



ELSEVIER

Journal of Econometrics 99 (2000) 139–171

JOURNAL OF  
Econometrics

www.elsevier.nl/locate/econbase

# Modeling long memory in stock market volatility

Ming Liu\*

*Department of Finance, The Chinese University of Hong Kong, Shatin, N.T.,  
Hong Kong, People's Republic of China*

Received 22 March 1996; received in revised form 1 February 2000; accepted 13 March 2000

---

## Abstract

Inspired by the idea that regime switching may give rise to persistence that is observationally equivalent to a unit root, we derive a regime switching process that exhibits long memory. The feature of the process that generates long memory is a heavy-tailed duration distribution. Using this process for volatility, we obtain a regime switching stochastic volatility (RSSV) model that we fit to daily S&P returns from 1928 through 1995 by means of the efficient method of moments estimation (EMM) method. Forecasts of RSSV volatility given past returns can be generated by reprojection, as we illustrate. The RSSV model is accepted according to the EMM chi-squared statistic. Using this statistic, we also evaluate several other models that have been proposed in the literature and some modifications to them. We find that models that exhibit long memory in volatility and heavy tails conditionally, as does the RSSV model, fit the data, whereas models without these characteristics do not. We also find weak evidence that suggests the presence of an additional short memory component of volatility over and above the long memory component. © 2000 Elsevier Science S.A. All rights reserved.

*JEL classification:* C15; C22; G12

*Keywords:* Long memory; Regime switching; Efficient method of moments; Stochastic volatility model

---

\* Tel.: + 86-852-2609-7757; fax: + 86-852-2603-6586.

*E-mail address:* mliu@se.cuhk.edu.hk (M. Liu).

It is now more or less commonly understood that financial time series can exhibit significant persistence in volatility. To model this persistence, we have most notably a large family of ARCH models, starting from Engle (1982). Recently, people have begun to realize that the observed persistence can be better captured by long memory processes as shown by the studies of Ding et al. (1993), Harvey (1993), Baillie et al. (1996), Bollerslev and Mikkelsen (1996) and Breidt et al. (1998), among others. All of these studies suggest that the correlation of the volatility of the return series decays slowly in a hyperbolic way, instead of exponentially as implied by any traditional model in the ARCH family.

The existence of long memory, especially in an economic time series, has been understood to arise from the aggregation of a cross section of time series. Granger (1980) proposed that the aggregation of a cross section of time series with different persistence levels would introduce long memory. His argument was used by Haubrich and Lo (1991) in explaining the long memory pattern in business cycles. Recently, the argument was used by Andersen and Bollerslev (1997) in studying the relationship between long memory in volatility and the aggregation of different information flows. The ingenuity of the Granger argument lies in the Beta distribution that he imposed on the distribution of the different persistence levels of the aggregated time series. As the number of aggregated series goes to infinity, Granger (1980) showed that the aggregation series exhibits the long memory pattern.

In this paper, we propose regime switching as another explanation for the observed long memory. As the regime switches in a heavy-tail manner, that is, when the duration of the regimes has a heavy-tail distribution, the long memory pattern appears in the autocorrelation function of the time series. This idea is applied to the case of stock market volatility. We construct a model, which we call the regime switching stochastic volatility (RSSV) model, to model the long memory pattern in stock market volatility. We argue that the arrival of major news triggers volatility jumps or switches in stock market volatility. In particular, when different news arrive at the market in a heavy-tail fashion, we observe long memory in the stock market volatility.

Diebold (1986) and Lamoureux and Lastrapes (1990), among others, pointed out that exogenous deterministic structural change could give rise to persistence observationally equivalent to that of the GARCH model. Hamilton and Susmel (1994) and Cai (1994) attach an endogenous regime switching component to the original ARCH models to study the potential impact of regime switching on the stock return dynamics. With both the ARCH component and the regime switching component in their model, their studies suggest that most of the observed persistence in the volatility process may actually come from the persistence of the regime.

Our argument can be considered an extension of the above regime switching argument. The regime switching model we use, however, differs from the

Hamilton regime switching model. Instead of modeling regime switching as a transition according to a Markov transition matrix, we model regime switching as a transition across i.i.d. regimes with the duration of each regime distributed according to a certain probability law. Under the assumption that the duration of a regime is distributed as any heavily tailed distribution,<sup>1</sup> we show that we have long memory with a magnitude determined by the tail index of this distribution. As a result, RSSV model proposed in this paper generates the long memory behavior in volatility as observed in the stock market.

Although algebraically quite maneuverable, the inference of the RSSV model is not simple. A simulation-based indirect inference method, specifically the EMM estimation, is employed to investigate the empirical relevance of the proposed model. The result of our estimation is quite positive. We find that the duration of a regime is considerably heavy-tail distributed and we cannot reject the hypothesis that the observed stock price dynamics actually arise from the RSSV model. The estimated heavy-tail index is also compared with the estimated long memory coefficient and we cannot reject the proposition that the observed long memory is a consequence of regime switching.

The paper is organized as follows. In Section 1, based on a mathematical argument of Levy (1983), we present the basic assumptions and the mechanism by which a regime switching model can imply long memory. In Section 2 we propose the regime switching stochastic volatility model and derive the functional limit of its integrated series as the sample size goes to infinity. In Section 3 we go over the basics of the EMM estimation and show our empirical results. Section 4 concludes the paper.

## 1. Introduction: Long memory and regime switching

In this section, based on a mathematical argument of Levy (1983), we propose that regime switching gives rise to the observed long memory phenomenon. The condition under which regime switching can exhibit long memory is stated. Some comparisons between this model and the existing regime switching model are also made. While the mathematical argument in the proof of Theorem 1.1 is taken from Levy (1983), all other results are ours. The model we give in this section is not intended to be specific so that it can be applied in other different circumstances where long memory is of concern.

---

<sup>1</sup> Heavy tail usually means a probability distribution with a tail probability larger than the normal distribution in the literature. However, here by heavy tail we mean a distribution with a polynomial tail as defined in Assumption 1.1.

1.1. Long memory – a consequence of regime switching

Long memory has been used to model the persistence of stationary economic data ever since the work of Granger and Joyeux (1980). From the perspective of the time domain, we define a long memory process as follows:

*Definition 1.1.* A weakly stationary process has long memory if its autocorrelation function  $\rho(\cdot)$  has a hyperbolic decay

$$\rho(t) \sim L(t)t^{2d-1} \quad \text{as } t \rightarrow \infty, \quad 0 < d < \frac{1}{2}, \quad L(t) \text{ is slow varying.} \quad (1)$$

In contrast, a short memory time series will have an autocorrelation function geometrically bounded as follows:

$$|\rho(t)| \leq Cr^{|t|} \quad \text{for some } C > 0, \quad 0 < r < 1. \quad (2)$$

It is easy to see that most stationary ARMA models such as the ARMA( $p, q$ ) model or the Markov chain regime switching model (as shown in Theorem 1.2) will only have a short memory.

There has been a vast literature on regime switching models from the mid-1980s onward due to the intuitive appeal of such models. As in the literature, here a regime may very well be related to certain latent state variables, which are relatively stable compared with the economic variables we are concerned with. The latent state variables may very well be taken simply as environmental parameters when economic agents try to make their more transient decisions. In the context of financial economics, this regime-specific variable may correspond to some specific monetary policy, in the case of interest rates; or correspond to market uncertainty levels as laid out by various pieces of major market news, in the case of stock market volatility. These regime-specific variables are relatively stable compared with the transience of everyday life, yet they change over regimes. Note here that by modeling the world as if there were only regime switchings, we intentionally abstract ourselves from the dynamics within a regime and focus ourselves on the dynamics across regimes.

In the regime switching model, discrete calendar time is divided according to different regimes. The regimes are related to some latent state variable  $W$  which takes on some regime-specific value  $W_k$  in the  $k$ th regime. We will call the time that the  $k$ th regime lasts the  $k$ th interarrival time and will denote it as  $T_k$ .

We assume that the duration of the regime (interarrival time, waiting time)  $T_k$  is i.i.d.<sup>2</sup> and that it has the tail probability behavior of Assumption 1.1. As we

---

<sup>2</sup> The i.i.d. assumption enables us to take the process as a renewal process and is essential for the analytical argument to take effect. And we believe a minor deviation from this assumption will not affect the correctness of the result.

will see from Theorem 1.1, this heavy-tail interarrival time distribution is the only type of distribution hypothesis that gives rise to the long memory behavior if long memory is truly related to regime switching.

*Assumption 1.1.* The interarrival time  $T_k$  is i.i.d. with a heavy-tail stationary distribution in the form of  $P(T_k > t) \sim t^{-\alpha}h(t)$ , as  $t \rightarrow \infty$  where  $1 < \alpha < 2$  and  $h(t)$  is slow varying.

This kind of distribution implies a possible clustering of regime switching; that is, in the view of an agent or market observer, there may appear a long-lasting tranquil period followed by a period of frequent regime switchings.<sup>3</sup> This kind of interarrival distribution can be endogenized by the crossing behavior of a random walk. When a certain key economic time series hits its threshold, it may trigger a jump in certain environmental parameters and thus cause regime switching. This kind of threshold rationale suggests that regimes tend to switch in a heavy-tail way.<sup>4</sup> The rise of regime switching suggests that there can be connections between the interarrival time  $T_k$  and the level of the variable  $W_k$  in that regime. The following assumption is therefore obviously an abstraction and a convenient starting point for a much more thorough study.

*Assumption 1.2.* The regime switching variable  $W_k \in \mathcal{W} \subset R$  is constant in a regime and i.i.d. with a 0 mean<sup>5</sup> and a finite second moment  $\sigma_W^2$  across regimes. Also, it is independent of the interarrival time  $T_k$ .

As we will see, the above two assumptions provide another way of endogenizing the regime switching without resorting to the Markov chain. We note that Assumption 1.2 does not rule out the case when the support of  $W_k$  contains a finite number of values.

---

<sup>3</sup> This can be seen from the possibility of some very long-lasting regimes and the relatively short mean of the interarrival times. Another way of looking at this is through duration dependence. This distribution implies negative duration dependence for regimes lasting long enough. See Diebold et al. (1993) for related literature.

<sup>4</sup> See Theorems 8.9.12, 8.9.13 and 8.9.14 of Bingham et al. (1987) on fluctuation theory. The basic idea is that, for most random walks, the ladder epoch will be distributed as a stable distribution with an index of  $\frac{1}{2}$ , and if we think of our interarrival time as determined by 2 or 3 independent ladder epoches superimposed together, then we could get an interarrival distribution as in our assumption. Of course, more involved triggering mechanics are possible and the empirical study of them in an economic time series could be challenging. Also see the qualitative threshold ARCH model as proposed in Gouriéroux and Monfort (1992).

<sup>5</sup> We assume a 0 mean only for analytical convenience and we can eliminate this zero mean assumption without doing any harm to our argument.

Given a sequence of i.i.d. interarrival times  $\{T_k\}$  with mean  $\mu$ , we can define a renewal process  $\{S_k\}$  as follows:

$$S_k = \begin{cases} S_0 + T_1 + \dots + T_k, & k = 1, 2, \dots, \\ S_0, & k = 0, \end{cases} \tag{3}$$

where we take  $S_0$  to denote the time between the start of our sample and the epoch when we observe the first regime switching. It is independent of the  $T_k$ 's and distributed as

$$P(S_0 = u) = \mu^{-1}P(T_k \geq u + 1) \quad \text{for } u = 0, 1, \dots \tag{4}$$

With the 0th renewal epoch having a steady-state distribution, we achieve the stationarity of  $\{S_k\}$ . It is easy to see that  $S_0$  is distributed as stable with an index  $\alpha - 1$ ; namely  $P(S_0 > t) \sim \alpha^{-1}\mu^{-1}t^{1-\alpha}h(t)$  by the Karamata Theorem. Imposing  $S_0$  as the starting point of our renewal sequence makes sense because, for any sample, the epoch 0 is not necessarily the exact point of regime switching; that is, we are not necessarily starting afresh with a new regime when the sample starts. Our sample is more likely to start from a regime which has already been in place for a long while, and the time between epoch 0 and the first observed break point may simply correspond to the concept of forward waiting time, which has the distribution of  $S_0$ .

We take the regime switching variable  $w_t$ , where  $w_t = W_k$  if time  $t$  lies in the  $k$ th regime, as a flow variable. In certain contexts, we may also be interested in studying the behavior of the first-order integration of  $w_t$ , i.e. the stock variable  $\bar{w}_t$ ,

$$\bar{w}_t = \sum_{s=1}^t w_s. \tag{5}$$

Note that  $\bar{w}_t$  is no longer regime switching.

With this notation in hand, we can state the main theorem.<sup>6</sup>

*Theorem 1.1.* Under the heavy-tail regime switching mechanism with an index  $\alpha$ , the regime switching variable will exhibit long memory property with  $d = 1 - \alpha/2$  in its autocorrelation function, i.e. the covariance between  $w_t$  and  $w_s$  will be of the form  $\sigma_w^2 \mu^{-1}(t - s)^{1-\alpha}h(t - s)$ , as  $t - s \rightarrow \infty$  and given  $t > s$ .

*Proof.* By the assumption of the independence of  $W_i$  and  $W_j$ ,  $w_t$  will be correlated with  $w_s$  only when they are in the same regime, i.e. when no regime switching happens during the period  $[s, t]$ . Thus, if we use  $C(t - s)$  to denote the

---

<sup>6</sup>The proof is adapted from a mathematical argument of Levy (1983).

covariance between the two epochs  $s$  and  $t$ ,<sup>7</sup> we have

$$\begin{aligned} C(t - s) &= Ew_t w_s \\ &= \sum_t E[w_k^2 I(S_{k-1} < t \leq S_k) I(S_{k-1} < s \leq S_k)] \\ &= \sigma_w^2 \sum_k P(S_{k-1} < s < t \leq S_k) \\ &= \sigma_w^2 P(S_0 \geq t - s). \end{aligned}$$

This last equation comes from an argument from pp. 369–370 of Feller (1971), and the theorem is established given the probability structure of  $S_0$ .  $\square$

Now we consider how long memory preserves itself when we apply a functional transformation to the original series  $w_t$ ,  $z_t = f(w_t)$ . It is easy to see that  $z_t$  is still i.i.d. across regimes. If  $z_t$  still has a finite variance, from the proof of the lemma it is apparent that the long memory property is invariant to this transformation and is uniquely determined by the underlying renewal structure. We state this observation as Lemma 1.1.

*Lemma 1.1. Given a time series  $w_t$  with a heavy-tail regime switching mechanism with index  $\alpha$ , any functional transformation of the original process preserving the property of finite variance also preserves the property of long memory, i.e. it has a long memory of the magnitude  $d = 1 - \alpha/2$ .*

As we can see from the above proof, we can have a corresponding correlation structure for any given distribution of regime duration. The converse question can be asked: Do we have a unique distribution of regime duration given that we know a particular autocorrelation structure? This question has been partially addressed by Parke (1995) in the case of the ARFIMA process. More generally, it is easy to see from the proof of the above lemma,

$$C(t - s) = \sigma_w^2 P(S_0 \geq t - s). \tag{6}$$

Hence,

$$P(S_0 = t) = \frac{1}{\sigma_w^2} [C(t) - C(t + 1)], \tag{7}$$

---

<sup>7</sup>The stationarity of process  $w_t$  is obvious.

implying

$$P(T_k = t) = \frac{\mu}{\sigma_w^2} [C(t - 1) - C(t)] - [C(t) - C(t + 1)]. \tag{8}$$

So if both the autocorrelation and its first difference are decreasing in time,<sup>8</sup> we can have a correspondingly unique distribution for the duration of a regime.<sup>9</sup>

Under Assumption 1.2,  $\alpha$  is in the interval (1, 2), and correspondingly the long memory coefficient is in the region  $(0, \frac{1}{2})$ . A similar argument to the above can show that, when  $\alpha$  is in the region (2, 3), we have a series with  $d$  in  $(-\frac{1}{2}, 0)$ .

The model in this section generates a covariance structure with long memory characteristics, and for some very general forms of long memory covariance structure, we can find a corresponding regime switching structure. This, however, should not be taken to mean that all long memory models have an *equivalent* counterpart in the regime switching form or vice versa. While two models may share the same covariance structure, they may differ significantly. As there are many jumps in the regime switching model, this model is more volatile than its ARFIMA counterpart. As shown in Levy (1983), the functional limit of  $\bar{w}_t$  is the Levy motion, which can be thought of as an aggregation of a diffusion part and a jump part. This limit is different from the temporal aggregation limit of the Gaussian ARFIMA model, which has been shown by Sowell (1990) as fractional Brownian motion. For more along these lines, see Levy (1983) or Taqqu and Levy (1986).

### 1.2. Comparison with the Markov regime switching model

In this subsection, we will compare our model with the popular Markov regime switching model. We will first state some theoretical results that demonstrate the inability of the Markov regime switching model to generate long memory behavior.

We start with a simplified version of the Markov chain regime switching model proposed in Hamilton (1989). There are finite states in this setup,

---

<sup>8</sup> To state loosely, if the a.c.f. is decreasing and convex.

<sup>9</sup> The corresponding duration probability structure of the simple integrated series  $(1 - L)^d w_t = \varepsilon_t$  has been derived by Parke (1995) as

$$P(T_k > t) = \frac{\Gamma(t + d)\Gamma(2 - d)}{\Gamma(t + 2 - d)\Gamma(d)}. \tag{9}$$

As can be easily seen from Sterling’s formula, it converges to  $t^{-(2-2d)}$  as  $t \rightarrow \infty$  and thus the duration probability structure is heavy tail with  $\alpha = 2 - 2d$ . Also, we want to note that for the simple integrated series, the mean of the duration is tied to the tail index in the corresponding renewal structure.



$\{W_1, W_2, \dots, W_n\}$ . The regime switching variable  $w_t$  takes on the above values according to a Markov transition matrix  $P = (p_{i,j})$ . We assume that the chain is irreducible and recurrent, and that there exists a stationary probability for the chain as  $P_0 = (p_1, p_2, \dots, p_n)'$ . The next theorem will show that, under the above conditions, the Hamilton Markov chain regime switching model is in the class of short memory models.

*Theorem 1.2. If the Markov chain is stationary, then the Markov chain regime switching model is in the class of short memory models.*

*Proof.* Given that there exists a stationary distribution, namely  $p_i \geq 0$ , then by 13.III of Romanovskii (1970), we know that the transition matrix  $P$  is regular. If we denote the eigenvalues of  $P$  as  $\{1, \lambda_1, \lambda_2, \dots, \lambda_k\}$ , then we know  $\lambda_i < 1 \forall i$ , and the eigenvalue 1 is simple.

Now let us consider the covariance between  $w_t$  and  $w_{t+s}$ , assuming that, without loss of generality,  $Ew_t = 0$ . We have

$$\text{cov}(w_t, w_{t+s}) = \sum_{i=1}^n W_i \sum_{j=1}^n p_{i,j}^s W_j, \tag{10}$$

where  $p_{i,j}^s$  denotes the  $ij$ th element of the  $s$  step transition matrix,  $P^{(s)} = P^s$ . Yet by Perron's formula (see Romanovskii, 1970), we know that  $p_{i,j}^s = p_j + \sum_{l=1}^k Q_{ijl}(s)\lambda_l^s$ , where  $Q_{ijl}(s)$  is some polynomial of finite degree in  $s$  so that

$$\text{cov}(w_t, w_{t+s}) = \sum_{i=1}^n \sum_{j=1}^n W_i W_j \sum_{l=1}^k Q_{ijl}(s)\lambda_l^s. \tag{11}$$

Since all  $\lambda_l < 1$ , it follows that we can find  $C$  and  $r$  such that  $\text{cov}(w_t, w_{t+s}) < Cr^s$ . This places it in the class of short memory models.  $\square$

Now suppose  $w'_t = w_t + u_t$  as in the Hamilton (1989) model. We assume, as in his model, independence between the regime switching part  $w_t$  and the random error part  $u_t$ . Then  $\text{cov}(w'_t, w'_{t+s}) = \text{cov}(w_t, w_{t+s}) + \text{cov}(u_t, u_{t+s})$ . It is easy to see that, as long as the error part  $u_t$  is in the short memory class,  $w'_t$  will also be in the short memory class.

Another important class of model we shall consider is regime switching in slope model. Here we will restrict ourselves to the following simple model:

$$x_t = w_t x_{t-1} + \varepsilon_t, \quad \text{with } \varepsilon_t, \text{ i.i.d. } (0, 1) \text{ variable} \tag{12}$$

and  $w_t$  switches among a finite number of values  $W_k$  all less than 1 in absolute value. Given the value of  $\{w_t\}$ , we have  $\text{cov}(x_t, x_{t+s}) = \prod_{i=t+1}^{t+s} w_i \text{ var } x_t$ . Since we can find some  $W$  greater than  $W_k \forall k$ , and less than 1 in absolute value, it is

obvious that this regime switching in slope model is also in the class of short memory models.

Despite the fact that the Markov chain regime switching model cannot give rise to the long memory phenomenon asymptotically, considerable persistence can be generated by assuming persistence in some regimes. When the sample size is small, both the switching behavior and the implied autocorrelation may be similar for the Markov regime switching model and the model proposed in the previous subsection.

## 2. Long memory and regime switching in volatility

Andersen and Bollerslev (1997) conjecture that long memory in volatility comes from an aggregation of an underlying ‘news’ arrival process with different persistence levels in the manner of Granger (1980), but they do not attempt to refute the key assumption of a Beta distribution. Some researchers (Backus and Zin, 1993) believe that long memory in a financial time series is spread from the aggregate variables, such as inflation rate, with their long memory again due to aggregation. Yet no study has been done along these lines.

We propose in this paper that regime switching causes long memory in stock market volatility. This argument is an extension of the argument in Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994), and Cai (1994), among others. All of these papers suggest that regime switching may be the main reason for the persistence of the volatility. We carry their argument a little further and suggest that persistence in the form of long memory is also caused by regime switchings. Based on the regime switching argument, we build a model that contains a regime switching part, and show that this model can yield persistence in the form of long memory in its volatility. We will examine the empirical relevance of the model in the next section.

Since the discovery of long memory in volatility, a variety of models have been used to model this long memory by various researchers. Among the most prominent of these models is the FIEGARCH model by Bollerslev and Mikkelsen (1996) and the long memory stochastic volatility model (LMSV) by Harvey (1993). The regime switching understanding of long memory allows a totally different way of modeling long memory in the volatility of stock market returns. Instead of employing the ARFIMA process in the volatility, in this paper, we adapt the regime switching argument in Section 1 to the stochastic volatility model. If we denote the innovation of price changes as  $u_t$ , and the market volatility as  $v_t = e^{w_t}$ , we will have

$$u_t = e^{w_t} \varepsilon_t, \quad (13)$$

where  $w_t$  is the regime switching variable bearing an interarrival structure satisfying Assumption 1.2 and independent of  $\varepsilon_t$ .  $\varepsilon_t$  is i.i.d. with mean 0 and finite

variance. Here we assume, for analytical convenience, that the interarrival time is independent of the volatility regime switching variable  $w_t$ . We may occasionally denote  $\varepsilon_t$  as  $\varepsilon_{k,s}$ , where  $k$  denotes that the epoch  $t$  is in the  $k$ th regime, and  $s$  denotes the spent waiting time of the current regime. Subsequently, we will call the above model the regime switching stochastic volatility (RSSV) model.

Because of the independence of  $\varepsilon_t$ , it is easy to see that the correlation between different time points of the time series  $u_t$  is 0, and the price series integrated from  $u_t$  is a martingale. Applying Lemma 1.1, the model has the ability to yield long memory in volatility if  $w_t$  satisfies Assumption 1.1. As in Ding et al. (1993) and in Harvey (1993), we can also look at the autocorrelation of the time series,  $|u_t|$ .  $|u_t|$  can be written as follows:

$$|u_t| = e^{w_t} E(|\varepsilon_t|) + e^{w_t} (|\varepsilon_t| - E(|\varepsilon_t|)). \quad (14)$$

Because the correlation between the first term and the second term is 0, and the second term has a time series autocorrelation 0, the autocorrelation of  $|u_t|$  depends on the autocorrelation of the first term  $e^{w_t} E(|\varepsilon_t|)$ . Lemma 1.2 implies that  $e^{w_t}$  is a regime switching variable satisfying Assumption 1.1, as  $w_t$  is such a variable. Therefore, the absolute series exhibits long memory with  $d = 1 - \alpha/2$ . This line of argument can easily be extended to other power series of  $u_t$  or some other transformation of the series. The above discussion is put in Theorem 2.1.

*Theorem 2.1. If the regime duration of our RSSV model satisfies Assumptions 1.1 and 1.2, there is long memory in the volatility series with a magnitude of  $d = 1 - \alpha/2$ .*

The rest of this section will be devoted to the temporal aggregation of the above model. It is interesting to see that, regardless of regime switching and even in a quite peculiar way, we still have Brownian motion as our limit instead of any jump process as the sample size goes to infinity. Also, even when we have long memory in volatility, the long term dependence does not show at all with the temporal aggregation limit.

It has been shown by Diebold (1988) and Drost and Nijman (1993) that the temporal aggregation of any ARCH model should roughly converge to Brownian motion, although with high-frequency data the ARCH structure gives us high leptokurtosis (see the result by De Haan et al., 1989). An interesting question in the current context is whether temporal aggregation can preserve the high leptokurtosis with the regime switching stochastic volatility model. A tentative suggestion is that the temporal aggregation should be stable motion given that the mean counterpart converges to stable motion, and the following theorem says the contrary.

*Theorem 2.2.* Given a regime switching stochastic volatility model satisfying Assumption 1.1, the stock variable

$$\bar{u}_t = \sum_{i=0}^t u_i = \sum_{i=1}^n v_i \varepsilon_i \quad \text{for } t = 1, 2, \dots, T \tag{15}$$

converges to Brownian motion in finite-dimensional distribution<sup>10</sup> if we normalize it by the square root of the sample size, i.e.  $\sigma_\varepsilon^{-1} T^{-1/2}$ .

*Proof.* The proof is carried out in two steps. In the first step, we show that  $\bar{u}_{[Tt_2]} - \bar{u}_{[Tt_1]}$  can be equalized with a summation of  $T(t_2 - t_1)/\mu$  terms of the independent variables with a structure as  $\sum_{i=\delta_i}^{\delta_i+1} v_i \varepsilon_i$  in probability. Then since the above term can be considered as the summation of i.i.d. terms with finite variance  $\mu$ , which is the mean of the interarrival time, we can employ the central limit theorem to prove that it converges to a normal variable. The independence follows from the 0 correlation.

(1) It is clear that

$$\begin{aligned} \bar{u}_t = & \left( v_0 \sum_{i=1}^{S_0} \varepsilon_{0,i} + \sum_{k=1}^{k(t)-1} v_k \sum_{i=1}^{T_k} \varepsilon_{k,i} + v_{k(t)-1} \sum_{i=1}^{t-S_{k(t)-2}} \varepsilon_{k(t)-1,i} \right. \\ & \left. - v_{k(t)-1} \sum_{i=1}^{T_{k(t)-1}} \varepsilon_{k,i} \right) I(S_0 < T) + v_0 \sum_{i=1}^T \varepsilon_{0,i} I(S_0 \geq T). \end{aligned}$$

We can prove that all the other terms vanish except the term  $R_2(t) = \sum_{k=1}^{k(t)-1} v_k \sum_{i=1}^{T_k} \varepsilon_{k,i}$  when they are normalized by  $T^{-1/2}$ . We will denote the above five parts as  $R_1, R_2, R_3, R_4,$  and  $R_5$ , respectively. And we will prove that only  $R_2$  remains in probability after normalization.

Firstly,  $R_1$  converges to 0 in mean square,

$$\text{var } R_1 = \sigma_v^2 \frac{E[S_0 I(S_0 < T)]}{T} \rightarrow T^{1-\alpha}. \tag{16}$$

The convergence follows from the probability structure of  $S_0$ . See Feller (1971, pp. 311–315). And  $R_1$  does converge to 0.

In a similar fashion,  $R_3$  and  $R_4$  can be proved to be negligible.  $R_5$  converges to 0 as  $I(S_0 \geq T)$  is a negligible event. We have

$$\text{var } R_5 = \frac{\sigma_v^2}{P(S_0 \geq T)} \rightarrow T^{1-\alpha}. \tag{17}$$

---

<sup>10</sup> We have tried to establish the weak convergence without success.

If we define  $Q_j(T) = \sum_{k=k([s_{j-1}T]+1)}^{k([s_jT])} v_k I(S_0 \leq s_{j-1}T) \sum_{i=1}^{T_k} \varepsilon_{k,i}$ , the fact that there is no difference between a corresponding  $R_2$  term and  $Q_j(T)$  enables us to see that the finite-dimensional distribution of  $x([s_jT]) - x([s_{j-1}T])$  is determined by  $Q_j(T)$ .

If we define that  $Y_N = \sum_{k=1}^N v_k \sum_{i=1}^{T_k} \varepsilon_{k,i}$  and  $Q'_j(T) = (Y_{[\mu^{-1}s_jT]} - Y_{[\mu^{-1}s_{j-1}T]})I(S_0 \leq s_{j-1}T)$  and  $\varepsilon'_j(T) = Q_j(T) - Q'_j(T)$ , then, if we apply Theorem 7.3.2 of Chung (1973),  $\varepsilon'_j(T)$  vanishes.

Because the  $Q'_j(T)$  are independent and since  $\bar{u}_{[T_{s_j}]} - \bar{u}_{[T_{s_{j-1}}]}$  do not differ in distribution, we have

$$T^{-1/2}(\bar{u}_{[T_{s_1}]}, \bar{u}_{[T_{s_2}]}, \dots, \bar{u}_{[T_{s_n}]}) \xrightarrow{\Delta} B(s_1, s_2, \dots, s_n), \tag{18}$$

where  $\xrightarrow{\Delta}$  denotes convergence in finite distribution.  $\square$

### 3. Estimation and empirical results

The RSSV model is theoretically capable of giving rise to long memory in volatility, but is this theoretical result empirically relevant? First, we want to know how well this RSSV model explains the overall complex dynamics of stock prices, especially when its performance is compared to other models with the same degree of complexity. Second, and more specifically, we want to know how relevant our key assumption – Assumption 1.1 – is, and how this assumption relates to the observed long memory pattern in the volatility of stock returns. These inference problems rely on the estimation and inference of the RSSV model.

Unfortunately, we are not able to use MLE to estimate the RSSV model, as in Hamilton and Susmel (1994) and Cai (1994). First, the current state is a latent variable. Second, the probability of taking on the current state depends on the length of time stayed in the current state, and this length of time is also a latent variable. Then the time when we have a new regime is also latent. Consequently, the conditional probability depends on an infinite past. We can use the QMLE approach, as in Sowell (1992) and Bollerslev and Mikkelsen (1996), or the frequency QMLE approach as in Breidt et al. (1998), to estimate our model based on the implied covariance structure of the distribution of the regime duration. This approach only uses the data information up to the second-order moments. A full examination of the model, however, confronts the model with a full set of dynamics as contained in the data set. The simulation-based efficient method of moments (EMM) method as proposed by Gallant and Tauchen (1996) enables us to do exactly such a full examination.

In the following section, we will first set up the econometrics problem and go over the essentials of EMM. We will then introduce the auxiliary model, the benchmark models and the various extensions of the RSSV model. We use benchmarks to show that our model can yield comparable performance and is a serious contender. Finally, we will present our empirical results.

### 3.1. EMM estimation

In EMM, the GMM technique is combined with simulation-based methods to estimate the coefficients of complicated nonlinear structural models. In particular, EMM provides a systematic and efficient way of choosing moments by employing the scores of the auxiliary model. It also sets up a systematic and meaningful way of drawing inferences with a set of diagnostics. The methodology is related to the more general Indirect Inference, as proposed by Gourieroux et al. (1993), although in that work their moments are generated from the parameter estimates. In matching the model with data along the dimension of scores defined by an auxiliary model, in addition to the delivery of a set of estimators, EMM also provides some econometrically meaningful metrics of the extent of the success of the examined model. It can, therefore, pinpoint any possible inherent merits and drawbacks of the model, which is exactly what we are concerned with. When the auxiliary model is properly chosen, the EMM estimation is as efficient as the maximum likelihood estimation.

Given a stationary process  $\{y_t\}$  generated from the underlying time-invariant structural model  $p(y_{-L}, \dots, y_0 | \rho)$ ,  $\rho \in R^{p_\rho}$ , with parameter  $\rho^0$ , the EMM estimator  $\hat{\rho}_n$  is computed as follows. Use the auxiliary model

$$f(y_t | y_{t-L}, \dots, y_{t-1}, \theta), \quad \theta \in R^{p_\theta} \tag{19}$$

which is specified in the sequel and the data  $\{\tilde{y}_t\}_{t=-L}^n$  to compute the maximum likelihood estimate

$$\tilde{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} \frac{1}{n} \sum_{t=1}^n \log[f(\tilde{y}_t | \tilde{y}_{t-L}, \dots, \tilde{y}_{t-1}, \theta)] \tag{20}$$

and the corresponding estimate of the information matrix

$$\tilde{\mathcal{J}}_n = \frac{1}{n} \sum_{t=1}^n \left[ \frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right] \left[ \frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n) \right]', \tag{21}$$

where

$$\tilde{x}_{t-1} = (\tilde{y}_{t-L}, \dots, \tilde{y}_{t-1}). \tag{22}$$

Define

$$m(\rho, \theta) = \int \cdots \int \frac{\partial}{\partial \theta} \log[f(y_0|y_{-L}, \dots, y_{-1}, \theta)] p(y_{-L}, \dots, y_0|\rho) dy_{-L} \cdots dy_0 \tag{23}$$

which is computed by averaging over a long simulation

$$m(\rho, \theta) \doteq \frac{1}{N} \sum_{t=1}^N \frac{\partial}{\partial \theta} \log[f(y_t|y_{t-L}, \dots, y_{t-1}, \theta)] \tag{24}$$

as described above. The EMM estimator is

$$\hat{\rho}_n = \underset{\rho \in R}{\operatorname{argmin}} m'(\rho, \tilde{\theta}_n)(\tilde{\mathcal{J}}_n)^{-1} m(\rho, \tilde{\theta}_n). \tag{25}$$

For the asymptotics of the estimator, please see Gallant and Tauchen (1996) and Gallant and Long (1996).

Under the null hypothesis that the model we are investigating is the correct model, EMM criterion

$$C = nm'(\hat{\rho}_n, \tilde{\theta}_n)(\tilde{\mathcal{J}}_n)^{-1} m(\hat{\rho}_n, \tilde{\theta}_n) \tag{26}$$

is asymptotic chi-square with  $p_\theta - p_\rho$  degrees of freedom.

One criticism of the EMM approach in the estimation of a nonlinear system with partially observed state is that filtered volatility could not be recovered from the estimates. Gallant and Tauchen (1998) further develop their approach to address this criticism with the reprojection idea.

Given the EMM estimate of system parameter  $\hat{\rho}_n$ , we should like to elicit the dynamics of the implied conditional density for observables

$$\hat{p}(y_0|y_{-L}, \dots, y_{-1}) = p(y_0|y_{-L}, \dots, y_{-1}, \hat{\rho}_n).$$

Define

$$\hat{\theta}_K = \underset{\theta \in \mathcal{R}^{p_K}}{\operatorname{argmax}} \mathcal{E}_{\hat{\rho}_n} \log f_K(y_0|y_{-L}, \dots, y_{-1}, \theta),$$

where  $f_K(y_0|y_{-L}, \dots, y_{-1}, \theta)$  is the density function for the estimated auxiliary model. Let

$$\hat{f}_K(y_0|y_{-L}, \dots, y_{-1}) = f_K(y_0|y_{-L}, \dots, y_{-1}, \hat{\theta}_K).$$

Theorem 1 of Gallant and Long (1996) states that

$$\lim_{K \rightarrow \infty} \hat{f}_K(y_0|y_{-L}, \dots, y_{-1}) = \hat{p}(y_0|y_{-L}, \dots, y_{-1}).$$

Convergence is with respect to a weighted Sobolev norm that they describe. We, therefore, can use  $\hat{f}_K$  to approximate  $\hat{p}$ . Given this approximation of  $\hat{p}$ , the reprojected volatility is the one-step-ahead standard deviation evaluated at data values; that is, the square root of

$$\begin{aligned} \text{var}(y_0|y_{-L}, \dots, y_{-1}) \\ = \int [Y_0 - \mathcal{E}(y_0|x_{-1})] \times [Y_0 - \mathcal{E}(y_0|x_{-1})]' f_K(y_0|x_{-1}, \hat{\theta}_K) dy_0, \end{aligned}$$

with  $(y_{-L}, \dots, y_{-1}) = (\tilde{y}_{t-L}, \dots, \tilde{y}_{t-1})$  for  $t = 0, \dots, n$ .

### 3.2. The auxiliary model and the score

To implement the EMM estimator we need an auxiliary model  $f(y|x)$  that fits the data well. The auxiliary model has been estimated using the seminon-parametric (SNP) method developed by Gallant and Nychka (1987) and has been applied to many studies on the stock price movements such as Gallant et al. (1992), and Tauchen et al. (1996), among others.

The SNP density is a member of a class of parameterized conditional densities

$$\mathcal{H}_K = \{f_K(y|x, \theta): \theta = (\theta_1, \theta_2, \dots, \theta_{t_k})\} \tag{27}$$

which expands  $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \dots$  as  $K$  increases. In practice, the  $K$ th model on the hierarchy is given by

$$f_K(y_t|x_{t-1}, \theta) = \frac{\{P_K[r_{x_{t-1}}^{-1}(y_t - \mu_{x_{t-1}}), x_{t-1}]\}^2 \phi[r_{x_{t-1}}^{-1}(y_t - \mu_{x_{t-1}})]}{|r_{x_{t-1}}|^{1/2} \int [P_K(u, x_{t-1})]^2 \phi(u) du}, \tag{28}$$

where  $P_K(\cdot, \cdot)$  is a Hermite polynomial given as follows:

$$P_K(z, x) = \sum_{\alpha=0}^{K_x} \left( \sum_{|\beta|}^{K_x} a_{\alpha\beta} x^\beta \right) z^\alpha, \tag{29}$$

where  $\mu_x$  and  $r_x$  are the location and the scale function, respectively, and  $\phi(z) = (2\pi)^{-1/2} e^{-z^2/2}$ .

The location function  $\mu_x$  is affine in  $x$ :

$$\mu_{x_{t-1}} = b_0 + b'x_{t-1}. \tag{30}$$



The scale function  $r_x$  is affine in the absolute values of  $x$ :

$$r_{x_{t-1}} = \rho_0 + \rho'|x_{t-1}|. \tag{31}$$

The vector  $\theta$  contains the coefficients  $A = [a_{\alpha\beta}]$  of the Hermite polynomial, the coefficients  $[b_0, b]$  of the location function, and the coefficients  $[\rho_0, \rho]$  of the scale function. To achieve identification, the coefficient  $a_{0,0}$  is set to 1. The tuning parameters are the number of lags in the location function ( $L_\mu$ ), the scale function ( $L_r$ ), the Hermite polynomial ( $L_p$ ), and the degrees of the polynomial in  $z$  ( $K_z$ ) and in  $x$  ( $K_x$ ).

### 3.3. Estimation model and benchmarks

Since a financial time series usually exhibits some autocorrelation even in its mean, we will adjust this by a mean equation

$$y_t = \mu_0 + \mu_1 y_{t-1} + \mu_2 y_{t-2} + u_t. \tag{32}$$

In all models that we fit, including the RSSV models and the other three types of benchmark models, we will impose the same mean equation with the parameters  $\mu_0$  and  $\mu_1$  and  $\mu_2$ .

The RSSV model we have proposed so far can be written as

$$u_t = e^{w_t} \varepsilon_t, \tag{33}$$

where  $w_t$  is the regime switching variable, with the probability law of duration of regime governed by some heavy-tail distribution  $L_T$ ; and the regime switching variable taking values according to some distribution  $L_w$ ;  $\varepsilon_t$  is distributed according to law  $L_\varepsilon$ . All three parts are independent. Sections 1 and 2 give us fairly general conditions for long memory behavior, here we are more specific about these laws in order to do estimations. Since we do not have much prior knowledge on how to choose among different laws of distribution for this new model, we begin with some standard probability distributions, and explore the potential of the RSSV model. Throughout our EMM exercise, we thus specify

$$w \sim N(\mu_w, \sigma_w^2), \tag{34}$$

$$\varepsilon \sim N(0, 1), \tag{35}$$

$$\text{Prob}(T_k \geq t) = (1 + ct)^{-\alpha}, \quad 0 < \alpha < \infty, \quad t \text{ is a integer greater than } 0, \tag{36}$$

where  $T_k$  denotes the duration of the regime and all the above quantities are independent. In the duration distribution,  $\alpha$  frames the tail behavior of the distribution, and  $c$  controls the scale. As there are many other probability

distributions with potential heavy tails, we settle for this particular probability distribution because of its numerical convenience. This distribution family allows for the possibilities of both heavy tail and non-heavy tail. When  $\alpha \leq 1$ , the duration has no mean; when  $1 < \alpha < 2$ , it has no variance and the resulting time series exhibits the long memory property; and when  $\alpha \geq 2$ , the distribution belongs to the normal attraction, and thus has less tail than in the case  $\alpha < 2$ . In total there are seven coefficients to be estimated in the model, namely,  $\rho = (\mu_0, \mu_1, \mu_2, \mu_w, \sigma_w, c, \alpha)$ .

Three classes of benchmark models are used to make comparisons. We use benchmarks to highlight points where the RSSV model is successful and to pinpoint any possible drawbacks. They also help us to identify the differences between different long memory modeling techniques. We note however that it is not our intention in this paper to establish any dominance of our model over other contenders.

The first benchmark model we use is the FIEGARCH model proposed by Bollerslev and Mikkelsen (1996). With the AR(2) mean equation (32), the rest of the model is as follows:

$$u_t = e^{w_t} \varepsilon_t, \quad \varepsilon_t \text{ is i.i.d. } N(0, 1), \tag{37}$$

$$w_t = \mu_w + (1 + \psi_1 \mathcal{L})(1 - \phi_1 \mathcal{L})^{-1}(1 - \phi_2 \mathcal{L})^{-1}(1 - \mathcal{L})^{-d}g(\varepsilon_{t-1}), \tag{38}$$

$$g(\varepsilon_t) = \theta \varepsilon_t + \gamma[|\varepsilon_t| - E(|\varepsilon_t|)]. \tag{39}$$

Since the model has a Gaussian mean innovation  $\varepsilon_t$ , we call it FIEGARCH with Gaussian error model. For contrast we also use a slight variation of the FIEGARCH model, the FIEGARCH with spline error model, where, instead of Gaussian error,  $\varepsilon_t$  will be distributed as the spline transformation of a series of  $N(0, 1)$  variates  $z_t$ ,

$$\varepsilon_t = T(z_t), \tag{40}$$

$$T(z_t) = b_{z0}(b_c, b_d) + b_{z1}(b_c, b_d)z_t + b_{z2}(b_c, b_d)z_t^2 + b_{z3}(b_c, b_d)z_t \max(0, z_t). \tag{41}$$

To achieve identification, the constraint that  $T(z_t)$  has a mean of 0 and a variance of 1 is imposed.<sup>11</sup> When  $b_c$  and  $b_d$  are equal to 0,  $b_{z0}(b_c, b_d) = b_{z2}(b_c, b_d) = b_{z3}(b_c, b_d) = 0$  and  $b_{z1}(b_c, b_d) = 1$ . Given that the estimation

---

<sup>11</sup>  $b_{z0}(b_c, b_d), b_{z1}(b_c, b_d), b_{z2}(b_c, b_d), b_{z3}(b_c, b_d)$  can be seen to be functions of  $b_c$  and  $b_d$  as:  $b_{z0} = (b_c + 0.5b_d)/s, b_{z1} = 1/s, b_{z2} = b_c/s,$  and  $b_{z4} = b_d/s,$  where  $s = (b_c + 0.5b_d)^2 + 1 + 3b_c^2 + 1.5b_d^2 + 2[(b_c + 0.5b_d)b_c + 0.5(b_c + 0.5b_d)b_d + 0.7979b_d + 1.5b_c b_d]$ .

method is simulation based, this particular form of error terms is flexible enough to incorporate different features of the error terms, yet causes no numerical difficulty. For a successful use of the spline transformation errors see Gallant et al. (1997).

The second benchmark model to be used is the long memory stochastic volatility model as proposed in Harvey (1993). The particular form of the model is taken from Gallant et al. (1997) and Liu and Zhang (1997), with the mean equation as Eq. (32),

$$u_t = e^{w_t} \varepsilon_t, \quad \varepsilon_t \text{ is i.i.d. } N(0, 1), \quad (42)$$

$$w_t - \mu_w = a_1(w_{t-1} - \mu_w) + a_2(w_{t-2} - \mu_w) + (1 - \mathcal{L})^{-d} r_w (\tilde{\varepsilon}_t + \gamma \varepsilon_t), \quad (43)$$

where  $\tilde{\varepsilon}_t$  is an  $N(0, 1)$  error term independent of  $\varepsilon_t$ . This model is denoted as LMSV with Gaussian error, due to the Gaussian error used for the mean innovation  $\varepsilon_t$ . A slightly different form of the model which uses the spline errors is denoted as LMSV with spline error of the form (41).

The third benchmark, as suggested by one of the referees, is a finite-state regime switching extension of the second benchmark in the spirit of Hamilton and Susmel (1994). That is, a finite-state Markov regime switching structure is added to the volatility term of the second benchmark

$$u_t = e^{w_t} \varepsilon_t, \quad \varepsilon_t \text{ is i.i.d. } N(0, 1), \quad (44)$$

$$w_t - w_t^r = a_1(w_{t-1} - w_{t-1}^r) + a_2(w_{t-2} - w_{t-2}^r) + (1 - \mathcal{L})^{-d} r_w (\tilde{\varepsilon}_t + \gamma \varepsilon_t), \quad (45)$$

where  $\tilde{\varepsilon}_t$  is an  $N(0, 1)$  error term independent of  $\varepsilon_t$ , with  $w_t^r$  as the regime switching variable taking on three potential values ( $s_0, s_1, s_2$ ) as in Hamilton and Susmel (1994) with a Markov transition matrix as following:

$$p = \begin{pmatrix} p_{11} & p_{12} & 1 - p_{11} - p_{12} \\ p_{21} & p_{22} & 1 - p_{21} - p_{22} \\ p_{31} & p_{32} & 1 - p_{31} - p_{32} \end{pmatrix}.$$

The parameters in the transition matrix are positive numbers taking values between 0 and 1. In addition, the sum of the two parameters in the same row must be less than or equal to 1. This benchmark allows a direct comparison of heavy-tail regime switching with a Markov chain regime switching model similar to that of Hamilton and Susmel (1994). Also, as the model contains both long memory and regime switching components, it can address the question of whether long memory is still a pervasive phenomenon with the presence of

Markov regime switching in the model. Breidt et al. (1997), for example, hypothesized that the observation of long memory can arise as a result of a Markov chain regime switching model. Later we will call this the LMSV-Markov-RS model when the fractional component is present and SV-Markov-RS model when the fractional component is not present.

Besides a comparison with the different benchmarks, various extensions of the RSSV model will also be examined. In the RSSV model, the volatility autocorrelation functions both at the short term and the long term are determined by the regime switching structure. This can potentially cause a conflict between the modeling of long-term persistence and short-term dynamics. Thus, the first type of extension aims to provide a more flexible form to better accommodate the short-term dynamics. For this purpose, we include an additional AR term in the volatility equation and an additional term representing the leverage effect. Second, even in modeling the long-term dynamics, the modeling technique here differs from the ARFIMA modeling in Bollerslev and Mikkelsen (1996) and Harvey (1993). As we remarked at the end of Section 1, the regime switching model can generate a more volatile series than the ARFIMA model. Given the differences between these two types of modeling techniques, and the fact that the ARFIMA modeling of the long memory in volatility is successful, there may be some unique elements in the ARFIMA modeling which our regime switching modeling cannot capture. We thus also include, in our second extension, a long memory term modeled in the ARFIMA way. An examination of these extensions suggests venues upon which the RSSV model can be improved, and also provides some assessment of the other modeling techniques with the heavy-tail regime switching present.

The first extension of the RSSV model incorporates the features of the stochastic volatility benchmark. We add a term  $w_t^s$  in the volatility to model short-term dynamics,

$$u_t = e^{w_t + w_t^s} \varepsilon_t, \quad \varepsilon_t \text{ is i.i.d. } N(0, 1), \tag{46}$$

$$w_t^s = a_1 w_{t-1}^s + r_w^s (\tilde{\varepsilon}_t + \gamma \varepsilon_{t-1}), \tag{47}$$

where  $\tilde{\varepsilon}_t$  is  $N(0, 1)$  variates and is independent of  $\varepsilon_t$ ,  $w_t$  denotes the regime switching variable as in the RSSV model given by (36) and (38), and the probability distribution in  $w_t^s$  is independent of the probability distribution in  $w_t$ . We will call this extension the RSSV AR-V model. In the second extension, in addition to the above modification, we will add the fractional integration term

$$w_t^s = a_1 w_{t-1}^s + (1 - \mathcal{L})^{-d} r_w^s (\tilde{\varepsilon}_t + \gamma \varepsilon_{t-1}) \tag{48}$$

and we will call it the RSSV ARFIMA-V model.

### 3.4. Empirical results

We fit the regime switching stochastic volatility model to the daily Standard and Poor's composite price index. It spans a period from the beginning of 1928 to the end of 1995 and consists of 18,149 observations. The earlier portion of this dataset is taken from the widely cited studies of stock market dynamics of Gallant et al. (1992, 1993). We have extended their dataset from 1987 to 1995. The raw data is first-differenced in its logarithm and normalized by 100 to obtain a price change series. This series is then adjusted for systematic calendar effects in location and scale, as described in Gallant et al. (1992). Fig. 1 gives a plot of the adjusted time series and the autocorrelation pattern for the absolute price return, which gives the now well-recognized long memory pattern.

For the auxiliary SNP model, we have used the auxiliary model used by Liu and Zhang (1997). This model has 29 parameters with the following tuning parameters,  $L_\mu = 0$ ,  $L_r = 25$ ,  $L_p = 0$ ,  $k_z = 18$ , and  $K_x = 0$ . Twenty-six parameters of the model are used in modeling the volatility persistence and two are used in modeling the higher-order dynamics. It differs from some previously

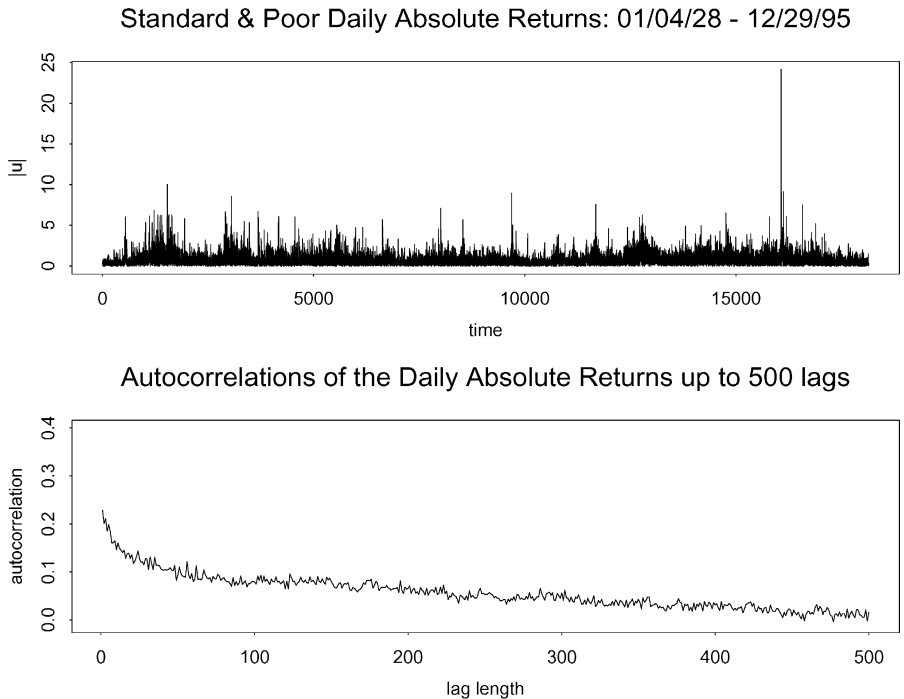


Fig. 1. The upper panel shows the absolute value of the innovation of the market returns. The lower panel shows the autocorrelation pattern of the absolute stock market returns.

Table 1  
EMM tests of different long memory models<sup>a</sup>

Model	$l_\rho$	$\chi^2$	df	$p$ -value	Ljung–Box
FIEGARCH, Gaussian error	10	46.336	19	< 0.001	234.138
FIEGARCH, spline error	12	18.928	17	0.333	234.139
Long memory SV, Gaussian error	9	39.348	20	0.006	240.640
Long memory SV, spline error	10	12.077	19	0.882	227.954
SV-Markov-RS	16	33.400	13	0.001	487.900
LMSV-Markov-RS	17	18.010	12	0.115	223.296
RSSV	7	26.483	22	0.232	223.727
RSSV and AR(1) volatility term	10	18.983	19	0.458	234.138
RSSV and ARFIMA volatility term	11	18.354	18	0.433	218.361

<sup>a</sup> $l_\rho$  denotes the number of parameters of the structure models. The column marked  $\chi^2$  gives the value of the EMM criterion, which is a  $\chi^2$  statistic with degree of freedom of df. The  $p$ -value of these statistics is shown in the next column. The column marked Ljung–Box gives the Ljung–Box statistics on the absolute value of the model residual. The degree of freedom of the Ljung–Box statistics is 200.

used SNP auxiliary models (see Gallant et al. (1997) for other possible SNP auxiliary models) by including some higher-order dynamics while keeping the very long lag of those models in its ARCH expression. It also suppresses some moments to keep the auxiliary model simple enough. The model passes the diagnostic test of Liu and Zhang (1997). For more details regarding the selection of this auxiliary model, please refer to Liu and Zhang (1997). Andersen et al. (1999) have shown with Monte Carlo experiments that EMM inference is robust to the choice of auxiliary model in reasonably large sample. For more on the selection of auxiliary model, please refer to their study.

Table 1 presents the fitting results for both the various RSSV models and for the benchmark models. While the FIEGARCH with Gaussian error model and LMSV with Gaussian error model are both rejected for the given score generator, we cannot reject the RSSV model at the 10% confidence level. However, both the FIEGARCH model and the LMSV model fit the stock market dynamics well, as long as we allow thick tail error in the mean innovation. See Fig. 2 for the estimated spline error distribution in the FIEGARCH with spline error model.<sup>12</sup> This corresponds well with the study of Gallant et al. (1997) and Chib et al. (1998), where such thick tail errors have been found to be very useful in studying stock return dynamics with the stochastic volatility model and the GARCH model. As there are inherent jumps in the RSSV models, they do not

<sup>12</sup> The bimodal feature of the distribution may appear too outlandish and suggest some kind of regime switching would help in improving the fit.

## autocorrelation implied by different models

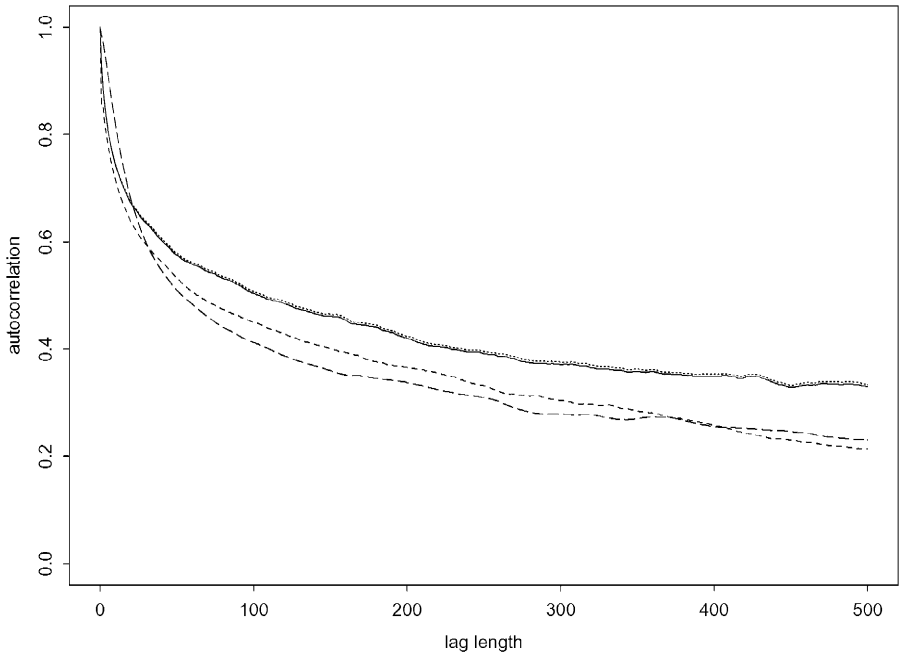


Fig. 2. The estimated distribution of the spline errors in the FIEGARCH with spline error model. The distribution can be seen to be asymmetric. A similar spline error distribution is also found in the LMSV with spline error model.

require thick tail errors to give volatile dynamics. The SV-Markov-RS model yields a very large EMM criterion and is rejected as a model to fit the stock return dynamics. The LMSV-Markov-RS model, a variation of SV-Markov-RS incorporating the long memory feature, cannot be rejected however based on the EMM criterion.

While the EMM criteria indicate the overall goodness-of-fit of the models, the Ljung–Box Statistics reported in Table 1 give an indication of goodness-of-fit on persistence of the volatility as implied by the different models. A model, fitted well in persistence and badly in high-order moments, can very well give good Ljung–Box statistics but not a good EMM criterion. The reported Ljung–Box statistics were generated based on the autocorrelations (up to 200 lags) of the time series consisting of the absolute value of the model residuals ( $\varepsilon_t$ 's).<sup>13</sup> It has

<sup>13</sup> The model residuals are generated using the reprojection idea of Gallant and Tauchen (1998).

### The estimated spline distribution

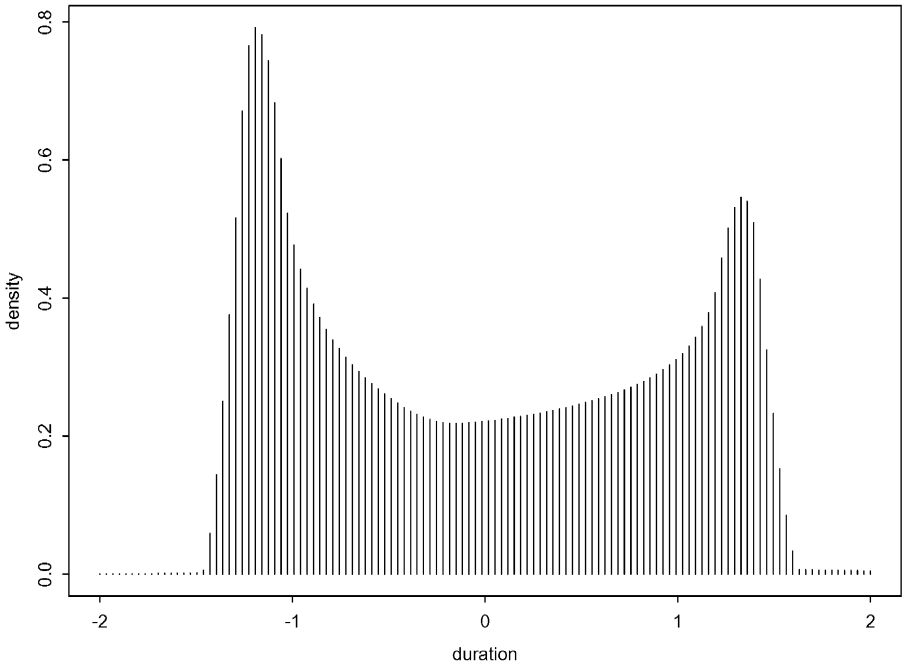


Fig. 3. Autocorrelation patterns of volatility  $v_t$  of different models, plotted based on simulation sample size of 150,000. Solid: RSSV model; dotted: LMSV-Markov-RS; short dashed: LMSV with Gaussian error; long dashed: FIEGARCh with Gaussian error. We have omitted the autocorrelation patterns of the other models because of their similarities with those depicted in the figure.

a chi-square distribution with 200 degrees of freedom. Most of the models presented in the table fit the persistence well enough with the exception of the SV-Markov-RS model, which again shows the inability of this model to fit the volatility persistence. In Fig. 3, we plot the autocorrelation pattern of the volatility of four different models. They do not differ too much and all exhibit the typical long memory feature. Other models with long memory terms also show a similar autocorrelation pattern in the volatility.

While the RSSV model fits the data reasonably well, the introduction of short memory dynamics such as the AR(1) component in the RSSV AR-V model also seems to be helpful. As a matter of fact, if we hypothesize that the short-term dynamics part is unnecessary in the RSSV AR-V model, then the hypothesis can be rejected at the 10% confidence interval. The EMM criterion difference between the RSSV model and the RSSV AR-V model gives a value of 7.500 for a  $\chi^2$  of degree 3, which in turn gives a  $p$ -value of 0.058. The introduction of the



fractional integration term, however, does not seem to help much. This can be seen from the trivial improvement of the EMM criterion in Table 1.

Table 2 gives the parameter estimates for each individual model and their criteria-difference confidence interval. The estimates for the long memory coefficients are of most interest, including the parameter  $d$  in the LMSV models, the FIEGARCH models and the RSSV ARFIMA-V model and  $\alpha$  in the RSSV models. The long memory coefficients  $\alpha$ , in the RSSV models, are highly significant and have a similar magnitude as  $2 - 2d$  in both the FIEGARCH model and the LMSV model, as predicted by Theorem 1.1. Based on a 95% confidence interval, we are not able to reject the existence of heavy-tail regime switching in the volatility. This suggests that regime switching with a heavy-tail modeling of the long memory is able to deliver the same magnitude of long memory as the ARFIMA modeling in the volatility. Also, with a heavy-tail regime switching volatility structure, the persistence level, as characterized by  $a_1$  in the RSSV AR-V model and  $a_1$  and  $d$  in the RSSV ARFIMA-V model, is no longer significant.

We now turn to a discussion of the potential relationship between Markov regime switching and long memory. We first look at the parameter estimates of the SV-Markov-RS model. A very high transition probability of returning to the same state is found with 2 out of 3 volatility states. In addition, even with this very high persistence in different volatility states, the persistence level in the volatility equation as characterized in  $a_1$  is greater than 0.9. According to the Ljung–Box statistics in Table 1, the model does not fit the observed persistence in market volatility. This result stands in contrast to the result in Hamilton and Susmel (1994), where a smaller  $a_1$  is reported with a similar model. We note, however, that this study uses almost 10 times more data points than (about three times more in terms of length of time and daily series) the study of Hamilton and Susmel (1994). And a large sample size typically improves inferences regarding persistence. In the LMSV-Markov-RS model, the long memory feature is added to the SV-Markov-RS model. With the addition of long memory, the LMSV-Markov-RS model yields an EMM criterion with a  $p$ -value around 10%, and the regimes no longer appear as persistent when compared with the SV-Markov-RS model. Therefore, the drastic improvement of the EMM criterion with the addition of the long memory term provides empirical evidence in support of Theorem 1.2, where the Markov regime switching model is proven unable to generate a long memory pattern in large sample.

Fig. 4 shows the regime switching pattern as implied by the RSSV model. This particular regime switching pattern is generated from the RSSV AR-V model. We do not present other regime switching patterns because they are similar to those in Fig. 4. One notable feature of the duration distribution is the size of the tail. There are a few long-lasting regime durations that generate the tail of the long memory pattern in Fig. 4 according to Theorem 2.1. There are also many short durations. As a matter of fact, the median of all the durations is only

Table 2  
Fitted parameter estimates for various models<sup>a</sup>

Regime switching stochastic volatility models			
	RSSV	RSSV AR-V	RSSV ARFIMA-V
	$\hat{\rho}$	$\hat{\rho}$	$\hat{\rho}$
	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$
$\mu_0$	0.041	0.054	0.051
$\mu_1$	0.067	0.048	0.050
$\mu_2$	0.029	0.035	0.035
$\mu_w$	0.021	0.025	0.024
$\sigma_w$	0.461	0.463	0.463
$c$	0.605	0.605	0.605
$\alpha$	1.047	1.050	1.046
$a_1$		0.130	0.101
$\gamma$		-50.808	-50.846
$r_w^s$		1.310e-4	1.266e-4
$d$			0.111
			(-0.230, 0.562)
			(-145.404, 25.300)
			(3.4e-7, 0.006)
			(-0.130, 0.260)

Different long memory stochastic volatility models			
	Long memory SV with Gaussian error	Long memory SV with spline error	
	$\hat{\rho}$	$\hat{\rho}$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$
	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	
$\mu_0$	0.043	0.053	(0.053, 0.121)
$\mu_1$	0.086	-0.088	(-0.296, -0.087)
$\mu_2$	-0.062	0.072	(-0.029, 0.173)
$\mu_w$	-0.067	-0.027	(-0.111, -0.027)
$r_w$	-0.060	0.034	(0.034, 0.034)

$\gamma$	0.004	(-0.016, 0.025)	0.072	(-0.029, 0.173)
$b_c$			0.561	(0.560, 0.573)
$b_d$			-0.899	(-0.912, -0.887)
$\phi_1$	0.876	(0.869, 0.882)	0.748	(0.647, 0.798)
$\phi_2$	-0.410	(-0.486, -0.354)	0.202	(0.102, 0.252)
$d$	0.422	(0.395, 0.458)	0.505	(0.455, 0.524)

Different FIEGARCH models

FIEGARCH with Gaussian error		FIEGARCH with spline error		
$\hat{\rho}$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	$\hat{\rho}$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	
$\mu_0$	0.014	(0.013, 0.021)	0.075	(0.041, 0.154)
$\mu_1$	0.024	(0.020, 0.025)	0.069	(-0.050, 0.201)
$\mu_2$	0.019	(0.007, 0.021)	0.011	(0.010, 0.012)
$\mu_w$	-0.021	(-0.021, -0.021)	-0.072	(-0.347, 0.160)
$\theta$	-0.125	(-0.125, -0.124)	-0.040	(-0.035, 0.044)
$\gamma$	-0.002	(-0.003, 0.008)	0.299	(0.080, 0.501)
$b_c$			0.560	(0.460, 0.601)
$b_d$			-0.860	(-0.859, -0.829)
$\phi_1$	0.776	(0.776, 0.798)	0.893	(0.340, 0.945)
$\phi_2$	-0.524	(-0.524, -0.451)	-0.392	(-0.546, -0.093)
$\psi_1$	0.308	(0.303, 0.308)	-0.394	(-0.422, -0.341)
$d$	0.522	(0.514, 0.537)	0.238	(0.133, 0.475)

Table 2. *Continued*

	Without long memory		With long memory	
	$\hat{\rho}$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$	$\hat{\rho}$	$(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$
$\mu_0$	0.052	(0.034, 0.054)	0.049	(0.043, 0.057)
$\mu_1$	0.010	(0.076, 0.096)	-0.005	(-0.005, 0.005)
$\mu_2$	0.012	(-0.074, -0.058)	0.006	(0.005, 0.015)
$s_0$	-0.098	(-0.090, -0.043)	-0.036	(-0.042, -0.030)
$s_1$	0.065	(0.031, 0.073)	0.044	(0.036, 0.052)
$s_2$	0.212	(0.163, 0.293)	0.282	(0.274, 0.286)
$p_{11}$	0.952	(0.921, 0.989)	0.840	(0.760, 0.873)
$p_{12}$	0.012	(0.005, 0.016)	0.000	(0.000, 0.001)
$p_{21}$	0.009	(0.002, 0.018)	0.228	(0.157, 0.270)
$p_{22}$	0.958	(0.940, 0.978)	0.221	(0.123, 0.290)
$p_{31}$	0.165	(0.102, 0.193)	0.197	(0.167, 233)
$p_{32}$	0.292	(0.260, 327)	0.260	(0.208, 299)
$r_w$	-0.067	(-0.063, -0.056)	-0.054	(-0.058, -0.054)
$\gamma$	-0.082	(-0.016, 0.025)	-0.071	(-0.084, -0.060)
$\phi_1$	0.959	(0.869, 0.882)	0.819	(0.813, 0.823)
$\phi_2$	0.086	(-0.486, -0.354)	0.157	(0.103, 0.211)
$d$			0.399	(0.388, 0.411)

<sup>a</sup>Column  $\hat{\rho}$  shows the parameter estimates for the concerned parameters. Column  $(\hat{\rho}_{0.025}, \hat{\rho}_{0.975})$  gives the confidence intervals of the estimates. These are generated by inverting the EMM criterion difference. For example, a point associated with  $\hat{\rho}_{0.975}$  is given by the point such that the difference of EMM criteria evaluated at different points gives a value of  $\chi_{0.975}(1)$ .

## Duration distribution inferred from RSSV AR-V

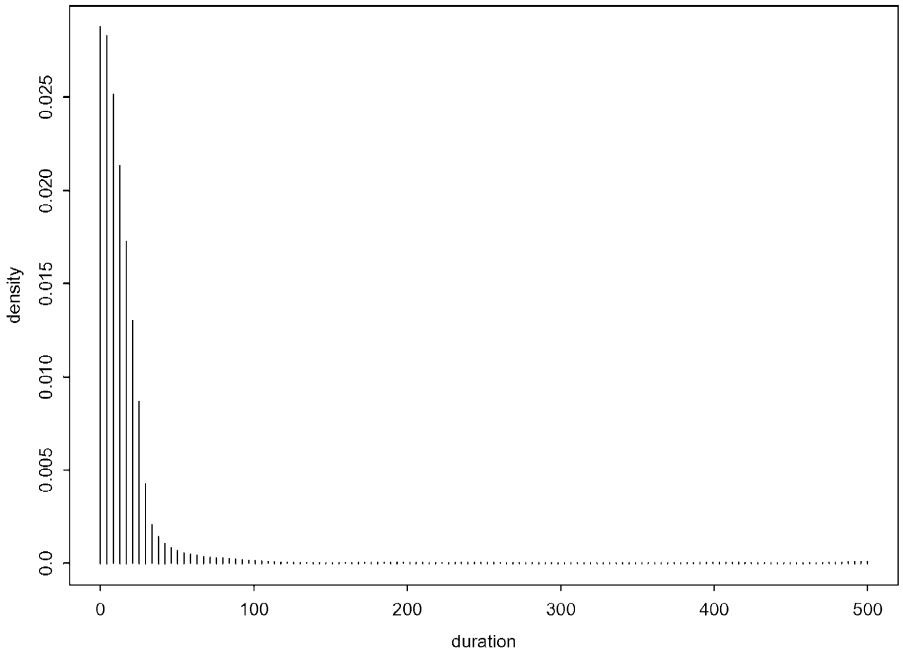


Fig. 4. The projected/reprojected volatility estimate with the estimated model: (a) equally weighted MA(4) of squared AR(2) residuals; (b) the auxiliary SNP model; (c) the FIEGARCH model with Gaussian error; (d) the FIEGARCH model with spline error; (e) the long memory stochastic volatility model with Gaussian error; (f) the long memory stochastic volatility model with spline error; (g) the long memory stochastic volatility with Markov regime switching model; (h) the regime switching stochastic volatility model.

2 days. The existence of these many short durations suggests that it is rather difficult for a volatility regime to establish itself and should not cause much alarm. As we can see from Theorem 2.1, in the RSSV model, the persistence, or the impact of a particular regime on the persistence, arises through the length of that regime. The large number of short regimes is weighted with the very short durations of these regimes. Consequently, even though there are many short-duration regimes, only those long-duration regimes count in generating the long memory pattern of persistence.

In practice, people often like to forecast volatility with a model. Within the setting of an RSSV model, this can be done by applying the idea of reprojection. The reprojected volatility figure can be used in any setting when a volatility number is needed. In Fig. 5, we show the volatility estimate of different models together with an equally weighted MA(26) estimate of the volatility. As we can

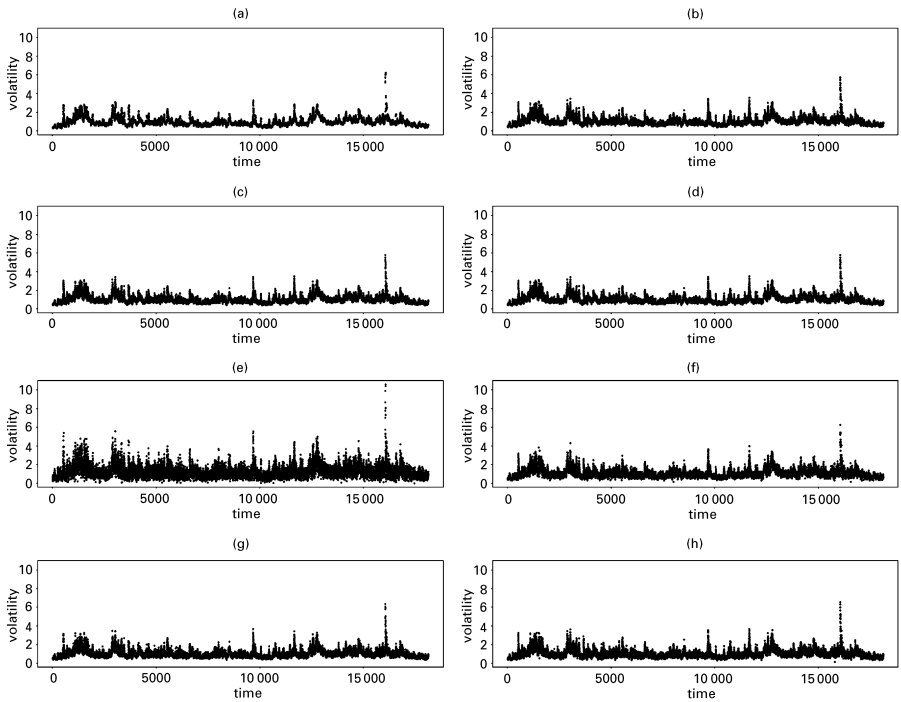


Fig. 5. The empirical distribution of the regime duration, plotted according to a simulation size of 150,000.

see, RSSV, LMSV-Markov-RS, and FIEGARCH all generate similar volatility estimates. The heavy-tailed regime switching in the volatility of the RSSV model can therefore generate a good fit while providing enough structure to yield a useful volatility prediction.

#### 4. Conclusion and future research

As Mandelbrot (1963) stated: “... large changes tend to be followed by large changes – of either sign – and small changes by small changes...”. This is the so-called volatility clustering phenomenon. The persistence characteristic of this phenomenon has even been shown as long memory in several recent studies. While previous researchers tended to use the fractional integration structure to model this phenomenon, we interpret the observed persistence as arising from volatility regime switchings, which are in turn triggered by different news arrivals. We show that when the duration of the regime has a heavy-tail distribution, we do indeed have the long memory behaviour. This paper thus

provides an interesting and intuitively appealing interpretation of the observed long memory behavior in stock market volatility.

The above rationale for generating long memory is then examined using the regime switching stochastic volatility (RSSV) model. With the model containing a regime switching part in its volatility, the empirical work could potentially refute or fail to refute our regime switching conjecture. Using the newly proposed efficient method of moments, for the S&P composite return series, we find evidence in support of the assumption of a heavy-tail distribution with the duration of a regime. The model is found to fit the dynamics of the stock prices extremely well and the estimated tail index is highly significant. We thus cannot reject either the hypothesis that all the observed characteristics of the data as captured by the flexible seminonparametric model is actually generated by the RSSV model or the proposition that the observed long memory pattern is related to the regime switching.

Long memory exists in other dimensions of the economic system as well. Exchange rates, interest rates and the prices of different commodities have all been shown to exhibit long memory behavior (see Baillie, 1996). It will be interesting to see the empirical relevance of the regime switching argument in explaining the persistence of the above series. More work towards understanding the mechanism of regime switching, especially the duration structures of regimes along the lines of Gouriéroux and Monfort (1992), will be interesting as well.

## 5. For further reading

Liu, 1996.

## Acknowledgements

This paper is part of my Ph.D. Dissertation. I thank my advisor George Tauchen for introducing me to the idea of long memory and for his constant advice and encouragement during the research. I also thank Mark An, Ravi Bansal, Ngai Hang Chan, John Coleman, Frank Diebold, Ron Gallant, Stephen Gray, Jim Hamilton and seminar participants at the University of Pennsylvania and Tilburg University for their valuable comments. The current version of the paper is a substantial improvement over older versions, thanks to the many suggestions of the referees and editors. All remaining errors are mine. In this paper, the EMM programs of Gallant and Tauchen (1996) are used.

## References

Andersen, T.G., Bollerslev, T., 1997. Heterogenous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Journal of Finance* 52 (3), 975–1005.

- Andersen, T.G., Chung, H.J., Sorensen, B.E., 1999. Efficient method of moments estimation with a stochastic volatility model: a Monte Carlo study. *Journal of Econometrics* 91, 61–87.
- Backus, D.K., Zin, S.E., 1993. Long memory inflation uncertainty: evidence from the term structure of interest rates. *Journal of Money, Credit and Banking* 25 (3), 681–700.
- Baillie, R.T., 1996. Long memory process and fractional integration in econometrics. *Journal of Econometrics* 73 (1), 5–59.
- Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74 (1), 3–30.
- Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular variation. In: *Encyclopedia of Mathematics and its Application*, Vol. 27. Cambridge University Press, Cambridge.
- Bollerslev, T., Mikkelsen, H.O., 1996. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* 73 (1), 151–184.
- Breidt, F.J., Crato, N., de Lima, P., 1997. Modeling the persistent volatility of asset returns. *Proceedings of the IEEE/IAFE, Computational Intelligence for Financial Engineering*, pp. 266–272.
- Breidt, F.J., Crato, N., de Lima, P., 1998. Modeling long memory stochastic volatility. *Journal of Econometrics* 83, 325–348.
- Cai, J., 1994. A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics* 12 (3), 309–316.
- Chib, S., Kim, S., Shephard, N., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65 (3), 361–393.
- Chung, K.L., 1973. *A Course in Probability Theory*, 2nd Edition. Academic Press, New York.
- De Haan, L., Resnick, S.I., Rootzen, H., De Vries, C.G., 1989. Extremal behavior of solutions to a stochastic difference equation with application to ARCH process. *Stochastic Processes and their Applications* 32, 213–224.
- Diebold, F.X., 1986. Modeling the persistence of conditional variances: a comment. *Econometric Reviews* 5 (1), 51–56.
- Diebold, F.X., 1988. *Empirical Modeling of Exchange Rates*. Springer, New York.
- Diebold, F.X., Rudebush, G.D., Sichel, D.E., 1993. Further evidence of business-cycle duration dependence. In: Stock, J.H., Watson, M.W. (Eds.), *Business Cycles, Indicators, and Forecasting*. NBER, The University of Chicago Press, Chicago.
- Ding, Z., Granger, C., Engle, R.F., 1993. Long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.
- Drost, F.C., Nijman, T.E., 1993. Temporal aggregation of GARCH process. *Econometrica* 61, 909–928.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50 (4), 987–1007.
- Feller, W., 1971. *An Introduction to Probability Theory and its Applications*, 2nd Edition, Vol. 2. Wiley, New York.
- Gallant, A.R., Long, J.R., 1996. Estimating stochastic differential equations efficiently by minimum chi-square. *Biometrika* 84, 125–141.
- Gallant, A.R., Nychka, D.W., 1987. Semi-nonparametric maximum likelihood estimation. *Econometrica* 55, 363–390.
- Gallant, A.R., Hsieh, D.A., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics* 81 (1), 159–192.
- Gallant, A.R., Rossi, P.E., Tauchen, G., 1992. Stock prices and volume. *The Review of Financial Studies* 5, 199–242.
- Gallant, A.R., Rossi, P.E., Tauchen, G., 1993. Nonlinear dynamic structures. *Econometrica* 61 (4), 871–907.
- Gallant, A.R., Tauchen, G., 1996. Which moment to match? *Econometric Theory* 12, 657–681.



- Gallant, A.R., Tauchen, G., 1998. Reprojecting partially observed systems with application to interest rate diffusions. *Journal of the American Statistical Association* 93, 10–24.
- Gourieroux, C., Monfort, A., 1992. Qualitative threshold ARCH models. *Journal of Econometrics* 52, 159–199.
- Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. *Journal of Applied Econometrics* 8, S85–S118.
- Granger, C.W.J., 1980. Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14, 227–238.
- Granger, C.W.J., Joyeux, R., 1980. An introduction to long memory time series and fractional differencing. *Journal of Time Series Analysis* 1, 14–29.
- 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., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics* 64, 307–333.
- Harvey, A.C., 1993. Long memory in stochastic volatility. Working Paper, London School of Economics.
- Haubrich, J.G., Lo, A.W., 1991. The sources and nature of long memory in the business cycle. Working Paper No. 9116, Federal Reserve Bank of Cleveland.
- Lamoureux, C.G., Lastrapes, W.D., 1990. Persistence in variance, structural change, and the GARCH model. *Journal of Business and Economic Statistics* 8, 225–234.
- Levy, J.B., 1983. High variability and long range dependence: a modeling approach based on renewal sequences. M.S. Thesis, Cornell University, Ithaca, NY.
- Liu, M., 1996. Essays in long memory and stock market volatility. Ph.D Dissertation, Duke University, unpublished.
- Liu, M., Zhang, H.H., 1997. Specification tests in the efficient method of moments framework with application to the stochastic volatility models. Manuscript, Carnegie Mellon University.
- Mandelbrot, B., 1963. The variation of certain speculative prices. *Journal of Business* 36, 394–419.
- Parke, W.R., 1995. What is a fractional unit root? Working Paper, University of North Carolina, Chapel Hill.
- Romanovskii, V.I., 1970. *Discrete Markov Chains*. Wolters-Noordhoff, Groningen.
- Sowell, F., 1990. The fractional unit root distribution. *Econometrica* 58, 495–505.
- Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53, 165–188.
- Taqqu, M.S., Levy, J.B., 1986. Using renewal processes to generate long range dependence and high variability. In: Eberlein, E., Taqqu, M.S. (Eds.), *Dependence in Probability and Statistics: A Survey of Recent Results*. Birkhauser, Boston.
- Tauchen, G., Zhang, H., Liu, M., 1996. Volume, volatility, and leverage: a dynamic analysis. *Journal of Econometrics* 74, 177–208.