

# The Structural Break in the Equity Premium

**Chang-Jin KIM**

Department of Economics, Korea University, Seoul, 136-701, Korea (*ckim@korea.ac.kr*)

**James C. MORLEY**

Department of Economics, Washington University, St. Louis, MO 63130 (*morley@economics.wustl.edu*)

**Charles R. NELSON**

Department of Economics, University of Washington, Seattle, WA 98195 (*cnelson@u.washington.edu*)

This article uses Bayesian marginal likelihood analysis to compare univariate models of the stock return behavior and test for structural breaks in the equity premium. The analysis favors a model that relates the equity premium to Markov-switching changes in the level of market volatility and accommodates volatility feedback. For this model, there is evidence of a one-time structural break in the equity premium in the 1940s, with no evidence of additional breaks in the postwar period. The break in the 1940s corresponds to a permanent reduction in the general level of stock market volatility. Meanwhile, there appears to be no change in the underlying risk preferences relating the equity premium to market volatility. The estimated unconditional equity premium drops from an annualized 12% before to the break to 9% after the break.

**KEY WORDS:** Bayes factor; Equity premium; Marginal likelihood; Markov switching; Structural break; Volatility feedback.

## 1. INTRODUCTION

The equity premium is the expected excess return on a market portfolio over the risk-free interest rate. Although one of the most important variables in financial economics, it is not directly observable and must be estimated, typically from historical stock return data. The standard approach to estimation is to use the average excess return over a given sample period. This approach is reasonable to the extent that excess returns behave as if drawn from a stable distribution for the entire sample. However, beyond the simple iid case, using the sample average to estimate the equity premium does not exploit all of the information inherent in a fully specified probabilistic model of excess stock returns. It also does not account for possible structural (permanent) breaks in the equity premium during the sample period under investigation.

In this article, we use Bayesian marginal likelihood analysis to compare univariate models of excess stock returns and test for structural breaks in the equity premium. In the classical framework, model comparison and tests for structural breaks with unknown breakpoints are complicated because of the presence of nuisance parameters under alternative hypotheses. However, in the Bayesian framework, nuisance parameters do not pose any special problem as long as they, along with the other parameters, can be integrated out of the likelihood function to solve for the marginal likelihood of a given model of interest. This is the approach that we take in this article. In particular, we follow Chib's (1995) procedure for directly calculating the marginal likelihood based on the output of the Gibbs sampler, which we also use to obtain inferences about the distributions of model parameters. Then we construct Bayes factors based on the marginal likelihoods, allowing us to make model comparisons and to test for structural breaks. Meanwhile, the Gibbs sampler makes it computationally feasible to estimate models that feature both Markov switching and structural breaks with unknown breakpoints.

We consider four models of excess stock returns. Model I assumes a constant equity premium and a constant level of

market volatility within each structural regime (i.e., excess returns are iid normal within each subsample between structural breaks). Model II assumes a constant equity premium within each structural regime, but allows two-state Markov-switching market volatility. Model III assumes that the equity premium changes in response to the Markov-switching volatility within each structural regime. Model IV assumes that the equity premium responds to volatility within each structural regime and also accommodates the volatility feedback effects of unanticipated changes in the equity premium on excess returns. For each model, we test for structural breaks with unknown breakpoints in model parameters corresponding to market volatility and the equity premium. We use monthly excess return data for a value-weighted portfolio of all NYSE stocks from the CRSP files for the sample period 1926–1999.

Our main findings can be summarized as follows. The empirical Bayes factors strongly favor the three models that incorporate Markov-switching volatility (models II–IV) over the simple iid model (model I). The “most preferred” model (model IV) relates the equity premium to Markov-switching changes in the level of market volatility and also accommodates the volatility feedback effects of unanticipated changes in the equity premium on excess returns. For all four models, there is evidence of a permanent reduction in the general level of stock market volatility in the 1940s. Furthermore, because the “most preferred” model directly relates the equity premium to the level of market volatility, the permanent reduction in volatility in the 1940s corresponds to a structural break in the equity premium. There is mild evidence against a change in the underlying risk preferences relating the equity premium to market volatility and stronger evidence against additional structural breaks. The estimated annualized equity premium drops from around 12% before the structural break to about 9% after the break.

These findings represent somewhat of a departure from what has been found elsewhere in the literature on structural breaks and the equity premium. Pástor and Stambaugh (2001), who also used a Bayesian approach to test for structural breaks in the equity premium, reported a large number of structural breaks (15 in total) over a longer sampling period (1834–1999), including a sizeable break in the 1990s. However, they assumed that excess returns were iid normal within structural regimes. [It should be noted that Pástor and Stambaugh (2001) also allowed for a transition period between structural regimes that essentially accounts for volatility feedback after a permanent change in the level of market volatility.] In contrast, we allow for temporary changes in volatility and the equity premium, and our preferred model allows us to separate out whether the structural break in the equity premium is driven by a change in the volatility process or by a change in risk preferences. Fama and French (2002) found a much lower estimate of 2.5% for the equity premium using the average dividend yield plus the average dividend growth rate for the S&P index over the period 1951–2000. The logic of their approach is that if the dividend yield is stationary, then the average dividend growth rate provides a reasonable estimator of the expected capital gain. Furthermore, unlike the average capital gain, their estimator does not suffer from a small-sample bias due to large capital gains/losses generated by unanticipated changes in the equity premium. They argued that the lower estimate for the equity premium reflects an elimination of this bias. On the other hand, they conceded that the dividend growth rate approach can suffer from a small-sample bias of its own if there is a large shift among firms away from dividend payments toward share repurchases. [Fama and French (2002) also considered earnings data instead of dividends. Earnings data produce a higher estimated equity premium than dividends, although the estimate of 4.3% is still well below our estimate of 9% for the corresponding sample period. However, as with dividends, there is a bias in their approach if the relationship between prices and earnings undergoes permanent changes over the sample period.] In contrast, we address the bias in using the average capital gain directly by accounting for the volatility feedback effects on share prices of unanticipated changes in the equity premium due to changes in the level of market volatility. Our approach avoids the bias introduced by shifts in dividend policies or any other factors that might cause permanent changes in the dividend yield.

The rest of the article is organized as follows. Section 2 presents the details of the four models considered in our Bayesian model comparison. Section 3 discusses how we incorporate structural breaks with unknown breakpoints into these models to allow for permanent changes in market volatility and the equity premium. Section 4 provides an overview of the Bayesian approach used in the article, and Section 5 presents our empirical analysis. Section 6 concludes.

## 2. MODELS

The most basic model of excess stock returns assumes a constant equity premium and constant market volatility:

Model I: *Constant mean and variance*

$$r_t = \mu + \varepsilon_t$$

and

$$\varepsilon_t \sim \text{iid } N(0, \sigma^2),$$

where  $r_t$  is an excess return on a market portfolio,  $\mu$  is the equity premium, and  $\varepsilon_t$  is market news at time  $t$  with constant variance  $\sigma^2$ . Given this model, a diffuse prior, and a quadratic loss function, the sample average would provide the optimal Bayesian estimate of the equity premium. However, the higher moments of monthly stock returns are not consistent with this model. In particular, the historical excess returns on a value-weighted NYSE portfolio examined in this article are characterized by negative sample skewness ( $-.4910$  vs.  $0$  for normality) and excess kurtosis ( $10.2907$  vs.  $3$  for normality). The corresponding Jarque–Bera test of normality yields a  $p$  value of  $< .00001$ .

A standard way of capturing the excess kurtosis in stock returns is to allow for time-varying volatility. The second model that we consider allows market volatility to switch between two discrete levels (high and low):

Model II: *Constant mean and switching variance*

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \\ \varepsilon_t | S_t &\sim N(0, \sigma_t^2), \\ \sigma_t^2 &= \sigma_L^2 + (\sigma_H^2 - \sigma_L^2) S_t, \\ S_t &= \{0, 1\}, \quad \text{and} \quad \sigma_L^2 < \sigma_H^2, \end{aligned}$$

where the variance of market news equals either  $\sigma_L^2$  or  $\sigma_H^2$ , depending on the latent first-order Markov-switching state variable  $S_t$ , which evolves according to transition probabilities

$$\Pr[S_t = 0 | S_{t-1} = 0] = q$$

and

$$\Pr[S_t = 1 | S_{t-1} = 1] = p.$$

Note that the unconditional variance  $\sigma^2 \equiv E[\sigma_t^2]$  for this model is

$$\sigma^2 = \sigma_L^2 + (\sigma_H^2 - \sigma_L^2) \frac{1-q}{2-q-p}.$$

For this model, the sample average no longer provides the optimal estimate of the equity premium. Intuitively, returns drawn from the high variance distribution are less informative and should receive less weight than returns drawn from the low variance distribution in making inferences about the equity premium.

Because a trade-off between risk and return is one of the central tenets of financial economics, it seems natural to allow the possibility that time-varying volatility could generate a time-varying equity premium. The third model that we consider assumes that the equity premium is a linear function of the expected level of market volatility:

Model III: *Time-varying mean and switching variance*

$$r_t = \mu_t + \varepsilon_t$$

and

$$\mu_t = \alpha + \beta E_{t-1}[\sigma_t^2],$$

where  $\mu_t$  is the time-varying equity premium. The intercept  $\alpha$  allows the marginal price of risk, measured by the slope  $\beta$ ,

differ from the average price of risk. As discussed by Wang (2001), a difference could arise due to the existence of a constant premium to compensate for the general risk of regime change in addition to the time-varying premium reflecting the current level of volatility. Indeed, the intercept can account for any time-invariant factor that generates an equity premium. Motivated by financial theory, we constrain  $\alpha$  and  $\beta$  to be non-negative, although we also consider the sensitivity of our findings to these constraints. [We note that Merton (1980) proposed nonnegativity constraints as reasonable prior beliefs to be used in estimating the equity premium, as opposed to regularities that one might verify with the data. In practice, the constraint on the intercept term is not binding. Likewise, a nonnegativity constraint on the constant equity premium in models I and II is not binding. Meanwhile, as discussed later, the constraint on the marginal price of risk is not binding if it is linked to volatility feedback.] The specifications for  $\varepsilon_t$  and  $\sigma_t^2$  are the same as in the previous model. The unconditional variance  $\sigma^2$  is also the same. In this case the unconditional equity premium  $\mu \equiv E[\mu_t]$  is

$$\mu = \alpha + \beta\sigma^2.$$

Again, the sample average does not provide the optimal estimate of the unconditional equity premium for this model. As before, the information content of a given return depends on the level of market volatility.

The fourth and last model that we consider assumes a linear relationship between the equity premium and the expected level of market volatility and also allows for volatility feedback effects of unanticipated changes in the equity premium:

Model IV: *Time-varying mean and switching variance with volatility feedback*

$$r_t = \mu_t + f_t + \varepsilon_t$$

and

$$f_t = \delta(\sigma_t^2 - E_{t-1}[\sigma_t^2]),$$

where  $f_t$  is the volatility feedback term and  $\delta$  is the volatility feedback coefficient. The specifications for  $\mu_t$ ,  $\varepsilon_t$ , and  $\sigma_t^2$  are the same as in the previous model. The unconditional equity premium,  $\mu$ , and unconditional variance,  $\sigma^2$ , are also the same. As with the previous two models, the sample average does not provide the optimal estimate of the unconditional equity premium. In this case, the information content of a given return depends on both the level of market volatility and the volatility feedback term.

Volatility feedback is the idea that an exogenous change in the level of market volatility generates a large price adjustment in response to new information about future discounted expected returns (see Kim, Morley, and Nelson 2004 and references therein for a thorough discussion of volatility feedback). In terms of the foregoing model, the new information is measured by the deviation between the expected level of volatility at the beginning of a trading period,  $E_{t-1}[\sigma_t^2]$ , and the actual realized level of volatility,  $\sigma_t^2$ . This specification is from Turner, Startz, and Nelson (1989) and was also used by Kim et al. (2001, 2004). Given a persistent volatility process ( $p + q > 1$ ), stock prices should initially move in the opposite direction to a change in the equity premium. Because the price movement

generated by the change in volatility reflects the effect on all future discounted expected returns, rather than just the partial effect on the current period equity premium, the sign of the marginal price of risk is easier to identify from the volatility feedback coefficient than from the slope coefficient  $\beta$ . Indeed, if we link the marginal price of risk to the volatility feedback by constraining the sign of the volatility feedback effect to be opposite to the change in the equity premium (i.e.,  $\beta \cdot \delta \leq 0$ ), then the nonnegativity constraints on  $\alpha$  and  $\beta$  do not bind. That is, the estimated volatility feedback effect makes it clear that the price of risk is positive. We discuss this point in more detail in Section 5.

### 3. STRUCTURAL BREAKS

In our empirical analysis of the four models just introduced, we allow some or all of the model parameters to undergo structural (permanent) breaks with unknown breakpoints. The timing of a breakpoint  $\tau_i$ ,  $i = 1, \dots, n$ , where  $n$  is the total number of breakpoints, is determined by a latent variable,  $D_t$ , which we model as following a  $(n + 1)$ -state Markov-switching process with constrained transition probabilities as suggested by Chib (1998) and used previously to test for structural breaks in a Bayesian context by Kim and Nelson (2001). In particular, the transition probabilities for  $D_t$  are given by

$$\Pr[D_t = j | D_{t-1} = j] = \kappa_j,$$

$$\Pr[D_t = j + 1 | D_{t-1} = j] = 1 - \kappa_j,$$

and

$$\Pr[D_t = j' | D_{t-1} = j] = 0, \quad j' \neq j, j + 1,$$

where  $j = 1, \dots, n + 1$ ,  $0 \leq \kappa_i \leq 1$ , and  $\kappa_{n+1} = 1$ , which implies that the last structural regime is “absorbing” in the sense that there are no additional structural breaks within the sample.

Denoting the set of model parameters affected by structural breaks as  $\theta_{D_t}$ , we link these parameters to the latent variable  $D_t$ ,

$$\theta_{D_t} = \sum_{j=1}^{n+1} D_{jt} \theta_j$$

and

$$D_{jt} = \begin{cases} 1, & \text{if } D_t = j \\ 0, & \text{otherwise,} \end{cases}$$

where  $\theta_j$  is the set of parameter values that prevails during the structural regime associated with  $D_t = j$ .

There is an important practical issue in modeling structural breaks with unknown breakpoints. It turns out to be necessary to set a lower bound on the length of each structural regime to avoid any irregularities in the likelihood function that may occur when a subsample is too small. For example, outliers such as the 1987 stock market crash can easily be overfit by frequently changing means and variances. Also, structural regimes need to last a significant number of periods to identify transitory Markov-switching regime shifts within the structural regimes. Thus, in our estimation we consider a lower bound for the length of structural regimes of 88 months, corresponding to 10% of the 1926–1999 sample.

For every model, we consider three different assumptions about structural breaks: (A) There is no structural break; (B) only parameters associated with the volatility process are affected by structural breaks; and (C) all of the parameters (except the break probabilities  $\kappa_i$ ,  $i = 1, \dots, n$ ) are affected by structural breaks. Note that for model I, only  $\sigma^2$  is associated with the volatility process, whereas for models II–IV, the parameters  $\sigma_L^2$ ,  $\sigma_H^2$ ,  $q$ , and  $p$  are all associated with the volatility process. For model II, both structural break assumptions (B) and (C) correspond to structural breaks in the unconditional variance  $\sigma^2$ , whereas for models III and IV, both assumptions (B) and (C) correspond to structural breaks in the unconditional mean  $\mu$  and unconditional variance  $\sigma^2$ . However, for models III and IV and assumption (B), structural breaks in the unconditional mean  $\mu$  are driven entirely by changes in the unconditional variance, not by changes in the risk preference parameters  $\alpha$  and  $\beta$ .

It should be noted that the character of the latent variable  $D_t$  is somewhat different than that of the latent variable  $S_t$  in models II–IV. Not only are the states for  $D_t$  “terminal” in the sense that on exit they never recur, but also the model parameters associated with the volatility process, the unconditional variance, and the unconditional mean change when  $D_t$  switches. In contrast, only the parameter associated with the current level of volatility changes with  $S_t$  (between  $\sigma_L^2$  and  $\sigma_H^2$ ). For  $S_t$ , the volatility process, the unconditional variance, and the unconditional mean remain fixed. Although it is possible to think of models II–IV with  $S_t$  as restricted versions of model I with a large number of states for  $D_t$ , these restrictions are important because they greatly reduce the number of model parameters. Also, the restriction that structural regimes last at least 10% of the sample creates a clear distinction between permanent changes and transitory Markov-switching regime shifts.

#### 4. THE BAYESIAN APPROACH

In the classical framework, testing for Markov switching and/or structural breaks with unknown breakpoints is hindered by the presence of nuisance parameters under the alternative hypotheses (see Hansen 1992 and Garcia 1998 on testing for Markov switching within the classical framework). Specifically, the transition probabilities  $q$  and  $p$  are nuisance parameters for Markov-switching models and the unknown breakpoints  $\tau_i$  are nuisance parameters for models with structural breaks. In a Bayesian framework, however, nuisance parameters do not pose any special problem for inference as long as they can be integrated out of the likelihood function for each model under consideration. Thus, in this article we cast the problem of making inferences about Markov switching and structural breaks into a Bayesian model selection framework.

First, we use the Gibbs sampler to obtain the marginal posterior distributions for the model parameters (see Kim and Nelson 1999a for a direct comparison of Bayesian inference and classical inference for Markov-switching models). In Appendix A we provide a description of the Gibbs-sampling approach for the most general model specification under consideration. For model comparison, we use Bayes factors based on marginal likelihoods for each model. In particular, if we assume that the

data  $\tilde{Y}_T = \{r_1, \dots, r_T\}'$  have arisen from a given model specification according to probability function (marginal likelihood)  $m(\tilde{Y}_T|\omega)$ , where  $\omega$  is a model indicator parameter, then the Bayes factor in favor of model  $i$  over model  $j$  is defined as

$$BF_{ij} \equiv \frac{m(\tilde{Y}_T|\omega = i)}{m(\tilde{Y}_T|\omega = j)}, \quad i \neq j.$$

Various ways of calculating Bayes factors have been proposed in the literature. For example, Carlin and Polson (1991), George and McCulloch (1993), Geweke (1996), and Carlin and Chib (1995) provided procedures for model comparison based on the sensitivity of the posterior probability of the model indicator parameter  $\omega$  to the prior probability. Kim and Nelson (1999b) extended Carlin and Chib’s (1995) procedure to deal with tests of Markov switching in univariate and dynamic factor models. Verdinelli and Wasserman (1995) and Koop and Potter (1999) suggested a way to indirectly calculate the Bayes factor using the “Savage–Dickey” density ratio for nested models. Alternatively, Chib (1995) suggested a procedure for directly calculating the marginal likelihoods based on the Gibbs output, and Kass and Raftery (1995) provided a general discussion of Bayesian model comparison and the issues related to the calculation of Bayes factors. Kim and Nelson (2001) applied Chib’s (1995) procedure to test for a structural break in a Markov-switching model of the business cycle. In this article we use Chib’s (1995) procedure as implemented by Kim and Nelson (2001). Appendix B provides a description of how we use this procedure.

For model selection using Bayes factors, we adopt the following criteria, based on those of Jeffreys (1961) and Kass and Raftery (1995):

1.  $\ln(BF_{ij}) > 0$  implies that the evidence supports model  $i$ .
2.  $-1.15 < \ln(BF_{ij}) < 0$  corresponds to “very slight evidence” against model  $i$ , with the data up to three times more likely to have arisen from model  $j$ .
3.  $-2.30 < \ln(BF_{ij}) < -1.15$  corresponds to “slight evidence” against model  $i$ , with the data between 3 and 10 times more likely to have arisen from model  $j$ .
4.  $-4.61 < \ln(BF_{ij}) < -2.30$  corresponds to “strong evidence” against model  $i$ , with the data between 10 and 100 times more likely to have arisen from model  $j$ .
5.  $\ln(BF_{ij}) < -4.61$  corresponds to “very strong evidence” against model  $i$ , with the data more than 100 times more likely to have arisen from model  $j$ .

The  $\ln(BF_{ij})$  measure is calculated as the difference between the log marginal likelihoods for models  $i$  and  $j$ .

In addition to allowing for straightforward model comparison, the Bayesian approach also provides computationally feasible estimation of models that feature both Markov switching and structural breaks with unknown breakpoints. In particular, the hierarchical nature of the Gibbs sampler allows us to break down the problem of joint estimation into the more manageable separate problems of making inference about Markov switching conditional on the structural breaks and making inference about structural breaks conditional on the Markov switching.

### 5. EMPIRICAL ANALYSIS

The data used in this article are monthly excess returns (capital gains plus dividends) on the value-weighted portfolio of all NYSE-listed stocks over the yield on 1-month U.S. Treasury bills from the CRSP files. We take natural logarithms of the gross returns and multiply by 12 to obtain annualized continuously compounded returns. We consider the sample period 1926–1999.

Table 1 reports log marginal likelihoods for the four models presented in Section 2 under the three assumptions about the structural break discussed in Section 3. The first thing to notice about the results is the strong support for a Markov-switching specification of market volatility. For every structural break assumption, there is “very strong evidence” in favor of models II–IV (Markov-switching variance) over model I (constant variance). The implied log Bayes factors range from 39.24 for model III over model I under assumption (C) with two breaks for each model to 158.11 for model IV over model I under assumption (A) of no structural breaks. To put these numbers into some perspective, a log Bayes factor of 4.61 corresponds to “very strong evidence” in favor of one model over another according to the model selection criteria presented in Section 4.

Given the decisive evidence in favor of Markov-switching volatility over a constant volatility assumption, the next thing to notice about the results in Table 1 is that the evidence in favor of a time-varying equity premium depends crucially on whether volatility feedback is taken into account. Comparing model II (constant mean) and model III (no volatility feedback) shows no strong evidence in favor of a time-varying equity premium. If anything, the evidence is generally more supportive of constant equity premium (at least within structural regimes). The implied log Bayes factors range from .99 for model II over model III under assumption (B) with two breaks for each model to .26 for model III over model II under assumption (A) of no structural breaks. However, comparing model II (constant mean) with model IV (volatility feedback) gives “strong evidence” or “very strong evidence” for a time-varying equity premium in every case. The implied log Bayes factors range from 4.27 for model IV over model II under assumption (B) with two breaks for each model to 7.00 for model IV over model II under assumption (A) of no structural breaks. Meanwhile, the evidence favors the existence of volatility feedback.

Table 2. Posterior Moments for the “Most Preferred” Model

Parameter	Mean	Standard deviation	Median	90% posterior bands
$\alpha$	.079	.021	.080	(.042, .112)
$\beta$	.045	.043	.032	(.002, .132)
$\delta$	-.287	.145	-.278	(-.542, -.066)
$\sigma_{L,1}^2$	.269	.071	.257	(.189, .379)
$\sigma_{H,1}^2$	2.113	.453	2.046	(1.507, 2.950)
$\sigma_{L,2}^2$	.166	.015	.165	(.141, .191)
$\sigma_{H,2}^2$	.675	.202	.633	(.430, 1.058)
$q_1$	.948	.038	.956	(.889, .987)
$p_1$	.945	.037	.953	(.875, .990)
$q_2$	.959	.020	.962	(.920, .984)
$p_2$	.803	.088	.820	(.630, .918)
$\kappa$	.989	.008	.991	(.973, .998)

NOTE: Results are for model IV with structural break assumption (B) and one structural break.

Comparing models III and IV shows “very strong evidence” in favor of volatility feedback in every case. The implied log Bayes factors range from 4.75 for model IV over model III under assumption (C) with two breaks for each model to 6.74 for model IV over model III under assumption (A) of no structural breaks.

Table 2 reports posterior moments for model IV under assumption (B) with one structural break in the volatility process. The results make it clear why this is the “most preferred” model. In particular, although there is considerable uncertainty regarding the magnitude of the marginal price of volatility  $\beta$ , there is little doubt that the volatility feedback as captured by the parameter  $\delta$  is an important aspect of stock return dynamics. In particular, the 90% posterior bands for  $\delta$  (–.542, –.066) do not include 0. Meanwhile, the reduction in the level and persistence of high volatility episodes after the structural break (compare  $\sigma_{H,1}^2$  with  $\sigma_{H,2}^2$  and  $p_1$  with  $p_2$ ) explains the support for structural break assumption (B).

The strong evidence for a negative volatility feedback effect reflects the fact that volatility feedback is easier to identify than the marginal impact of a change in volatility on a single-period expected return because volatility feedback summarizes the present value impact of volatility on all expected future returns. As evidence of this, we note that when we consider an alternative basis for rejection sampling that only restricts the marginal impact parameter  $\beta$  to be the opposite sign of the volatility feedback parameter  $\delta$  (i.e.,  $\beta \cdot \delta \leq 0$ ), we find

Table 1. Model Comparison

Structural break assumption	Number of breaks	Log marginal likelihood			
		Model I: constant mean and variance	Model II: constant mean and switching variance	Model III: time-varying mean and switching variance	Model IV: time-varying mean and switching variance with volatility feedback
(A) No structural break	0	–886.12	–735.01	–734.75	–728.01
(B) Structural breaks in volatility process	1	–778.32	–732.31	–733.14	–727.27
	2	–776.72	–736.09	–737.08	–731.82
	3	–776.05			
	4	–778.96			
(C) Structural breaks in all parameters	1	–779.00	–733.36	–733.38	–728.28
	2	–777.84	–738.13	–738.60	–733.85
	3	–777.91			
	4	–782.00			

NOTE: For each model, additional breaks are considered until the last break is no longer preferred for either structural break assumption (B) or (C).

nearly identical results to when we restrict  $\beta$  to be nonnegative. The 90% posterior bands for  $\beta$  are (.002, .127), whereas the posterior bands for  $\delta$  are (-.540, -.068). However, the log marginal likelihood given this alternative restriction on the prior and posterior distributions is  $-727.63$ , which means that there is “very slight evidence” in favor of the nonnegativity constraint. We also consider the effect of relaxing the sign restrictions entirely. Again, the 90% posterior bands for the volatility feedback parameter (-.586, -.166) do not include 0. However, the posterior mean for the marginal price of risk parameter is negative (-.071) with 90% posterior bands of (-.224, .093). The log marginal likelihood is  $-727.12$  for the unrestricted model versus  $-727.27$  when the nonnegativity constraints are imposed. Thus, there is “very slight evidence” in favor of an unrestricted model according to marginal likelihood analysis. However, other considerations discussed at the end of this section when we present estimates of the equity premium lead us to impose the nonnegativity constraints as a strong prior.

Given our results, a reasonable question is whether other models of stock returns would be preferable to model IV. For example, we could consider more complicated univariate models that allow for a leverage effect and/or a third recurring volatility regime to capture the negative skewness and excess kurtosis of stock returns. Also, we could consider models with lagged returns or multivariate models to capture any apparent predictability in stock returns. However, Table 3 presents evidence suggesting that model IV captures many of the key features of stock return behavior with a relatively small number of parameters. First, we consider the sample skewness and kurtosis of the standardized residuals for the preferred specifications of each model under consideration. Not surprisingly, model I, with its iid normal assumption, fails miserably at capturing the distribution of stock returns, even though two structural breaks in variance are allowed. Models II and III do a much better job capturing the excess kurtosis of the raw data, but they do not appear to capture all of the negative skewness in returns. Model IV appears to explain most of the negative skewness and all of the excess kurtosis in the raw data. Second, we consider the modified Ljung–Box  $Q$ -statistic tests of residual serial correlation. As discussed by Kim et al. (2001), a positive relationship between the equity premium and the Markov-switching level of

market volatility implies a small degree of positive serial correlation in monthly returns. In particular, due to the persistence of the variance regimes, periods of above- (below-) average returns typically should be followed by more above- (below-) average returns. The results of the  $Q$ -tests confirm that model IV does a reasonable job explaining the apparent short-horizon predictability in excess returns. Kim et al. (2001) showed that model IV can also explain the apparent long-horizon return predictability (mean reversion in stock prices) reported by Fama and French (1988) and Poterba and Summers (1988). In light of these diagnostic test results, we argue that model IV is an appropriate model for conducting empirical analysis of structural changes in the distribution of stock returns.

Having established model IV as a reasonable empirical model of stock return behavior, we turn to the issue of structural breaks. Returning to Table 1, there is support for assumption (B) of at least one structural break in the volatility process for every model. Comparing assumption (B) with assumption (A) of no structural breaks, the implied log Bayes factors range from .74 for model IV with one break to 110.07 for model I with three breaks. Also, comparing assumption (B) with assumption (C) of structural breaks in all parameters (except break probabilities), the implied log Bayes factors range from .24 for model III with one break to 2.04 for model II with two breaks. Comparing the “most preferred” model with model IV under assumption (C) with one break, the implied log Bayes factor is 1.01. However, note that because the equity premium depends on the level of volatility for the “most preferred” model, the structural break in the volatility process still corresponds to a structural break in the equity premium. Meanwhile, there is “strong” or “very strong” evidence for only one break, except for model I. Comparing one break with two breaks for models II–IV, the implied log Bayes factors range from 3.78 for model II under assumption (B) to 5.57 for model IV under assumption (C). For model I, the evidence is strongest for three breaks. Comparing three breaks versus one break for model I under assumption (B), the implied log Bayes factor is 2.27. However, the support for multiple breaks for model I likely reflects the fact that, as with the model used by Pástor and Stambaugh (2001), the constant mean/constant variance specification does not allow for transitory changes in volatility. Thus the model may be prone to confusing transitory but persistent changes in volatility with structural breaks.

Just as important as the existence of a structural break in the equity premium is the timing and nature of the break. Figure 1 displays posterior distributions for the break dates for the “most preferred” model in panel (a) and for the preferred two-break model [model IV under assumption (B)] in panels (b) and (c). The posterior distributions in (a) and (b) suggest a break in the volatility process and, by implication, the equity premium sometime during the early 1940s. The corresponding cumulative distributions are very similar for the two models and are reasonably steep, implying a fair degree of certainty about the general timing of the structural break. In contrast, the cumulative distribution for the second break date for the preferred two-break model is reasonably flat, suggesting that the data are not particularly informative about a location for a second break date. Indeed, from the posterior distribution, the most probable second break date has a  $< 2\%$  probability. Given this finding, it is not surprising that the marginal likelihood analysis

Table 3. Diagnostics of Standardized Residuals for Preferred Model Specifications

	Model I	Model II	Model III	Model IV
Break assumption	(B)	(B)	(B)	(B)
Number of breaks	3	1	1	1
Skewness	-.50	-.17	-.16	-.09
Kurtosis	4.74	2.64	2.61	2.79
JB	148.71	8.77	9.17	2.88
	(<.01)	(.01)	(.03)	(.24)
Q(1)	5.18	1.41	1.50	.70
	(.02)	(.24)	(.22)	(.40)
Q(5)	11.03	9.61	10.51	8.50
	(.05)	(.09)	(.06)	(.13)
Q(24)	36.07	32.80	33.18	31.63
	(.05)	(.11)	(.10)	(.14)

NOTE: For each tests statistic, the  $p$ -value is reported in parentheses. JB stands for the Jarque-Bera test statistic for normality. The  $p$ -value is based on a chi-squared(2) distribution under normality.  $Q(k)$  stands for the Ljung–Box test statistic for serial correlation at up to  $k$  lags. The  $p$ -value is based on a chi-squared( $k$ ) distribution under the absence of serial correlation.

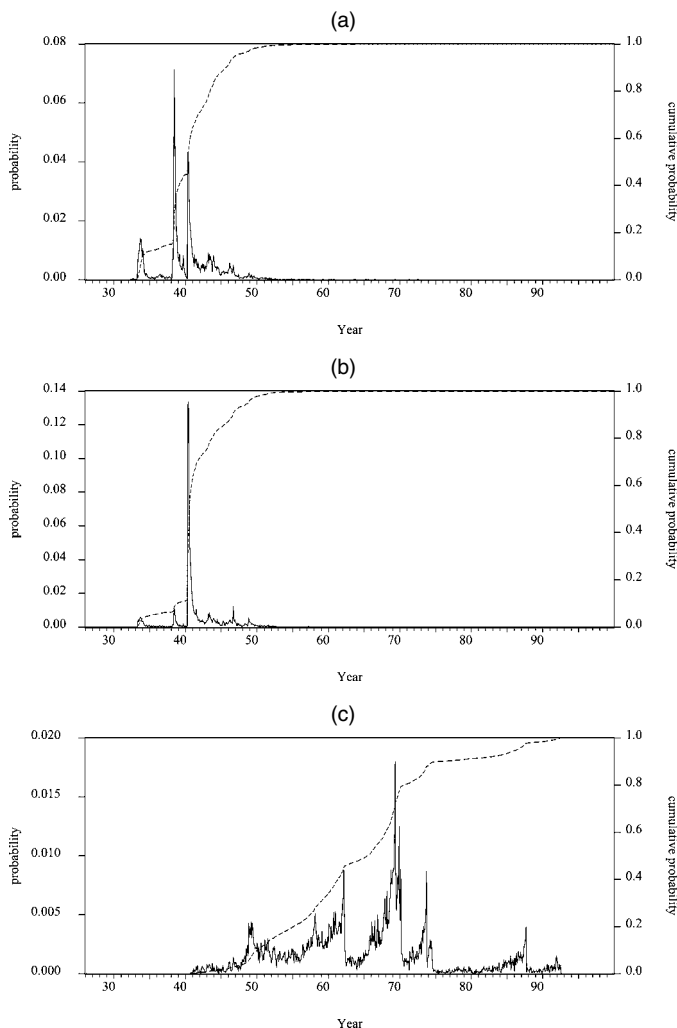


Figure 1. Posterior Distributions of Structural Breaks. The probability distribution (left scale) is the solid line and the cumulative probability distribution (right scale) is the dashed line. (a) The structural break in the “most preferred” model [model IV with structural break assumption (B) and one structural break]. (b) The first structural break in the preferred model given an assumption of two structural breaks [model IV with structural break assumption (B) and two structural breaks]. (c) The second structural break in the preferred model given an assumption of two structural breaks.

favors the one-break specification over the two-break specification for model IV. Comparing the “most preferred” model with the preferred two-break model, the implied log Bayes factor is 4.55. Also, as an additional check for a second structural break, we consider marginal likelihood analysis using the shorter sample period 1952–1999. The log marginal likelihood for model IV under assumption (A) of no structural break is  $-391.12$ , whereas the log marginal likelihood for model IV under assumption (B) with one structural break is  $-395.28$ . Thus the log Bayes factor in favor of no additional structural break in the postwar period is 4.16.

Figure 2 displays excess returns and estimates of the equity premium for the “most preferred” model. There is a clear reduction in the general level and persistence of volatility during the 1940s, which is reflected in the reduction in the unconditional equity premium and the persistence of the conditional equity premium over the same time period. Thus, as stocks

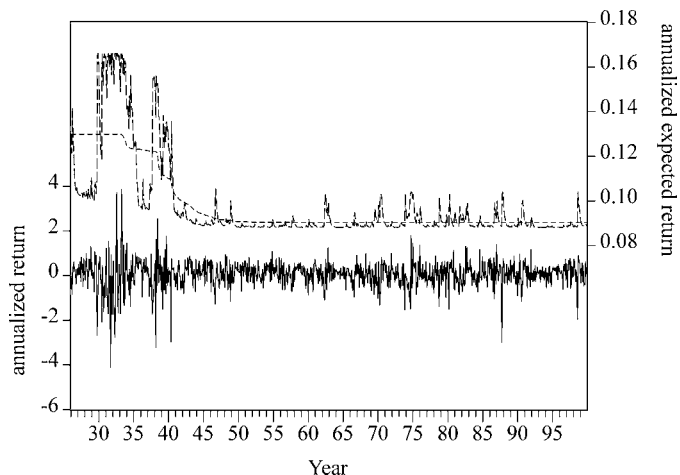


Figure 2. Excess Returns and the Equity Premium. Excess returns (left scale) are monthly value-weighted NYSE returns minus the yield on 1-month U.S. Treasury bills from CRSP. The measures of the equity premium (right scale) are from the “most preferred” model. The more volatile measure is the posterior mean of the equity premium conditional on the volatility regime. The smoother measure is the posterior mean of the unconditional equity premium.

became less risky after the structural break, the expected reward for investing in stocks declined as well. The reduced persistence of high-volatility episodes also suggests that volatility feedback may be less important in the postwar period, although the persistence of low-volatility regimes and the big negative returns in 1987 and 1998 suggest that it is still relevant. It is worth emphasizing, however, that the structural break in the equity premium appears to be driven entirely by the reduction in volatility (i.e., the “quantity of risk”), and not by a change in the risk preference parameters relating the equity premium to the level of volatility (i.e., the “price of risk”) or a change in volatility feedback. In particular, the marginal likelihood analysis favors constant risk preference parameters  $\alpha$  and  $\beta$  and the related volatility feedback parameter  $\delta$  across structural regimes. This result suggests that the prebreak data contain relevant information about the postbreak behavior of stock returns and the equity premium.

Finally, with the most preferred model specification in hand and an understanding of the nature and timing of the structural break, we turn to the equity premium itself. As displayed in Figure 2, the posterior mean of the unconditional equity premium  $\mu$  for the “most preferred” model is about 12% near the beginning of the sample and drops to about 9% for most of the postwar period. The posterior mean of the conditional equity premium jumps between 10% and 16% in the volatile prebreak period. In the less volatile postwar period, the posterior mean of the conditional equity premium jumps between 9% and 10%. There is a large amount of uncertainty surrounding these point estimates, however. Figure 3 displays the posterior mean of the unconditional equity premium for the “most preferred” model along with 90% posterior bands. The 90% posterior bands are wide at the beginning of the sample and include the possibility that the equity premium is as high as 20%. After the structural break, the 90% posterior bands converge to 6% and 12%, which includes the sample mean of 6.5%. We also report 90% posterior bands for an unrestricted version of model IV under assumption (B) with one structural break that does not impose any

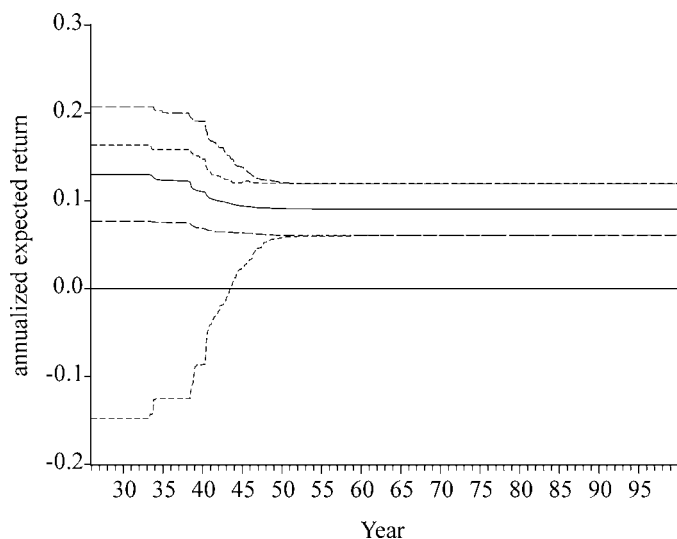


Figure 3. *Uncertainty About the Equity Premium.* The posterior mean and 90% posterior bands are for the “most preferred” model. The unrestricted posterior bands are for model IV with assumption (B) and one break estimated without nonnegativity constraints. For computational tractability, the posterior bands are based on every tenth draw from the Gibbs sampler (i.e., 1,000 highly independent draws). [— posterior mean; - - - - 90% lower; - - - - 90% upper; ····· 90% lower (unrestricted); ····· 90% upper (unrestricted).]

constraints on  $\alpha$  and  $\beta$ . The bands are much wider before the structural break and actually include negative values. We argue that this finding for the unrestricted model provides justification for our rejection sampling prior, because the rejection sampling produces results that are more in accordance with what we consider to be reasonable beliefs about the equity premium, namely that it is positive. However, we note that the rejection sampling has little impact on the 90% posterior bands after the structural break. Also, we note that the rejection sampling has no impact on the evidence in favor of a structural break in the volatility process. The log marginal likelihood for an unrestricted version of model IV under assumption (A) is  $-727.95$ , whereas the log marginal likelihood for an unrestricted version of model IV under assumption (B) with one break is  $-727.12$ . Therefore, the log Bayes factor in favor of a structural break of this form is .83, compared with .74 when the nonnegativity restrictions are imposed. Thus we conclude from our analysis that the postwar unconditional equity premium is likely somewhere between 6% and 12%, with 9% as our point estimate.

An interesting question is why our point estimate is so much higher than the sample mean for excess returns and other estimates reported in recent studies (e.g., Fama and French 2002). In terms of the sample mean, we note that allowing for heteroscedasticity in stock returns generally produces higher estimates of the equity premium, because less weight is put on higher variance returns in estimation. It is well known that stock returns and market volatility are negatively correlated (see, e.g., French, Schwert, and Stambaugh 1987). As discussed by Kim et al. (2004), volatility feedback provides an explanation for this negative correlation given a positive underlying relationship between the equity premium and market volatility. Thus, by accounting for the volatility feedback effect, our model implicitly puts less weight on highly volatile returns associated

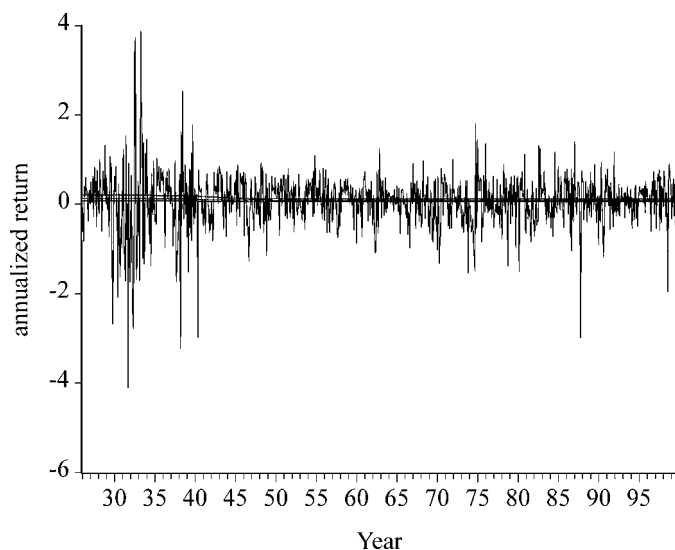


Figure 4. *The Equity Premium in Perspective.* Excess returns are monthly value-weighted NYSE returns minus the yield on 1-month U.S. Treasury bills from CRSP. The posterior mean of the unconditional equity premium and 90% posterior bands are for the “most preferred” model.

with changes in the volatility regime when estimating the equity premium. In terms of our higher estimate than the values reported by Fama and French (2002), we note that if any of the decline in the dividend yield or increase in the price/earnings ratio over the past 50 years is permanent, then their point estimates of 2.5% and 4.3% understate the true equity premium.

A related concern is that the high estimated value for the equity premium might somehow be implausible given other estimates and theoretical considerations. We argue that, on the contrary, our estimate is quite plausible. First, we note that unconditional sample standard deviation of excess returns is .653. Thus the classical 90% confidence bands for the equity premium based on the sample mean and standard deviation cover the wide range of 3%–10%. Therefore, there is little sense from a purely econometric perspective that a point estimate of 9% is implausible just because the sample mean is around 6.5% and other estimates are lower. Another way to see this point is to notice that the scales in Figure 2 are vastly different. Figure 4 plots excess returns, the posterior mean of the unconditional equity premium, and the 90% posterior bands all on the same scale. From this perspective, the equity premium appears essentially constant, and deciphering 6% from 12% is not particularly easy. Meanwhile, from a theoretical standpoint, the very possibility of structural change and the existence of transitory Markov-switching shifts in volatility could explain a high equity premium. As discussed by Wang (2001), shifts in volatility and the volatility process might generate a constant positive “jump-risk premium” making the unconditional equity premium higher than it would be given a stable iid normal distribution.

## 6. CONCLUSIONS

We find evidence of a structural break in the equity premium using data for excess returns on a value-weighted portfolio of NYSE stocks over the period 1926–1999. The break,



which likely occurred in the early 1940s, appears to be driven by a reduction in the general level and persistence of market volatility, not by a change in risk preferences. This finding suggests that although the overall distribution of excess returns has changed, there is still useful information about current behavior embedded in the prebreak data. The finding of a break in the equity premium arises from a model of stock returns that relates the equity premium to the Markov-switching level of volatility and allows for volatility feedback in the event of unanticipated changes in the volatility. The Bayesian model selection used in this article strongly favors this model over three more basic models of excess returns that either do not relate the equity premium to changes in the level of volatility or do not account for volatility feedback. Our analysis also suggests that our “most preferred” model captures most of the negative skewness, excess kurtosis, and short-horizon serial correlation in monthly stock returns. According to the model, the postbreak unconditional equity premium is about 9%. Finally, there is strong evidence against a second structural break during the postwar period, including during the 1990s. In particular, the large capital gains in the early and mid 1990s appear to be related to an unusually long period of low market volatility reminiscent of the 1950s and 1960s. Indeed, by the end of the sample in the late 1990s, it is clear that the extended period of low volatility was only temporary, rather than the result of a permanent change in the volatility process.

### ACKNOWLEDGMENTS

The authors have received helpful comments from Jeremy Piger, two anonymous referees, and participants in the CSSS seminar at the University of Washington and the “State-Space and Unobserved Components Models” academic colloquium in honour of James Durbin at the Royal Netherlands Academy of Arts and Sciences in Amsterdam. Responsibility for any errors is entirely their own. Support from the Ford and Louisa Van Voorhis Endowment at the University of Washington is gratefully acknowledged.

### APPENDIX A: THE GIBBS SAMPLING APPROACH

To make inferences about model IV under assumption (C) with two structural breaks in all parameters, we need marginal posterior distributions for the following:

$$\begin{aligned} \tilde{\mu} &= \{\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \delta_1, \delta_2, \delta_3\}'; \\ \tilde{\sigma}^2 &= \{\sigma_{L,1}^2, \sigma_{L,2}^2, \sigma_{L,3}^2, \sigma_{H,1}^2, \sigma_{H,2}^2, \sigma_{H,3}^2\}'; \\ \tilde{\mathbf{D}}_T &= \{D_1, \dots, D_T\}'; \quad \tilde{\mathbf{S}}_T = \{S_1, \dots, S_T\}'; \\ \tilde{\mathbf{S}}_T^e &= \{S_1^e, \dots, S_T^e\}', \quad \text{where} \\ S_t^e &\equiv E[S_t | \tilde{\mathbf{Y}}_{t-1}], \quad \text{with } \tilde{\mathbf{Y}}_t = \{r_1, \dots, r_t\}'; \\ \tilde{\mathbf{p}} &= \{q_1, q_2, q_3, p_1, p_2, p_3\}'; \end{aligned}$$

and

$$\tilde{\kappa} = \{\kappa_1, \kappa_2\}'.$$

In principle, the marginal posterior distribution of a set of parameters can be obtained by integrating the other parameters out of the joint posterior distribution

$$p(\tilde{\mu}, \tilde{\sigma}^2, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa} | \tilde{\mathbf{Y}}_T).$$

However, the hierarchical nature of the model allows us to use Gibbs sampling to obtain the marginal posterior distributions of interest. This is done by successively sampling from the full conditional densities. The Gibbs sampling procedure involves the following steps:

1. Generate  $\tilde{\mathbf{S}}_T$  and  $\tilde{\mathbf{S}}_T^e$  from  $p(\tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e | \tilde{\mu}, \tilde{\sigma}^2, \tilde{\mathbf{D}}_T, \tilde{\mathbf{p}}, \tilde{\mathbf{Y}}_T)$ , where, conditional on  $\tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T$  and  $\tilde{\mathbf{S}}_T^e$  are independent of  $\tilde{\kappa}$ .
2. Generate  $\tilde{\mathbf{D}}_T$  from  $p(\tilde{\mathbf{D}}_T | \tilde{\mu}, \tilde{\sigma}^2, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\kappa}, \tilde{\mathbf{Y}}_T)$ , where, conditional on  $\tilde{\mathbf{S}}_T, \tilde{\mathbf{D}}_T$  is independent of  $\tilde{\mathbf{p}}$ .
3. Generate  $\tilde{\mathbf{p}}$  from  $p(\tilde{\mathbf{p}} | \tilde{\mathbf{S}}_T)$ , where, conditional on  $\tilde{\mathbf{S}}_T, \tilde{\mathbf{p}}$  is independent of the other parameters.
4. Generate  $\tilde{\kappa}$  from  $p(\tilde{\kappa} | \tilde{\mathbf{D}}_T)$ , where, conditional on  $\tilde{\mathbf{D}}_T, \tilde{\kappa}$  is independent of the other parameters.
5. Generate  $\tilde{\mu}$  from  $p(\tilde{\mu} | \tilde{\sigma}^2, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{Y}}_T)$ , where, conditional on  $\tilde{\mathbf{D}}_T$  and  $\tilde{\mathbf{S}}_T, \tilde{\mu}$  is independent of  $\tilde{\mathbf{p}}$  and  $\tilde{\kappa}$ .
6. Generate  $\tilde{\sigma}^2$  from  $p(\tilde{\sigma}^2 | \tilde{\mu}, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{Y}}_T)$ , where, conditional on  $\tilde{\mathbf{D}}_T$  and  $\tilde{\mathbf{S}}_T, \tilde{\sigma}^2$  is independent of  $\tilde{\mathbf{p}}$  and  $\tilde{\kappa}$ .

This procedure is a straightforward extension of Albert and Chib’s (1993) Bayes inference via Gibbs sampling for an autoregressive time series subject to Markov-switching mean and variance. Kim and Nelson (1999b) extended the procedure to incorporate a one-time structural break in the Markov-switching parameters. Note that as a byproduct of generating  $\tilde{\mathbf{D}}_T$ , we can get the marginal posterior distribution of the break dates  $\tau_1$  and  $\tau_2$ .

The prior distributions are given as follows:

*Parameters subject to structural breaks*

$$\begin{aligned} \alpha_j &\sim N(0, .04)_{[\alpha \geq 0]}, & \beta_j &\sim N(0, .04)_{[\beta \geq 0]}, \\ \delta_j &\sim N(0, .04), \\ \frac{1}{\sigma_{L,j}^2} &\sim \Gamma\left(\frac{4}{2}, \frac{1}{2}\right), & \frac{1}{\sigma_{H,j}^2} &\sim \Gamma\left(\frac{2}{2}, \frac{1}{2}\right)_{[\sigma_H^2 > \sigma_L^2]}, \\ q_j &\sim \text{beta}(4, 1), & p_j &\sim \text{beta}(4, 1), \quad j = 1, 2, 3, \end{aligned}$$

and

*Structural break parameters*

$$\kappa_i \sim \text{beta}(20, .1), \quad i = 1, 2.$$

Here the subscript  $[\cdot]$  function denotes a truncated distribution obtained via rejection sampling. Additional rejection sampling occurs unless simulated structural regimes last at least 88 months, corresponding to 10% of the sample. All inferences are based on the last 10,000 of 12,000 Gibbs simulations (i.e., the first 2,000 simulations are discarded). Sensitivity analysis suggests that the qualitative results are robust with respect to a wide range of priors. We also examine convergence of the Gibbs sampler by trying different starting values for the parameters and by comparing reported inferences to those from an additional 10,000 draws. We find that estimates are robust to the reported number of decimal places. We also note that the serial correlation of draws from the Gibbs sampler dies out monotonically, with simulated values more than 10 draws apart having less than 5% correlation.

## APPENDIX B: CALCULATING THE MARGINAL LIKELIHOOD

Define  $\theta = \{\tilde{\mu}', \tilde{\sigma}^{2'}, \tilde{\mathbf{p}}', \tilde{\kappa}'\}'$  to be a vector of the model parameters. Then, following Chib (1995), we can write the marginal density of  $\tilde{\mathbf{Y}}_T$ , by virtue of being the normalizing constant of the posterior density, as

$$m(\tilde{\mathbf{Y}}_T) = \frac{f(\tilde{\mathbf{Y}}_T|\theta)\pi(\theta)}{\pi(\theta|\tilde{\mathbf{Y}}_T)},$$

where the numerator is the product of the sampling density and the prior, with all integrating constants included, and the denominator is the posterior density of  $\theta$ . Because the foregoing identity holds for any  $\theta$ , we may evaluate  $m(\tilde{\mathbf{Y}}_T)$  at the posterior mean  $\theta^*$ . Taking the logarithm of the foregoing equation for computational convenience, we have

$$\ln m(\tilde{\mathbf{Y}}_T) = \ln f(\tilde{\mathbf{Y}}_T|\theta^*) + \ln \pi(\theta^*) - \ln \pi(\theta^*|\tilde{\mathbf{Y}}_T).$$

The log-likelihood function and the log of the prior density at  $\theta = \theta^*$  can be evaluated relatively easily. First, the log-likelihood is given by

$$\ln f(\tilde{\mathbf{Y}}_T|\theta^*) = \sum_{t=1}^T \ln \left( \sum_{S_t=0}^1 \sum_{D_t=0}^1 p(S_t, D_t|\tilde{\mathbf{Y}}_{t-1}, \theta^*) \times f(r_t|\tilde{\mathbf{Y}}_{t-1}, S_t, D_t, \theta^*) \right).$$

Second, the log of the prior density is given by

$$\ln \pi(\theta^*) = \ln \pi(\tilde{\mu}^*) + \ln \pi(\tilde{\sigma}^{2*}) + \ln \pi(\tilde{\mathbf{p}}^*, \tilde{\kappa}^*),$$

where it is assumed that the parameter subsets are independent of each other.

For evaluating the posterior density at  $\theta = \theta^*$ , we consider the following decomposition of the posterior density:

$$\pi(\theta^*|\tilde{\mathbf{Y}}_T) = \pi(\tilde{\mu}^*|\tilde{\mathbf{Y}}_T)\pi(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\mathbf{Y}}_T)\pi(\tilde{\mathbf{p}}^*, \tilde{\kappa}^*|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T),$$

where

$$\begin{aligned} \pi(\tilde{\mu}^*|\tilde{\mathbf{Y}}_T) &= \int \pi(\tilde{\mu}^*|\tilde{\sigma}^2, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa}, \tilde{\mathbf{Y}}_T) \\ &\quad \times \pi(\tilde{\sigma}^2, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mathbf{Y}}_T) d\tilde{\sigma}^{2*} d\tilde{\mathbf{D}}_T d\tilde{\mathbf{S}}_T d\tilde{\mathbf{S}}_T^e d\tilde{\mathbf{p}} d\tilde{\kappa}, \end{aligned}$$

$$\pi(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\mathbf{Y}}_T)$$

$$\begin{aligned} &= \int \pi(\tilde{\sigma}^2|\tilde{\mu}^*, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa}, \tilde{\mathbf{Y}}_T) \\ &\quad \times \pi(\tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\mathbf{Y}}_T) d\tilde{\mathbf{D}}_T d\tilde{\mathbf{S}}_T d\tilde{\mathbf{S}}_T^e d\tilde{\mathbf{p}} d\tilde{\kappa}, \end{aligned}$$

and

$$\begin{aligned} \pi(\tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T) &= \int \pi(\tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{Y}}_T) \\ &\quad \times \pi(\tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T) d\tilde{\mathbf{D}}_T d\tilde{\mathbf{S}}_T d\tilde{\mathbf{S}}_T^e. \end{aligned}$$

This decomposition of the posterior density suggests that  $\pi(\tilde{\mu}^*|\tilde{\mathbf{Y}}_T)$  can be calculated based on draws from the full Gibbs

run and  $\pi(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\mathbf{Y}}_T)$  and  $\pi(\tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T)$  can be calculated based on draws from reduced Gibbs runs,

$$\begin{aligned} \hat{\pi}(\tilde{\mu}^*|\tilde{\mathbf{Y}}_T) &= \frac{1}{G} \sum_{g=1}^G \pi(\tilde{\mu}^*|\tilde{\sigma}^{2g}, \tilde{\mathbf{D}}_T^g, \tilde{\mathbf{S}}_T^g, \tilde{\mathbf{S}}_T^{eg}, \tilde{\mathbf{p}}^g, \tilde{\kappa}^g, \tilde{\mathbf{Y}}_T), \\ \hat{\pi}(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\mathbf{Y}}_T) &= \frac{1}{G} \sum_{g_2=1}^G \pi(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\mathbf{D}}_T^{g_2}, \tilde{\mathbf{S}}_T^{g_2}, \tilde{\mathbf{S}}_T^{eg_2}, \tilde{\mathbf{p}}^{g_2}, \tilde{\kappa}^{g_2}, \tilde{\mathbf{Y}}_T), \end{aligned}$$

and

$$\begin{aligned} \hat{\pi}(\tilde{\mathbf{p}}^*, \tilde{\kappa}^*|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T) &= \frac{1}{G} \sum_{g_3=1}^G \pi(\tilde{\mathbf{p}}^*, \tilde{\kappa}^*|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{D}}_T^{g_3}, \tilde{\mathbf{S}}_T^{g_3}, \tilde{\mathbf{S}}_T^{eg_3}, \tilde{\mathbf{Y}}_T), \end{aligned}$$

where the superscript “g” refers to the gth draw from the full Gibbs run and the superscript “g<sub>i</sub>,”  $i = 2, 3$ , refers to the g<sub>i</sub>th draw from the appropriate reduced Gibbs runs. Thus, apart from the usual  $G$  iterations for the full Gibbs run, we need additional  $2 \times G$  iterations for the appropriate reduced Gibbs runs. For example, to calculate  $\pi(\tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{Y}}_T)$ , we need output from an additional  $G$  iterations for the following reduced Gibbs run:

1. Generate  $\tilde{\mathbf{p}}$  and  $\tilde{\kappa}$  from  $p(\tilde{\mathbf{p}}, \tilde{\kappa}|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{D}}_T, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{Y}}_T)$ .
2. Generate  $\tilde{\mathbf{D}}_T$  from  $p(\tilde{\mathbf{D}}_T|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e, \tilde{\mathbf{p}}, \tilde{\kappa}, \tilde{\mathbf{Y}}_T)$ .
3. Generate  $\tilde{\mathbf{S}}_T$  and  $\tilde{\mathbf{S}}_T^e$  from  $p(\tilde{\mathbf{S}}_T, \tilde{\mathbf{S}}_T^e|\tilde{\mu}^*, \tilde{\sigma}^{2*}, \tilde{\mathbf{D}}_T, \tilde{\mathbf{p}}, \tilde{\kappa}, \tilde{\mathbf{Y}}_T)$ .

Note that throughout the reduced Gibbs run,  $\tilde{\mu}$  and  $\tilde{\sigma}^{2*}$  are not generated, but are set equal to their posterior means  $\tilde{\mu}^*$  and  $\tilde{\sigma}^{2*}$ .

[Received November 2002. Revised June 2004.]

## REFERENCES

- Albert, J. H., and Chib, S. (1993), “Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts,” *Journal of Business & Economic Statistics*, 11, 1–15.
- Carlin, B. P., and Chib, S. (1995), “Bayesian Model Choice via Markov Chain Monte Carlo Methods,” *Journal of the Royal Statistical Society*, 57, 473–484.
- Carlin, B. P., and Polson, N. G. (1991), “Inference for Nonconjugate Bayesian Models Using the Gibbs Sampler,” *Canadian Journal of Statistics*, 19, 399–405.
- Chib, S. (1995), “Marginal Likelihood From the Gibbs Output,” *Journal of the American Statistical Association*, 90, 1313–1321.
- (1998), “Estimation and Comparison of Multiple Change-Point Models,” *Journal of Econometrics*, 86, 221–241.
- Fama, E. F., and French, K. R. (1988), “Permanent and Temporary Components of Stock Prices,” *Journal of Political Economy*, 96, 246–273.
- (2002), “The Equity Premium,” *Journal of Finance*, 57, 637–659.
- French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987), “Expected Stock Returns and Volatility,” *Journal of Financial Economics*, 19, 3–29.
- Garcia, R. (1998), “Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models,” *International Economic Review*, 39, 763–788.
- George, E. I., and McCulloch, R. E. (1993), “Variable Selection via Gibbs Sampling,” *Journal of the American Statistical Association*, 88, 881–889.
- Geweke, J. (1996), “Variable Selection and Model Comparison in Regression,” in *Bayesian Statistics 5*, eds. J. O. Berger, J. M. Bernardo, A. P. Dawid, and A. F. M. Smith, Oxford, U.K.: Oxford University Press, pp. 609–620.
- Hansen, B. E. (1992), “The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov-Switching Model of GNP,” *Journal of Applied Econometrics*, 7, S61–S82.
- Jeffreys, H. (1961), *Theory of Probability* (3rd ed.), Oxford, U.K.: Clarendon Press.

- Kass, R. E., and Raftery, A. E. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773–795.
- Kim, C.-J., Morley, J. C., and Nelson, C. R. (2001), "Does an Intertemporal Trade-off Between Risk and Return Explain Mean Reversion in Stock Prices?" *Journal of Empirical Finance*, 8, 403–426.
- (2004), "Is There a Positive Relationship Between Stock Market Volatility and the Equity Premium?" *Journal of Money, Credit, and Banking*, 36, 336–360.
- Kim, C.-J., and Nelson, C. R. (1999a), *State-Space Models With Regime Switching: Classical and Gibbs-Sampling Approaches With Applications*, Cambridge, MA: MIT Press.
- (1999b), "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *Review of Economics and Statistics*, 81, 608–616.
- (2001), "A Bayesian Approach to Testing for Markov Switching in Univariate and Dynamic Factor Models," *International Economic Review*, 42, 989–1013.
- Koop, G., and Potter, S. M. (1999), "Bayes Factors and Nonlinearity: Evidence From Economic Time Series," *Journal of Econometrics*, 88, 251–281.
- Merton, R. C. (1980), "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Finance*, 8, 323–361.
- Pástor, L., and Stambaugh, R. F. (2001), "The Equity Premium and Structural Breaks," *Journal of Finance*, 56, 1207–1239.
- Poterba, J. M., and Summers, L. H. (1988), "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22, 27–59.
- Turner, C. M., Startz, R., and Nelson, C. R. (1989), "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market," *Journal of Financial Economics*, 25, 3–22.
- Verdinelli, I., and Wasserman, L. (1995), "Computing Bayes Factors Using a Generalization of the Savage–Dickey Density Ratio," *Journal of the American Statistical Association*, 90, 614–618.
- Wang, Z. (2001), Discussion of "The Equity Premium and Structural Breaks," by L. Pástor and R. F. Stambaugh, *Journal of Finance*, 56, 1240–1246.