ELSEVIER

Journal of Empirical Finance 5 (1998) 131–154

Journal of EMPIRICAL FINANCE

Testing for mean reversion in heteroskedastic data based on Gibbs-sampling-augmented randomization [1]

Chang-Jin Kim [a,2], Charles R. Nelson [b,*], Richard Startz [b,3]

[a] *Department of Economics, Korea University, Anam-dong, Seongbuk-ku, Seoul 136-701, South Korea*
[b] *Department of Economics, University of Washington, Box 353330, Seattle, WA 98195, USA*

Abstract

Previous work reported that heteroskedasticity did not affect the sampling distribution of the variance ratio, or had assumed that the investigator knew a priori the pattern of heteroskedasticity. This paper uses the Gibbs sampling approach in the context of a three state Markov-switching model to show how heteroskedasticity affects inference and suggest two strategies for valid inference. We find that test procedures which ignore the pattern of heteroskedasticity are biased, rejecting the null hypothesis of no mean reversion too often. We present a resampling strategy that standardizes historical returns, using the Gibbs sampling approach to allow for uncertainty in parameters and states while conditioning on the information in the data. A second strategy is to estimate the VR from standardized data. Again, Gibbs sampling makes appropriate use of the information in the data. For CRSP stock returns 1926–86 we find that evidence of mean reversion is substantially weakened. Gibbs sampling estimates of VRs are closer to unity, and the pattern shifts to shorter lags. There is weak evidence of mean reversion in returns on equal-weighted portfolios and essentially none for the value-weighted. Thus, when returns from high variance periods are

* Corresponding author. E-mail: cnelson@u.washington.edu.

[1] C.-J.K. acknowledges assistance from NON-DIRECTED RESEARCH FUND, Korea Research Foundation, 1996.

[2] E-mail: cjkim@kuccnx.korea.ac.kr.
[3] E-mail: startz@u.washington.edu.

PII S0927-5398(97)00015-7

appropriately standardized, the sample evidence for mean reversion changes in pattern and significance.

Keywords: Gibbs sampling; Randomization; Heteroskedasticity; Mean reversion; Variance ratio

1. Introduction

The variance ratio (VR) statistic was introduced by Cochrane (1988) to measure the relative importance of the random walk component of GDP. In using the VR to test the null hypothesis that a series is a random walk against the alternative that it is mean reverting, one makes use of the fact that the variance of the K-period difference of a random walk is simply proportional to K. In the asset return context, let V_t denote the log of the market value of a risky portfolio so that the K-period difference is the K-period return

$$R_t^K = V_t - V_{t-K}. \tag{1}$$

If V_t is a random walk, or, equivalently, one period returns are serially random, then we have

$$V_t - V_{t-1} = u_t; \quad u_t \sim iid(\mu, \sigma^2) \tag{2}$$

and since the K-period return is the accumulation of K successive u_t,

$$\mathrm{Var}(R_t^K) = K\sigma^2. \tag{3}$$

The VR statistic is defined in the asset context as

$$VR(K) = \frac{\mathrm{Var}(R^K)}{\mathrm{Var}(R^1)}\frac{1}{K} \tag{4}$$

which will be unity under the random walk hypothesis.

If the series exhibits mean reversion, so that changes in either direction tend to be offset over time by movement back toward the starting point, then $\mathrm{Var}(R^K)$ will be less than K times as large as $\mathrm{Var}(R^1)$, so the VR will be less than unity. One explanation for mean reversion would be the presence of a transitory component in asset prices. To judge whether a sample VR is significantly below unity one needs to know the sampling distribution of the VR under the null hypothesis.

Poterba and Summers (1988), hereafter PS, and Lo and MacKinlay (1989), used the VR to test for mean reversion in stock prices and concluded that a transitory component accounted for a substantial fraction of the variance in stock returns over horizons of several years. Inference was based on Monte Carlo simulation of the sampling distribution of the VR under the null hypothesis of serially random returns. Recognizing the well-documented heteroskedasticity in stock returns over their sample period, PS compared sampling distributions of the VR for homoskedastic and heteroskedastic data generating processes and found no meaningful difference. Their data generating process was intended to preserve the persistence of heteroskedasticity that is observed in historical returns but not the specific pattern. Kim et al. (1991), hereafter KNS, estimated the sampling distribution of the VR by randomizing actual returns and also suggested a 'stratified randomization' that preserved the historical pattern of high and low volatility periods. The fact that the latter revealed substantially weaker evidence of mean reversion than the former suggests that the specific pattern of heteroskedasticity in the sample period may play an important role in inference. However, the approach of KNS assumes that the econometrician has certain knowledge of the pattern, and does not exploit any information from the pattern of heteroskedasticity in the estimation of the VR. Further, resampling of returns is limited by segregation into sub-periods according to volatility.

This paper suggests two changes in the way we interpret the VR in the presence of heteroskedasticity. One is to utilize fully the information in the data about the pattern of heteroskedasticity in simulating the sampling distribution of the VR without pretending to have prior knowledge of the pattern. The second suggests a modification of the VR statistic to make more efficient use of the information in the data about mean reversion by weighting observations appropriately based on the information in the data about the timing and magnitude of volatility changes. The model is a three state Markov-switching process estimated by an extension of the Bayesian Gibbs sampling approach of Albert and Chib (1993) in which the parameters as well as the unobserved states are viewed as random variables for which we obtain conditional distributions given the data.

The format of the paper is as follows. Section 2 presents the three state Markov-switching model and discusses the Gibbs sampling approach to estimation. Results for monthly stock returns 1926–86 suggest that the model adequately captures the persistence structure of volatility. In Section 3 we report a sequence of Monte Carlo experiments designed to investigate the influence of heteroskedasticity, its persistence, and specific pattern on the sampling distribution of the VR based on estimates of the parameters from Section 2. Motivated by the finding that it is important to condition on information in the data about the specific pattern of heteroskedasticity, Section 4 suggests two ways to incorporate that finding. In Section 4.1 we present a strategy for estimating the sampling distribution of the VR that randomizes appropriately standardized returns using Gibbs sampling to capture uncertainty about the parameters and the states, conditioning on the information in the data. In Section 4.2 we propose a modified VR computed from appropriately standardized returns, again using the Gibbs sampling approach to both incorporate uncertainty and preserve the information contained in the data. Section 5 presents empirical results for the data set studied by PS and KNS. Section 6 concludes the paper.

2. A Markov-switching model of heteroskedasticity for stock returns and Gibbs sampling

Fig. 1A and B show that stock returns tend to exhibit nonnormal unconditional sampling distributions, in the form of skewness and excess kurtosis, a fact known at least since Fama (1963) and Mandlebrot (1963). The pronounced peak and heavy tails in the distribution of stock returns, as mentioned in Turner et al. (1989), are typical of unconditional densities of normal observations subject to heteroskedasticity.

The specification that has been most commonly used in the study of stock return volatility or heteroskedasticity is the GARCH model developed by Engle (1982) and Bollerslev (1986), as extensively surveyed in Bollerslev et al. (1992).

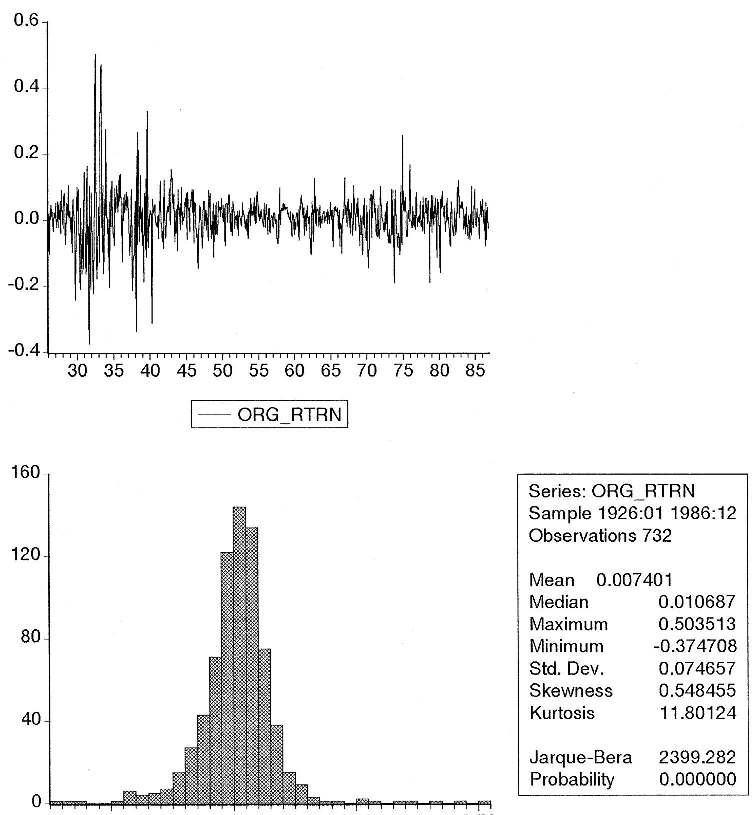

Fig. 1. (A) Plot of historical returns: monthly excess returns, equal-weighted portfolio, 1926–1986. (B) Distribution of historical returns: monthly excess returns, equal–weighted portfolio, 1926–1986.

An alternative approach to modeling stock returns would be to assume that the return is drawn from a mixture of normal densities.

Recently, Hamilton and Susmel (1994) propose a SWARCH (Switching ARCH) model in which they allow parameters of an ARCH process to come from one of several different regimes. While the long run dynamics are governed by regime shifts in unconditional variance according to a first order Markov-switching process, the short run dynamics within a regime are governed by an ARCH process. They apply the SWARCH model to weekly stock returns, and they find that ARCH effects die out almost completely after a month.[4] This suggests that volatility or heteroskedasticity in monthly stock returns may be modeled as a pure Markov-switching variance model, as in Turner et al. (1989). This leads us to consider the following three state Markov-switching model of monthly stock returns:

$$y_t \sim N(0, \sigma_t^2), \tag{5}$$

$$\sigma_t^2 = \sigma_1^2 S_{1t} + \sigma_2^2 S_{2t} + \sigma_3^2 S_{3t}, \tag{6}$$

$$S_{kt} = 1 \text{ if } S_t = k, \text{ and } S_{kt} = 0, \quad \text{otherwise; } k = 1, 2, 3 \tag{7}$$

$$\Pr[S_t = j | S_{t-1} = i] = p_{ij}, \quad i, j = 1, 2, 3 \tag{8}$$

$$\sum_{j=1}^{3} p_{ij} = 1 \tag{9}$$

$$\sigma_1^2 < \sigma_2^2 < \sigma_3^2 \tag{10}$$

where y_t is demeaned stock return, S_t is an unobserved state variable which evolves according to a first order Markov process with transition probabilities in Eq. (8).

The above model could be estimated using the maximum likelihood estimation method of Hamilton (1989) and Hamilton and Susmel (1994) or the EM algorithm of Hamilton (1990). In this paper, we employ an extended version of Albert and Chib's (1993) Bayesian Gibbs sampling approach to estimate the model. Gibbs sampling[5] is an iterative Monte Carlo technique that generates a simulated sample

---

[4] A simplified version of SWARCH model estimated by Hamilton and Susmel (1994) is given by: $y_t = \sigma_{s_t} u_t$, $u_t = h_t v_t$, $v_t \sim$ i.i.d. with Student t Distribution, $h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \beta d_{t-1} u_{t-1}^2$, where σ_{s_t} captures Markov-switching variances and the d_{t-1} is a dummy variable introduced to capture leverage effects. In the above specification, an ARCH(2) process is incorporated within a given volatility regime. Their estimates of the model using weekly stock returns (1962–1987) show that $\hat{\lambda} = \hat{\alpha}_1 + \hat{\alpha}_2 = 0.48$. Note that $\hat{\lambda}^4 = 0.05$, and this means that the volatility effects captured by u_t or by ARCH effects die out almost completely after a month, suggesting that no ARCH term may be necessary in modeling monthly stock returns.

[5] For a general introduction to Gibbs sampling, readers are referred to Geman and Geman (1984), Gelfand and Smith (1990), and Gelfand et al. (1990).

from the joint distribution of a set of random variables by generating successive samples from their conditional distributions. In the Gibbs sampling approach, all the parameters of the model are treated as random variables with appropriate prior distributions. Thus, random variables to be drawn in the present case are given by $\tilde{S}=\{S_t,\ t=1,\ 2,\ \ldots,\ T\}$, σ_1^2, σ_2^2, σ_3^2, and $\tilde{p}=\{p_{11},\ p_{12},\ p_{21},\ p_{22},\ p_{31},\ p_{32}\}$. Starting from arbitrary initial values of the parameters, Gibbs sampling proceeds by taking:

Step 1: a drawing from the conditional distribution of $\tilde{S}$ given the data, σ_1^2, σ_2^2, σ_3^2, and $\tilde{p}$; then

Step 2: a drawing from the conditional distribution of σ_1^2 given the data, $\tilde{S}$, σ_2^2, σ_3^2, and $\tilde{p}$; then

Step 3: a drawing from the conditional distribution of σ_2^2 given the data, $\tilde{S}$, σ_1^2, σ_3^2, and $\tilde{p}$; then

Step 4: a drawing from the conditional distribution of σ_3^2 given the data, $\tilde{S}$, σ_1^2, σ_2^2, and $\tilde{p}$; and then

Step 5: a drawing from the conditional distribution of $\tilde{p}$ given the data, $\tilde{S}$, σ_1^2, σ_2^2, and σ_3^2.

By successive iteration, the procedure simulates a drawing from the joint distribution of all the state variables and parameters of the model, given the data. It is straightforward then to summarize the marginal distributions of any of these, given the data. For more details of the Gibbs sampling approach to a 3-state Markov-switching variance model given above, readers are referred to the Appendix A.

Advantages of the Gibbs sampling approach over the maximum likelihood method are discussed in detail in Albert and Chib (1993). Among others, the Gibbs sampling approach provides us with a way to deal with uncertainty associated with underlying parameters and unknown states of the model. As we will see in later sections, an appropriate treatment of heteroskedasticity in the VR tests of mean reversion is associated with standardizing the historical returns. But the variance of stock returns at each point in time has to be estimated and the parameters that govern the dynamics of the stock return variance act like nuisance parameters in the estimation of the empirical distribution of the VR statistic. Employing the Gibbs sampling approach in the standardization step of the new tests to be introduced in this paper is equivalent to integrating out these nuisance parameters.

Data that we use in this section are monthly excess returns calculated from monthly total returns on all NYSE stocks from the CRSP files for equal-weighted portfolios and the one-month Treasury-bill rate from Ibbotson Associates for the sample period 1926.01–1986.12. To make inferences from the model, Gibbs sampling is run such that 1000 draws are discarded and the next 10,000 are recorded. We employ diffuse or non-informative priors for all the parameters of the model. For details of the priors employed, readers are referred to Appendix A. Table 1 presents the Bayesian posterior distributions of the parameters that result

Table 1
Bayesian inferences on parameter estimates: A three-state Markov-switching model of heteroskedasticity for CRSP excess returns, equal-weighted portfolio, 1926–1986

Parameter	Posterior			
	mean	SD	MD	95% interval
p_{11}	0.9700	0.0204	0.9746	(0.9164, 0.9954)
p_{12}	0.0271	0.0199	0.0225	(0.0020, 0.0791)
p_{21}	0.0302	0.0233	0.0250	(0.0035, 0.0878)
p_{22}	0.9598	0.0248	0.9649	(0.8995, 0.9907)
p_{31}	0.0116	0.0139	0.0066	(0.0001, 0.0510)
p_{32}	0.0250	0.0242	0.0183	(0.0005, 0.0879)
σ_1^2	0.0013	0.0002	0.0013	(0.0009, 0.0017)
σ_2^2	0.0039	0.0007	0.0038	(0.0028, 0.0054)
σ_3^2	0.0258	0.0044	0.0253	(0.0187, 0.0351)

Non-informative priors were given for all the parameters of the model; SD and MD refer to standard deviation and median, respectively.

from Gibbs sampling. At the end of each run of Gibbs sampling, we have a simulated set of $\{S_t,\ t=1,\ 2,\ \dots,\ T\}$, and thus, of $\{S_{kt},\ t=1,\ 2,\ \dots,\ T,\ k=1,\ 2,\ 3\}$, σ_k^2, $k=1,\ 2,\ 3$, and $\tilde{p}$. Then using these particular realizations of the states and the parameters for each run of Gibbs sampling, we can calculate σ_t^2 for $t=1,\ 2,\ \dots,\ T$ using Eq. (6). Using these, we can also standardize returns by $y_t^* = (y_t/\sigma_t)$. Thus, when all the iterations are over, we have 10,000 sets of realized return variances $\tilde{\sigma}^2 = \{\sigma_t^2,\ t=1,\ 2,\ \dots,\ T\}$ and 10,000 sets of realized standardized returns $\tilde{y}^* = \{y_t^*\ t=1,\ 2,\ \dots,\ T\}$, each of which is associated with different realizations of the parameters and the state variable. Fig. 2 plots the

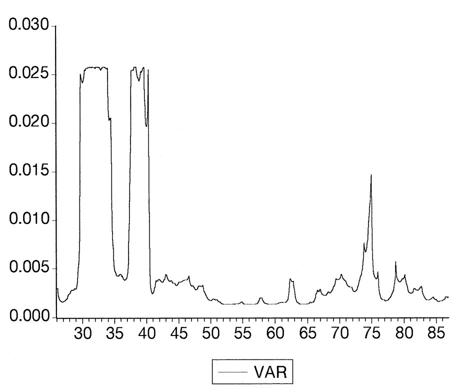

Fig. 2. Estimated variance of historical returns (3-State Markov- switching, Gibbs sampling): monthly excess returns, equal-weighted portfolio, 1926–1986.

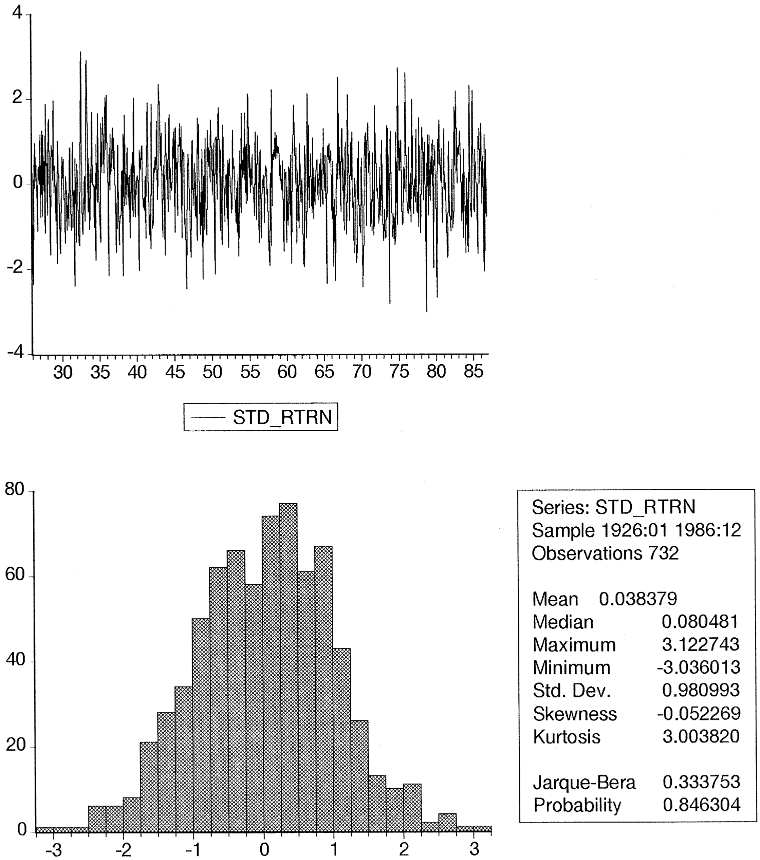

Fig. 3. (A) Plot of standardized returns (average of 1000 realizations from Gibbs sampling): monthly excess returns, equal-weighted portfolio, 1926–1986. (B) Distribution of standardized returns (from Gibbs sampling): monthly excess returns, equal-weighted portfolio, 1926–1986.

average of 10,000 sets of $\tilde{\sigma}^2$, which are our estimates of the stock return variance. Fig. 3A and B show the plot and the sampling distribution of the standardized returns, which is the average of 10,000 realized sets of $\tilde{y}^*$ from Gibbs sampling.

To check whether the three state Markov-switching variance we employ captures most of the dynamics in the stock return variance, we applied ARCH tests to the above standardized returns. [6] No ARCH effects were found. This is

[6] ARCH tests were performed on the average of 10,000 sets of standardized returns from Gibbs sampling. The LM test statistics at lags 1 through 5 were 0.001, 0.824, 0.657, 0.754, and 0.678, respectively.

consistent with Hamilton and Susmel (1994), who show that all the ARCH effects that show up in weekly stock returns data die out almost completely after 4 weeks (a month), after allowing for Markov-switching variance. In addition, the standardized returns show little excess kurtosis, with a p-value of 0.846 for the Jarque-Bera joint test of Normality. These suggest that the three-state Markov-switching variance model is a reasonable approximation to the heteroskedasticity in the stock returns for the period 1926.01–1986.12.

3. Historical pattern of heteroskedasticity and the sampling distribution of the variance ratio statistic

In measuring the statistical significance of the VR statistic, PS use estimates of the standard error based on Monte Carlo simulations assuming independently and normally distributed returns. However, as noted above, stock returns are unconditionally nonnormal and heteroskedastic with high persistence. As a justification for their use of the estimates of the standard error based on Monte Carlo simulations of i.i.d. normal returns, PS show that the empirical distribution of the variance ratio statistic with heteroskedasticity is no different from that with homoskedasticity.

Instead of employing Monte Carlo experiments which require a distributional assumption, KNS employ randomization methods in order to estimate the unknown distribution of the VR. Randomization focuses on the null hypothesis that one variable is distributed independently of another. To estimate the distribution of the VR statistic under the null, they shuffle the data to destroy any time dependence and then recalculate the test statistic for each reshuffling. In the presence of persistent heteroskedasticity, however, the usual randomization method may fail because the errors are not interchangeable since randomization also destroys any time dependence in variance. KNS also present results for a 'stratified randomization' that preserves the historical pattern of heteroskedasticity. [7]

Poterba and Summers (1988) report a Monte Carlo experiment which mimics the actual persistence of volatility but does not preserve the historical pattern. [8] This may be valid when the particular historical pattern of heteroskedasticity does not affect the sampling distribution of the variance ratio statistic.

In this Section, we re-examine the effects of heteroskedasticity on the sampling distribution of the variance ratio statistic, in particular the persistence or specific

[7] Their stratified randomization provides a way to retain information in historical heteroskedasticity in returns. However, their division of the sample into low and high variance states is arbitrary and limited.

[8] The volatility process they used for simulation is $\log(\sigma_t^2) = \alpha + \beta \log(\sigma_{t-1}^2) + \omega_t$, $\omega_t \sim i.i.d.N(0, \sigma_\omega^2)$, where α, β and σ_ω^2 are estimated from historical data as in French et al. (1987).

pattern of heteroskedasticity. We experiment by incorporating different amounts of information about heteroskedasticity in historical returns when generating artificial histories, and the empirical distribution of the VR for each case is compared to that of the homoskedastic case. All the results below are based on 10,000 sets of generated monthly returns with 732 observations.

3.1. Monte Carlo experiment #1: No persistence in heteroskedasticity

3.1.1. Data generating process

$$y_t \sim N(0, \sigma_t^2),$$
$$\sigma_t^2 = \sigma_1^2 S_{1t} + \sigma_2^2 S_{2t} + \sigma_3^2 S_{3t},$$
$$S_{kt} = 1 \text{ if } S_t = k, \text{ and } S_{kt} = 0, \quad \text{otherwise; } k = 1, 2, 3$$
$$\Pr[S_t = j] = p_j, \quad j = 1, 2, 3$$

where the unobserved state variable, S_t, at time t evolves independently of past realizations. The values of p_1, p_2, and p_3 used for data generation are the steady state probabilities calculated from estimates of the transition probabilities in the three-state Markov-switching variance model of Section 2, based on historical data. Values for the other parameters, σ_1^2, σ_2^2, and σ_3^2 are also taken from estimation of the same model based on historical data. In the above data generating process, we allow variance of the returns to be heteroskedastic and regime switching, but switches between regimes are independent of the previous regime. The generated returns contain part of the information in historical returns: the proportion that each regime occurs out of the whole sample. They do not retain

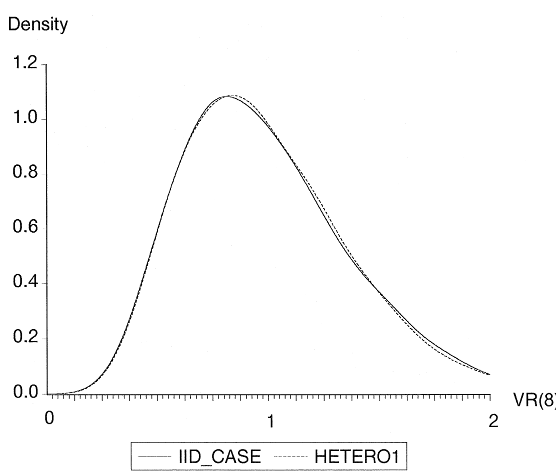

Fig. 4. Empirical distribution of 96-month variance ratio statistic with homoskedastic returns and heteroskedastic returns from Monte Carlo experiment #1.

information on either the persistence or the pattern of heteroskedasticity in historical returns.

Fig. 4 shows the empirical distribution function for the 96-month VR statistic in both the homoskedastic and heteroskedastic cases. When returns are independent, heteroskedasticity does not seem to affect the empirical distribution function of the VR.

3.2. Monte Carlo experiment #2: Heteroskedasticity with historical persistence

3.2.1. Data generating process

$$y_t \sim N(0, \sigma_t^2),$$
$$\sigma_t^2 = \sigma_1^2 S_{1t} + \sigma_2^2 S_{2t} + \sigma_3^2 S_{3t},$$
$$S_{kt} = 1 \text{ if } S_t = k, \text{ and } S_{kt} = 0, \quad \text{otherwise}; \ k = 1, 2, 3$$
$$\Pr[S_t = j | S_{t-1} = i] = p_{ij}, \quad i, j = 1, 2, 3$$
$$\sum_{j=1}^{3} p_{ij} = 1$$
$$\sigma_1^2 < \sigma_2^2 < \sigma_3^2$$

which is the model described in Section 2. All the parameters that are needed to generate returns are from the Gibbs sampling estimates based on historical data. Using the estimates of transition probabilities, $\{S_t, t = 1, 2, \dots, T\}$ are generated first. Then along with estimates of σ_k^2, $k = 1, 2, 3$, we can easily calculate $\{\sigma_t^2, t = 1, 2, \dots, T\}$ using Eqs. (6) and (7), which are used to generate returns. In this way, information in historical returns is conveyed to generated returns only

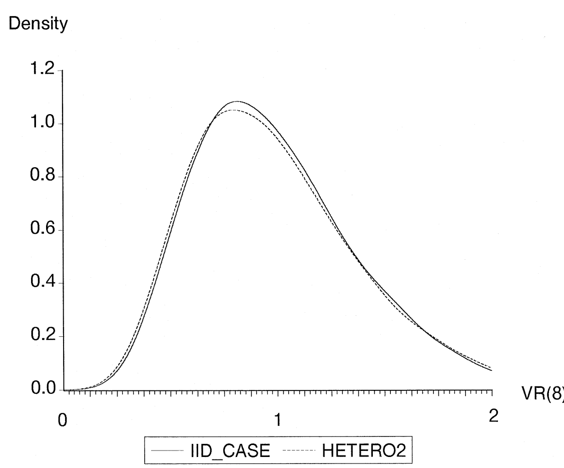

Fig. 5. Empirical distribution of 96-month variance ratio statistic with homoskedastic returns and heteroskedastic returns from Monte Carlo experiment #2.

through the parameter estimates of the model used for return generation. Returns generated in this way retain additional information in historical returns in addition to that in experiment #1: persistence of historical return variance. However, they do not retain the specific pattern of heteroskedasticity in historical returns.

The results are shown in Fig. 5. Even when we incorporate persistent heteroskedasticity in our data generating process, the empirical distribution of the VR does not seem to be affected very much. The results in Figs. 4 and 5 are consistent with those of PS. They seem to suggest that the degree of persistence in heteroskedasticity does not affect the distribution of the variance ratio statistic very much, when the pattern of heteroskedasticity is not fixed when data are generated.

3.3. Monte Carlo experiment #3: Heteroskedasticity with historical persistence and pattern

As we see in Fig. 2, we have a period of unusually high volatility in the earlier portion of our sample associated with the Great Depression. The question one may ask is: if the history goes on, how likely is it that we may have another episode like that again. If the answer is negative, it may suggest that such unusually rare events should be controlled for in our Monte Carlo experiments. For this purpose, while maintaining the data generating process in experiment #2, we retain both the persistence and the pattern of heteroskedasticity in historical returns in generating data for our new Monte Carlo experiments.

The Steps 2 through 5 of Section 5 are replaced by the final parameter estimates. Conditional on data and final parameter estimates of the model, only

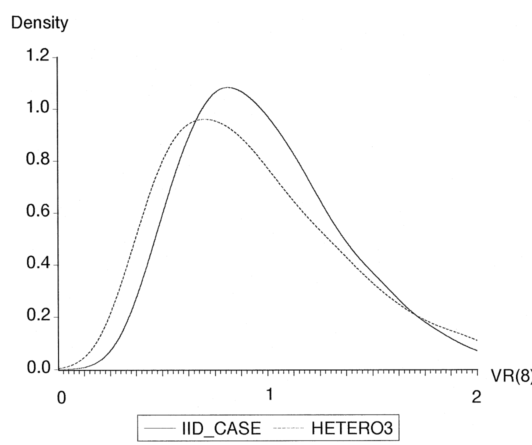

Fig. 6. Empirical distribution of 96-month variance ratio statistic with homoskedastic returns and heteroskedastic returns from Monte Carlo experiment #3.

Step 1 is repeated 10,000 times to generate 10,000 sets of $\tilde{S} = \{S_t,\ t = 1, 2, \dots, T\}$. Once $\tilde{S}$ is generated, along with estimates of σ_k^2, $k = 1, 2, 3$, it is straightforward to generate artificial histories and the rest is the same as in Monte Carlo experiments #1 and #2. Each set of artificial returns generated in this way will on average retain the same historical pattern of heteroskedasticity plotted in Fig. 2.

The results are shown in Fig. 6. Unlike the previous two cases, when the pattern of heteroskedasticity is the historical one, the empirical distribution of the variance ratio is much different from that in the homoskedastic case.[9] The distribution has wider variance and is more skewed than in the homoskedastic case. This suggests that the variance ratio tests of PS based on Monte Carlo experiments and those of KNS based on the usual randomization method have the wrong size, rejecting the null of random returns in favor of mean reversion too often.

4. New tests of mean reversion based on variance ratio statistics

In this section, we consider two new tests of mean reversion that condition on the information that the data contain the historical pattern of heteroskedasticity. KNS carried out a stratified randomization of the data in which returns from the high-variance period 1930–1939 were placed in a separate urn when generating artificial histories. Their division of the whole sample into only high-variance and low-variance periods, however, is arbitrary and limited and fails to reflect the uncertainty inherent in estimating state changes. The Gibbs sampling approach allows us to obtain valid significance levels. Then, we consider a potentially more efficient test for mean reversion based on the variance ratio of appropriately standardized returns. Randomization methods, instead of Monte Carlo experiments, are employed to allow for stock returns to be nonnormal within a state.

4.1. Tests based on the variance ratio of original returns

Assume that

$$y_t \sim \left(0,\ \sigma_t^2(\theta)\right),$$

$$\theta = \left\{\sigma_1^2,\ \sigma_2^2,\ \sigma_3^2,\ p_{11},\ p_{12},\ p_{21},\ p_{22},\ p_{31},\ p_{32}\right\},$$

where demeaned return y_t is heteroskedastic with variance σ_t^2 and θ is a vector of parameters that describes the dynamics of σ_t^2. If we randomize the original return y_t, we lose information on the pattern and persistence of heteroskedasticity

[9] Intuitively, in a regression model with heteroskedastic disturbances, we know that the covariance matrix of OLS estimator is different from that in the case of homoskedastic disturbances.

in historical returns. A natural way to randomize returns without losing time dependence in historical returns would be:

Step 1: standardize y_t to get $\{y_t^* = y_t/\sigma_t,\ t=1, 2, \ldots, T\}$,

Step 2: randomize the standardized returns to get $\{y_t^{**},\ t=1, 2, \ldots, T\}$, and finally,

Step 3: de-standardize y_t^{**} to get $\{\tilde{y}_t^{**} = y_t^{**} \times \sigma_t,\ t=1, 2, \ldots, T\}$, which is to be used to estimate the empirical distribution of the variance ratio statistic under the null.

If σ_t^2, the variance of return for each time point, or the parameters (θ) that govern the evolution of σ_t^2 were known, the above procedure would be straightforward. In practice, σ_t^2 or the parameters (θ) associated with σ_t^2 have to be estimated using historical data. As θ is subject to parameter uncertainty, the standardized returns themselves are subject to sampling variation. Thus, the empirical distribution of the variance ratio statistic with heteroskedasticity would have to account for both the effect of parameter uncertainty in θ and the effect of randomization.

In order to incorporate the effect of uncertainty in the parameters associated with variance of returns, we augment the Gibbs sampling approach introduced in Section 2 with the standardizing step of the above procedure. As in Section 2, each run of Gibbs sampling based on historical returns provides us with particular realizations of the set, $\{S_t,\ t=1, 2, \ldots, T\}$ and $\{\sigma_1^2, \sigma_2^2, \sigma_3^2\}$, which are used to calculate σ_t^2 according to Eq. (6). Using σ_t^2, $t=1, 2, \ldots, T$ simulated in this way we can proceed with Steps 1 through 3. If the above procedure is repeated, say, 10,000 times, with each iteration augmented by simulations of σ_t^2 from each run of Gibbs sampling, we have 10,000 sets of randomized returns. These artificial histories condition on the information about the pattern of heteroskedasticity contained in the historical returns, incorporate parameter uncertainty, and are consistent with the null of mean reversion due to randomization. For each of these 10,000 sets of artificial histories, the variance ratio statistic is calculated, which can be used to estimate the empirical distribution of the variance ratio statistic. To estimate the significance level, we count how many times the variance ratios for the artificial histories fall below the variance ratios for original historical returns.

4.2. Tests based on the variance ratio of standardized returns

In the context of a regression equation with heteroskedastic disturbances, the test of mean reversion considered in the previous section is analogous to running OLS and then using the heteroskedasticity-consistent covariance matrix for inference. In this section, we develop a potentially more efficient estimate of the variance ratio, which is analogous to the GLS estimate in the regression context, and its sampling distribution.

By standardizing historical returns before calculating the variance ratio test statistic, appropriate weights can be assigned to observations depending on their volatility. An additional complication of this approach is that, unlike the test based on original returns, the test statistic itself is subject to sampling variation due to uncertainty in the parameters that describe the dynamics of heteroskedasticity. Thus, we compare two distributions: the distribution (due to parameter uncertainty) of the variance ratio test statistic for standardized historical returns, and the distribution of the variance ratio test statistic under the null hypothesis estimated from randomizing the standardized returns. As before, in standardizing historical returns, we employ the Gibbs sampling described in Section 2.

As in Section 4.1, at the end of each run of the Gibbs sampling, we have the simulated set $\{\sigma_t^2,\ t=1,\ 2,\ \dots,\ T\}$. We follow the following steps to perform tests based on the VR of standardized returns:

Step 1: Standardize y_t to get $\{y_t^* = y_t \times (1/\sigma_t),\ t=1,\ 2,\ \dots,\ T\}$, which is a particular realization of standardized historical returns.

Step 2: Calculate VR's for standardized historical returns y_t^*, denoted by $VR^*(k),\ k=1,\ 2,\ \dots,\ K$.

Step 3: Randomize the standardized returns from Step 1, to get $\{y_t^{**},\ t=1,\ 2,\ \dots,\ T\}$.

Step 4: Calculate VR's for randomized returns y_t^{**} from Step 3, denoted by $VR^{**}(k),\ k=1,\ 2,\ \dots,\ K$.

Step 5: Compare $VR^*(k)$ and $VR^{**}(k)$.

These steps are repeated, say, 10,000 times to get the posterior distribution of the VR for standardized historical returns, $VR^*(k)$, and the empirical distribution of the VR under the null of no mean reversion, $VR^{**}(k)$. To estimate the significance level for the test of mean reversion, we count how many times the variance ratio for the standardized and randomized returns $(VR^{**}(k))$ from Gibbs-sampling-augmented randomization falls below the variance ratio for standardized historical returns $(VR^*(k))$ from Gibbs sampling.

5. Empirical results and comparison with prior literature

The data set consists of monthly total returns on NYSE stocks from the CRSP files for both value weighted (VW) and equal weighted portfolios. The one month T bill return from Ibbotson Associates is subtracted to obtain the excess return. For comparability with PS and KNS the sample period is 1926–1986.

The sample values of VRs for original returns reported in Table 2 are necessarily the same as those reported by KNS in their Tables 1 and 2 for the same sample period and essentially the same as those reported by PS in their Table 2 for 1926–1985. The sample estimates point to mean reversion at a horizon of 7 years in VW returns and at 9 years for EW returns. What differs in this paper is the method of obtaining p-values.

Based on Monte Carlo methods PS report a smallest p-value of 0.08 for VW and 0.005 for EW excess returns. They found that heteroskedasticity that mim-

Table 2
Variance ratios for original monthly CRSP excess returns, 1926–1986

	Lag K (years)							
	2	3	4	5	6	7	8	9
	Sample VR for value-weighted portfolio returns							
	1.035	0.980	0.919	0.849	0.775	0.682	0.671	0.709
Sampling distribution based on Gibbs-sampling-augmented randomization								
Mean	0.999	1.006	1.012	1.011	1.004	0.994	0.982	0.971
Median	0.992	0.987	0.978	0.961	0.939	0.920	0.899	0.879
SD	0.142	0.227	0.293	0.348	0.392	0.427	0.454	0.476
p-value	0.614	0.489	0.419	0.365	0.316	0.255	0.273	0.332
	Sample VR for equal-weighted portfolio returns							
	1.009	0.923	0.877	0.783	0.646	0.487	0.427	0.421
Sampling distribution based on Gibbs-sampling-augmented randomization								
Mean	0.998	1.004	1.008	1.005	0.996	0.985	0.972	0.959
Median	0.996	0.984	0.974	0.952	0.926	0.901	0.878	0.860
SD	0.149	0.236	0.302	0.356	0.402	0.439	0.468	0.490
p-value	0.532	0.396	0.373	0.300	0.197	0.091	0.074	0.088

p-value is the frequency with which simulated VR's smaller than the historical sample value were observed in the Gibbs-sampling-augmented randomization under the null hypothesis.

icked the persistence but not the specific pattern of volatility had no impact on the sampling distribution of the VR. KNS report results of a stratified randomization that preserves the apparent historical pattern of volatility but does not recognize uncertainty in dating the change. The smallest p-values are 0.135 at lag seven years for VW and 0.059 at lag eight years for EW returns. In Table 2 of this paper the smallest p-values reported are 0.255 at lag 7 years for VW and 0.074 at 8 years for EW returns. Thus, when we use the Gibbs sampling procedure of Section 4.1 to account for parameter uncertainty and to allow resampling of the whole data series rather than only subsamples, we find little if any evidence of mean reversion in VW returns and weaker evidence for mean reversion in EW returns.

In Section 4.2 we proposed that the VR be computed after weighting the returns appropriately based on information contained in the data about the pattern of heteroskedasticity. These results are reported in Table 3 of this paper. We no longer can report a single number for the sample VR statistic, rather the evidence from the data is summarized in the form of a posterior distribution. Looking at the results for VW returns first, note that the departure of the VR from one is no longer greatest at lag 8 years, rather the smallest posterior mean is observed at lag 5 years. The p-values are obtained as described in Section 4.2; briefly, each iteration of the Gibbs sampling is compared with the corresponding realization of the randomization. The strongest evidence of mean reversion in VW returns, corresponding to the smallest p-values, is actually at a lag of only 4 years. Thus, the standardized returns approach to estimating the VR suggests that mean

Table 3
Posterior sampling distributions of variance ratios of standardized monthly CRSP excess returns, 1926–1986

	Lag K (years)							
	2	3	4	5	6	7	8	9
	Posterior distribution of the VR for standardized VW returns							
Mean	0.962	0.883	0.829	0.826	0.848	0.843	0.901	1.002
Median	0.961	0.882	0.826	0.824	0.848	0.844	0.902	1.004
SD	0.024	0.036	0.042	0.051	0.064	0.078	0.092	0.106
Sampling distribution of the VR for standardized randomized returns								
Based on Gibbs-sampling-augmented randomization								
Mean	0.994	0.993	0.992	0.992	0.992	0.993	0.994	0.995
Median	0.993	0.985	0.975	0.964	0.954	0.945	0.931	0.921
SD	0.107	0.174	0.228	0.275	0.316	0.354	0.390	0.424
p-value	0.385	0.278	0.248	0.298	0.361	0.380	0.460	0.571
	Posterior distribution of the VR for standardized EW returns							
Mean	0.916	0.819	0.769	0.763	0.755	0.706	0.710	0.742
Median	0.916	0.818	0.768	0.760	0.749	0.700	0.701	0.733
SD	0.034	0.053	0.067	0.078	0.091	0.106	0.121	0.137
Sampling distribution of the VR for standardized randomized returns								
Based on Gibbs-sampling-augmented randomization								
Mean	0.995	0.994	0.993	0.992	0.992	0.992	0.993	0.995
Median	0.993	0.987	0.976	0.965	0.954	0.945	0.936	0.923
SD	0.108	0.177	0.230	0.276	0.317	0.355	0.391	0.424
p-value	0.243	0.174	0.179	0.218	0.247	0.229	0.263	0.313

p-value is the frequency with which realizations of the Gibbs sampling of the posterior distribution were smaller than the corresponding realization under the null hypothesis.

reversion, if it is present, occurs at much shorter lags than had been reported previously. It also suggests that the evidence for mean reversion is weaker than previously reported; the smallest *p*-value in Table 3 is 0.248 for VW returns, compared with the smallest *p*-value of 0.135 reported by KNS and 0.08 by PS.

Corresponding results for the EW returns are reported in the lower panel of Table 3. Again, we have a posterior distribution for the historical sample VR, and its mean departs farthest from one at lag 7 years. However, the departure from one is not nearly as large as in the case of the original VR. The *p*-values tell an even more different story. The smallest occur at a lag of 3 years, with a value of 0.174. That compares with a smallest *p*-value of 0.059 at lag 8 years reported by KNS and 0.005 by PS. Thus the Gibbs sampling approach to estimating the VR both shortens the lag at which evidence of mean reversion is apparent and suggests that the evidence for mean reversion is weaker than previously reported. It seems clear from these results that making use of the information in the data about the pattern

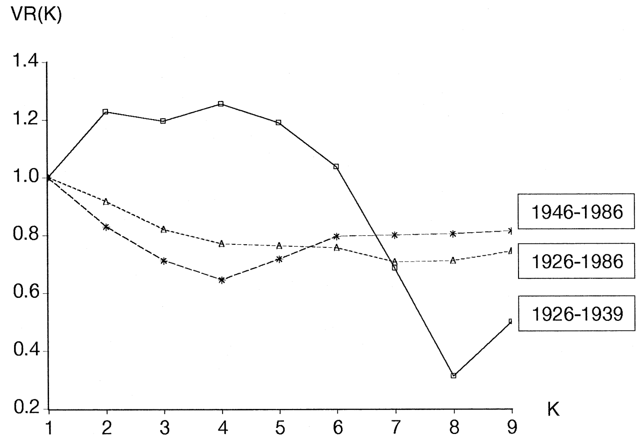

Fig. 7. Mean of VR for standardized EW returns (1926–1986, 1926–1939, 1946–1986).

of heteroskedasticity, both in estimating the VR and in obtaining *p*-values, can substantially affect the inferences that are drawn from the data about the degree and lag dynamics of mean reversion. In addition, by comparing the standard deviations (SD's) of the sampling distributions of VR's reported in Tables 2 and 3, one may argue that the tests in Section 4.2 based on standardized returns are in general more efficient than those in Section 4.1 based on original returns.

In Figs. 7 and 8, posterior means of the VR's for the standardized returns for two sub-periods (1926–1939 and 1946–1986) are plotted against those for the full sample (1926–1986). Focusing, for example, on the equal weighted portfolios, for the subsample that includes Great Depression, the *p*-value was lowest with 0.06 at

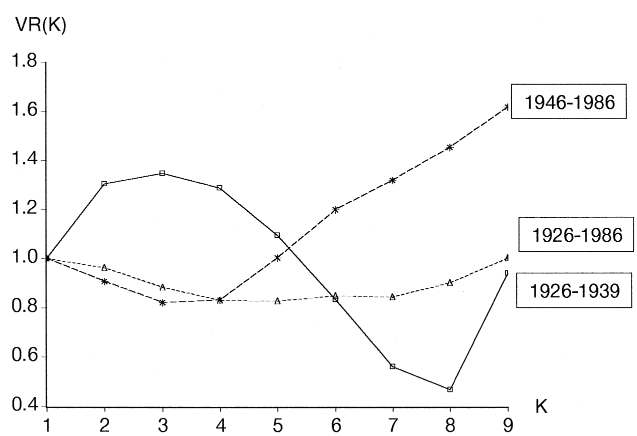

Fig. 8. Mean of VR for standardized VW returns (1926–1986, 1926–1939, 1946–1986).

lag 8 years. For the post-World War II subsample, the p-value was lowest with 0.103 at lag 3 years. The VR's for the full sample look like some kind of weighted average of those for the two subsamples. If the period of highly volatile returns tends to show stronger evidence of mean reversion and if the period of far less volatile returns shows less evidence of mean reversion, the VR test results based on original returns may be dominated by the period of highly volatile returns, even though the length of such period is short relative to the full sample. Our test results based on the VR of original returns tend to confirm this. The information content of returns during the period of high volatility may be small and using the standardized returns in the VR tests is analogous to assigning less weight on returns during the period of high volatility. Then the test results tend to be dominated by the relatively longer period of low volatility period. Test results based on the VR of standardized returns tend to confirm this.

6. Summary and conclusion

The variance ratio statistic has become a standard tool of time series analysis, widely applied in testing for mean reversion, while heteroskedasticity is pervasive in economic data. Thus it would seem to be important to be able to make appropriate allowance for heteroskedasticity when basing inference on the VR statistic. Previous work had either reported that heteroskedasticity did not affect the sampling distribution of the VR, or had assumed that the investigator knew a priori the pattern of heteroskedasticity. This paper uses the Gibbs sampling approach in the context of a three state Markov-switching model to show how heteroskedasticity influences the sampling distribution of the VR and then suggests two avenues of valid inference.

We find that the sampling distribution of the VR is affected by the particular pattern of heteroskedasticity during the sample period, and that this effect is substantively important in the case of monthly stock returns 1926–86. Simulation methods that assume heteroskedasticity or allow for persistence in heteroskedasticity but do not condition on the particular pattern of the historical period produce a biased test, leading the investigator to reject the null hypothesis of no mean reversion too often.

Building on this finding, we present a resampling strategy for estimating the sampling distribution of the VR that standardizes historical returns using estimated variances. Instead of conditioning on estimates of these variance and the dates of regime switches, the Gibbs sampling approach is used to allow for uncertainty in these parameters and states while conditioning on the information in the data. We then estimate how likely it is that a VR as extreme as that observed in the historical sample would have occurred under the null hypothesis.

If the previous procedure can be thought of as finding the heteroskedastic consistent standard error for an OLS estimator, then the GLS analog is to estimate the VR itself from standardized data. Again, Gibbs sampling makes appropriate use of the information in the data about the pattern of heteroskedasticity while taking into account uncertainty about parameters and states. The output is a posterior distribution for the VR rather than a single point estimate. The *p*-value is no longer the fractile of a sampling distribution corresponding to a sample estimate, rather we must compare the posterior distribution of the sample VR with the distribution under the null hypothesis that returns are serially random. At the end of each iteration of the Gibbs sampling we compare the estimate of the VR from the standardized historical data with the corresponding VR from randomized, standardized data. We estimate a *p*-value by counting how often the former falls below the latter.

Turning to the monthly CRSP stock returns 1926–86 studied in several previous papers, we find that inference based on the original VR is substantially affected by taking into account the historical pattern of heteroskedasticity. While the point estimates of the VR, which are not affected, still suggest mean reversion at lags of several years, the *p*-values, which are affected by the Gibbs sampling approach, are larger than reported previously based on Monte Carlo methods that did not preserve the pattern of heteroskedasticity. Evidently, the evidence for mean reversion depends more heavily on periods, such as the 1930s, which were also noisier. When Gibbs sampling is used to estimate the VR at a given lag based on appropriately standardized returns, we obtain a posterior distribution. These do not deviate from unity as much as the conventional VR, and where they do, it is at shorter lags of 3 or 4 years. Comparing the posterior distribution with that under the null, we find only modest evidence of mean reversion in the case of returns on the equal-weighted CRSP portfolios and essentially none for the value-weighted portfolios. Speaking loosely, doing GLS instead of OLS does matter. When returns from the high variance periods are appropriately standardized, the evidence for mean reversion changes in its pattern and significance.

Since Gibbs sampling procedures are highly operational, we recommend that this approach be explored when investigating mean reversion in heteroskedastic data such as interest rates, asset returns, and exchange rates.

Appendix A

A.1. Generating S_t, $t = 1, 2, \ldots, T$, conditional on data and parameters of the model

Defining $\theta = \{\sigma_1^2, \sigma_2^2, \sigma_3^2, p_{11}, p_{12}, p_{21}, p_{22}, p_{31}, p_{32}\}$ and $\tilde{y}_t = [y_1 \ \ldots \ y_t]'$, the set $\tilde{S}_t = [S_1 \ \ldots \ S_t]'$ can be generated based on the following distribution, as in Kim and Nelson (1997):

$$p\left(\tilde{S}_T | \tilde{y}_T, \theta\right) = p\left(S_T | \tilde{y}_T, \theta\right) \prod_{t=1}^{T-1} p\left(S_t | \tilde{y}_t, \theta, S_{t+1}\right) \tag{A.1}$$

In order to simulate $\tilde{S}_T$ from the above distribution, we first run Hamilton's (1989) basic filter for the model to get $p(S_t|\tilde{y}_t, \theta)$ and $p(S_t|\tilde{y}_{t-1}, \theta)$, for $t=1, 2, \dots, T$ and save them. The last iteration of the filter provides us with $p(S_T|\tilde{y}_T, \theta)$, from which S_T is generated. Then, we can successively generate S_t from $p(S_t|\tilde{y}_t, \theta, S_{t+1})$, for $t=T-1, T-2, \dots, 1$, using the following results:

$$p\left(S_t|\tilde{y}_t, \theta, S_{t+1}\right) = \frac{p\left(S_{t+1}|S_t\right) p\left(S_t|\tilde{y}_t, \theta\right)}{p\left(S_{t+1}|\tilde{y}_t, \theta\right)} \tag{A.2}$$

As S_t is a three state Markov-switching variable, we need to pay special attention to its generation based on the uniform distribution. Conditional on $S_{t+1}=k$, $k=1, 2, 3$, define $p_j = p_{jk} \times p(S_t = j|\tilde{y}_t, \theta)$, $j=1, 2, 3$, where p_{jk} is the transition probability. We first generate a random number from the uniform distribution. If the generated number is greater than or equal to $p_1/(p_1+p_2+p_3)$, we set $S_t=1$; if it is less than $p_1/(p_1+p_2+p_3)$, we generate another random number from the uniform distribution. And then, if the generated number is greater than or equal to $p_2/(p_2+p_3)$, we set $S_t=2$; if it is less than $p_2/(p_2+p_3)$, we set $S_t=3$.

A.2. Generating σ_j^2, $j=1, 2, 3$, conditional on data, S_t, $t=1, 2, \dots, T$, and other parameters of the model

In order to give a constraint that $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$, we may re-define σ_2^2 and σ_3^2 in the following way:

$$\sigma_2^2 = \sigma_1^2(1+h_2) \quad \text{and} \quad \sigma_3^2 = \sigma_1^2(1+h_2)(1+h_3), \tag{A.3}$$

where $h_2 > 0$ and $h_3 > 0$. We first generate σ_1^2, then generate $1+h_2$ and $1+h_3$.

First, to generate σ_1^2, we transform Eq. (5) as follows:

$$Y_{1t} = \frac{y_t}{\sqrt{(1+S_{2t}h_2)(1+S_{3t}h_2)(1+S_{3t}h_3)}} \tag{A.4}$$

By choosing the inverse gamma distribution as the prior (IG$(\nu_1/2, \delta_1/2)$), one can show that the conditional distribution from which σ_1^2 is generated is given by:

$$\left[\sigma_1^2|\tilde{Y}_{1T}, \tilde{S}_T, \tilde{\theta}_{j\neq\sigma_1^2}\right] \sim \mathrm{IG}\left(\frac{\nu_1+T}{2}, \frac{\delta_1+\sum_{t=1}^{T} Y_{1t}^2}{2}\right), \tag{A.5}$$

where $\tilde{\theta}_{j\neq\sigma_1^2}$ represents a vector of parameters of the model that excludes σ_1^2.

Second, to generate $\bar{h}_2 = 1+h_2$, and thus, σ_2^2, we transform Eq. (5) as follows:

$$Y_{2t} = \frac{y_t}{\sqrt{\sigma_1^2(1+S_{3t}h_3)}} \tag{A.6}$$

Here, we note that the likelihood function of h_2 depends on the values of y_t for which $S_t = 2$ *or* 3. By defining $T_2 = \{t:S_t = 2 \text{ or } 3\}$ and choosing the inverse gamma distributions for the priors of $\bar{h}_2$ (IG$(\nu_2/2,\ \delta_2/2)I_{[\bar{h}_2 > 1]}$), one can show that the complete conditional is given by:

$$\left[\bar{h}_2|\tilde{Y}_{2T},\ \tilde{S}_T,\ \tilde{\theta}_{j \neq h_2}\right] \sim \mathrm{IG}\left(\frac{\nu_2 + N_2}{2},\ \frac{\delta_2 + \sum_{t=1}^{T_2} Y_{2t}^2}{2}\right) I_{[\bar{h}_2 > 1]}, \tag{A.7}$$

where $\tilde{\theta}_{j \neq h_2}$ represents a vector of parameters of the model that excludes h_2; I is the indicator function on $[\bar{h}_2 > 1]$; N_2 are cardinalities of T_2 and the sum is over the elements T_2.

Finally, to generate $\bar{h}_3 = 1 + h_3$, and thus, σ_3^2, we transform Eq. (5) as follows:

$$Y_{3t} = \frac{y_t}{\sqrt{\sigma_1^2 (1 + S_{3t} h_2)}} \tag{A.8}$$

Here, we note that the likelihood function of h_3 depends only on the values of y_t for which $S_t = 3$. By defining $T_3 = \{t:S_t = 3\}$ and choosing the inverse gamma distributions for the priors of $\bar{h}_3 = 1 + h_3$ (IG$(\nu_3/2,\ \delta_3/2)I_{[\bar{h}_3 > 1]}$), one can show that the complete conditional is given by:

$$\left[\bar{h}_3|\tilde{Y}_{3T},\ \tilde{S}_T,\ \tilde{\theta}_{j \neq h_3}\right] \sim \mathrm{IG}\left(\frac{\nu_3 + N_3}{2},\ \frac{\delta_3 + \sum_{t=1}^{T_3} Y_{3t}^2}{2}\right) I_{[\bar{h}_3 > 1]}, \tag{A.9}$$

where $\tilde{\theta}_{j \neq h_3}$ represents a vector of parameters of the model that excludes h_3; I is the indicator function on $[\bar{h}_3 > 1]$; N_3 are cardinalities of T_3 and the sum is over the elements T_3.

The quantity ν_i, $i = 1, 2, 3$, represent the strength of the priors of σ_1^2, $\bar{h}_2$, and of $\bar{h}_3$. For our application of the approach to CRSP excess returns in Section 2, we employ $\nu_i = 0$ and $\delta_i = 0$ for $i = 1, 2, 3$.

A.3. Generating transition probabilities (p_{11}, p_{12}, p_{21}, p_{22}, p_{31}, p_{32}) conditional on data, S_t, $t = 1, 2, \ldots, T$, and other parameters of the model

Conditional on $\tilde{S}_T$, the transition probabilities are independent of $\tilde{y}_T$ and other parameters of the model, as in Albert and Chib (1993). For a two-state Markov-switching model, Albert and Chib (1993) derive the full conditional distributions of the transition probabilities as a product of independent beta distributions. For a three-state Markov-switching model in this paper, we need a slight modification of their approach.

Given $\tilde{S}_T$ and the initial state, let n_{ij}, $i,j = 1, 2, 3$, be the total number of transitions from state $S_{t-1} = i$ to $S_t = j$, $t = 1, 2, \ldots, T$. Define $\bar{p}_{ii} = \Pr(S_t \neq i|S_{t-1} = i)$ and $\bar{p}_{ij} = Pr(S_t = j|S_{t-1} = i,\ S_t \neq i)$, $i = 1, 2, 3$, $j = 1, 2$. Correspondingly, we have $p_{ij} = \bar{p}_{ij} \times (1 - p_{ii})$ for $i \neq j$. Similarly, define $\bar{n}_{ii}$ to be the number of transitions from state $S_{t-1}=i$ to $S_t \neq i$ and $\bar{n}_{ij}$ to be the number of transitions from state $S_{t-1}=i$ to state $S_t=j$, conditional on $S_t \neq i$.

Then, as in Albert and Chib (1993), by taking the beta family of distributions as conjugate priors, it can be shown that the posterior distributions of p_{ii} are given by

$$\left[p_{ii}|\tilde{S}_T \right] \sim beta\left(u_{ii}+n_{ii}, \bar{u}_{ii}+\bar{n}_{ii} \right), \quad i=1,2,3, \tag{A.10}$$

where u_{ii} and $\bar{u}_{ii}$ are the hyperparameters of the prior. Once p_{ii}, $i=1, 2, 3$, are generated from the above distribution, generation of the other parameters is straightforward. For example, given that p_{ii} is generated, p_{ij} can be calculated by $p_{ij}=\bar{p}_{ij}\times(1-p_{11})$, where $\bar{p}_{ij}$ can be generated from the following beta distribution:

$$\left[\bar{p}_{ij}|\tilde{S}_T \right] \sim beta\left(u_{ij}+n_{ij}, u_{ik}+n_{ik} \right), \quad i \neq j \neq k, \tag{A.11}$$

where u_{ij} and u_{ik} are the hyperparameters of the prior.

For our application of the approach to CRSP excess returns in Section 2, the hyper-parameters that we employ are: $u_{ii}=0.5$ and $\bar{u}_{ii}=0.5$; and $u_{ij}=0.5$ and $u_{ik}=0.5$ for $i \neq j \neq k$. These hyper-parameters imply almost noninformative priors for the transition probabilities in the sense that they imply that expected duration of a state is two months with very high standard deviation.

References

Albert, J.H., Chib, S., 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), 1–15.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–317.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics 52, 5–59.

Cochrane, J.H., 1988. How big is the random walk in GNP?. Journal of Political Economy 96, 893–920.

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Fama, E.F., 1963. Mandlebrot and the stable Paretian hypothesis. Journal of Business 4, 420–429.

French, K.R., Schwert, W., Stambaugh, R., 1987. Expected stock returns and stock market volatility. Journal of Financial Economics 19, 3–30.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association, Theory and Methods 85 (410), 398–409.

Gelfand, A.E., Hills, S.I., Racine-Poon, A., Smith, A.F.M., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. Journal of the American Statistical Association, Applications and Case Studies 85 (412), 972–985.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741.

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

Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39–70.

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

Kim, C.-J., Nelson, C.R., 1997. Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime-Switching. Review of Economics and Statistics, forthcoming.

Kim, M.J., Nelson, C.R., Startz, R., 1991. Mean reversion in stock prices? A reappraisal of the empirical evidence. Review of Economic Studies 58, 515–528.

Lo, A.W., MacKinlay, A.C., 1989. The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. Journal of Econometrics 40, 203–238.

Mandlebrot, B., 1963. The variation of certain speculative prices. Journal of Business 4, 394–419.

Poterba, J.M., Summers, L.H., 1988. Mean reversion in stock prices: Evidence and implications. Journal of Financial Economics 22, 27–59.

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