Graflund, Andreas

Working Paper

A Bayesian Inference Approach to Testing Mean Reversion in the Swedish Stock Market

Working Paper, No. 2000:8

Provided in Cooperation with:
Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Graflund, Andreas (2000) : A Bayesian Inference Approach to Testing Mean Reversion in the Swedish Stock Market, Working Paper, No. 2000:8, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:
https://hdl.handle.net/10419/259836

A Bayesian Inference Approach to Testing Mean Reversion in the Swedish Stock Market

Andreas Graflund
Department of Economics
Lund University
Box 7082
S-220 07 LUND
Sweden

Phone: +46 (0) 46 222 79 19
Fax: +46 (0) 46 222 41 18
Email: andreas.graflund@nek.lu.se

Department of Economics Working Paper Series 2000:8
JEL: G10 C11 C15

Keywords: *Market efficiency, variance ratio, Gibbs sampling, hidden Markov chains.*

January, 2002

Abstract

In this paper we use a Bayesian approach to test for mean reversion in the Swedish stock market on monthly data 1918-1998. By simply account for the heteroscedasticity of the data with a two-state hidden Markov model of normal distributions and taking estimation bias into account via Gibbs sampling we cannot find support of mean reversion. This is a contradiction to previous result from Sweden. We find that a tranquil and a volatile regime can characterize the Swedish stock market and within the regimes the stock market is random. This finding of randomness is in line with recent evidence for the U.S. stock market.

1 Introduction

This paper addresses the question of whether or not the Swedish stock market is mean reverting. Previous research by Frennberg and Hansson (1993) concludes that this is the case. Utilizing the variance ratio test, hereafter VR, of Cochrane (1988) they find evidence that the Swedish stock market is mean reverting with increasing investment horizon. In other words the stock market is less risky in the long run. This have implications for portfolio selection as well as the pricing of options. However, later research by Berg and Lyhagen (1998) has questioned their findings. Notwithstanding, the evidence of mean reversion via variance ratio is controversial because the test of the null hypothesis of random walk is only valid under the assumption of constant expected return. The return series from financial markets are well known to exhibit time variation, especially in volatility (for Sweden see Hansson and Hördahl (1997))

Poterba and Summers (1988) and Lo and MacKinlay (1989) use Monte Carlo simulation to show that the distribution of the VR statistic is unaffected by heteroscedasticity in returns. Kim et al (1991) points out that Monte Carlo simulation while retaining, the degree persistence in the heteroscedasticity in monthly stock returns destroys the historical pattern of heteroscedasticity.

Kim et al (1991, 1998a) questions the often used assumption of homoscedastic volatility and argues that the significant divergences some times found when using VR statistic might in fact be explained by the historical pattern of variance shifts. Malliaropulos and Priestley (1999) utilize a bootstrap approach to account for the small sample distributions of variance ratio test of Southeast Asian stock markets. They find that mean reversion is due to time-variation and they point out the danger of testing market efficiency without adjusting returns for time-variation in expected return. Hence, mean reversion might be explained by the historical pattern of time-variation, or regime switches in volatility, and taking this aspect into consideration might influence the VR test statistic.

This study differs from previous studies on the Swedish stock market in that we employ a Bayesian approach to test for mean reversion on standardized excess returns as suggested by Kim et al (1998a). The idea is to capture the time variation in the variance by a regime switching model, also known as hidden Markov model, of Gaussian mixtures. Thus we assume two regimes: low and high volatility.

Goldfeld and Quandt (1973) introduced the Markov switching models in economics but its breaktrough in economics and finance came with Hamilton's (1989) seminal paper. The drawback with regime switching models is that ordinary optimization of the likelihood

function can be difficult.[1] Albert and Chib (1993) address this problem with a Gibbs sampling approach in order to estimate the two-state regime switching model suggested by Hamilton (1989).[2] The Bayesian framework in combination with Geman and Geman's (1984) numerical integration technique of Gibbs sampling is very advantageous. First, we can use prior information in the estimation of the conditional distribution of the parameters without estimation of a likelihood function. Second, all inferences in Gibbs sampling are made from joint distributions of the variates and the unknown parameters of the model. Thus, we are able to account for the parameter uncertainty of the underlying parameters in the model.

In our analysis we find no support of mean reversion and our two-state regime switching model of normal distributions suggests that mean reversion if found in the Swedish stock market can be explained by the historical pattern of time variation in the volatility.

The outline of the chapter is as follows: The underlying assumptions of the variance ratio test are presented in section 2. In section 3 we describe the regime switching model and give a brief presentation of Bayesian statistics. The Gibbs sampler and the prior distributions are specified in this section along with a presentation of the Bayesian resampled variance ratio tests of Kim et al (1998a). Section 4 presents the data and the results and section 5 concludes the chapter.

2 Variance ratio

The variance ratio test, VR, of Cochrane (1988) has been frequently used as a test of mean reversion. The variance ratio is a test of linear dispersion of the asset price and the asset price is said to be a random walk if the variance is linearly increasing with time. If the VR is less than unity the dispersion is less than in the random walk case and this is referred to as mean reversion. The advantage of the test is that it allows us to study if returns follow a random walk and if this property changes with the investment horizon q. The q period return $y_{q,t}$ is computed as the q period difference between the log of the monthly prices of the portfolio I_t and I_{t-q}, in our case the Swedish stock market portfolio:

$$y_{q,t} = \ln I_t - \ln I_{t-q} \tag{1}$$

[1] Ordinary optimization algorithms often fail to estimate the true HMM correct. Another approach is to employ the simulated annealing, SA, algorithm. This is also a MCMC approach and thus, computer intensive.

[2] Kim et al (1998a, 1998b) extended Albert and Chib's model to a three-state HMM.

Let $y_{1,t}$ be the monthly return including dividends of the market portfolio. The asset price, I_t, is assumed to be a random walk and this implies that the arithmetic return being a drift μ plus a white noise term ε_t. In this context the q-month arithmetic return is:

$$y_{q,t} = q\mu + \varepsilon_t + \ldots + \varepsilon_{t-q+1} \tag{2}$$

$$y_{q,t} = y_{q,t-1} + \varepsilon_t \tag{3}$$

The expected q period return is equal to the monthly mean return times the holding period q and the variance of the q period return is q times the variance of monthly returns.

$$E\left[y_{q,t}\right] = q\mu, \quad Var\left[y_{q,t}\right] = q\sigma^2 \tag{4}$$

The variance ratio statistic, VR, is defined as:

$$VR\left(q\right) = \frac{Var\left[y_{q,t}\right]}{q \cdot Var\left[y_{1,t}\right]} \tag{5}$$

$$= 1 \text{ under random walk}$$

Under the null hypothesis, the $VR(q)$ statistic is equal to unity for all q and $VR(q)$ is asymptotically normal distributed. In our investigation we have chosen the investment horizon q to range from two to twelve months and yearly up to ten years. This enables us to study the random walk hypothesis, with respect to dispersion, both in the short-run and the long-run.

3 Methodology

3.1 Hidden Markov Model

Let the monthly de-meaned excess stock returns y_t be described as a k-state hidden Markov model (HMM) of Gaussian mixtures. Where S_t is an unobserved state variable following a Markov process.

$$y_t \sim N\left(0, \sigma_i^2\right) \tag{6}$$

$$\sigma_t^2 = \sum_{i=1}^{k} \sigma_i^2 S_{it} \tag{7}$$

subject to the restriction:

$$\sigma_1 < \sigma_2 < ... < \sigma_k \tag{8}$$

The probability for the Markov process to move from one state i at time $t-1$ to state j at time t is called transition probability, $p_{ij} = \Pr[S_t = j | S_{t-1} = i]$. The transition probabilities p_{ij} are collected in the transition matrix $\mathbf{P}$, which forms the nucleus of the Markov model.

$$\Pr[S_t = j | S_{t-1} = i] = p_{ij};\ i, j = 1, ..., k \tag{9}$$

and

$$\sum_{J=1}^{k} p_{ij} = 1, \quad i = 1, ..., k \tag{10}$$

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{bmatrix} \tag{11}$$

This is a standard Markov switching model or regime switching model of Hamilton (1994). In our case we have chosen two states (k=2).[3]

3.2 Bayesian statistic

The fundamental idea behind Bayesian statistic is to condition on the observed data, Y, and regard the parameters, θ, as random variables. Suppose that $p(\theta)$ is a probability distribution of the parameter θ.

$$p(Y \mid \theta)\, p(\theta) = p(Y, \theta) = p(\theta \mid Y)\, p(Y). \tag{12}$$

The probability distribution of θ conditional on the observed data is expressed by Bayes theorem:

[3]We have also done estimations using three-state hidden Markov model. The results suggest that a two-state hidden Markov model being more appropriate. The results of the estimations are available on request.

$$p(\theta \mid Y) = \frac{p(Y \mid \theta)\, p(\theta)}{p(Y)}. \tag{13}$$

where $p(\theta)$ is the prior probability density function and describes the information in θ without any knowledge about the data, Y. $p(\theta \mid Y)$ is the posterior probability density *function* and gives a description of what is known about θ given the data, Y. Given the data, Y, the conditional probability distribution $p(Y \mid \theta)$ can be seen as a function of the parameters θ and this is the likelihood function of θ, $L(Y \mid \theta)$. As $p(Y)$ is constant the posterior probability density function is proportional to the likelihood function times the prior probability density function.

$$p(\theta \mid Y) \propto L(Y \mid \theta)\, p(\theta). \tag{14}$$

This yields an appealing property of the Bayesian approach as we do not need a specification of the likelihood function to sample from the marginal distributions of the parameters. In general, the joint posterior distribution, $p(\theta \mid Y)$, is unknown, but can be simulated using Gibbs sampling

3.3 The Gibbs sampler

Gibbs sampling is a special case of the Metropolis-Hastings algorithm, see Metropolis et al (1953) and Hastings (1970), the difference being that in Gibbs sampling we always accept the candidates. Its breakthrough came with the papers by Gefland and Smith (1990) and Gefland et al (1990).[4] The Gibbs sampler provides the analyst with the tools to sample from the marginal distribution of the parameters of interest. The idea behind the algorithm is to sample from the conditional distribution of the parameter space $\{\theta_1, \theta_2, \ldots, \theta_k\}$. After specifying initial values $\sigma_1^{(0)}$, $\sigma_1^{(0)}$, $\mathbf{P}$ and augment the data with a randomly generated state vector $\mathbf{S}$, the parameters are generated recursively by cycling through the conditional distributions.

Step 1: Specify arbitrary initial values, $\left(\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_k^{(0)}\right)$, and set $n = 1$.
Step 2: Cycle through the full conditionals by drawing:

(1) $\theta_1^{(n)}$ from $\left[\theta_1 \mid \theta_2^{(n-1)}, \ldots, \theta_k^{(n-1)}\right]$

(2) $\theta_2^{(n)}$ from $\left[\theta_2 \mid \theta_1^{(n)}, \theta_3^{(n-1)}, \ldots, \theta_k^{(n-1)}\right]$

$\vdots$

(k) $\theta_k^{(n)}$ from $\left[\theta_k \mid \theta_1^{(n)}, \ldots, \theta_{k-1}^{(n)}\right]$

[4] See also Casella and George (1992) for an explanation of the Gibbs sampler

`Step 3: set` $n = n + 1$, `and go to step 2 until` $n = N$.

The simulated series are ergodic Markov chains, and so after a large number of iterations, the simulated series represent drawings from their respective marginal distributions. The recursion is continued in order to generate samples of each parameter from their marginal distributions. In our case N is set to 20.000 iterations and we obtain the sample values $\left\{\theta_i^{(N)}\right\}_{i=1}^{k}$.[5] The first M iterations when the chains have not converged are discarded leaving us with a sample of $(N - M)$ useful iterations. For a large number, $(N - M)$, the simulated values, $\left\{\theta_i^{(N-M)}\right\}_{i=1}^{k}$, can be treated as an approximate sample from marginal distribution of the parameters, see Tierney (1994).

3.4 Priors and prior distributions

We use conjugate prior distributions and the specification of the prior parameters and their distributions follows from Albert and Chib (1993), Tanner (1996), Kim et al (1998a) and Robert and Casella (1999).[6] Each row of the transition probability matrix $\mathbf{P}$ is generated as random draws from a Dirichlet distribution.[7]

$$\mathbf{P}(i) \sim D\left(u_{i1} + n_{i1}, u_{i2} + n_{i2}\right), \qquad i = 1, 2 \tag{15}$$

where n_{ij}, is the number of transitions from state i to state j. We consider u_{ij}, $i = 1, 2$, $j = 1, 2$, as non-informative priors and set them equal to 1.

In order to satisfy the constraint, $\sigma_1^2 < \sigma_2^2$, we need to first generate σ_1^2 and re-define σ_2^2 conditional on σ_1^2.

$$\sigma_2^2 = \sigma_1^2 \left(1 + h\right) \tag{16}$$

where $h > 0$. Where σ_1^2 and $\overline{h} = (1 + h)$ are random draws from the inverse-gamma, IG,

[5] This is a computer intensive simulation. All simulations are done in MATLAB and the estimation time is approximately 6 hours on a standard Intel PII 450 MHz.

[6] See also Gilks et al (1996) *"Markov Chain Monte Carlo in Practice"*.

[7] The Dirichlet density function has the property that it can assume a large number of various shapes in the sample space [0, 1]. Another property of the multivariate Dirichlet distribution is that the sampled probabilities sum to unity. This makes the Dirichlet distribution family very suitable in representing any experiments on multivariate continuous random variables in the [0, 1] space.

distribution family.[8]

$$Y_{1t} = \frac{y_t}{\sqrt{\sigma_1^2 \left(1 + S_{2t} h\right)}} \tag{17}$$

$$\left[\sigma_1^2 \mid \widetilde{Y}_{1T}, \widetilde{S}_T, \widetilde{\theta}_{j \neq \sigma_1^2}\right] \sim IG\left(\frac{v_1 + T}{2}, \frac{\delta_1 + \sum_{t=1}^{T} Y_{1t}^2}{2}\right), \tag{18}$$

$$Y_{2t} = \frac{y_t}{\sqrt{\sigma_1^2}} \tag{19}$$

We define N_2 as the number of times state 2 occurs $N_2 = \{t : S_t = 2\}$ and T_2 is the sum of the elements in N_2.

$$\left[\overline{h} \mid \widetilde{Y}_{2T}, \widetilde{S}_T, \widetilde{\theta}_{j \neq h}\right] \sim IG\left(\frac{v_2 + T_2}{2}, \frac{\delta_2 + \sum_{t=1}^{N_2} Y_{2t}^2}{2}\right) I_{[\overline{h} > 1]}, \tag{20}$$

We use non-informative priors and set v_1, v_2, δ_1, and δ_2 equal 1.[9]

3.5 Missing data simulation

As we cannot observe the two regimes we have to regard the states as a missing data problem. However, we can compute the probability of a given observation y_t belong to state i, $i = 1, 2$, and from this information construct forecast probabilities of which state j, $j = 1, 2$, observation y_{t+1} belong to. The probabilities are computed for all observations y_t, $t = 1...T$, with the local updating algorithm of Robert (1993).[10] This is repeated for every Gibbs sweep. The local updating algorithm is a forward algorithm in which each state is simulated from the full conditional ($1 \leqslant i \leqslant k$).

$$p\left(S_1 = i \mid S_2, ..., \mathbf{P}\right) \propto \rho_i p_{iS_2} f\left(y_1 \mid 0, \sigma_i\right) \tag{21}$$

$$p\left(S_1 = i \mid ..., S_{t-1}, S_{t+1}, ..., \mathbf{P}\right) \propto p_{S_{t-1}i} p_{iS_{t+1}} f\left(y_j \mid 0, \sigma_i\right), \quad (1 < t < T) \tag{22}$$

[8] A draw from any (inverse)gamma distribution is always positive. This makes them an ideal distribution family for generating second order moments.

[9] Non-informative prior refers to a prior with little influence on the shape of a prior distribution.

[10] This is a considerable more efficient algorithm than the forward backward algorithm suggested by Kim and Nelson (1998).

$$p\left(S_T = i \mid ..., S_{T-1}, \mathbf{P}\right) \propto p_{S_{T-1}i} f\left(y_T \mid 0, \sigma_i\right) \tag{23}$$

Where $(\rho_i, ..., \rho_k)$ is the stationary distribution of the transition matrix $\mathbf{P}$ and $f\left(\cdot \mid 0, \sigma_i\right)$ denotes the density of the normal distribution, see Hamilton (1994). Thus, the ρ_i's are computed from the transition matrix, $\mathbf{P}$, at each sweep of the Gibbs sampler. Using the probabilities from the local updating algorithm we generate the two states $S = 1, 2$, from a two point distribution. The states are generated by drawing random numbers from a uniform distribution. We set the state $S_t = 1$; if the generated number is less or equal to $p_1/\left(p_1 + p_2\right)$. If it is greater than $p_1/\left(p_1 + p_2\right)$, we set $S_t = 2$. This is repeated for all observations $t = 1...T$.

3.6 A Bayesian approach to variance ratio test

The following two resampled based variance ratio tests have been suggested by Kim et al (1998a). At the end of each sweep of the Gibbs-sampling algorithm the following procedure is computed:

Step 1: We divide the monthly returns y_t by the standard deviation σ_t in order to get the standardized returns y_t^*.

Step 2: Scramble the standardized returns y_t^* to yield a new randomized vector y_t^{r*}.

Step 3: Create a new series of de-standardized randomized monthly returns y_t^r by scaling the randomized-standardized returns y_t^{r*} by the standard deviation σ_t.

We now have four return series, first *original returns* y_t, second *standardized original returns* y_t^*, third *randomized standardized returns* y_t^{r*} and fourth *randomized de-standardized returns* y_t^r. Next we calculate the q-month variance ratio for the four return series.

3.6.1 Is the Variance Ratio test sensitive to randomization?

If we scramble a time series its time series properties will be destroyed and by construction a typical randomized series will behave as a random walk. The idea of the Gibbs sampler is to preserve the historical pattern of the time variation in volatility of the data. However as this is repeated 20.000 times the volatility and the volatility structure, or state vector, is subject to sampling variation. Thus, the re-standardized series are subject to both

randomization and parameter uncertainty. Computing the VR and repeating 20.000 times will result in a distribution of VR_q^r statistics representing the null hypothesis of mean reversion due to randomization. These values are compared to the variance ratio statistic computed from the original data, VR_q.

3.6.2 Is the Variance Ratio test sensitive to randomization and standardization?

The second test is based upon the standardized returns. We first compute the variance ratio test on the standardized returns, VR_q^*. This is a variance ratio test statistic filtered by the historical pattern of the volatility. However as mentioned above each sweep of the Gibbs sampler provides a new sample of parameters and after a large number of iterations we have an empirical distribution of the variance ratios VR_q^* computed on the standardized returns. This distribution is compared with the empirical distribution of VR_q^{r*} computed on the standardized randomized returns. The later distribution is representing the null hypothesis as a randomized series will behave as a random walk. Hence, if the VR_q^*-test is not sensitive to the filtering of volatility and the volatility structure, the distribution of the VR_q^* will be below the distribution of the VR_q^{r*} on randomized standardized returns.

The significance levels of the two one-sided VR-tests, of H_{02}: no mean reversion against H_1: mean reversion, are estimated as the fraction of VR for the artificial returns that fall below the VR of the original historical returns. Thus we will have two tests for every q-month horizon. First, a test based on original returns,

$$P(H_0) = \frac{\#(VR_q < VR_q^r)}{(N-M)} \tag{24}$$

Second, a test based on standardized returns,.

$$P(H_0) = \frac{\#(VR_q^* < VR_q^{r*})}{(N-M)} \tag{25}$$

At the end of the Gibbs sampling we will have 20.000 realizations of each of the two tests for each of the 20 q-month test horizons. An advantage with our Bayesian approach is that we are able to account for the parameter uncertainty in θ as well as the effect of the randomization.

4 Empirical results

4.1 Data

We use 80 years of monthly value weighted Swedish stock market returns including dividends and the Swedish risk-free rate from December 1918 to December 1998. All data are from the Frennberg and Hansson (1998) database. From this data set we compute the monthly excess return of the Swedish stock market and subtract the mean of the excess return to get de-meaned excess return.

4.2 Bayesian inference on parameter estimates

The convergence of the Gibbs sampler or burn in time is determined via monitoring techniques. We run several Gibbs sequences and use different values of the priors in order to reveal possible slow mixing of the Markov chain. We monitor all parameters of the Gibbs sequence, Figure 2, and the convergence is based on the worst scenario, the parameter with the slowest mixing. Figure 2 displays the convergence, or mixing, as the average parameter value versus the number of iterations, for the transition probabilities, p_{11} and p_{22}, and the two variances. The variance parameters converge quickly, but the transition probabilities exhibits slow convergence. Thus the burn in time is based on the later and M is set to 8.000 iterations, leaving 12.000 Gibbs sequences from which to make statistical inference.

The stability of the states is quite clear from Figure 3a and Figure 3b. The graph, Figure 3a, is called assignment map and plot the assignment of the states at a given observation against the iterations as black for state 1 and white for state 2, see Robert and Mengersen (1998).[11] If there is no information at all the state vectors are random and the assignment map blurred. Figure 4 is a vizualization of the non-informative case and presents randomly generated states vectors. Figure 4a is the assignment map and Figure 4b the probability of state 1. If the Gibbs sampling algorithm has problems identifying the states the assignment map will have horizontal stripes. However, if the Gibbs sampler at each sweep assigns the same state to the same observation the assignment map will have vertical bars, see Figure 3a compared to Figure 4a.

[11] Robert and Mengersen refer to *allocation maps*. Recent literature (Bilio, Monfort and Robert (1999)) calls them *assignment maps*.

Our Gibbs sampler is able to find stable assignments for the data set, see Figure 3a. Thus, we have quite clear allocation of the low-volatility state and a bit blurred picture of the allocations to the high-volatility state. This is also confirmed by Figure 3b, the probabilities of a specific observation being allocated to state 1.

The mean, median and the 2.5 upper and lower percentiles of the posterior distribution of the transition probabilities are presented in Table 1. Given that we are in a specific regime S we can compute the expected duration of the regime by $1/\left(1-p_{ij}\right)$ conditional on $i=j$, see Kim and Nelson (1999) p 71-72. The last column in Table 1 shows the persistence or duration of a state. The expected duration of the states is 2.6 months and 1.7 months for state 1 and state 2. Both the duration of the states, Table 1, and the assignment map, Figure 3, indicates that the model frequently switches between the regimes with different volatility.

Table 1
Transition probabilities.

Parameter	Posterior mean	median	duration
p_{11}	0.620 [0.620, 0.620]	0.621	2.639
p_{22}	0.413 [0.413, 0.413]	0.413	1.704

Note: 2.5 and 97.5 percentiles within brackets

The mean, median and the 2.5% upper and lower percentiles of the conditional distributions of the estimated volatility parameters are presented in Table 2. There is a significant difference in the volatility between the two-states with 8.0% for state 1 and 36.6% for state 2. The posterior distributions of the volatility parameters are presented in Figure 5.

Table 2
Volatility.

Parameter	Posterior mean	median
σ_1	7.954 [6.677, 9.631]	7.876
σ_2	36.622 [30.749, 44.341]	36.259

Note: 2.5 and 97.5 percentiles within brackets

4.3 Variance Ratios

We will exemplify the sampled distributions of the different variance ratios using histograms of the results from the five-year horizon, $q = 60$ months. Figure 6 shows the distributions of the variance ratio test computed for the five-year horizon on the randomized standardized returns, the randomized de-standardized returns and the standardized original returns. The mean, median and 95% interval of the variance ratios for all twenty-investment horizons is presented in Table 3 and Table 4.

Table 3
Variance ratios of de-standardized returns.

Investment horizon, q (months)	Variance ratio VR(q)		
	Original	Scrambled	Prob. Value
2	1.164	1.001 [0.939, 1.068]	0.999
3	1.202	1.001 [0.909, 1.100]	0.999
4	1.237	1.001 [0.885, 1.128]	0.999
5	1.255	1.001 [0.869, 1.150]	0.999
6	1.265	1.001 [0.850, 1.168]	0.997
7	1.297	1.001 [0.833, 1.188]	0.998
8	1.324	1.001 [0.819, 1.206]	0.998
9	1.350	1.001 [0.806, 1.222]	0.998
10	1.389	1.001 [0.795, 1.235]	0.998
11	1.435	1.000 [0.783, 1.249]	0.999
12	1.484	1.000 [0.773, 1.265]	0.999
24	1.623	0.993 [0.677, 1.383]	0.998
36	1.532	0.982 [0.604, 1.467]	0.984
48	1.428	0.967 [0.545, 1.534]	0.951
60	1.269	0.951 [0.501, 1.592]	0.869
72	1.093	0.935 [0.457, 1.638]	0.733
84	0.901	0.918 [0.421, 1.684]	0.536
96	0.779	0.902 [0.387, 1.735]	0.413
108	0.763	0.888 [0.355, 1.769]	0.423
120	0.765	0.875 [0.329, 1.801]	0.449

Note: 2.5 and 97.5 percentiles within brackets

The probability values of the VR-test decrease as the horizon q increase. This is expected as the randomization of the returns leads to flatter distributions of the VR as

the investment horizon q increases. The maximum and minimum values of the original VR are 1.623 at 24 months and 0.763 at 108 months. This is an unexpected result especially as the high VR occur at 12, 24, and 36 months. Thus, it justifies our approach of utilizing computations of monthly VR with short-run horizons of 2-12 months and long-run horizons of 1 to 10 years.

Table 4
Variance ratios of standardized returns.

Investment horizon,	Variance ratio VR(q)		
q (months)	Standardized	Scrambled standardized	Prob. Value
2	1.179 [1.140, 1.218]	0.999 [0.936, 1.062]	1.000
3	1.247 [1.178, 1.317]	0.997 [0.906, 1.094]	1.000
4	1.303 [1.203, 1.403]	0.996 [0.883, 1.117]	0.999
5	1.342 [1.219, 1.470]	0.995 [0.865, 1.138]	0.999
6	1.371 [1.224, 1.525]	0.994 [0.847, 1.154]	0.999
7	1.418 [1.252, 1.593]	0.992 [0.831, 1.172]	1.000
8	1.467 [1.280, 1.663]	0.991 [0.817, 1.188]	1.000
9	1.513 [1.306, 1.727]	0.990 [0.807, 1.202]	1.000
10	1.567 [1.343, 1.799]	0.989 [0.796, 1.214]	1.000
11	1.628 [1.392, 1.877]	0.988 [0.782, 1.228]	1.000
12	1.691 [1.440, 1.954]	0.986 [0.771, 1.242]	1.000
24	2.004 [1.613, 2.418]	0.973 [0.668, 1.350]	1.000
36	2.049 [1.588, 2.548]	0.961 [0.599, 1.435]	0.999
48	2.010 [1.527, 2.548]	0.948 [0.542, 1.504]	0.997
60	1.853 [1.386, 2.390]	0.934 [0.489, 1.559]	0.988
72	1.644 [1.189, 2.187]	0.921 [0.452, 1.610]	0.960
84	1.444 [0.982, 2.002]	0.906 [0.416, 1.652]	0.899
96	1.321 [0.860, 1.890]	0.891 [0.384, 1.695]	0.845
108	1.309 [0.841, 1.886]	0.877 [0.354, 1.739]	0.841
120	1.310 [0.835, 1.896]	0.862 [0.329, 1.773]	0.843

Note: 2.5 and 97.5 percentiles within brackets

A general result is that the p-values from the standardized returns are lower then the p-values computed from the VR-test of the original returns. Our lowest p-value is 0.413 at 96 months horizon for standardized returns to be compared with the p-value of 0.845 for original returns. Our highest p-values are all from the short run horizons and the p-values

decay with the investment horizon. However, we cannot reject the null hypothesis of random walk for any horizon q and this result is robust to VR computed on standardized or de-standardized returns.

Frennberg and Hansson (1993) find support of mean reversion in the Swedish stock market and that the mean reversion increases with the length of the investment horizon. This they conclude indicates that the risk in the Swedish stock market decreases with the holding period. Our analysis offsets their result. By simply accounting for the heteroscedasticity of the data and taking estimation bias into account we can find no support of mean reversion. On the contrary the historical pattern of heteroscedacity in Swedish stock market excess returns can be characterized by two regimes, a tranquil and a volatile. We find that the historical pattern of heteroscedasticity affects the variance ratio test this finding is in line with what Kim et al (1998a) finds for the U.S. stock market 1926-1986. Thus, accounting for time-variation in volatility and estimation bias improves the variance ratio test.

5 Conclusion

This paper addresses the question if the Swedish stock market is subject to mean reversion. Previous studies find support of mean reversion in the Swedish stock market and that the mean reversion increases with the length of the investment horizon. However the results of these studies are controversial as they ignore the assumption of constant expected return. Resent research have found that the historical pattern of heteroscedasticity seriously affects the probability of the variance ratio test to reject the null hypothesis of random walk.

We model the well-known heteroscedasticity of the Swedish stock market return with a two-state hidden Markov model of normal mixtures. We use a Bayesian approach and estimate the model by Gibbs sampling, a computer intensive Markov chain Monte Carlo method. Our two-state hidden Markov model is clearly specified along with the priors and prior distributions employed in the Gibbs sampler. Further we use the information at each run of Gibbs sampler to compute variance ratios test on standardized as well as de-standardized returns.

Our analysis finds no support for mean reversion and we cannot reject the null hypothesis of random walk for any of the investment horizons. This result is robust to variance ratios computed on standardized or de-standardized excess returns. Our two-state regime

switching models of normal distributions capture the heteroscedasticity by a tranquil and a volatile state and suggests that mean reversion if found in the Swedish stock market can be explained by time-variation in volatility and estimation bias of the variance ratio test.

References

Albert, J. H., and Chib, S. (1993) "Bayesian 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.

Berg, L. and Lyhagen, J. (1998) "Short and Long Run Dependence in Swedish Stock Returns." *Applied Financial Economics*, **8**, 435-443.

Billio, M. Monfort, A. and Robert, C. P. (1999) "Bayesian Estimation of Switching ARMA Models." *Journal of Econometrics, **93**(2)* 229-255.

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

Cochrane, J. H. (1988) "How Big is the Random Walk in GNP?" *Journal of Political Economy, **96*** 893-920.

Frennberg, P. and Hansson, B. (1992) "Computation of a Monthly Index for Swedish Stock Returns 1919-1990." *Scandinavian Economic History Review,* **XL**(1) 3-27.

Frennberg, P. and Hansson, B. (1993) "Testing the Random Walk Hypothesis on Swedish Stock Prices: 1919-1990." *Journal of Banking and Finance, **17*** 175-191.

Gefland, A.E. and Smith, A.F.M. (1990) "Sampling Based Approaches to Calculating Marginal Densities." *Journal of American Statistical Association*, **85**. 398-409.

Gefland, A.E. Smith, A.F.M. and Lee, T.M. (1990) "Bayesian Analysis of Constrained Parameters and Truncated Data Problems using Gibbs Sampling." *Journal of American Statistical Association,* **87** 523-532.

Geman, S. and Geman, D. (1984) "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images." *IEEE Trans. Pattern Anal. Mach. Intell.*, **6** 721-741.

Gelman, A. and Rubin, D. B. (1992) "Inference from Iterative Simulation Using Multiple Sequences." *Statistics Science,* ***7*** 457-472.

Gilks, W. R. Richardson S. and Spiegelhalter D. J. (1996) *"Markov Chain Monte Carlo in Practice."* Chapman and Hall 1.st Ed.

Goldfeld, S. M. and Quandt, R. E. (1973) "A Markov Model for Switching Regressions." *Journal of Econometrics,* ***1*** 3-16.

Hamilton, D. J. (1989) "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica,* ***57****(2)* 357-384.

Hamilton, D. J.. (1994) *"Time Series Analysis."* Princeton University Press.

Hastings, W.K. (1970) "Monte Carlo Sampling Methods Using Markov Chains and Their Application." *Biometrika,* **57** 97-109.

Kim, C. J. Nelson , C. R. and Startz, R. (1998) "Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented Randomization." *Journal of Empirical Finance,* ***5****(2)* 131-154.

Kim, C. J. Nelson , C. R. and Startz, R. (1991) "Mean Reversion in Stock Prices? A Reappraisal of the Empirical Evidence." *Review of Economic Studies,* ***58*** 515-528.

Kim, C. J. and Nelson , C. R. (1998) "Testing for Mean Reversion in Heteroskedastic Data II: Autocorrelation Tests based on Gibbs-Sampling-Augmented Randomization." *Journal of Empirical Finance,* ***5****(2)* 385-396.

Lo, A. W. and MacKinlay, A. C. (1988) "Stock Market Prices do not follow Random Walks: Evidence from a Simple Specification Test." *The Review of Financial Studies* ***1*** 41-66.

Lo, A. W. and MacKinlay, A. C. (1989) "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation." *Journal of Econometrics* ***45*** 203-238.

Malliaropulos, D. and Priestley, R. (1999) "Mean Reversion in Southeast Asian Stock Markets." *Journal of Empirical Finance,* **6** 355-384.

Metropolis, N. Rosenbluth, A.W. Rosenbluth, M.N. Teller, A.H. and Teller, E. (1953) "Equations of State Calculations by Fast Computing Machines." *Journal of Chemical Physics*, **21** 1087-1092.

Poterba, J. and Summers, L. (1988) "Mean Reversion in Stock returns: Evidence and Implications." *Journal of Financial Economics*, **22** 27-60.

Robert, P. C. and Casella, G. (1999) "*Monte Carlo Statistical Methods."* Springer-Verlag, New York, Inc..

Robert, C P. Celeaux, G. and Diebolt, J. (1993) "Bayesian Estimation of Hidden Markov Models: a Stochastic Implementation." *Statistic and probability letters, **64*** 327-355.

Robert, C. P. Rydén, T and Titterington, D. M. (1999) "Convergence Controls for MCMC Algorithms, with Applications to Hidden Markov Models." *Journal of Statistics and Computational Simulations, **16*** 327-355.

Robert, C. P. and Mengersen, K. L. (1998) "Reparameterization Issues in Mixture Estimation and Their Bearings on the Gibbs Sampler." *Comput. Stat. Data Ana., **29*** 325-343.

Robert, C. P. (1996) "*Mixtures of Distributions: Inference and Estimation. Markov Chain Monte Carlo in Practice.*" Editors: Gilks, W., R., S., Richardson and D., J., Spiegelhalter. Chapman and Hall. 1:st Ed.

Tanner, A. (1996) *"Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions."* Springer Series in Statistics. Springer Verlag.

Tanner, A. and Wong, W. H. (1987) "The Calculation of Posterior Distributions by Data Augmentation" *Journal of American Statistical Association, **79*** 709-711.

Tierny, L. (1994) "Markov Chains for Exploring Posterior Distributions." *Technical Report No. 560 (Revisited), School of Statistics. University of Minnesota.*

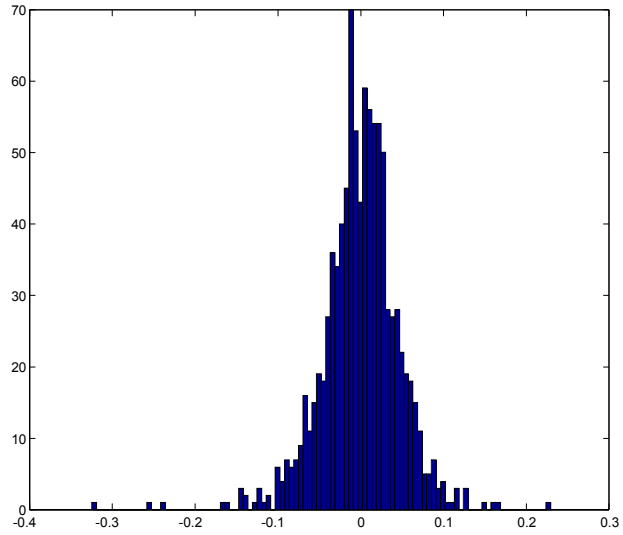

Figure 1: Percentile histogram of de-meaned monthly stock market excess returns 1919-1998.

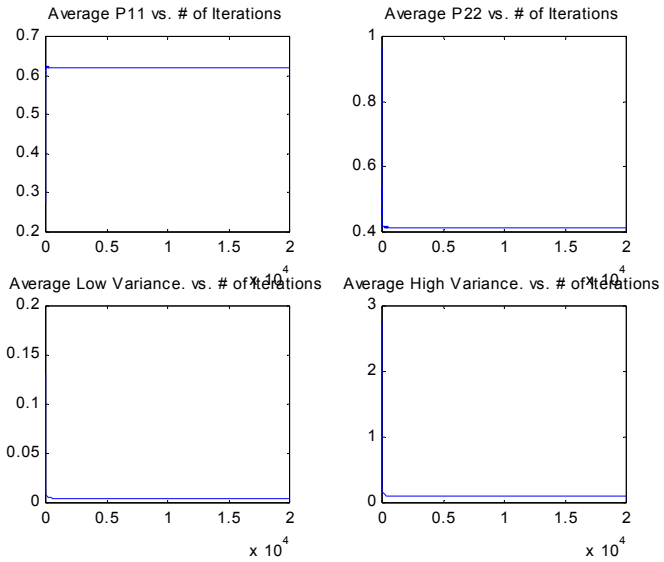

Figure 2: Ergodic avererage of estimated parameters vs. # itererations

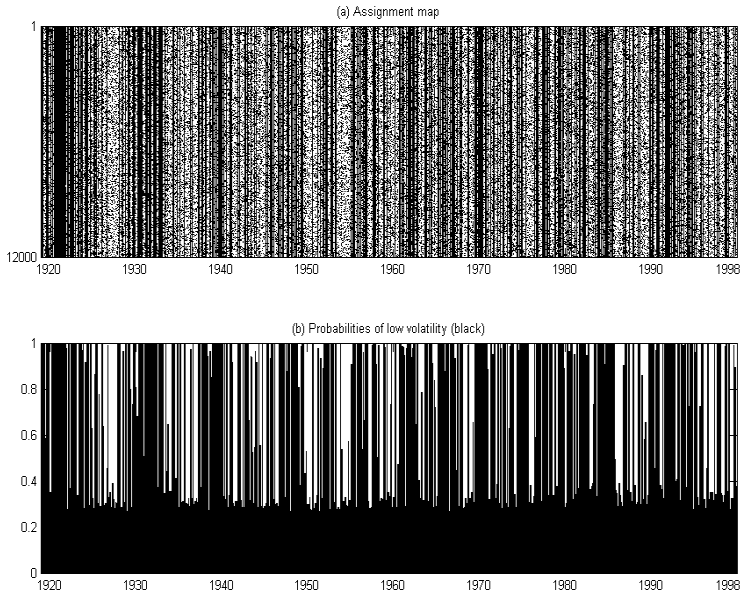

Figure 3: **(a)** Assignment map: Stacked state vectors after burn in time, state 1 black. **(b)** Probabilities of state 1 prevailing, state 1 black.

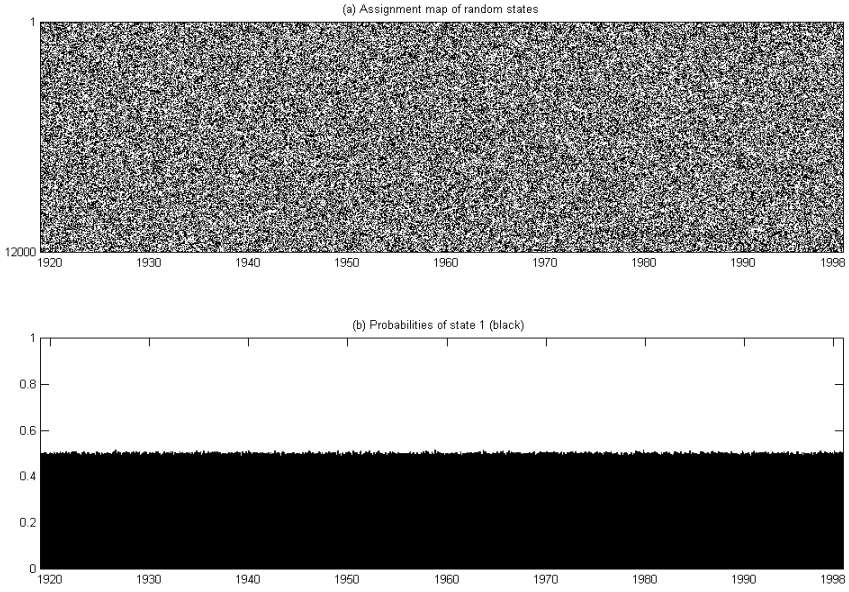

Figure 4: **(a)** Assignment map: Non informative randomly generated states. **(b)** Probabilities of a state 1 prevailing, state 1 black.

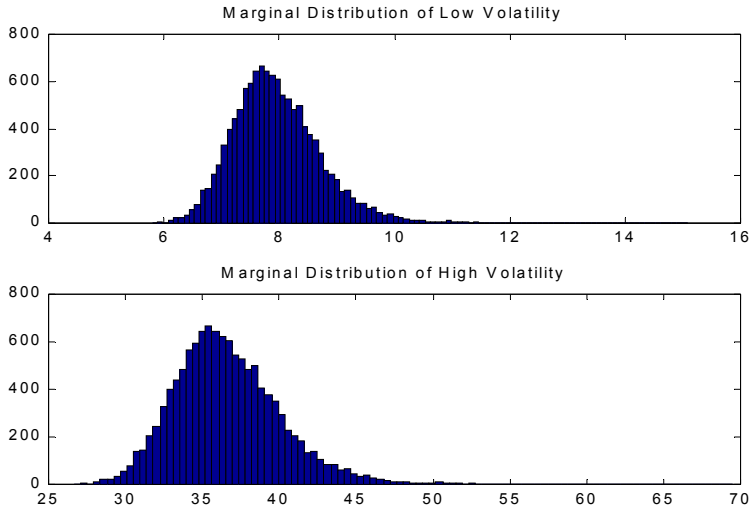

Figure 5: Posterior distribution of low and high volatility for Sweden

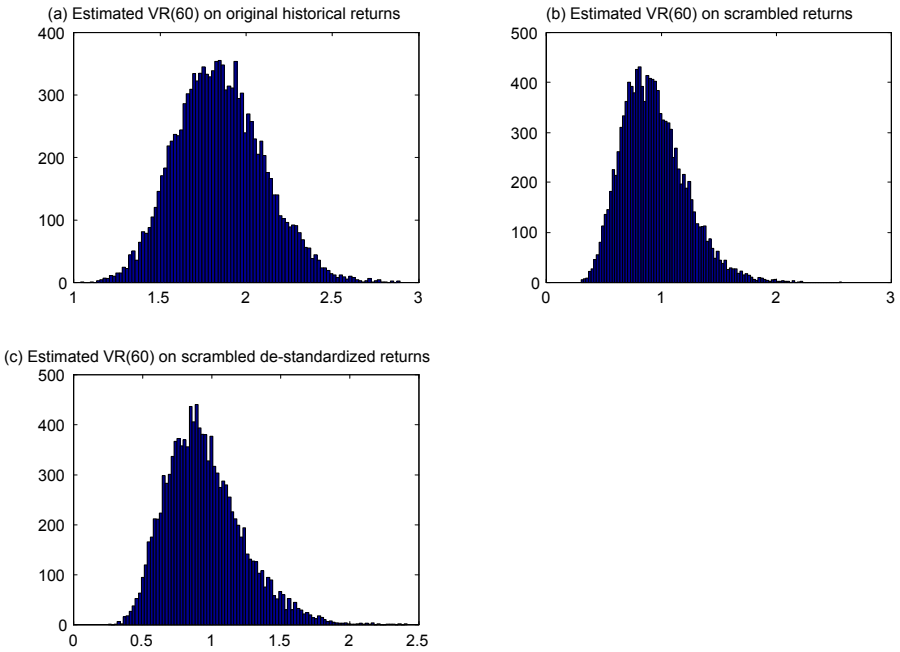

Figure 6: Posterior distribution of 5-year VR for Sweden