Markov Switching and Long Memory: A Monte Carlo Analysis

Wei-Choun Yu*

March 1, 2007

Abstract

This paper finds the close relationship between long memory and some forms of Markov-switching models. The simulation results suggest: (1) when the transition probabilities are closer to unity, it is more likely to generate long memory process; (2) magnitude of regime-switching plays an important role in generating long memory; and (3) process with switching in variance (disturbance) is much less likely to explain long memory process than switching in mean (intercept) and autoregressive coefficient. Therefore, given the observed high persistence in financial volatility data, volatility modeling by switching in mean and AR coefficient is preferred to that by switching in variance.

JEL classification: C22; C13; C15; C50

Keywords: Markov-switching model, long memory, Monte Carlo experiment, GPH test, SEMIFAR test

* Economics and Finance Department, Winona State University, Somsen 319E, Winona, MN 55987. Email: wyu@winona.edu. Tel: +1-507-457-2982. Fax: +1-507-457-5697.

1. Introduction

Since Hamilton (1989), the Markov-switching model, which could explore occasional but recurrent and endogenous in-sample structural change of macroeconomic and financial variables governed by latent states, has become a very popular nonlinear time series model. Since Granger and Joyeux (1980) and Ding, Granger, and Engle (1993), on the other hand, the long memory model (fractional integrated), which could capture the persistent dynamics of time series, has been proved not only a flexible linear model compared to *I*(1) or *I*(0), but also a robust out-of-sample forecasting model, especially in financial volatility. Nevertheless, the relationship and confusion between these two models have been getting more and more attention. For example, the persistence of asset return volatility may be overstated with the presence of structural change. Diebold and Inoue (2001) point out that Markov-switching model could more likely produce long memory persistence when the transition probability is closer to unity.

This paper uses Monte Carlo simulation to examine the relationship between Markov-switching and long memory model, in particular to see how several types of Markov-switching models could produce long memory behavior. We enlarge the scope of Markov-switching models considered in Diebold and Inoue (2001). Their simulation model is simply mean switching model

$$y_t = \mu_{St} + e_t \qquad (1)$$

where $e_t \sim N(0,\sigma^2)$ and $\mu_0 = 0, \mu_1 = 1$. Our generalized model includes not only switching in mean, but also switching in AR (1) coefficients and variance, from which the model could incorporate the short memory (AR (1) coefficients) and parameter uncertainty (variance) in the model.

There are at least two reasons for understanding the relationship between Markov-switching and long memory model. First, although the popularity of Markov-switching model for investigating the richer in-sample dynamics, like other nonlinear-model family, it suffers from the poor out-of-sample forecasting accuracy (Stock and Watson, 1999; and Terasvirta, 2005). The long memory model, on the other hand, has been proved robust in out-of-sample forecasts perspective by Andersen et al. (2003) and Granger and Hyung (2004). Therefore, it would be of interest to know what kinds of Markov-switching models could be substituted by long memory model in out-of-sample forecasts.

Second, in the literature of Markov-switching volatility models, generally there are two types of switching: switching in means, i.e. the fundamental model in Hamilton and Susmel (1994)[1]; switching in variance (disturbance), i.e. the fad models in Turner, Startz, and Nelson (1989) and Kim and Kim (1996)[2]. It would be of interest to see what kind of model could produce the observed long memory property in reality.

2. The Model and the Results

For the Monte Carlo experiment, we use the general autoregressive (1) 2-state Markov-switching model

$$\begin{aligned} y_t &= \mu_{St} + \phi_{St} y_{t-1} + e_t \\ e_t &\sim N(0, \sigma_{St}^2) \\ \mu_{St} &= \mu_1 S_{1t} + \mu_2 S_{2t} \\ \sigma_{St}^2 &= \sigma_1^2 S_{1t} + \sigma_2^2 S_{2t} \\ \phi_{St} &= \phi_1 S_{1t} + \phi_2 S_{2t} \end{aligned} \tag{2}$$

[1] They propose the Markov switching ARCH model where ARCH effect represents short memory. They find that stock market volatility can be explained by regime switching model triggered by fundamental factors, i.e. general business cycle.
[2] Kim and Kim (1996) argue that the presence of volatility regime switching is better explained by the short-lived fad factors, such as speculative bubbles, finance crisis and crash.

where $S_{1t}=1$, $S_{2t}=0$ if $S_t=1$ and $S_{1t}=0$, $S_{2t}=1$ if $S_t=2$. And the diagonal elements of transition probability $\Pr[S_t=1 \mid S_{t-1}=1]=p_{11}$ and $\Pr[S_t=2 \mid S_{t-1}=2]=p_{22}$ are generated by first-order Markov-switching process. In this paper, we consider eight situations of (p_{11}, p_{22}): (0.6,0.6), (0.7,0.7), (0.8,0.8), (0.9,0.9), (0.95,0.95), (0.98,0.98), (0.99,0.99), and (0.999,0.999).[3] First, we simulate 10,000 realizations of Markov-switching in mean (intercept) for three pairs of intercepts: (0.5, 1), (0.5, 2), and (0.5,3) where we fix the AR(1) coefficient = 0.5 and σ =1. Figure 1 presents the simulated series for intercept pair (0.5, 2) under the (p_{11}, p_{22}) = (0.99, 0.99). Second, we conduct the long memory parameter *d* estimations proposed by Geweke and Porter-Hudak (1983), henceforth GPH. The long memory parameter *d* could be estimated by the least-squared regression

$$\ln f(w_j) = \beta - d \ln[4\sin^2(\frac{w_j}{2})] + e_j \quad (3)$$

where *j*=1, 2, ..*m* and $f(w_j)$ is the spectral density at the *j*th Foirier frequency. The least squares estimate $\hat{d}$ is normally distributed in large sample. We consider $(10,000)^{0.7}$ and $(10,000)^{0.8}$ periodogram ordinates, which means that α = 0.7 and 0.8 respectively. Furthermore, to allow for the data-driven distinction of long memory, short memory, stochastic trends, and deterministic trends without any prior knowledge, Beran and Ocker (2001) proposed a semiparametric fractional autoregressive (SEMIFAR) model

$$\phi(L)(1-L)^{\delta}((1-L)^m y_t - g(i_t)) = \varepsilon_t \quad (4)$$

where δ is the long memory parameter, and $g(i_t)$ is a smooth trend function on [0,1] with $i_t = t/T$. The y_t must be differenced to achieve stationarity by parameter $d=\delta+m$.

[3] When the transition probability is closer to 1, it means that the state is more persistent. For simplicity, we only consider the situation when $p_{11}=p_{22}$. The result would not be affected if $p_{11} \neq p_{22}$.

The m determines whether the trend should be estimated from the original data (when m = 0) or the first difference (when m = 1). Finally, the Ljung-Box autocorrelation test for twentieth order serial correlation, Q(20) is also computed.

Table 1 reports the long memory estimation results. The findings suggest two things. First, when the transition probabilities are closer to unity, it is more likely to find evidence for long memory. This result is consistent with Diebold and Inoue (2001). Second, when the difference of regime's means is larger, it produces higher persistence. For example, when the (p_{11}, p_{22}) = (0.999, 0.999), three pairs of intercept (0.5, 1), (0.5, 2), and (0.5, 3) produce GPH (α = 0.7) d = 0.28, 0.43 and 0.56, respectively.

We also analyze switching in AR (1) coefficient (ϕ) for three pairs of coefficients: (0.4, 0.6), (0.4, 0.8) and (0.4, -0.4). Figure 2 shows the simulated series for ϕ pair (0.4, 0.8) under the (p_{11}, p_{22}) = (0.99, 0.99). Table 2 illustrates the estimation results. First, we find that switching in short memory parameters (for (0.4, 0.6) and (0.4, 0.8)), indeed, produces long memory. Nevertheless, it produces less degree of long memory compared with switching in mean. And when the AR(1) coefficients are more differentiated, it produces higher persistence. For example, when the (p_{11}, p_{22}) = (0.999, 0.999), two pairs of AR(1) (0.4, 0.6), (0.4, 0.8) produce GPH (α = 0.7) d = 0.10, and 0.31, respectively. But the pair of AR(1): (0.4, -0.4) could not produce any long memory.

Finally, we analyze switching in error term (e_t) by three pairs of σ : (1, 1.5), (1,3), (1,6). Figure 3 displays the simulated series for σ pair (1, 3) under the (p_{11}, p_{22}) = (0.99, 0.99). Table 3 shows the estimation results. For instance, when the (p_{11}, p_{22}) = (0.999, 0.999), we get the GPH (α = 0.7) d estimations of 0.02, 0.01 and 0.01 for the three pairs of σ (0.5, 1), (0.5, 2), and (0.5, 3), respectively. The findings suggest that switching-in-variance model is less likely to produce long memory dynamics. From Table 1 and Table 3's results, the paper suggests that Hamilton and Susmel's (1994) fundamental model explains observed long-memory volatility better than Kim and Kim's (1996) fad model.

It is worth nothing that Beran and Ocker's SEMIFAR model gives inconsistent estimation for our Monte Carlo simulation, i.e. Table 1: intercept pair (0.5, 2), (0.5, 3) and Table 3. It implies that their model, which mainly pays attention to detecting the linear trend in the series, does relatively poor job on estimating Markov-switching type of data generating process (DGP).

3. Conclusions

Our Monte Carlo experiment sheds light on the close relationship between long memory and some forms of Markov-switching models. When out-of-sample forecasting ability of Markov-switching model is still problematic, long memory model could provide an easy alternative for forecasting the true Markov-switching DGP. The findings also suggest: (1) when transition probabilities are closer to unity, it is more likely to generate long memory process; (2) magnitude of regime-switching plays a role in generating long memory; and (3) processes with switching in mean (intercept) and autoregressive coefficients are more likely to explain long memory process than switching in variance. Therefore, given the observed high persistence in financial volatility data, volatility modeling by switching in mean and AR coefficient is preferred to that by switching in variance.

References

Andersen, T. G., T. Bollerslev, F.X. Diebold, and P. Labys. (2003). "Modeling and Forecasting Realized Volatility," *Econometrica,* 71, pp. 529-626.

Beran, J., and D. Ocker. (2001). "Volatility of Stock-Market Indexes—An Analysis Base on SEMIFAR Models," *Journal of Business and Economic Statistics,* 19:1, pp 103-116.

Diebold, F. X., and A. Inoue. (2001). "Long Memory and Regime Switching," *Journal of Econometrics,* 105:1, pp. 131-59.

Ding, Z., C. W. J. Granger, and R. F. Engle. (1993). "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance,* 1, pp. 83-106.

Geweke, J., and S. Porter-Hudak. (1983). "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis,* 4, pp 221-237.

Granger, C.W.J., and R. Joyeux. (1980). "An Introduction to Long Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis,* 1, pp 15-39.

Granger, C.W.J., and N. Hyung. (2004). "Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns," *Journal of Empirical Finance,* 11, pp. 399-421.

Hamilton, J. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica,* 57:2, pp. 357-84.

Hamilton, J. and R. Susmel. (1994). "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, pp. 307-333.

Stock, J. H. and M. W. Watson. (1999). "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series," in *Cointegration, Causality, and Forecasting*, R. F. Engle and H. White. (Eds). Oxford University Press, Oxford.

Terasvirta, T. (2006). "Forecasting Economic Variables with Nonlinear Models," in *Handbook of Economic Forecasting*, G. Elliott, C.W.J. Granger and A. Timmermann. (Eds). Elsevier, Amsterdam.

Turner, C. M., R. Startz, and C. R. Nelson. (1989). "A Markov Model of Heteroscedasticity, Risk, and Learning in the Stock Market," *Journal of Financial Economics,* 25, pp. 3-22.

Table 1. Regime Switching in Intercept and Long Memory Parameter

Switching Factor: Intercept (μ_S)	Transition Probability	GPH Long Memory Parameter		SEMIFAR Parameters		
		GPH, a=0.7	GPH, a=0.8	d	AR(1)	Q(20)
	0.6, 0.6	0.02 (0.88)	0.14 (8.994)**	-0.013 (-0.54)	0.53	3600
	0.7, 0.7	0.01 (0.38)	0.11 (7.28)**	-0.05 (-2.13)*	0.55	3326
	0.8, 0.8	0.05 (2.05)*	0.17 (10.11)**	-0.01 (-0.35)	0.53	3788
0.5, 1	0.9, 0.9	0.07 (2.61)**	0.20 (12.25)**	0.04 (1.82)*	0.51	5081
	0.95, 0.95	0.16 (6.17)**	0.24 (14.47)**	0.12 (5.95)**	0.43	5623
	0.98, 0.98	0.20 (7.68)**	0.26 (15.82)**	0.20 (10.66)**	0.36	7616
	0.99, 0.99	0.24 (9.24)**	0.27 (16.63)**	0.20 (11.66)**	0.32	8233
	0.999, 0.999	0.28 (10.77)**	0.26 (15.84)**	0.42 (53.24)**	N/A	10302
	0.6, 0.6	0.01 (0.52)	0.13 (8.39)**	-0.027 (-1.06)	0.58	4297
	0.7, 0.7	-0.00 (-0.00)	0.17 (10.52)**	-0.01 (-0.51)	0.62	5435
	0.8, 0.8	0.06 (2.28)*	0.24 (14.78)**	-0.00 (-0.09)	0.68	8162
0.5, 2	0.9, 0.9	0.19 (7.39)**	0.38 (23.56)**	-0.16 (-8.76)*	0.87	14336
	0.95, 0.95	0.32 (12.12)**	0.48 (29.36)**	-0.26 (-20.63)**	0.94	25001
	0.98, 0.98	0.50 (19.05)**	0.52 (31.92)**	-0.34 (34.20)**	0.98	43144
	0.99, 0.99	0.53 (20.26)**	0.51 (30.73)**	-0.41 (-45.71)**	0.99	61404
	0.999, 0.999	0.43 (16.62)**	0.36 (21.78)**	0.36 (23.83)**	0.19	77367
	0.6, 0.6	0.06 (2.30)*	0.17 (10.50)**	-0.00 (-0.05)	0.6	4937
	0.7, 0.7	0.03 (1.18)	0.20 (12.19)**	0.04 (1.29)	0.63	6790
	0.8, 0.8	0.08 (3.02)**	0.28 (17.26)**	0.11 (4.01)**	0.67	10790
0.5, 3	0.9, 0.9	0.19 (7.36)**	0.44 (26.69)**	0.10 (3.91)**	0.77	19736
	0.95, 0.95	0.34 (12.87)**	0.56 (34.32)**	-0.05 (-3.35)**	0.9	34016
	0.98, 0.98	0.53 (20.34)**	0.62 (38.01)**	-0.22 (-21.74)**	0.98	70959
	0.99, 0.99	0.62 (23.94)**	0.63 (37.67)**	-0.30 (-32.80)**	0.99	91801
	0.999, 0.999	0.56 (21.26)**	0.45 (27.28)**	0.43 (30.34)**	0.14	128345

1. In parentheses are t-statistics. * means 5% significant and ** means 1% significant.
2. Transition probability denotes the diagonal elements of transition matrix.
3. GPH test is based on Geweke and Porter-Hudak (1983).
4. SEMIFAR (Semiparametric Fractional Autoregressive) model is based on Beran and Ocker (2001). $\phi(L)(1-L)^{\delta}[(1-L)^{m} y_t - g(i_t)] = \varepsilon_t$. By using a nonparametric kernel estimate of $g(i_t)$ instead of constant term μ. The method uses BIC to choose the short memory parameter *p*.
5. Ljung-Box test statistics for twentieth order serial correlation, Q(20).

Table 2. Regime Switching in AR(1) and Long Memory Parameter

Switching Factor: AR(1) coefficient (ϕ_S)	Transition Probability	GPH Long Memory Parameter		SEMIFAR Parameters		
		GPH, a=0.7	GPH, a=0.8	d	AR(1)	Q(20)
0.4, 0.6	0.6, 0.6	-0.03 (-1.05)	0.12 (7.33)**	-0.066 (-2.58)**	0.58	3552
	0.7, 0.7	0.03 (1.28)	0.12 (6.80)**	0.40 (50.90)**	N/A	3107
	0.8, 0.8	-0.01 (-0.27)	0.13 (7.81)**	-0.05 (-1.94)*	0.55	3457
	0.9, 0.9	0.07 (2.69)**	0.15 (9.04)**	-0.01 (-0.28)	0.53	3840
	0.95, 0.95	0.10 (3.68)**	0.20 (12.20)**	0.09 (4.26)**	0.43	4507
	0.98, 0.98	0.10 (3.58)**	0.26 (15.82)**	0.07 (3.09)**	0.45	4225
	0.99, 0.99	0.09 (3.46)**	0.19 (11.26)**	0.06 (2.96)**	0.45	3824
	0.999, 0.999	0.10 (4.00)**	0.20 (12.12)**	0.07 (3.18)**	0.46	4736
0.4, 0.8	0.6, 0.6	0.13 (5.12)**	0.23 (14.22)**	0.014 (0.51)	0.6	6228
	0.7, 0.7	0.07 (2.67)**	0.23 (14.10)**	0.01 (0.44)	0.63	6954
	0.8, 0.8	0.15 (5.80)**	0.29 (17.66)**	0.02 (0.77)	0.66	9103
	0.9, 0.9	0.15 (5.76)**	0.33 (19.87)**	-0.05 (-2.00)*	0.75	11316
	0.95, 0.95	0.22 (8.55)**	0.40 (24.39)**	-0.23 (-15.11)**	0.91	15694
	0.98, 0.98	0.32 (12.20)**	0.47 (28.34)**	-0.24 (-17.84)**	0.93	23585
	0.99, 0.99	0.33 (12.51)**	0.45 (27.19)**	0.36 (19.66)**	0.35	25195
	0.999, 0.999	0.31 (12.03)**	0.43 (25.90)**	-0.29 (-20.98)**	0.93	29582
0.4, -0.4	0.6, 0.6	0.017 (0.65)	0.02 (0.95)	0.03 (3.77)**	N/A	42.97
	0.7, 0.7	0.01 (0.24)	0.05 (3.22)**	0.09 (7.19)**	-0.07	99.15
	0.8, 0.8	0.01 (0.33)	0.05 (2.91)**	0.09 (7.40)**	-0.09	128
	0.9, 0.9	0.03 (1.21)	0.09 (5.37)**	0.18 (15.48)**	-0.14	331
	0.95, 0.95	0.09 (3.49)**	0.10 (6.21)**	0.19 (16.55)**	-0.17	370
	0.98, 0.98	0.12 (4.42)**	0.13 (7.66)**	0.24 (21.88)**	-0.2	660
	0.99, 0.99	0.22 (8,54)**	0.19 (11.26)**	0.24 (23.04)**	-0.27	879
	0.999, 0.999	0.12 (4.72)**	0.12 (7.25)**	0.19 (16.42)**	-0.15	680

1. In parentheses are t-statistics. * means 5% significant and ** means 1% significant.
2. Transition probability denotes the diagonal elements of transition matrix.
3. GPH test is based on Geweke and Porter-Hudak (1983).
4. SEMIFAR (Semiparametric Fractional Autoregressive) model is based on Beran and Ocker (2001). $\phi(L)(1-L)^{\delta}[(1-L)^{m} y_t - g(i_t)] = \varepsilon_t$. By using a nonparametric kernel estimate of $g(i_t)$ instead of constant term μ. The method uses BIC to choose the short memory parameter p.
5. Ljung-Box test statistics for twentieth order serial correlation, Q(20).

Table 3. Regime Switching in Variance and Long Memory Parameter

Switching Factor: Sigama (σ_S)	Transition Probability	GPH Long Memory Parameter		SEMIFAR Parameters		
		GPH, a=0.7	GPH, a=0.8	d	AR(1)	Q(20)
	0.6, 0.6	0.05 (1.83)*	0.16 (9,98)**	0.00 (0.11)	0.51	3571
	0.7, 0.7	0.03 (1.16)	0.11 (6.68)**	0.00 (0.21)	0.5	3288
	0.8, 0.8	0.04 (1.43)	0.12 (7.18)**	0.40 (51.20)**	N/A	3153
1, 1.5	0.9, 0.9	0.06 (2.12)*	0.12 (7.02)**	0.01 (0.38)	0.49	3384
	0.95, 0.95	0.05 (1.88)*	0.14 (8.29)**	0.40 (51.14)**	N/A	3072
	0.98, 0.98	0.03 (1.29)	0.15 (8.91)**	-0.02 (-0.73)	0.53	3580
	0.99, 0.99	0.05 (1.72)	0.12 (7.10)**	-0.02 (-0.79)	0.52	3374
	0.999, 0.999	0.02 (0.64)	0.15 (8.85)**	0.40 (51.64)**	N/A	3333
	0.6, 0.6	0.03 (0.95)	0.12 (7.52)**	0.40 (51.36)**	N/A	3207
	0.7, 0.7	0.03 (1.35)	0.16 (9,56)**	-0.01 (-0.46)	0.53	3613
	0.8, 0.8	0.05 (1.82)*	0.16 (9,65)**	-0.03 (-1.13)	0.53	3391
1, 3	0.9, 0.9	0.07 (2.85)**	0.16 (9.48)**	0.00 (-0.17)	0.5	3312
	0.95, 0.95	0.01 (0.23)	0.12 (7.60)**	0.41 (51.77)**	N/A	3240
	0.98, 0.98	-0.03 (-1.27)	0.13 (8.08)**	-0.08 (-3.19)**	0.6	3686
	0.99, 0.99	0.02 (0.71)	0.15 (8.84)**	0.02 (0.78)	0.49	3468
	0.999, 0.999	0.01 (0.18)	0.13 (8.04)**	-0.08 (-3.19)**	0.59	3500
	0.6, 0.6	0.04 (1.50)	0.16 (9.45)**	0.02 (0.91)	0.49	3698
	0.7, 0.7	0.012 (0.45)	0.12 (7.23)**	-0.08 (-2.95)**	0.57	3364
	0.8, 0.8	0.05 (1.84)*	0.14 (8.64)**	0.39 (49.54)**	N/A	3135
1, 6	0.9, 0.9	0.04 (1.38)	0.13 (7.92)**	-0.01 (-0.4)	0.51	3154
	0.95, 0.95	0.07 (2.58)**	0.13 (8.17)**	0.39 (49.34)**	N/A	2919
	0.98, 0.98	0.08 (2.99)**	0.14 (8.37)**	0.01 (0.45)	0.49	3322
	0.99, 0.99	0.04 (1.69)	0.13 (7.68)**	-0.01 (-0.31)	0.5	3308
	0.999, 0.999	0.01 (0.32)	0.11 (6.42)**	0.39 (50.09)**	N/A	3105

1. In parentheses are t-statistics. * means 5% significant and ** means 1% significant.
2. Transition probability denotes the diagonal elements of transition matrix.
3. GPH test is based on Geweke and Porter-Hudak (1983).
4. SEMIFAR (Semiparametric Fractional Autoregressive) model is based on Beran and Ocker (2001). $\phi(L)(1-L)^{\delta}[(1-L)^{m} y_t - g(i_t)] = \varepsilon_t$. By using a nonparametric kernel estimate of $g(i_t)$ instead of constant term μ . The method uses BIC to choose the short memory parameter p.
5. Ljung-Box test statistics for twentieth order serial correlation, Q(20).

Figure 1

Simulated Series: Switching:Intercept(0.5,2), p11=0.99, p22=0.99

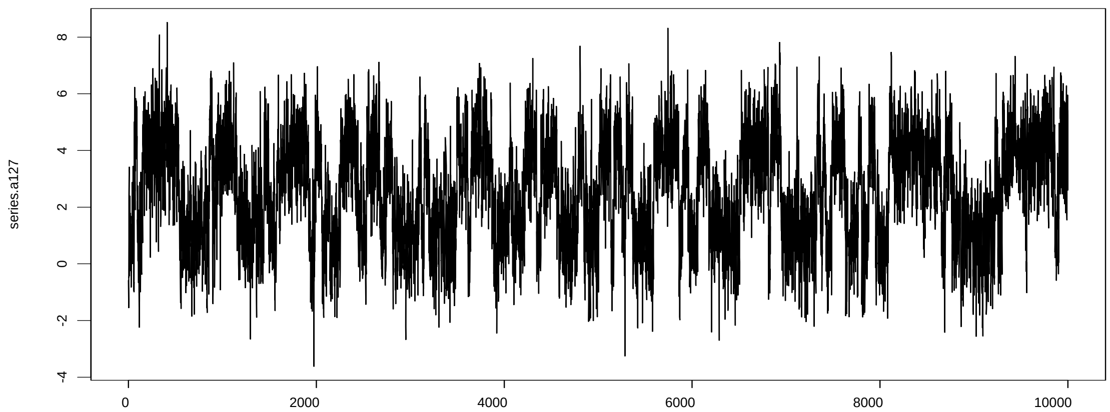

Figure 2

Simulated Series: Switching:AR(1)(0.4,0.8), p11=0.99, p22=0.99

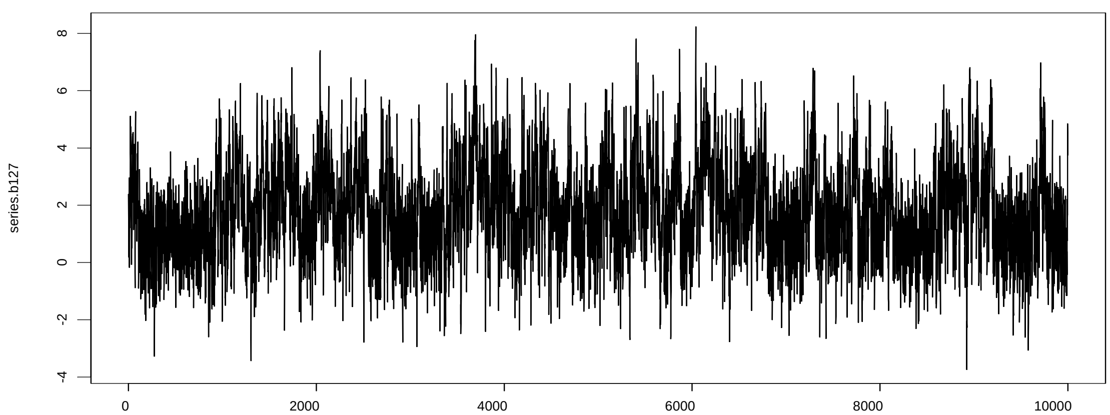

Figure 3

Simulated Series: Switching:sigma(1,3), p11=0.99, p22=0.99