table II reports the results for the 32 model pairs. The model pairs are distinguished by the predictors included. For example SRCTD stands for the model pair that includes the following variables (in addition to the constant): lagged stock returns (S), one-month Treasury bill rate (R), credit spread (C), term premium (T), and dividend yield (D). The Bayes factors are shown on the natural logarithm scale and multiplied by 2 to be consistent with the classification made by Kass and Raftery (1995), who suggest the following interpretations.

2loge(BF10)	Evidence against H0
0 to 2	Not worth more than a bare mention 2
to 6	Positive
6 to 10	Strong
> 10	Very Strong

Based on this classification, 27 out of the 32 model pairs show very strong evidence against linear model specifications. Only in two cases (SCTD and SRCTD), the regime switching models do not appear to have an edge over their linear counterparts. Moreover, the evidence against linear models appears very substantial since the majority of Bayes factors are well above the threshold levels set by Kass and Raftery. We also notice that the support for regime-switching models appears to decline as the number of predictors becomes large. Overall the results from table II are consistent with earlier findings (e.g. BH and PT) and demonstrate the importance of including nonlinear models when tackling stock return predictability problems.

C. Posterior Model Probabilities

While Bayes factors work great for pairwise model comparison, they do not reveal which one is the best-performing predictor or individual model. To this end, we need to calculate posterior model probabilities as defined in equation (26). Assuming equal prior model probabilities, it is easy to see that $p(M_i|y)$, the posterior model probability for model i, is the ratio of its own marginal likelihood over the sum of marginal likelihood for all the models.

To check the performance of various predictors, we report the cumulative posterior probabilities (CPP) of the five predictive variables in Table III. The Bayesian model averaging approach addresses the uncertainty in model specifications by taking a posterior probability weighted average model. CPP is the quantity that indicates the probabilities that each of the predictive variables appears in the weighted prediction model. To illustrate, suppose credit spread receives a CPP of 30%. This means that credit spread should appear in the weighted prediction model with a probability of 30%. Hence it is a very useful metric that summarizes the empirical support for each of the predictive variables. It is also used by Avramov (2002) to gauge the performance of linear forecasting models in the presence of model uncertainty.

We report the CPP for three cases. The first case takes the subset of linear forecasting models. The second case focuses on the subset of regime-switching models. The third case is the all-inclusive model space. In the linear model case, we find that the lagged one-month Treasury bill rate receive almost all the weight (> 99%). Thus if constrained within the set of linear models, we would only pick the short-term risk-free rate as a reliable predictor.

Turning to the regime-switching model case, we find some interesting results. First of all, CPP suggests that both lagged excess CRSP index returns and the lagged Treasury-bill rates be included in the optimal prediction model with a probability of one. By comparison, the other three predictors (credit spread, term premium, and dividend yield) receive nearly zero probabilities in the weighted model, which differs from the linear model case. Therefore, it appears that an econometrician's prior on the scope of models to be investigated has a nontrivial impact on our inference regarding the role of predictive variables. This leaves us to wonder which set of conclusions we should trust. The puzzle is solved by looking at the third case, which shows that the CPP for the all-inclusive case is in complete congruence with the regime-switching model subset. In other words, the linear models receive almost zero posterior model probability in the overall optimal weighted model.

Table IV reports the posterior model probability and the ranking of the 64 models under various priors. In panels A, B, C, and D, we show the results under the base prior, the no-predictability prior, the prior based on the 10-year training sample from 01/1934 to 12/1943, and the prior based on the 20-year training sample from 01/1934 to 12/1953, respectively. We note that in all cases, regardless of the priors implemented, the posterior model probability exclusively concentrates on the regime-switching model that include both the lagged CRSP index return and lagged Treasury bill rate as predictors. The remaining 63 models practically receive a posterior model probability of zero. Therefore, conditional on the set of models and predictors considered in this paper, we find the following individual model that uses lagged value-weighted CRSP index return (VW) and Treasury bill rate (TB) as its predictors is essentially equivalent to the overall weighted prediction model with its posterior model probability equal to one.

$$VW_t = b_0^0(1-S_t) + b_0^1 S_t + b_1^0(1-S_t)VW_{t-1} + b_1^1 S_t VW_{t-1} + b_2^0(1-S_t)TB_{t-1} + b_2^1 S_t TB_{t-1} + \varepsilon_t \quad (29)$$

Table V reports the posterior means and standard deviations for the various model parameters in equation (29). The posterior density plots are also shown in Figure 1. Because of the fact that this best individual model is virtually identical to the weighted prediction model, these results can be interpreted as those of the optimal weighted model as well. Several interesting findings emerge from these results.

First, we notice that our posterior parameter estimates for the transition probabilities indicate the regimes are not as persistent as those reported by Tu (2010). We attribute the difference to the fact that Tu imposes a comparatively strong prior on these parameters. For example, his prior is P = 91.67% and Q = 83.33%. In contrast, we are reluctant to impose such a strong prior and prefer to choose a non-informative prior and let the data speak for itself.

Second, we find that the constant term under regime zero is significantly negative, and vice versa under regime one. This result, combined with the fact that q is relatively small, indicates regime zero can be interpreted as the infrequent and less persistent “bear market” regime. On the other hand, regime one can be interpreted as a more long-lasting “bull market” regime. For instance, the average duration of the “bull market” regime is approximately 4.6 months, whereas the “bear market” regime only lasts for an average of 1.6 months. This interpretation is further confirmed by inspecting Figure 2, which shows that most of the time, the bull market regime plays a dominant role. However, during periods of uncertainty the bear market regime appears to take over. For example, the “bull market” regime appears to be interrupted by episodes of market uncertainty such as the oil crisis in the 1970s, the 1987 market crash, the liquidity crisis triggered by the fallout of the hedge fund Long-term Capital Management, and burst of the internet bubble after the year 2000.

Third, the slope coefficient for lagged stock market returns under regime zero is significantly positive whereas the coefficient under regime one is slightly negative. This is consistent with the evidence that stock returns in bear markets are more persistent. On the other hand, under the bull market regime, the CRSP index seems to be characterized by reversal rather than momentum in its returns as shown by the posterior parameter values. If we look at the coefficients for the lagged Treasury bill rate, we find that the relation between stock market return and short-term interest rate is negative in both regimes. This is not surprising given the fact that stock market often reacts favorably to announcements of interest rate cuts and unfavorably to news of rate hikes by the Federal Reserve Bank.

IV. Economic Significance

Kandel and Stambaugh (1996) (KS) study stock return predictability from an asset-allocation perspective. They show that it is important to consider not only statistical evidence but the economic significance of predictability as well. Motivated by their study, we consider a risk-averse Bayesian investor with a one-month investment horizon. The investor's objective is to maximize the expected utility by allocating his initial wealth WT in the value-weighted CRSP index and the risk-free asset at the end of month T . Let ω denote the vector of investment weights. Following Pastor and Stambaugh (2000), we assume the investor will choose w so as to maximize the mean-variance objective function:

$$\max_{w} \left(\omega' E - \frac{A}{2}\omega' V \omega\right) \tag{30}$$

where A is interpreted as the investor's coefficient of relative risk aversion. To capture different degrees of risk aversion, we study two cases where A = 5 and 10 in our empirical work. E and V denote the first two moments of predictive density function of asset returns. For models presented in this article, the predictive density cannot be solved analytically. However, it is straightforward to use the posterior outputs from the Gibbs sampler and draw samples from the predictive density. Details of the algorithm are explained in Albert and Chib (1993). It is well known that the solution to (30) is given by$\omega = \frac{1}{A}V^{-1}E$.

The metric for gauging the economic significance of modeling regime switching in fore-casting stock returns is defined as follows. Let ω_{RS} denote the optimal asset allocation under the propose regime switching model. Likewise, let ω_{Linear} denote the optimal asset allocation under an alternative linear forecasting model of stock market returns. Note that we use the same prior belief on both the regime switching and the linear models. CE_{RS} and CE_{Linear} are the certainty equivalent returns associated with ω_{RS} and ω_{Linear} respectively, where

$$CE_{RS} = \omega'_{RS}E - \frac{A}{2}\omega'_{RS}V\omega_{RS} \tag{31}$$

$$CE_{Linear} = \omega'_{Linear}E - \frac{A}{2}\omega'_{Linear}V\omega_{Linear} \tag{32}$$

Note that the E and V are predictive moments under the regime switching model. The proposed metric is the difference $CE_{RS} - CE_{Linear}$. It can be viewed as the certainty equivalent loss to an investor who would like to allocate his asset in accordance with the regime-switching model but is forced to accept the allocation under an alternative (suboptimal) linear model. This metric is first proposed by KS and has been widely applied in many studies.

Table VI reports the asset allocations under the optimal regime switching model as shown in equation (29) as well as the two alternative linear forecasting model. The first linear model has lagged CRSP index returns and lagged short-term interest rate as two predictive variables. Note that this is the linear counterpart of the optimal regime-switching model. For the second linear model, we choose the best model (as ranked by posterior model probability) when only linear models are considered. In this case, the lagged short-term interest rate turns out to be the only predictor (in addition to the constant term). We also present results under two priors: the base prior and the no-predictability prior.

First, we notice that under both models the risk-averse Bayesian investor chooses to take a short position in the stock market, which is attributable to the fact that our model is predicting a negative return in the coming month. However the main difference is that under the regime-switching model, the investor is willing to commit to a substantially larger position than under the alternative linear forecasting models. Hence it appears that incorporating the possibility of regime-switching has a significant impact on investor's portfolio decisions. In our view, this result could be driven by the fact that if regime switches in stock returns arepredictable, then the investors could afford to take a larger position, knowing that the regime risks are under control. On the other hand, an investor who are forced to use the linear model has to face the uncertainty introduced by the regime switches. In this suboptimal case, a natural way to hedge the regime risks is to hold more risk-free assets, which explains the observed smaller allocation to the stock market.

Second, we find that modeling regime switches for the purpose of predicting stock returns is economically very important. Recall that our measure of economics significance can be interpreted as utility loss perceived by investors who are forced to ignore regime switches and instead allocate funds based upon a suboptimal linear model. The metric is expressed in terms of an annual certainty equivalent risk-free rate. From table VI, we find under the base prior the annualized certainty equivalent loss from ignoring regime switches can range from 17.75% to more than 38%. Even when the investor has a strong prior belief that stock returns are unpredictable, the certainty equivalent loss can still be quite significant and range from 1.88% to 4.69%. It is important to note that in panels B and D the alternative linear forecasting model has already taking into account the effects of model uncertainty (in the space of all linear models). Hence the certainty equivalent metric reported here reflects the incremental economic benefits of allowing for regime switches beyond model uncertainty. This demonstrates that modeling regime switches is economically meaningful for stock return predictability.

V. Conclusion

Previous studies in the stock return predictability literature find evidence of model nonstationarity. To model the time-variation in model parameters, we propose to use the regime switching model in this article. We find important evidence, both statistically and economically, in support of modeling regime switches in stock return predictability.

To sum up, we find statistical evidence as showcased in the posterior odds ratios and posterior model probabilities strongly favors regime-switching models. In addition, the best predictors picked by linear models are different from those chosen by nonlinear models, which highlights in the importance of accounting for model uncertainty and indicates that focusing on linear models alone could lead to potentially misleading inferences. Moreover, the support for regime switches appears economically significant as measured by the utility loss perceived by investors who are forced to use an alternative linear forecasting model even after taking into account model uncertainty (in the linear model space). Conditional on the set of predictors included in this article, we find substantial evidence in favor of a regime- switching model that uses lagged stock returns and Treasury bill rates as its predictors. Our conclusions are robust to the use of different priors.

Unlike previous studies that focus exclusively on linear models, our approach is more flexible in that it can account for both model uncertainty and nonlinearity in the underlying DGP. However, the tradeoff is that computational costs are significantly higher than the case of linear models as analytical solutions are no longer available. Hence, we can only include a modest number of predictors.

For future research, we can extend the model by incorporating regime switches in stock return volatilities. One benefit is that it can account for stock return heteroscedasticity. In addition it might also help us further explore the relation between return volatility and the regimes identified in this article.

REFERENCES

Albert, James H., and Siddhartha Chib, 1993, Bayes inference via gibbs sampling of autoregressive time series subject to markov mean and variance shifts, *Journal of Business and Economic Statistics* 11, 1–15.

Ang, Andrew, and Allan Timmermann, 2011, Regime changes and financial markets, *NBER working paper 17182*.

Ang, Andrew, and Geert Bekaert, 2002, International asset allocation with regime shifts, *The Review of Financial Studies* 15, 1137–1187.

Avramov, Doron, 2002, Stock return predictability and model uncertainty, *Journal of Financial Economics* 64, 423–458.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing statistical criteria to select return forecasting models: What do we learn?, *Review of Financial Studies* 12, 405–428.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Chib, S., 1996, Calculating posterior distributions and modal estimates in Markov mixture models, *Journal of Econometrics* 75, 79–98.

Cooper, Michael J., John J. McConnell, and Alexei V. Ovtchinnikov, 2006, The other January effect, *Journal of Financial Economics* 82, 315–341.

Cremers, K. J. Martijn, 2002, Stock return predictability: A Bayesian model section perspective, *Review of Financial Studies* 15, 1223–1249.

Dangl, Thomas, and Michael Halling, 2012, Predictive regressions with time-varying coefficients, *Journal of Financial Economics,* 106, 157-181.

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

Gelfand, A. E., and Dey D., 1994, Bayesian model choice: Asymptotics and exact calculations, *Journal of the Royal Statistical Society Series B* 56, 501–514.

Geweke, John, 1999, Using simulation methods for bayesian econometric models: Inference, development, and communication (with discussion and rejoiner), *Econometric Reviews* 18, 1–126.

Gray, Stephen F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27–62.

Hamilton, James D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357–384.

Hamilton, J. D., and Raul Susmel, 1994, Autoregressive conditional heteroscedasticity and changes in regime, *Journal of Econometrics* 64, 307–333.

Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Handa, P., and A. Tiwari, 2006, Does stock return predictability imply improved asset allocation and performance? Evidence from the U.S. stock market (1954-2002), *Journal ofBusiness* 79, 2423–2468.

Kandel, Shmuel, and Robert F. Stambaugh, 1996, On the predictability of stock returns: An asset-allocation perspective, *Journal of Finance* 51, 385–424.

Kass, Robert E., and Adrian E. Raftery, 1995, Bayes factors, *Journal of the American Statistical Association* 90, 773–795.

Kim, Chang-Jin, and Charles R. Nelson, 1998, Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching, *The Review of Economics and Statistics* 80, 188–201.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *Review of Financial Studies* 21 (4), 1607-1652.

Lo, Andrew, and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

Maheu, John M., and Thomas H. McCurdy, 2000, Identifying bull and bear markets in stock returns, *Journal of Business and Economics Statistics* 18, 100–112.

Pastor, L., and Robert F. Stambaugh, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

Pesaran, M. Hashem, and Allan Timmermann, 1995, Predictability of stock returns: Robustness and economic significance, *Journal of Finance* 50, 1201–1228.

Pontiff, Jeffrey, and Lawrence D. Schall, 1998, Book-to-market ratios as predictors of market returns, *Journal of Financial Economics* 49, 141–160.

Tu, Jun, 2010, Is regime switching in stock returns important in portfolio decisions? *Management Science* 56(7), 1198-1215.

——— , and Guofu Zhou, 2004, Data-generating process uncertainty: What difference does it make in portfolio decisions?, *Journal of Financial Economics* 72, 385–421.

Turner, Christopher M., Richard Startz, and Charles R. Nelson, 1989, A Markov model of heteroskedasticity, risk, and learning in the stock market, *Journal of Financial Economics* 25, 3–22.

Veronesi, Pietro, 1999, Stock market overreaction to bad news in good time: A rational expectations equilibrium model, *The Review of Financial Studies* 12, 957–1007.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley & Sons).

Table I

Descriptive Statistics

This table reports some descriptive statistics for the data set used in this study. The variables include: excess monthly returns of value-weighted CRSP index (S_t), lagged value-weighted CRSP index returns (S_{t-1}), one-month Treasury Bill rates (R), credit spread (C), term premium (T), and dividend yield (D). The sample period is from January 1954 to December 2005.

Variable	Mean	Std. Dev.	Skewness	Ex. kurtosis	ρ_1
S_t	0.58335	4.1474	−0.42845	1.9907	0.0531
S_{t-1}	0.59098	4.1514	−0.43055	1.9767	0.0536
R	0.015786	0.98316	−0.27572	3.7047	0.8154
C	0.94963	0.41612	1.4201	2.2034	0.9720
T	1.3999	1.1882	−0.11784	−0.13130	0.9532
D	3.1640	1.0379	−0.10912	−0.55437	0.9890

Table II

Bayes Factors for Pairwise Model Comparison

This table reports the (logarithm of) Bayes Factors for pairs of regime-switching and linear models that include the same predictors: lagged value-weighted CRSP index returns (S), one-month Treasury Bill rates (R), credit spread (C), term premium (T), and dividend yield (D). Constant refers to the models where no predictors are included (except for the constant).

Model	$2log_e(BF)$	Model	$2log_e(BF)$
Constant	101.4	SRC	78.8
S	110.2	SRT	102.0
R	54.4	SRD	91.0
C	15.0	SCT	89.4
T	3.2	SCD	14.0
D	37.6	STD	38.2
SR	156.4	RCT	94.8
SC	86.8	RCD	34.6
ST	113.8	RTD	31.6
SD	85.0	CTD	22.8
RC	29.2	SRCT	31.8
RT	73.0	SRCD	8.0
RD	18.6	SRTD	77.2
CT	46.0	SCTD	1.2
CD	22.4	RCTD	7.8
TD	34.0	SRCTD	1.2

Table III

Cumulative Posterior Probabilities of the Predictive Variables

This table reports cumulative posterior probabilities of the five predictive variables: lagged value-weighted CRSP index returns (S), one-month Treasury Bill rates (R), credit spread (C), term premium (T), and dividend yield (D).. We show the cumulative posterior probabilities for three different cases: a subset with only linear models, a subset with only regime-switching models, and the all-inclusive model space.

Variable	Linear Model	Regime-Switching Model	Overall
S	0.00008	1.00000	1.00000
R	0.99658	1.00000	1.00000
C	0.00000	0.00000	0.00000
T	0.00002	0.00000	0.00000
D	0.00000	0.00000	0.00000

Table IV
Sensitivity of Model Comparison to Prior Specifications

This table presents model rankings and posterior probabilities for 64 linear and regime switching models. Predictive variables include: lagged value-weighted CRSP index returns (S), one-month Treasury Bill rates (R), credit spread (C), term premium (T), and dividend yield (D). We also consider the case where only a constant is included in the model (Constant). Panels A, B, C, and D report the results for our base uninformative prior, the no-predictability prior, and priors based on two training samples.

Panel A: Base Prior

Linear Model	Prob	Rank	Regime-Switching Model	Prob	Rank
Constant		9	Constant		3
S		10	S		2
R		8	R		5
C		30	C		26
T		12	T		11
D		32	D		24
SR		14	SR	1	1
SC		34	SC		13
ST		17	ST		4
SD		39	SD		18
RC		33	RC		25
RT		16	RT		7
RD		35	RD		29
CT		28	CT		19
CD		54	CD		48
TD		49	TD		40
SRC		41	SRC		22
SRT		21	SRT		6
SRD		46	SRD		23
SCT		38	SCT		15
SCD		50	SCD		44
STD		53	STD		42
RCT		43	RCT		20
RCD		45	RCD		37
RTD		55	RTD		47
CTD		57	CTD		56
SRCT		36	SRCT		27
SRCD		51	SRCD		61
SRTD		52	SRTD		31
SCTD		60	SCTD		59
RCTD		62	RCTD		58
SRCTD		64	SRCTD		63

Panel B: No-predictability Prior

Linear Model	Prob	Rank	Regime-Switching Model	Prob	Rank
Constant		18	Constant		16
S		19	S		9
R		13	R		12
C		40	C		26
T		21	T		17
D		44	D		33
SR		22	SR	1	1
SC		42	SC		3
ST		29	ST		8
SD		50	SD		10
RC		38	RC		32
RT		24	RT		14
RD		43	RD		23
CT		39	CT		27
CD		60	CD		46
TD		56	TD		30
SRC		47	SRC		6
SRT		31	SRT		2
SRD		52	SRD		11
SCT		45	SCT		7
SCD		58	SCD		35
STD		57	STD		15
RCT		49	RCT		34
RCD		51	RCD		37
RTD		59	RTD		25
CTD		62	CTD		53
SRCT		41	SRCT		4
SRCD		54	SRCD		20
SRTD		55	SRTD		5
SCTD		63	SCTD		36
RCTD		61	RCTD		48
SRCTD		64	SRCTD		28

Table IV
(continued)

Panel C: Training Sample Priors: January 1934 - December 1943

Linear Model	Prob	Rank	Regime-Switching Model	Prob	Rank
Constant		64	Constant		32
S		63	S		2
R		62	R		30
C		61	C		29
T		60	T		28
D		59	D		27
SR		58	SR	1	1
SC		57	SC		25
ST		56	ST		24
SD		55	SD		23
RC		54	RC		22
RT		53	RT		21
RD		52	RD		20
CT		51	CT		19
CD		50	CD		18
TD		49	TD		17
SRC		48	SRC		16
SRT		47	SRT		15
SRD		46	SRD		3
SCT		45	SCT		13
SCD		44	SCD		12
STD		43	STD		11
RCT		42	RCT		10
RCD		41	RCD		9
RTD		40	RTD		8
CTD		39	CTD		7
SRCT		38	SRCT		6
SRCD		37	SRCD		26
SRTD		36	SRTD		4
SCTD		35	SCTD		31
RCTD		34	RCTD		14
SRCTD		33	SRCTD		5

Panel D: Training Sample Priors: January 1934 - December 1953

Linear Model	Prob	Rank	Regime-Switching Model	Prob	Rank
Constant		64	Constant		32
S		63	S		31
R		62	R		30
C		61	C		4
T		60	T		28
D		59	D		27
SR		58	SR	1	1
SC		57	SC		5
ST		56	ST		24
SD		55	SD		23
RC		54	RC		2
RT		53	RT		21
RD		52	RD		20
CT		51	CT		19
CD		50	CD		18
TD		49	TD		17
SRC		48	SRC		3
SRT		47	SRT		15
SRD		46	SRD		6
SCT		45	SCT		13
SCD		44	SCD		12
STD		43	STD		11
RCT		42	RCT		10
RCD		41	RCD		25
RTD		40	RTD		16
CTD		39	CTD		8
SRCT		38	SRCT		22
SRCD		37	SRCD		26
SRTD		36	SRTD		7
SCTD		35	SCTD		9
RCTD		34	RCTD		14
SRCTD		33	SRCTD		29

Table V

Posterior Results for the Best Individual Model

This table reports posterior results for the following best individual model that includes lagged value-weighted CRSP index return (VW) and Treasury bill rate (TB).

$$VW_t = b_0^0(1 - S_t) + b_0^1 S_t + b_1^0(1 - S_t)VW_{t-1} + b_1^1 S_t VW_{t-1} + b_2^0(1 - S_t)TB_{t-1} + b_2^1 S_t TB_{t-1} + \varepsilon_t$$

	Posterior Mean	Posterior Std. Dev.
b_0^0	-2.7947	0.94368
b_0^1	1.7780	0.30900
b_1^0	0.42396	0.10582
b_1^1	-0.19390	0.064647
b_2^0	-1.1535	0.51885
b_2^1	-0.76362	0.20475
σ^2	11.946	0.93903
q	0.38132	0.085343
p	0.78378	0.10681

Asset Allocations and Utility Loss

This table reports the asset allocations for stock market under the optimal regime switching model and alternative linear models. Panels A and B show results obtained under the base prior. Panels C and D show results under the no-predictability prior. Utility loss is computed as the loss in an annual certainty equivalent risk-free rate perceived by investors who are forced to ignore regime switches and instead allocate funds using a suboptimal linear model.

Panel A: Base Prior and Linear Model with Lagged Return and Short Rate

A	Asset Allocation: RS Model	Asset Allocation: Linear Model	Utility Loss (%)
5	-2.9824	-0.054971	38.383
10	-1.4912	-0.027485	19.192

Panel B: Base Prior and Linear Model with Lagged Short Rate

A	Asset Allocation: RS Model	Asset Allocation: Linear Model	Utility Loss (%)
5	-2.9824	-0.16705	35.501
10	-1.4912	-0.083526	17.750

Panel C: No-predictability Prior and Linear Model with Lagged Return and Short Rate

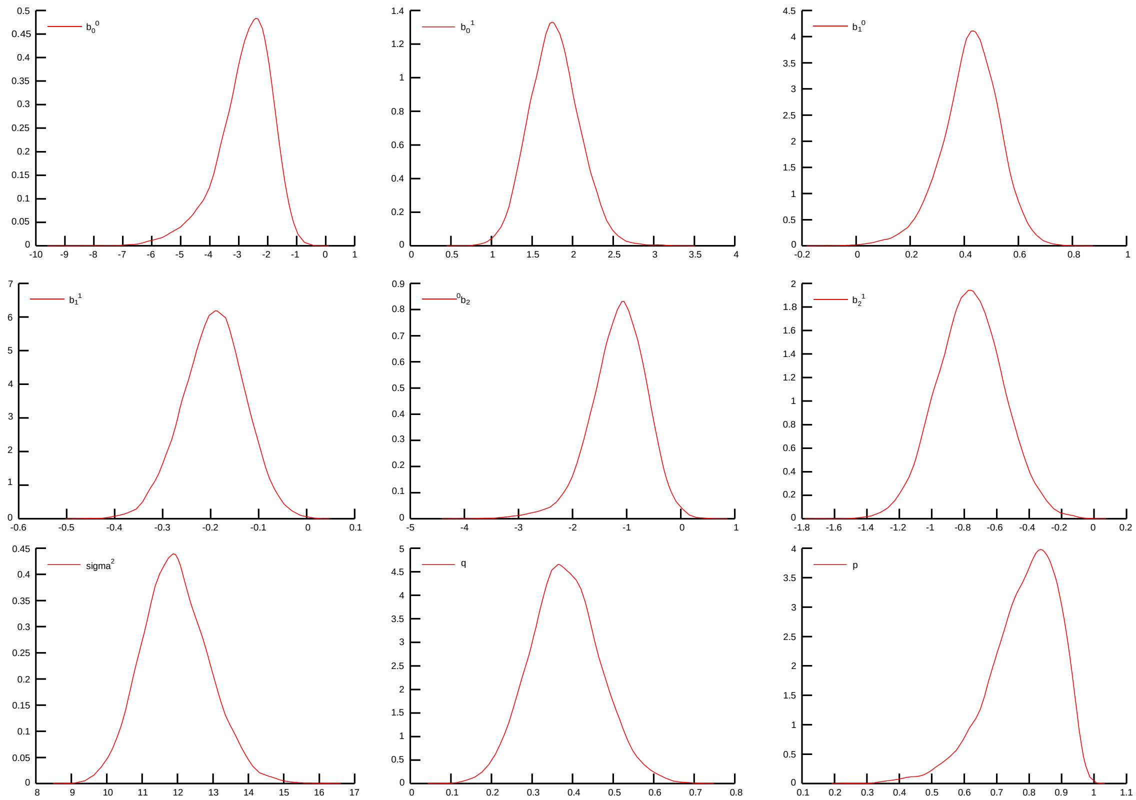
A	Asset Allocation: RS Model	Asset Allocation: Linear Model	Utility Loss (%)
5	-1.1275	-0.029494	4.6861
10	-0.56375	-0.014747	2.3431

Panel D: No-predictability Prior and Linear Model with Lagged Short Rate

A	Asset Allocation: RS Model	Asset Allocation: Linear Model	Utility Loss (%)
5	-1.1275	-0.14372	3.7618
10	-0.56375	-0.071860	1.8809

Figure 1

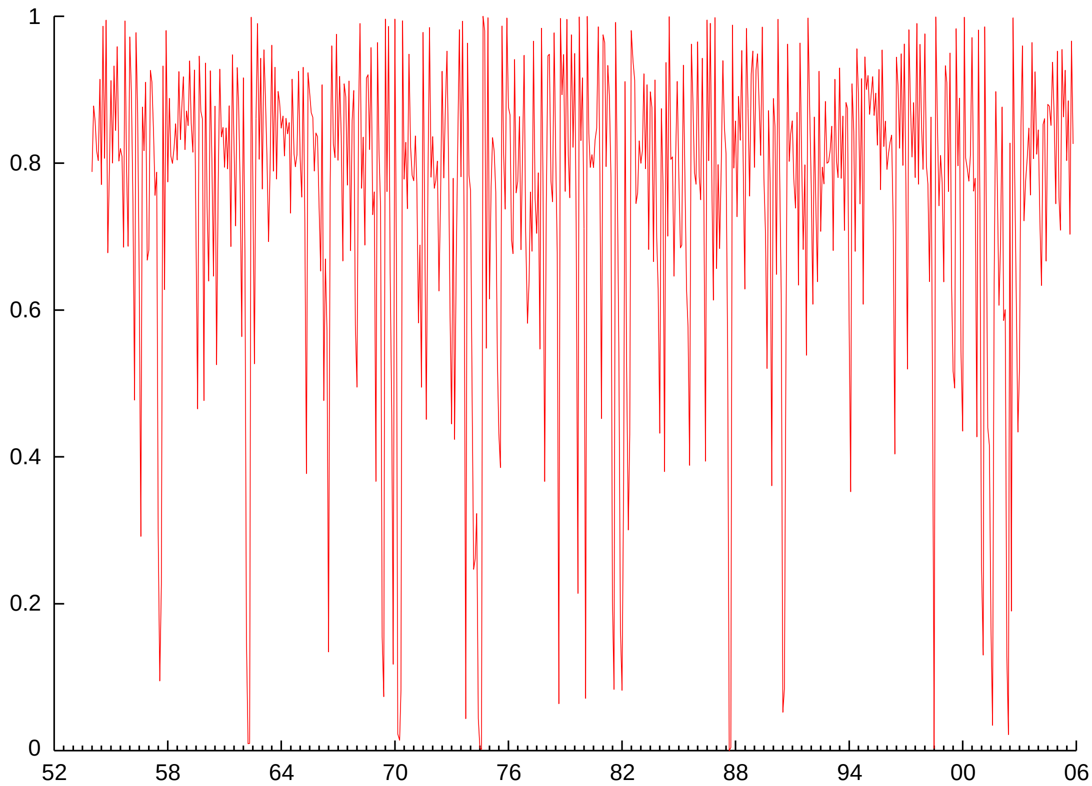
Posterior Density Plots



This figure plots the posterior densities for the model parameters as shown in equation (29).

Figure 2

Regime Probabilities



The figure plots the probabilities of the bull market regime.