

## **Bayesian Analysis of a Regime Switching In-Mean Effect for the Polish Stock Market**

Łukasz Kwiatkowski\*

Submitted: 18.02.2012, Accepted: 20.04.2012

### **Abstract**

The study aims at a statistical verification of breaks in the risk-return relationship for shares of individual companies quoted at the Warsaw Stock Exchange. To this end a stochastic volatility model incorporating Markov switching in-mean effect (SV-MS-M) is employed. We argue that neglecting possible regime changes in the relation between expected return and volatility within an ordinary SV-M specification may lead to spurious insignificance of the risk premium parameter (as being 'averaged out' over the regimes). Therefore, we allow the volatility-in-mean effect to switch over different regimes according to a discrete homogeneous two- or three-state Markov chain. The model is handled within Bayesian framework, which allows to fully account for the uncertainty of model parameters, latent conditional variances and state variables. MCMC methods, including the Gibbs sampler, Metropolis-Hastings algorithm and the forward-filtering-backward-sampling scheme are suitably adopted to obtain posterior densities of interest as well as marginal data density. The latter allows for a formal model comparison in terms of the in-sample fit and, thereby, inference on the 'adequate' number of the risk premium regimes.

**Keywords:** Markov switching, stochastic volatility, risk premium, in-mean effect, Bayesian analysis

**JEL Classification:** C50, C11, C22.

---

\*Cracow University of Economics, e-mail: kwiatkol@uek.krakow.pl

---

Łukasz Kwiatkowski

---

## 1 Introduction

Conceptually, the idea of a reward to a risk-averse investor for holding a risky asset appears theoretically sound and intuitively appealing. In a static framework a general agreement prevails that the more uncertain the investment, the higher the return expected by the investor. So far, however, a voluminous body of (mainly) empirical research has failed to unequivocally establish a dynamic version of the risk-return tradeoff, which postulates a rise of expected future income accompanying an increase in the expected volatility (risk) of some asset.

Overall, the literature provides mixed (if not contradictory) results. Some of the studies successfully pinpoint a positive relation (see e.g. French, Schwert and Stambaugh 1987; Bollerslev, Engle and Wooldridge 1988; Harvey 1989; Scruggs 1998; Harrison and Zhang 1999; Watanabe 1999; Guo and Whitelaw 2006; Ludvigson and Ng 2007; Smith 2007; Guo and Neely 2008). Others provide evidence of either no significant return-volatility link or a counterintuitive, negative relation (see e.g. Campbell 1987; Nelson 1991; Glosten, Jagannathan and Runkle 1993; Whitelaw 1994, 2000; Osiewalski and Pipień 2000; Pipień and Osiewalski 2001; Koopman and Hol Uspensky 2002; Li et al. 2005; Loudon 2006; Hibbert, Daigler and Dupoyet 2008; Kwiatkowski 2010; Abanto-Valle, Migon and Lachos 2011). The great variety of different approaches adopted in the cited papers indicates that the results on the risk premium phenomenon may heavily depend on the research methodology, as also pointed by Scruggs (1998), Harrison and Zhang (1999), Bollerslev and Zhou (2006) and Smith (2007). On the other hand, founded on some theoretical considerations, Backus and Gregory (1993) demonstrate that the theory does not exclude negative or even non-monotonic risk-return relationship - a stand also supported by Genotte and Marsh (1993), Veronesi (2000), Whitelaw (2000) and David and Veronesi (2009). Moreover, suggestions have also been made on dynamic instability of the risk premium (see e.g. French, Schwert and Stambaugh 1987; Harvey 1989; Loudon 2006; Darrat et al. 2011), some evidence for which has been provided by Chou, Engle and Kane (1992), Whitelaw (2000), Fiszeder and Kwiatkowski (2005a, b), Kim and Lee (2008) and Kwiatkowski (2010).

Much in line with the time-varying risk premium strand of literature we aim at empirical verification of Markovian breaks in the relation between expected return and volatility for Polish stock market data. Methodologically, the research builds heavily on our previous work (see Kwiatkowski 2010), in which we have introduced Bayesian two-state stochastic volatility with Markov switching in mean models (or SV-MS-M, in short). Similarly as before we follow a simple line of reasoning that if the 'true' risk premium should feature different states (regimes), then relying on a simple volatility-in-mean model with a constant risk premium parameter may spuriously imply its insignificance, as neglecting a possibly regime-changing pattern may lead to results that are somewhat 'averaged out' over the regimes.

Although we make no pretence of resolving the long-lasting risk premium puzzle, our paper contributes to the relevant literature on several counts. On the methodological

one, we allow for three rather than only two different risk premium regimes, and discuss the necessary modifications to estimation algorithm presented by Kwiatkowski (2010). Also, conditional upon the past moments of the  $K$ -state SV-MS-M process are derived. As regards the empirical facet of the present research, we follow the remarks concluding our previous study and shift the focus of interest from stock market indices to individual companies' share prices, since the relation between risk and return may be more evident in the case of individual stocks, rather than market aggregates. Adopting Bayesian methodology developed in the foregoing work, formal inference on the 'adequate' number of states is also carried out (though only for two- and three-state model specifications). Finally, an empirical justification for the superior data fit of the three-state SV-MS-M models is proposed.

The remainder of the paper is organized as follows. In the following section we generalise the two-state SV-MS-M model introduced by Kwiatkowski (2010), by allowing for  $K$  distinct risk premium regimes. Conditional moment structure of the generalised process is also presented. Section 3 addresses Bayesian estimation and model comparison issues, yet focusing mainly on their modifications (with respect to Kwiatkowski 2010) necessitated by the model extension. Empirical study, presented in Section 4, is divided into two subsections. Firstly, general results obtained for eleven Polish stock market datasets are discussed (Subsection 4.a). Then, we proceed with a more detailed investigation for a single company of Agora, the results for which appear most informative (Subsection 4.b). Finally, Section 5 concludes.

## 2 Stochastic volatility with Markov switching in mean process

Let the sequence  $\{S_t, t \in \mathbb{Z}\}$ , with  $\mathbb{Z}$  denoting the set of integers, constitute a  $K$ -state homogenous and ergodic Markov chain with a generic random variable  $S_t$  taking on values within the state space  $\mathbf{S} = \{1, 2, \dots, K\}$ . Transition probabilities are defined as  $p_{ij} \equiv Pr(S_t = j | S_{t-1} = i)$ , for  $i, j \in \mathbf{S}$ , and form the transition matrix  $P = [p_{ij}]_{i,j=1,2,\dots,K}$ . To ensure irreducibility and ergodicity of the chain we assume  $p_{ij} \in (0, 1)$  for  $i, j \in \mathbf{S}$ . The  $K$ -state stochastic volatility with Markov switching in mean process, SV-MS( $K$ )-M, is defined as follows.

**Definition 1** *A stochastic process  $\{y_t, t \in \mathbb{Z}\}$  follows the SV-MS( $K$ )-M process if for each  $t \in \mathbb{Z}$  the following conditions hold:*

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \gamma_{S_t} \sqrt{h_t} + \varepsilon_t \sqrt{h_t}, \quad (1)$$

$$\ln h_t = \mu + \varphi \ln h_{t-1} + \sigma \eta_t, \quad (2)$$

$$\{(\varepsilon_t, \eta_t)', t \in \mathbb{Z}\} \sim iiN^{(2)}(0_{(2 \times 1)}, I_2) \quad (3)$$

Łukasz Kwiatkowski

---

with  $iiN^{(2)}$  denoting independent and normally distributed bivariate random variables, and  $S_t$  representing the  $K$ -state homogenous and ergodic Markov process defined as above.

Equation (1), hereafter being referred to as the observation equation, defines a simple first-order autoregression on  $y_t$ , completed with the in-mean structure,  $\gamma_{S_t}\sqrt{h_t}$ , and the innovation term  $\varepsilon_t\sqrt{h_t}$ . From Equations (2) and (3) it follows that the innovations are governed by a basic stochastic volatility process (see Clark 1973; Taylor 1982, 1986; Pajor 2003). Random variable  $h_t$  is easily shown to be the conditional variance of  $y_t$ , once conditioning is made with respect to a  $\sigma$ -field generated by the lagged  $y_t$ 's, the current noise term  $\eta_t$  and the current state variable  $S_t$ , i.e.  $h_t = Var(y_t|\Psi_{t-1}, \eta_t, S_t)$ , where  $\Psi_{t-1}$  is the past information about the process  $\{y_t, t \in \mathbb{Z}\}$  up to time  $t-1$ . The above definition extends the one presented by Kwiatkowski (2010), though with regard solely to the number of regimes allowed for the switching parameter  $\gamma_{S_t}$ . The latter – much in vein of the ARCH-in-Mean (ARCH-M) model of Engle, Lilien and Robins (1987) – is intended to capture contemporaneous link between current volatility (conditional standard deviations) and return. Note that under  $\gamma_i = \gamma$  for each  $i \in \mathbf{S}$  the SV-MS-M process collapses to the SV-in-Mean structure (SV-M; cf. Koopman and Hol Uspensky 2002), whereas restricting all  $\gamma_i$ 's to equal zero yields the basic SV specification – both cases leaving the transition probabilities and state variables unidentified.

Although various functional forms of the in-mean structure, including the logarithm of  $h_t$  and identity map, are common in the literature, we adopt the square root of  $h_t$  as the original model specification developed by Engle, Lilien and Robins (1987), thereby assuming that - conditionally upon the current regime,  $S_t$  - an increase in the conditional variance transfers to a less than proportional rise in the expected rate of return. Under such a choice the observation equation admits the following form (see Kwiatkowski 2010):

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \xi_t \sqrt{h_t} \quad (4)$$

with  $\xi_t = \gamma_{S_t} + \varepsilon_t$ . Since the disturbances  $\{\varepsilon_t, t \in \mathbb{Z}\}$  are independent and normally distributed with zero mean and unit variance, it follows that - conditionally upon the parameters of the SV-MS-M process ( $\theta$ ) - distribution of  $\xi_t$  is a  $K$ -component mixture of Normals with state-dependent means:

$$p(\xi_t|\theta) = \sum_{i=1}^K \pi_i f_N^{(1)}(\xi_t|\gamma_i, 1) \quad (5)$$

with  $f_N^{(1)}(\cdot|a, b)$  denoting density of the univariate Normal distribution with mean  $a$  and variance  $b$ , and ergodic probabilities  $\pi_i \equiv Pr(S_t = i)$  serving as the 'weights'. For an ergodic  $K$ -state Markov chain the latter can be calculated according to the formula (cf. Krolzig 1997 p.17):

$$\pi = [\pi_1 \ \pi_2 \ \dots \ \pi_K] = \delta_K W_K^{-1}, \quad (6)$$

---

 Bayesian Analysis of a Regime Switching In-Mean Effect ...
 

---

where

$$\delta_K = [0 \ 0 \ \dots \ 0 \ 1]_{(1 \times K)},$$

$$W_K = \begin{bmatrix} I_{K-1} - P_{[1:(K-1),1:(K-1)]} & \iota_{K-1} \\ -P_{[K,1:(K-1)]} & 1 \end{bmatrix},$$

$$\iota_{K-1} = [1 \ 1 \ \dots \ 1]_{(1 \times (K-1))},$$

$P_{[m:n,p:q]}$  – block of  $P$ , comprised of rows  $m$  to  $n$  and columns  $p$  to  $q$  ( $m, n, p, q \in \{1, 2, \dots, K\}$ ,  $m \leq n$ , and  $p \leq q$ ) and  $I_{K-1}$  – the identity matrix of size  $K-1$ .

Investigating the  $S_{t-1}$ -conditional and unconditional (i.e. conditional only upon  $\theta$ ) moment structures of  $\xi_t$  in our previous work (see Kwiatkowski 2010) we inferred that – owing to (5) – the SV-MS-M process generates skewness in the past-conditional distribution of  $y_t$ , i.e.  $p(y_t | \Psi_{t-1}, \theta)$ . With straightforward derivations deferred to the Appendix, some basic characteristics of the relevant distribution are presented below.

**Theorem 2** Let  $\{y_t, t \in \mathbb{Z}\}$  follow the SV-MS( $K$ )-M process given by Definition 1, with the parameters collected in  $\theta$ . Denote the past-conditional moment of order  $r \in \mathbb{N}$  for random variable  $\gamma_{S_t}$  as  $\bar{\gamma}_{t|t-1}^{(r)} \equiv E(\gamma_{S_t}^r | \Psi_{t-1}, \theta) = E(\gamma_{S_t}^r | S_{t-1}, \theta) = \sum_{j=1}^K \gamma_{S_t=j}^r p_{S_{t-1}, S_t=j}$ . Then, past-conditional mean, variance and skewness coefficient (based on the third-order central moment) of  $y_t$  are given by:

$$E(y_t | \Psi_{t-1}, \theta) = \alpha_0 + \alpha_1 y_{t-1} + \bar{\gamma}_{t|t-1}^{(1)} \exp\left\{\frac{1}{2}(\mu + \varphi \ln h_{t-1}) + \frac{1}{8}\sigma^2\right\},$$

$$Var(y_t | \Psi_{t-1}, \theta) = \left[ \left( \bar{\gamma}_{t|t-1}^{(2)} + 1 \right) \exp\left(\frac{1}{4}\sigma^2\right) - \left( \bar{\gamma}_{t|t-1}^{(1)} \right)^2 \right] \exp\left(\mu + \varphi \ln h_{t-1} + \frac{1}{4}\sigma^2\right),$$

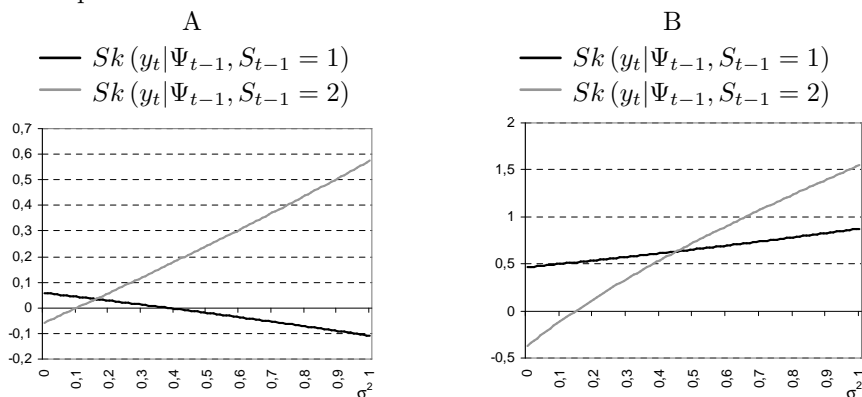
$$Sk(y_t | \Psi_{t-1}, \theta) = \frac{\left( \bar{\gamma}_{t|t-1}^{(3)} + 3\bar{\gamma}_{t|t-1}^{(1)} \right) \exp\left(\frac{3}{4}\sigma^2\right) - 3\bar{\gamma}_{t|t-1}^{(1)} \left( \bar{\gamma}_{t|t-1}^{(2)} + 1 \right) \exp\left(\frac{1}{4}\sigma^2\right) + 2\left( \bar{\gamma}_{t|t-1}^{(1)} \right)^3}{\left[ \left( 1 + \bar{\gamma}_{t|t-1}^{(2)} \right) \frac{1}{4}\sigma^2 - \left( \bar{\gamma}_{t|t-1}^{(1)} \right)^2 \right]^{\frac{3}{2}}}.$$

From Theorem 2 it follows that the SV-MS-M process features not only a time-varying past-conditional variance, but also mean. Even under no autoregressive part in the observation equation (i.e.  $\alpha_1 = 0$ ) conditional expectation of  $y_t$  evolves over time according to the Markov chain  $\{S_t\}$  (through  $\bar{\gamma}_{t|t-1}^{(1)}$ ) and the volatility process. Most importantly, though, the conditional skewness coefficient, as predicted by Kwiatkowski (2010), generally differs from zero, implying asymmetry of the corresponding density. Note that  $Sk(y_t | \Psi_{t-1}, \theta)$  depends on the volatility process neither through the intercept ( $\mu$ ) nor the 'elasticity of volatility' ( $\varphi$ ), but solely through variance of the innovations to  $\ln h_t$  ( $\sigma^2$ ). It follows that only the 'volatility of volatility', rather the mean volatility level, affects asymmetry of the past-conditional density. Complexity of the expression for skewness coefficient prevents one from further analytical investigation of  $Sk(y_t | \Psi_{t-1}, \theta)$  as a function of the parameters.

Łukasz Kwiatkowski

However, Figures 1 A and B – plotting the coefficient against different values of  $\sigma^2$  for two arbitrary SV-MS(2)-M specifications – suggest its monotonic dependence on the variability of volatility. Particularly, for a given state  $S_{t-1}$ , a shift in the asymmetry of the conditional density is possible at sufficiently large values of  $\sigma^2$ . (Notice that conditioning upon the past of the process  $\{y_t\}$  inherently involves conditioning on  $S_{t-1}$ . Hence  $Sk(y_t|\Psi_{t-1}, \theta) = Sk(y_t|\Psi_{t-1}, S_{t-1}, \theta)$ . Since  $S_{t-1} \in \mathbf{S}$ , one can consider  $K$  different conditional skewness coefficients in total). Moreover, the densities in question, being dependent on the Markov chain state at  $t - 1$ , may exhibit distinct asymmetry patterns (with respect to shape and intensity of skewness) for each  $S_{t-1} \in \mathbf{S}$ . Incidentally, under restriction of all  $\gamma_{S_t}$ 's being equal across the regimes, Theorem 2 yields valid expressions for conditional moments of the SV-M process. As hardly manageable as the formula for conditional skewness coefficients

Figure 1: Conditional skewness coefficients as a function of  $\sigma^2$  for some two-state SV-MS-M processes



$$\gamma_1 = -0.3; \gamma_2 = 0.6; p_{11} = p_{22} = 0.8 \quad \gamma_1 = -0.3; \gamma_2 = 2; p_{11} = 0.8; p_{22} = 0.95$$

may appear, Bayesian framework, equipped with MCMC techniques, enables one to obtain posterior densities of  $Sk(y_t|\Psi_{t-1}, S_{t-1} = i, \theta)$ ,  $i = 1, 2, \dots, K$ , with ease, thereby allowing statistical inference on asymmetries in the conditional data density.

### 3 Bayesian estimation for the $K$ -state SV-MS-M model

#### 3.1 General remarks

In the Section an estimation algorithm for Bayesian SV-MS( $K$ )-M models is presented. However, since our methodology is a direct generalisation of the one developed by

---

 Bayesian Analysis of a Regime Switching In-Mean Effect ...
 

---

Kwiatkowski (2010), we shall confine the following exposition only to the modifications necessitated by allowing for  $K$  rather than only two regimes.

We establish the notation first. Let  $y = (y_1, y_2, \dots, y_T)' \in Y \subseteq \mathbb{R}^T$  denote vector of  $T$  observations on logarithmic rates of return,  $h = (h_1, h_2, \dots, h_T)' \in H \subseteq \mathbb{R}_+^T$  – vector of latent conditional variances, and  $S = (S_1, S_2, \dots, S_T)' \in \mathbf{S}^T$  – vector of hidden Markov chain state variables. To ensure Markov chain states' identifiability, the risk premium parameter,  $\gamma_{S_t}$ , is parameterised as

$$\gamma_{S_t} = \gamma_1 + \sum_{j=2}^K \tau_j I(S_t \geq j), \quad (7)$$

with  $\gamma_1 \in \mathbb{R}$ ,  $\tau_j \leq 0$  for  $j \in \{2, 3, \dots, K\}$ , and  $I(\cdot)$  denoting the indicator function taking on the value of one if the event in parentheses occurs, and zero otherwise. Notice the cumulative nature of (7):

$$\gamma_{S_t} = \gamma_1 + \tau_2 + \tau_3 + \dots + \tau_{S_t} = \gamma_{S_{t-1}} + \tau_{S_t}, \quad S_t \in \{2, 3, \dots, K\},$$

indicating that the regime with a higher number of the state index features a lower in-mean effect. For estimation purpose we consider  $\gamma_1$  and contrasts  $\tau_j$  ( $j \in \{2, 3, \dots, K\}$ ), whereas posterior densities of  $\gamma_j$ 's are induced via (7). Parameters of the SV-MS-M model are arranged in  $\theta = (\alpha', \beta', \gamma_1, \tau, \sigma^2, q_1, q_2, \dots, q_K)' \in \Theta \subseteq \mathbb{R}^{K^2+5}$ , with  $\alpha = (\alpha_0, \alpha_1)'$ ,  $\beta = (\mu, \varphi)'$ ,  $\tau = (\tau_2, \tau_3, \dots, \tau_K)'$ , and  $q_i \equiv P_{[i,1:(K-1)]} = [p_{i1} \ p_{i2} \ \dots \ p_{i,K-1}]$  ( $i = 1, 2, \dots, K$ ) collecting free entries in the  $i$ -th row of matrix  $P$ . Note that  $\alpha$ ,  $\beta$  and  $\sigma^2$  are common to all switching and non-switching specifications.

Inference on all the unknown quantities of the model is based on the joint posterior distribution of  $\omega = (\theta', h', S')' \in \Omega = \Theta \times H \times \mathbf{S}^T$ , represented by its density:

$$p(\theta, h, S|y) \propto p(y|\theta, h, S) p(S|\theta) p(h|\theta) p(\theta), \quad (8)$$

where

$$p(y|\theta, h, S) \propto p(y_0) p(y|\theta, h, S) = p(y_0) \prod_{t=1}^T f_N(y_t | \alpha_0 + \alpha_1 y_{t-1} + \gamma_{S_t} \sqrt{h_t}, h_t), \quad (9)$$

$$p(S|\theta) \propto p(S_0) p(S|\theta) = p(S_0) \prod_{t=1}^T p(S_t | S_{t-1}, \theta) = p(S_0) \prod_{t=1}^T p_{S_{t-1}, S_t}, \quad (10)$$

$$p(h|\theta) \propto p(h_0) p(h|\theta) = p(h_0) \prod_{t=1}^T \left( \frac{1}{h_t} f_N(\ln h_t | \mu + \varphi \ln h_{t-1}, \sigma^2) \right), \quad (11)$$

$$p(\theta) = p(\alpha) p(\beta) p(\sigma^2) p(\gamma_1) \left( \prod_{j=2}^K p(\tau_j) \right) \left( \prod_{i=1}^K p(p_i) \right). \quad (12)$$

Łukasz Kwiatkowski

---

Note that in (12) we use  $p_i \equiv P_{(i,1:K)} = [p_{i1}, p_{i1}, \dots, p_{iK}]$  instead of  $q_i = [p_{i1}, p_{i1}, \dots, p_{i,K-1}]$ , which we find convenient for further considerations on prior assumptions. Initial conditions, i.e.  $y_0, h_0$  and  $S_0$ , are handled along the lines of Kwiatkowski (2010), and suppressed in the notation henceforth.

### 3.2 Prior structure

According to (12) mutual prior independence of the parameters is assumed. For all parameters not related directly to the in-mean structure, i.e.  $\alpha, \beta$ , and  $\sigma^2$ , we adopt the same prior specifications as in Kwiatkowski (2010), including truncated bivariate Normal distributions with zero correlations for  $\alpha$  and  $\beta$ , and the Inverse-Gamma distribution for  $\sigma^2$ :

$$p(\alpha) \propto f_N^{(2)}(\alpha|d_0, C_0^{-1}) I(|\alpha_1| < 1), \quad d_0 = 0_{(2 \times 1)}, C_0 = 0.01 \cdot I_2, \quad (13)$$

$$p(\beta) \propto f_N^{(2)}(\beta|b_0, A_0^{-1}) I(|\varphi| < 1), \quad b_0 = 0_{(2 \times 1)}, A_0 = 0.01 \cdot I_2, \quad (14)$$

$$p(\sigma^2) \propto f_{IG}(\sigma^2|\nu_1, \nu_2), \quad \nu_1 = 1.5, \nu_2 = 4, \quad (15)$$

where  $I_n$  stands for the  $(n \times n)$ -sized identity matrix and  $f_N^{(m)}$  – density of the  $m$ -variate Normal distribution. Truncations of the parameter space, made by the inequality restrictions in (13) and (14), are meant to guarantee non-explosiveness of the SV-MS-M process (or its second-order stationarity in the limiting case of  $t \rightarrow \infty$ ); see Kwiatkowski (2010). Densities (13)-(15) are shared across all model specifications considered in the paper, i.e. the SV, SV-M and SV-MS( $K$ )-M models. To complete the prior structure we specify densities for the model-specific parameters (i.e.  $\gamma_1, \tau, p_1, \dots, p_K$ ) as follows:

$$p(\gamma_1, \tau) \propto f_N^{(K)}(\gamma_1, \tau|\lambda_0, \Lambda_0^{-1}) \prod_{j=2}^K I(\tau_j < 0), \quad \lambda_0 = 0_{(K \times 1)}, \Lambda_0 = I_K, \quad (16)$$

$$p(p_i) \propto f_{Dir}(p_i|a_{i1}, a_{i2}, \dots, a_{iK}), \quad a_{ik} = 1, i, k \in \{1, 2, \dots, K\}, \quad (17)$$

with  $f_{Dir}$  standing for density of the Dirichlet distribution – a conditionally conjugate and therefore most common choice for  $p(p_i)$  in Bayesian (Markovian) mixture models. (By convention one of the coordinates in  $p_i$  is bound with others by an obvious restriction:  $\sum_{j=1}^K p_{ij} = 1$ ). It follows that transition probabilities in each row of  $P$  are uniformly distributed over a unit  $(K - 1)$ -simplex. Note that in the special case of a two-state model the above Dirichlet prior is equivalent to the Beta(1, 1) distribution. In the SV-M model a standard Normal prior for the risk premium parameter  $\gamma$  is assumed, reflecting our convictions on the magnitude of the in-mean effect. We emphasise that the truncated  $K$ -variate standard Normal in (16) induces a linear



decrease in the mean and a linear increase in the variance of priors for consecutive  $\gamma_k$ 's. Let  $\phi(x) \equiv f_N^{(1)}(x|0, 1)$ . Then, for  $k = 2, \dots, K$  we have:

$$E(\gamma_k) = E\left(\gamma_1 + \sum_{j=2}^k \tau_j\right) = -2(k-1)\phi(0) \approx -0.798(k-1), \quad (18)$$

and, owing to mutual prior independence among  $\gamma_1, \tau_2, \tau_3, \dots, \tau_k$ :

$$Var(\gamma_k) = Var(\gamma_1) + \sum_{j=2}^k Var(\tau_j) = 1 + (k-1) \left[1 - 4(\phi(0))^2\right] \approx 1 + 0.363(k-1). \quad (19)$$

The prior structure exposed in (13)-(17) is intended to represent our vague beliefs as of the model parameters. Particularly, (16) reflects our prior conviction as of the magnitude of the risk-return tradeoff, although the 'peculiarity' of priors induced for  $\gamma_j$ 's, manifesting itself through (18) and (19), may call for further search for a 'less informative' prior structure of the in-mean coefficients. This, however, is largely determined by the adopted parameterisation (see (7)), which is fairly common to Markov switching models entertained in the literature (see e.g. So, Lam and Li 1998; Kalimipalli and Susmel 2004; Shibata and Watanabe 2005).

### 3.3 Sampling algorithm

According to our previous study, obtaining posterior density presented in (8) requires use of Markov Chain Monte Carlo (MCMC) techniques, including the Gibbs sampler, the Metropolis-Hastings algorithm and the forward-filtering-backward-sampling scheme (FFBS); see Kwiatkowski (2010). Generating a pseudo-random sample from the joint posterior (8) is divided into three stages, in which the three: the parameters ( $\theta$ ), conditional variances ( $h_t, t = 1, 2, \dots, T$ ) and Markov chain state variables ( $S_t, t = 1, 2, \dots, T$ ), are sampled from their full conditional posterior distributions. A detailed description of each step for the two-state SV-MS-M model has already been presented by Kwiatkowski (2010). Therefore, only modifications to the MCMC algorithm, necessitated by the generalisation of the number of regimes, are discussed beneath.

#### 3.3.1 Sampling model parameters

As far as model parameters are concerned, introducing more than two states in the underlying Markov chain affects conditional posteriors only for the in-mean parameters ( $\gamma_1, \tau_j$ 's) and transition probabilities. We opt for joint simulation of all the coefficients in the observation equation (1) – along with  $\alpha_0$  and  $\alpha_1$  – as it is generally agreed that simultaneous (block) sampling facilitates convergence of the MCMC chain. Below, under convention of  $\theta_{-(\lambda)}$  denoting vector  $\theta$  with the  $\lambda$

Łukasz Kwiatkowski

coordinate removed, the full conditional posteriors for the relevant parameters are displayed:

$$p(\alpha, \gamma_1, \tau | \theta_{-(\alpha, \gamma_1, \tau)}, h, S, y) \propto f_N^{(4)}(\alpha, \gamma_1, \tau | g_*, G_*^{-1}) I(|\alpha_1| < 1) \prod_{j=2}^K I(\tau_j < 0),$$

where

$$G_* = G_0 + M'M, \quad g_* = G_*^{-1}(G_0 g_0 + M'u), \quad G_0 = \begin{bmatrix} C_0 & 0_{(K \times K)} \\ 0_{(K \times K)} & \Lambda_0 \end{bmatrix},$$

$$M = \begin{bmatrix} h_1^{-0.5} & y_0 h_1^{-0.5} & 1 & I(S_1 \leq 2) & I(S_1 \leq 3) & \cdots & I(S_1 = K) \\ h_2^{-0.5} & y_1 h_2^{-0.5} & 1 & I(S_2 \leq 2) & I(S_2 \leq 3) & \cdots & I(S_2 = K) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_T^{-0.5} & y_{T-1} h_T^{-0.5} & 1 & I(S_T \leq 2) & I(S_T \leq 3) & \cdots & I(S_T = K) \end{bmatrix},$$

$$g_0 = [d'_0 \ \lambda'_0]' \quad u = \begin{bmatrix} y_1 h_1^{-0.5} \\ y_2 h_2^{-0.5} \\ \vdots \\ y_T h_T^{-0.5} \end{bmatrix},$$

$$p(p_i | \theta_{-(p_i)}, h, S, y) \propto f_{Dir}(p_i | a_{i1}^*, a_{i2}^*, \dots, a_{iK}^*), \quad i = 1, 2, \dots, K,$$

$$a_{ij}^* = a_{ij} + n_{ij}, \quad \text{for } i, j = 1, 2, \dots, K$$

$$n_{ij} = \sum_{t=2}^T I(S_{t-1} = i) I(S_t = j)$$

Conditional posteriors for  $\beta$  and  $\sigma^2$  fully coincide with the ones provided by Kwiatkowski (2010), and are therefore omitted in the present paper.

### 3.3.2 Sampling conditional variances and Markov chain state variables

Generating latent variables in the SV-MS-M models is a more challenging task, which is attributable to their non-standard full conditional posteriors (see Kwiatkowski 2010). Nevertheless, adapting the Metropolis-Hastings algorithm, developed by Jacquier, Polson and Rossi (1994) for simple SV structures (see also Pajor 2003), Kwiatkowski (2010) notices that the very same Inverse-Gamma proposal density, from which candidate values of each  $h_t$  ( $t = 1, 2, \dots, T$ ) need to be drawn at each MCMC run, can successfully be utilised also in the case of the two-state SV-MS-M model. Noticing that the approach – requiring no further alterations – is also valid for the general  $K$ -state specification, we hereby refer the reader to the cited work for a more detailed description of the procedure.

Similarly, the forward-filtering-backward-sampling scheme (see Carter and Kohn 1994; Chib 1996), suitably tailored by Kwiatkowski (2010) to sample the switching process state variables, applies also in the present context of the SV-MS( $K$ )-M models. A single modification needs to be applied to the relevant formulae presented in the aforementioned paper, though. Namely, every summation over the state space ( $\mathbf{S}$ ) requires now to cover all  $K$  – instead of only two – states. Again, a detailed exposition of the FFBS routine is to be found in the cited paper.

## 4 Empirical study

The main objective of the following study is to examine several stock market individual companies, representative of Polish major stock market sectors, in search for possible Markovian breaks in the risk-return relationship. To this end four model specifications are estimated for each dataset: the basic stochastic volatility (SV), SV-in-Mean (SV-M), and the two- and three-state switching models (SV-MS(2)-M, SV-MS(3)-M). Formal Bayesian model comparison (via marginal data density) in terms of the in-sample fit hints at the 'adequate', data-supported number of distinct risk premium regimes, whereas a closer inspection of the parameters' posterior marginal moments and (bivariate) densities allows inference on the switching process dynamics and regime-specific in-mean effects, providing an accompanying picture of the statistical uncertainty.

Table 1: Mean and volatility equation specifications for the analysed models

Model	Observation equation	Log-volatility equation
$M_0$ : SV	$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t \sqrt{h_t}$	$\ln h_t = \mu + \varphi \ln h_{t-1} + \sigma \eta_t$
$M_1$ : SV-M	$y_t = \alpha_0 + \alpha_1 y_{t-1} + \gamma \sqrt{h_t} + \varepsilon_t \sqrt{h_t}$	
$M_2$ : SV-MS(2)-M	$y_t = \alpha_0 + \alpha_1 y_{t-1} + \gamma s_t \sqrt{h_t} + \varepsilon_t \sqrt{h_t}$	
$M_3$ : SV-MS(3)-M		

### 4.1 Datasets

In what follows we consider eleven times series of daily logarithmic rates of return (defined as  $y_t = 100 \ln \frac{x_t}{x_{t-1}}$ , with  $x_t$  denoting the asset closing price at time  $t = 1, 2, \dots, T$ ) on shares of selected companies quoted at the Warsaw Stock Exchange. All the equity prices have been obtained from [www.bossa.pl](http://www.bossa.pl).

Following previous studies by Nelson (1991), Fiszeder and Kwiatkowski (2005) and Pipień (2007), who observed that netting the data of a risk free rate of return bears little impact on the final results, we analyse nominal rather than excess rates. Table 2 contains a detailed list of the series under consideration, whereas Table 3 reports on basic descriptive statistics of the data. Figure 2 plots the series. The companies

---

 Łukasz Kwiatkowski
 

---

have been selected according to the following key. Primarily, a company featuring the highest capitalisation (as on October 12, 2010) within each sectoral subindex of the Warsaw Stock Exchange Main Index (WIG) has been chosen. Most data series start with the commencement of each company being quoted on the Warsaw Stock Exchange (although data available before January 3, 2001 has been trimmed). If the series' size turned up to be less than 1000 data points, then a following company - down the capitalisation ranking - was taken. All but one datasets end on October 11, 2010. The exception is made with regard to Kofola (with the sample ending as soon as on September 29, 2008), for which modelling the entire dataset available (including observations following September 29, 2008) results in a lack of the MCMC convergence (as monitored via standardised CUMSUM plots; see Bauwens and Lubrano 1998; Pajor 2003) for both the SV-MS-M and non-switching models. We found that what impedes the MCMC sampler is an abrupt rise in Kofola's share price by ca. 35% (in terms of logarithmic rates of return) on September 30, 2008, which was spurred with a call for subscription for the company's shares on previous day, and subsequently followed by a short period of abnormally low absolute returns (until November 3, 2008). Hence, the analysed sample path comprises observations only up to the date before the subscription call. Incidentally, other datasets - despite featuring numerous

Table 2: Dataset list

No.	Company	Sector	Sample range		Sample size ( $T$ )
			Start	End	
1	AGORA	media	Jan. 3, 2001	Oct. 11, 2010	2455
2	ASSECOPOL	IT	Jan. 3, 2001		2455
3	CIECH	chemicals	Feb. 11, 2005		1422
4	GTC	developers	May 7, 2004		1618
5	KOFOLA	food	Sep. 2, 2003	Sep. 29, 2008	1276
6	KOGENEREA	energy	Sep. 3, 2002	Oct. 11, 2010	2034
7	PBG	construction	July 7, 2004		1575
8	PKN ORLEN	oil & gas	Jan. 3, 2001		2455
9	PKO BP	banking	Nov. 12, 2004		1487
10	TPSA	telecommunication	Jan. 3, 2001		2455
11	TVN	media	Dec. 8, 2004		1469

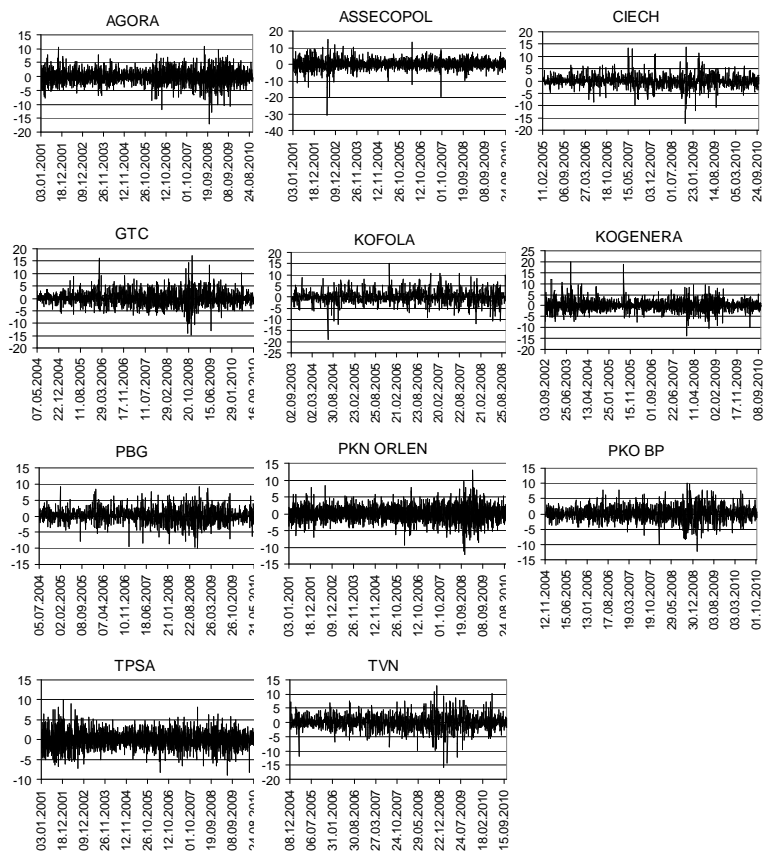
outliers (see Figure 2) - have not raised our concerns about the convergence of the MCMC algorithm. However, the switching structures required more burnt-in passes as compared with the simpler SV and SV-M models, which remains in accord with Kwiatkowski (2010). Note that two companies representing the very same media sector have been taken into account, i.e. Agora and TVN. Although it is the latter that satisfies the condition of maximal capitalisation within the sector (as on October 12, 2010), the former has proved to be an interesting empirical material in some previous research on the Polish financial market microstructure (see e.g. Doman 2008). Thus, based on the premise that it may also deliver some interesting insights

---

 Bayesian Analysis of a Regime Switching In-Mean Effect ...
 

---

Figure 2: Analysed series of daily logarithmic rates of return



into the risk-return relationship, we decided to include Agora in the present study as well.

## 4.2 General results

Below we present results obtained within four model specifications: SV, SV-M, SV-MS(2)-M and SV-MS(3)-M (see Table 1), estimated for each of the eleven companies (see Table 2). In each case posterior analysis is based on 1000000 MCMC samplings from the relevant joint posterior, preceded by 50000 transient runs for the non-switching models and 500000 passes for the switching structures. Calculations have been carried out with the author's own codes run under GAUSS 10.

We report on the model comparison first. Relevant quantities, including decimal logarithms of the marginal data density values and Bayes factors (against the

Łukasz Kwiatkowski

Table 3: Descriptive statistics for the analysed datasets

Company	min	max	average	standard deviation	skewness	excess kurtosis
AGORA	-16.919	10.851	-0.047	2.424	-0.225	2.781
ASSECOPOL	-30.561	14.622	-0.007	2.706	-0.736	11.064
CIECH	-17.313	13.604	-0.011	2.465	-0.008	6.076
GTC	-14.660	17.280	0.051	2.835	0.276	3.768
KOFOLA	-18.848	14.953	0.003	2.749	-0.030	4.297
KOGENERA	-13.976	20.067	0.105	2.394	0.784	7.259
PBG	-10.003	9.278	0.119	2.179	0.004	2.209
PKN ORLEN	-12.158	12.866	0.023	2.196	-0.024	1.955
PKO BP	-12.223	9.973	0.040	2.360	-0.014	1.955
TPSA	-9.022	10.178	-0.019	2.083	0.063	1.264
TVN	-15.932	12.859	0.063	2.599	-0.261	3.587

SV specification), are displayed in Table 4. Marginal data density of model  $M_l$  ( $l = 0, 1, 2, 3$ ) with all the parameters and latent variables collected in  $\omega_{(l)} \in \Omega_l$ , is defined as:

$$p(y|M_l) = \int_{\Omega_l} p(\omega_{(l)}, y|M_l) d\omega_{(l)} = \int_{\Omega_l} p(y|\omega_{(l)}, M_l) p(\omega_{(l)}|M_l) d\omega_{(l)},$$

and, along the lines of Kwiatkowski (2010), numerically evaluated via the Newton and Raftery (1994) technique, according to formula:

$$p(y|M_l) \approx \left[ \frac{1}{N} \sum_{q=M+1}^{M+N} \frac{1}{p(y|\omega_{(l)}^{(q)}, M_l)} \right]^{-1},$$

where  $M$  is the number of the burnt-in passes,  $N$  – the number of drawings from the joint posterior,  $q$  – the index of a single pass of the sampling procedure ( $q = 1, 2, \dots, M, \dots, M + N - 1, M + N$ ), and  $\omega_{(l)}^{(q)}$  – the outcome on  $\omega_{(l)}$  from the  $q$ -th cycle. A pairwise comparison of different model structures is performed by means of Bayes factors, calculated as:

$$BF_{k,l} = \frac{p(y|M_k)}{p(y|M_l)}$$

Under equal prior odds of each model (i.e.  $p(M_l) = \frac{1}{4}$ ),  $BF_{k,l}$  equals the posterior odds ratio of  $M_k$  against  $M_l$ . It appears that in all but one cases of the analysed series (TVN being an exception) the three-state SV-MS-M model overtakes the competition. Its data fit superiority is most evident for Ciech and Agora, for which the SV-MS(3)-M model is more probable *a posteriori* than the non-switching specifications (SV, SV-M) by as much as ca. 26 and 23 orders of magnitude, respectively. About a half of that is attained by the three-state model (as compared with the basic SV process) in the case of Assecopol, Kogenera and PBG (with  $\log BF_{3,0}$  hovering around 13). Interestingly,

## Bayesian Analysis of a Regime Switching In-Mean Effect ...

Table 4: Bayesian model comparison

Company	$M_0$ : SV		$M_1$ : SV-M			$M_2$ : SV-MS(2)-M			$M_3$ : SV-MS(3)-M		
	$\log p(y)$	R	$\log p(y)$	$\log BF_{1,0}$	R	$\log p(y)$	$\log BF_{2,0}$	R	$\log p(y)$	$\log BF_{3,0}$	R
AGORA	-2306.04	3	-2307.86	-1.81	4	-2303.82	2.23	2	-2282.61	23.43	1
ASSECOPOL	-2347.05	3	-2348.47	-1.42	4	-2344.01	3.04	2	-2332.72	14.32	1
CIECH	-1253.9	4	-1253.23	0.66	3	-1238.61	15.29	2	-1227.3	26.6	1
GTC	-1617.81	4	-1615.72	2.09	3	-1613.54	4.27	2	-1611.15	6.67	1
KOFOLA	-1224.32	3	-1223.61	0.71	2	-1224.77	-0.45	4	-1220.81	3.51	1
KOGENERA	-1788.17	4	-1783.49	4.68	3	-1778.55	9.62	2	-1775.02	13.15	1
PBG	-1409.74	4	-1406.26	3.47	3	-1396.34	13.4	2	-1395.9	13.84	1
PKN ORLEN	-2263.53	4	-2262.33	1.2	3	-2260.17	3.36	2	-2255.6	7.93	1
PKO BP	-1393.95	4	-1393.7	0.25	2	-1393.81	0.13	3	-1386.24	7.71	1
TPSA	-2216.74	4	-2212.58	4.16	3	-2211.56	5.18	2	-2207.01	9.73	1
TVN	-1427.94	2	-1431.52	-3.58	4	-1426.77	1.17	1	-1428.32	-0.39	3

Note: 'log' and 'R' stand for decimal logarithm and model rank, respectively;  $BF_{k,l}$  denotes Bayes factor defined as  $BF_{k,l} = p(y|M_k)/p(y|M_l)$ .

the two-state structure estimated for PBG performs comparably, indicating that introducing yet another risk premium state is superfluous. The two-state models estimated for Ciech and Kogenera, although favoured against simple SV by ca. 15 and 9 orders of magnitude, respectively, seem to leave some room for improvement achieved by allowing for three separate regimes. With regard to the other companies, values of the marginal data density may indicate that the SV-MS(2)-M model performs at least as good as the simplest SV process. However, due to inherent numerical instability of the Newton and Raftery (1994) algorithm, the corresponding values of  $\log BF_{2,0}$ , ranging from -0.45 (Kofola) to 5.18 (TPSA), should be perceived with caution.

As regards the specification featuring a constant risk premium parameter, only in the case of Kogenera, TPSA and PBG the SV-M model fit may somewhat imply a need for the risk-return relationship to be accounted for (since  $\log BF_{1,0} \approx 4$ ). However, the conjecture is hardly supported by the marginal characteristics of the in-mean parameter (see Table 5), with the posterior averages of  $\gamma$  being close to zero and the corresponding densities exhibiting relatively large dispersion. As a matter of fact, the aforesaid also holds for all other companies under consideration (see Table 5). Finally, bearing in mind the infinite variance of the Newton and Raftery (1994) estimator, it is legitimate to conclude that in the case of two equities: Kofola and TVN, all four model specifications perform equally well. Tables 6 and 7 contain posterior means and standard deviations of the parameters related to the Markov switching risk-return relationship within the SV-MS(2)-M and SV-MS(3)-M models, respectively. In the case of the two-state structures a common pattern emerges in the posterior averages of  $\gamma_1$  and  $\gamma_2$ , with  $E(\gamma_1|y, M_2) > 0$  and  $E(\gamma_2|y, M_2) < 0$  for each company, suggesting that the first and the second state are the ones of, respectively, positive and negative in-mean effect. However, the corresponding posterior distributions,  $p(\gamma_1|y, M_2)$  and  $p(\gamma_2|y, M_2)$ , exhibit a comparatively large dispersion, as indicated by

Łukasz Kwiatkowski

Table 5: Posterior means (and standard deviations) of the risk premium parameter in the SV-M models

Company	$E(\gamma y)$ ( $D(\gamma y)$ )	Company	$E(\gamma y)$ ( $D(\gamma y)$ )
AGORA	-0.024 (0.061)	PBG	0.064 (0.062)
ASSECOPOL	0.047 (0.061)	PKN ORLEN	-0.111 (0.081)
CIECH	0.043 (0.055)	PKO BP	-0.096 (0.079)
GTC	0.015 (0.064)	TPSA	-0.090 (0.077)
KOFOLA	0.131 (0.071)	TVN	-0.007 (0.082)
KOGENERA	0.068 (0.047)		

Table 6: Posterior means (and standard deviations) of the risk premium parameter in the SV-M models

Company	$p_{11}$	$p_{22}$	$\gamma_1$	$\gamma_2$	Company	$p_{11}$	$p_{22}$	$\gamma_1$	$\gamma_2$
AGORA	0.435 (0.237)	0.685 (0.266)	0.521 (0.472)	-0.252 (0.194)	PBG	0.507 (0.161)	0.811 (0.176)	1.061 (0.45)	-0.344 (0.205)
ASSECOPOL	0.49 (0.273)	0.679 (0.227)	0.587 (0.448)	-0.278 (0.298)	PKN ORLEN	0.486 (0.286)	0.541 (0.278)	0.139 (0.297)	-0.318 (0.24)
CIECH	0.522 (0.256)	0.795 (0.117)	1.144 (0.612)	-0.528 (0.263)	PKO BP	0.309 (0.274)	0.651 (0.249)	0.495 (0.476)	-0.323 (0.176)
GTC	0.378 (0.311)	0.656 (0.286)	0.701 (0.62)	-0.264 (0.322)	TPSA	0.3 (0.213)	0.682 (0.229)	0.549 (0.395)	-0.349 (0.179)
KOFOLA	0.622 (0.242)	0.606 (0.234)	0.562 (0.364)	-0.299 (0.311)	TVN	0.437 (0.282)	0.597 (0.279)	0.343 (0.383)	-0.206 (0.236)
KOGENERA	0.561 (0.292)	0.532 (0.276)	0.375 (0.376)	-0.258 (0.374)					

their posterior standard deviations. The same holds also for transition probabilities' densities,  $p(p_{11}|y, M_2)$  and  $p(p_{22}|y, M_2)$ , the means of which – fluctuating around a half – are hardly indicative of the regimes' persistence one would expect. On the other hand, as far as the two-state models are concerned, this evidence overturning the hypothesis of Markovian breaks within the risk premium is of rather little surprise in view of the results on the in-sample fit, in terms of which for most companies the SV-MS(2)-M and simple SV models perform comparably. The results obtained for the two exceptions: Ciech ( $\log BF_{2,0} = 15.29$ ) and PBG ( $\log BF_{2,0} = 13.40$ ), share the above characteristics of posterior distributions, although  $E(\gamma_1|y, M_2)$ 's are relatively more distant from zero than in the case of the other datasets.

For all but one companies under study posterior analysis of the marginal characteristics obtained within the three-state SV-MS-M models is, again, hardly



Bayesian Analysis of a Regime Switching In-Mean Effect ...

Table 7: Posterior means (and standard deviations) of the risk premium parameters and transition probabilities in the SV-MS(3)-M models

Company	AGORA			ASSECOPOL			CIECH			
	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	
$\gamma_{S_t}$	1.628 (0.332)	-0.304 (0.155)	-2.532 (0.691)	1.056 (0.486)	-0.046 (0.292)	-1.071 (0.912)	1.644 (0.559)	0.103 (0.513)	-0.817 (0.504)	
P	$S_{t-1} = 1$	0.296 (0.068)	0.619 (0.118)	0.085 (0.111)	0.323 (0.212)	0.454 (0.262)	0.223 (0.25)	0.326 (0.182)	0.256 (0.201)	0.418 (0.252)
	$S_{t-1} = 2$	0.115 (0.069)	0.807 (0.145)	0.077 (0.111)	0.251 (0.167)	0.489 (0.271)	0.26 (0.227)	0.31 (0.193)	0.39 (0.241)	0.3 (0.207)
	$S_{t-1} = 3$	0.382 (0.154)	0.39 (0.153)	0.228 (0.11)	0.29 (0.214)	0.256 (0.204)	0.453 (0.212)	0.124 (0.103)	0.242 (0.214)	0.634 (0.229)
Company	GTC			KOFOLA			KOGENERA			
	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	
$\gamma_{S_t}$	1.144 (0.536)	-0.053 (0.219)	-0.491 (0.398)	0.825 (0.386)	0.066 (0.331)	-0.571 (0.379)	0.663 (0.437)	0.03 (0.237)	-0.432 (0.339)	
P	$S_{t-1} = 1$	0.188 (0.164)	0.426 (0.27)	0.386 (0.274)	0.479 (0.24)	0.338 (0.218)	0.183 (0.182)	0.371 (0.226)	0.375 (0.235)	0.254 (0.215)
	$S_{t-1} = 2$	0.214 (0.181)	0.436 (0.268)	0.35 (0.247)	0.28 (0.201)	0.367 (0.236)	0.353 (0.23)	0.291 (0.222)	0.369 (0.249)	0.34 (0.234)
	$S_{t-1} = 3$	0.176 (0.176)	0.411 (0.255)	0.414 (0.255)	0.302 (0.195)	0.263 (0.209)	0.435 (0.234)	0.312 (0.216)	0.328 (0.229)	0.361 (0.226)
Company	PBG			PKN ORLEN			PKO BP			
	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$	
$\gamma_{S_t}$	1.378 (0.345)	-0.187 (0.374)	-0.799 (0.456)	0.482 (0.264)	-0.156 (0.233)	-0.574 (0.204)	0.96 (0.273)	-0.251 (0.188)	-0.699 (0.224)	
P	$S_{t-1} = 1$	0.424 (0.119)	0.418 (0.163)	0.158 (0.161)	0.269 (0.154)	0.296 (0.257)	0.435 (0.273)	0.163 (0.114)	0.67 (0.204)	0.167 (0.189)
	$S_{t-1} = 2$	0.137 (0.12)	0.254 (0.223)	0.608 (0.251)	0.377 (0.259)	0.3 (0.21)	0.323 (0.255)	0.123 (0.121)	0.354 (0.197)	0.523 (0.212)
	$S_{t-1} = 3$	0.209 (0.126)	0.438 (0.243)	0.352 (0.276)	0.248 (0.229)	0.389 (0.24)	0.363 (0.189)	0.393 (0.172)	0.231 (0.18)	0.377 (0.192)
Company	TPSA			TVN						
	$S_t = 1$	$S_t = 2$	$S_t = 3$	$S_t = 1$	$S_t = 2$	$S_t = 3$				
$\gamma_{S_t}$	0.733 (0.344)	-0.154 (0.277)	-0.542 (0.249)	0.588 (0.389)	-0.055 (0.215)	-0.398 (0.284)				
P	$S_{t-1} = 1$	0.221 (0.144)	0.391 (0.239)	0.387 (0.243)	0.284 (0.196)	0.37 (0.247)	0.346 (0.251)			
	$S_{t-1} = 2$	0.22 (0.184)	0.37 (0.241)	0.41 (0.253)	0.278 (0.218)	0.363 (0.243)	0.359 (0.244)			
	$S_{t-1} = 3$	0.264 (0.193)	0.342 (0.23)	0.394 (0.231)	0.28 (0.22)	0.354 (0.241)	0.366 (0.228)			

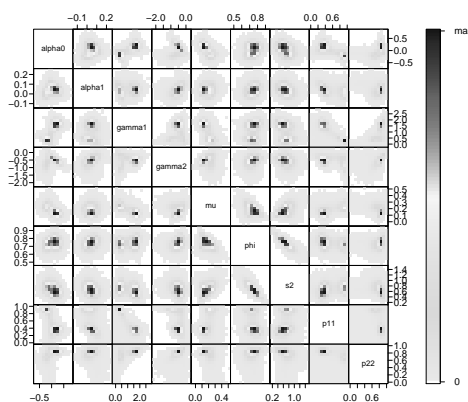
enlightening. Despite a common pattern of  $E(\gamma_1|y, M_3) > 0$ ,  $E(\gamma_3|y, M_3) < 0$  and  $E(\gamma_2|y, M_3)$  hovering around zero, which can easily be observed (notably for Agora, Ciech, Assecopol and PBG – datasets most supportive of the SV-MS(3)-M model), posterior means of the transition probabilities – usually staying in the vicinity of a third – along with a considerable posterior spread featured by  $\gamma_j$ 's and  $p_{ij}$ 's preclude any conclusive inferences on switches within the risk-return relation. The sole exception of Agora is discussed at length in the following subsection. In view of

Łukasz Kwiatkowski

the data fit superiority enjoyed by the three-state models, the above conclusions may be somewhat disappointing. However, we stress that posterior distributions obtained within both the two- and the three-state models may exhibit multimodality and other irregularities that render posterior averages quite uninformative. The issue is illustrated in Figures 3 and 6, depicting bivariate posterior densities of the parameters in the two- and the three-state SV-MS-M models, respectively, obtained for returns on Ciech's share prices – a dataset most supportive of the two switching specifications. Comparing the pictures outlined within the two- and the three-state model, it may appear that allowing for an additional risk premium state alleviates the multimodality found in  $p(\theta|y, M_2)$  – see the left panel in Figure 6. However, densities of the transition probabilities (see the right panel in Figure 6) contradict the impression.

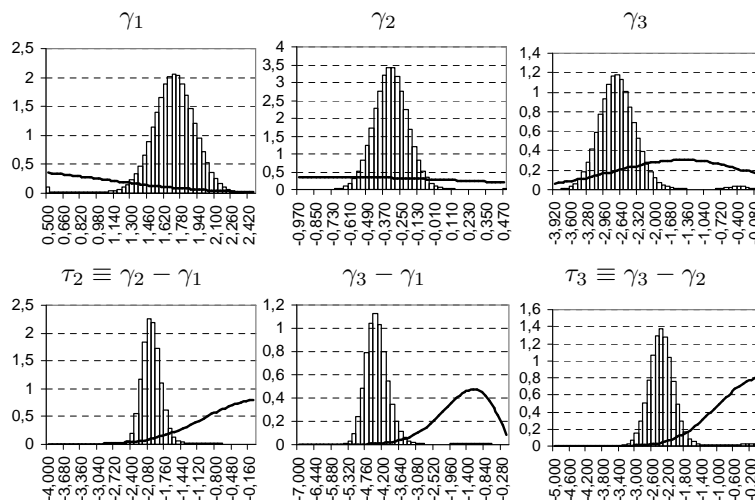
Although not presented in the paper, all plots obtained within the SV-MS(2)-M

Figure 3: Bivariate marginal posterior densities of the SV-MS(2)-M model's parameters for Ciech



models estimated for the other datasets emerge more regular (validating – to some extent – the use of basic moments for posterior inferences), whereas the picture revealed for Ciech's three-state model – the one of pronounced irregularities in the transition probabilities' (rather than the other parameters') posterior densities – is almost invariably shared across all the time series. (Relevant figures are available upon request). As gathered from Figure 7, yet again Agora represents a single exception, revealing – in the case of the SV-MS(3)-M model – regular and apparently unimodal bivariate marginal posteriors, according to which a more in-depth analysis – presented in the next subsection – is due.

Figure 4: Marginal posteriors (bars) and priors (solid line) of in-mean parameters within the SV-MS(3)-M model for Agora



### 4.3 Specific results for Agora

Since the results obtained for Agora appear most convincing (or, at least, intriguing), we devote the following subsection to discuss these in some more detail, with a particular focus on the inference within the three-state SV-MS-M model ( $M_3$ ).

#### 4.3.1 The SV-MS(3)-M model

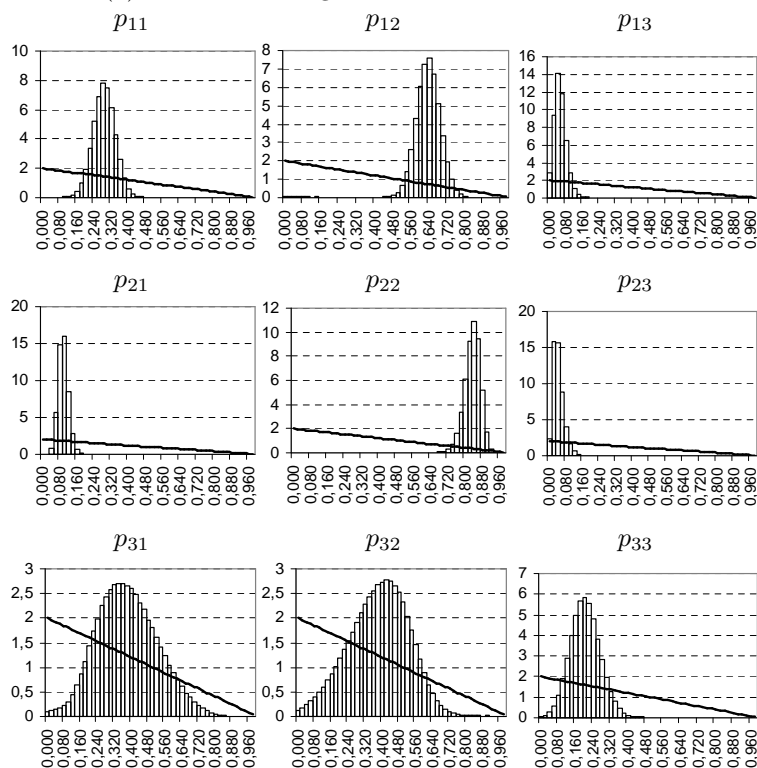
A closer inspection of the univariate marginal posteriors, presented in Figures 4 and 5, reveal some slight abnormalities – including minor additional modes and long tails - that made themselves somewhat inconspicuous in the bivariate plots (see Figure 7). In each case, however, posterior densities are sharply distinguishable from their prior counterparts, providing evidence of a strong data contribution to the inference. Particularly, as densities of the risk premium contrasts:  $\tau_2 \equiv \gamma_2 - \gamma_1$ ,  $\tau_3 \equiv \gamma_3 - \gamma_2$  and  $\gamma_3 - \gamma_1$ , are well separated from zero (see the bottom panels in Figure 4), it follows that the risk-return relationship is subject to switches over three quite distinct states, with the first and the third one pertaining to a strong positive and, respectively, a strong negative in-mean effect, and the second state – to a weakly negative relation (see Table 7). Note, however, that the two states of a pronounced in-mean effect ( $S_t = 1$  and  $S_t = 3$ ) are short-lived, with  $E(p_{11}|y)$  and  $E(p_{33}|y)$  hovering around 0.25 (see Table 7) and posterior densities of expected durations (defined as  $Dur_i = \frac{1}{1-p_{ii}}$ ,  $i = 1, 2, \dots, K$ ; see Hamilton, 1989, p.374) squeezed close to their means of  $E(Dur_1|y) = 1.43$  and  $E(Dur_3|y) = 1.30$  trading day; see Figure 9. On the other hand, the second regime prevails (on average) for five times as

---

 Łukasz Kwiatkowski
 

---

Figure 5: Marginal posteriors (bars) and priors (solid line) of transition probabilities within the SV-MS(3)-M model for Agora



much, with  $E(Dur_2|y) = 6.64$  and  $E(p_{22}|y) = 0.807$ . In view of quite an unusual pattern in the  $P$ 's diagonal posterior means, which implies a considerable variability of the regime-switching process, providing an empirical justification for the SV-MS(3)-M model's superior data fit seems quite a challenge. It may be that the untypical dynamics of the underlying Markov chain, combined with state-specific in-mean effects, enables the model to capture conditional skewness of the data. This, however, does not seem to be the case, for posterior distributions of the relevant coefficients,  $Sk(y_t|\Psi_{t-1}, S_{t-1} = i, \theta)$ ,  $i = 1, 2, 3$  (see Theorem 2) – with zero enjoying relatively high density values – do not reject conditional symmetry of the returns. (Incidentally, notice a left-hand skew in the prior densities of the coefficients. The asymmetry is attributable to the asymmetric priors of  $\gamma_j$ 's – somewhat enforced by the parameterisation in (7)).

Actually, the answer is provided by scatter plots in Figure 11, displaying posterior

Bayesian Analysis of a Regime Switching In-Mean Effect ...

Figure 6: Bivariate marginal posterior densities of the SV-MS(3)-M model's parameters for Ciech (top panel:  $\alpha_0, \alpha_1, \gamma_1, \gamma_2, \gamma_3, \mu, \varphi$  and  $\sigma^2$  coordinates; bottom panel: transition probabilities)

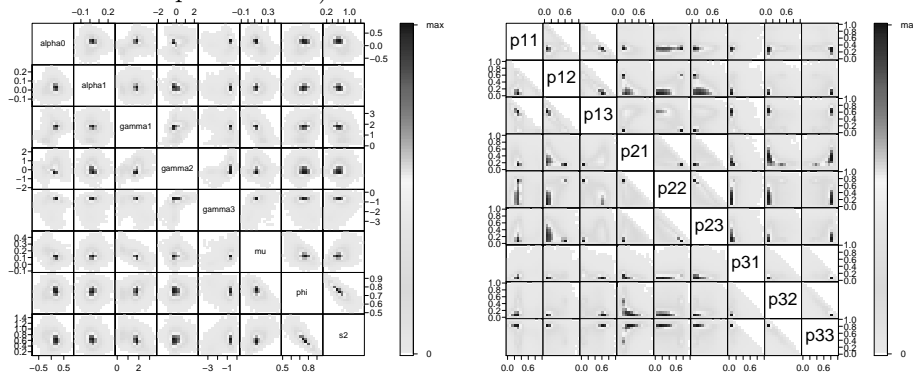
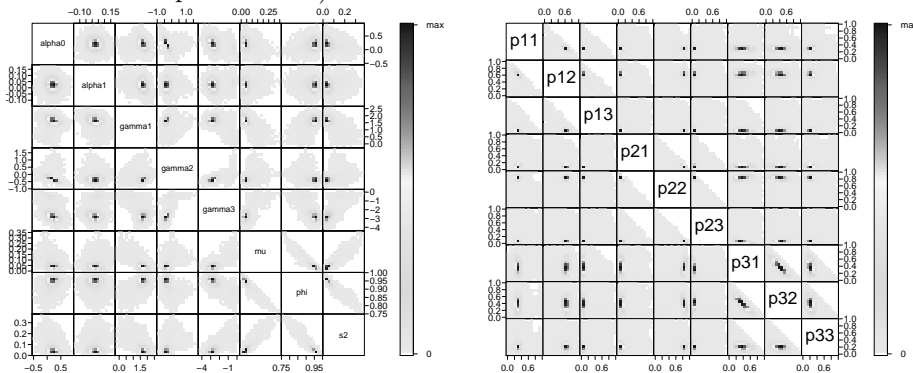


Figure 7: Bivariate marginal posterior densities of the SV-MS(3)-M model's parameters for Agora (top panel:  $\alpha_0, \alpha_1, \gamma_1, \gamma_2, \gamma_3, \mu, \varphi$  and  $\sigma^2$  coordinates; bottom panel: transition probabilities)



probabilities of each state at day  $t = 1, 2, \dots, T$ :

$$Pr(S_t = 1|y) \approx \frac{1}{N} \sum_{q=M+1}^{M+N} I(S_t^{(q)} = 1|y),$$

against the modelled log-returns on Agora's daily share prices,  $y_t$ . Apparently, the right-tail observations are assigned with high posterior probabilities of the first state, whereas the left-tail ones – with high probabilities of the third state (see the left- and the right-hand panels in Figure 11, respectively). For the 'middle zone' of the empirical data distribution it is the second state that appears most probable (see the

---

 Łukasz Kwiatkowski
 

---

mid-panel in Figure 11). In view of the above it is obvious now that the three-state SV-MS-M model exploits the underlying Markov switching in-mean mechanism to capture the outlying rates of return. Building upon the conclusion, as well as on the posterior averages of the transition probabilities (see Table 7) a general pattern in the switching process outlines as follows. Predominantly, the chain remains in the second state, with a probability of leaving it – most likely to the first state – totalling  $E(p_{21}|y) + E(p_{23}|y) = 0.115 + 0.077 \approx 0.19$  (see Table 7). A sharp rise in the equity price, resulting in a right-tail return, is captured as a shift to  $S_t = 1$ . Since  $E(p_{11}|y) = 0.30$ , there is (on average) a 30 percent chance of another consecutive strong price increase. On the other hand, a downright plunge to follow immediately an abnormal positive return – that is a switch directly from  $S_t = 1$  to  $S_t = 3$  – is hardly possible, as  $E(p_{13}|y) = 0.085$ . It appears that abrupt price soars are usually followed with 'typical' returns, for  $E(p_{12}|y) = 0.619$ . Still, if the price was to drop drastically (resulting in  $S_t = 3$ , regardless of the preceding state), then chances of the chain moving to any of the three states are quite similar ( $E(p_{31}|y) = 0.382$ ,  $E(p_{32}|y) = 0.390$ ,  $E(p_{33}|y) = 0.228$ ). Admittedly, the above results are rather non-standard, as they document some non-trivial pattern in dynamics of sharp price movements. It would merit a further research to compare the SV-MS(3)-M model with some common fat-tailed SV specifications (see e.g. Jacquier, Polson and Rossi 1999, 2004), which – by construction – neglect any dynamics in the occurrence of outliers. The issue goes beyond the current study, though, and shall be addressed elsewhere.

#### 4.3.2 Posterior analysis of the common parameters

In this final part of our study we gain an insight into the differences in posterior outcomes on the common parameters across the models. Since marginal posterior densities of these parameters (i.e.  $\alpha_0$ ,  $\alpha_1$ ,  $\mu$ ,  $\varphi$  and  $\sigma^2$ ) are of a very regular, unimodal shape (relevant figures – not presented in the paper – are available upon request), we restrict further considerations to basic posterior characteristics, including means and standard deviations (see Table 8).

With regard to parameters of the volatility equation (i.e.  $\mu$ ,  $\varphi$  and  $\sigma^2$ ; see Table 1) we notice that introducing a constant risk premium parameter to a simple SV specification does not affect the log-volatility dynamics. However, as long as allowing for the two-state switches in the risk-return relationship modifies the relevant posterior means only marginally, in the three-state SV-MS-M model they change quite remarkably. Firstly,  $E(\mu|y, M_3)$  drops by ca. 0.114 (with respect to SV and SV-M), amounting to around 70% of  $E(\mu|y, M_0)$  and  $E(\mu|y, M_1)$ . Such a change transfers to a decrease in the mean log-volatility level (in each model calculated as  $E(\ln h_t|\theta) = \frac{\mu}{1-\varphi}$ ,  $|\varphi| < 1$ ), from  $E[E(\ln h_t|\theta)|y, M_0] = E[E(\ln h_t|\theta)|y, M_1] = 1.409$  to  $E[E(\ln h_t|\theta)|y, M_3] = 0.881$  (see Table 9). One can also observe a sizeable rise in the volatility persistence –  $E(\varphi|y, M_3) = 0.945$  as opposed to  $E(\varphi|y, M_i) \approx 0.88$  (for  $i = 0, 1, 2$ ) – accompanied with a perceptibly lower variability of the log-volatility

---

 Bayesian Analysis of a Regime Switching In-Mean Effect ...
 

---

process (as implied by the results for  $\sigma^2$  and  $Var(\ln h_t|\theta) = \frac{\sigma^2}{1-\varphi^2}$ ,  $|\varphi| < 1$ ; see Tables 8 and 9). A smoother path of posterior averages of  $\ln h_t$ 's in the SV-MS(3)-M model (see Figure 8) arises from its capability of handling the outliers, with similar results having already been recognised in the fat-tailed SV literature (see e.g. Pajor 2003; Jacquier, Polson and Rossi 2004).

Table 8: Posterior means (and standard deviations) of the common parameters

Parameter	Model			
	SV	SV-M	SV-MS(2)-M	SV-MS(3)-M
$\alpha_0$	-0.028 (0.038)	0.013 (0.113)	0.074 (0.148)	0.215 (0.146)
$\alpha_1$	0.059 (0.022)	0.059 (0.022)	0.046 (0.03)	0.025 (0.034)
$\mu$	0.166 (0.035)	0.165 (0.035)	0.15 (0.036)	0.051 (0.031)
$\varphi$	0.882 (0.025)	0.883 (0.025)	0.886 (0.026)	0.945 (0.025)
$\sigma^2$	0.169 (0.041)	0.166 (0.041)	0.161 (0.042)	0.059 (0.037)

Table 9: Posterior means (and standard deviations) of the unconditional log-volatility expected value and variance

Parameter	Model			
	SV	SV-M	SV-MS(2)-M	SV-MS(3)-M
$E(\ln h_t \theta)$	1.409 (0.080)	1.409 (0.080)	1.312 (0.122)	0.881 (0.154)
$Var(\ln h_t \theta)$	0.757 (0.095)	0.755 (0.095)	0.745 (0.097)	0.530 (0.103)

Notice comparable posterior standard deviations of each volatility parameter across different model structures (see Table 8). However, this is no longer the case for the parameters common to the models' observation equations, i.e.  $\alpha_0$  and  $\alpha_1$ , with respect to which a few comments can be made. Firstly, introducing a constant in-mean term to the basic SV structure does not bear any influence on posterior moments of the autoregression coefficient,  $\alpha_1$ , since  $E(\alpha_1|y, M_1) = E(\alpha_1|y, M_0) = 0.059$  and  $D(\alpha_1|y, M_1) = E(\alpha_1|y, M_0) = 0.022$ ; see Table 8. Secondly, allowing for Markovian breaks in the risk-return relationship decreases posterior mean of  $\alpha_1$ , with the largest drop noted within the three-state SV-MS-M model. We conjecture that it is due to a hidden Markov process, with the second state featuring a relative persistence ( $E(p_{22}|y, M_3) = 0.807$ ), explaining a 'part' of the log-returns' autocorrelation.

Łukasz Kwiatkowski

Figure 8: Log-returns on Agora's share prices (top) along with posterior averages of  $\ln h_t$  (bottom)

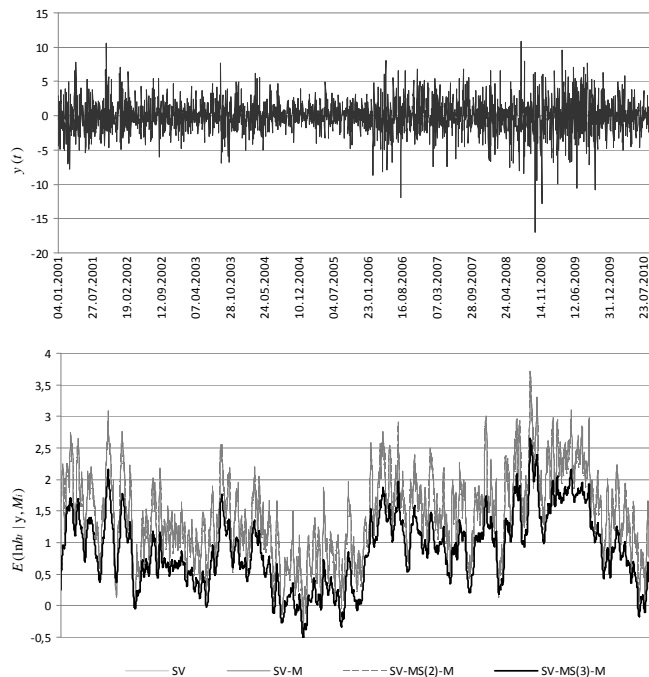
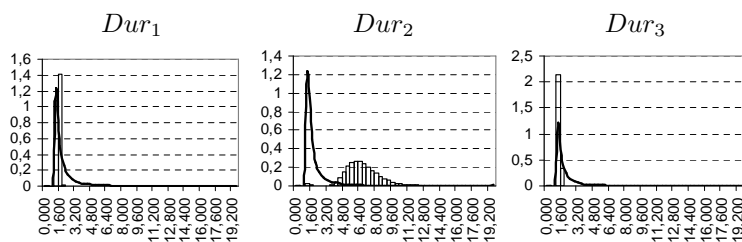


Figure 9: Marginal posteriors (bars) and priors (solid line) of expected durations within the SV-MS(3)-M model for Agora



Thirdly, as compared with the SV and SV-M specifications, results obtained within the switching models exhibit slightly more dispersion of the parameter in question, which may be attributable to a higher number of parameters involved in the observation equation. Finally, posterior characteristics of the intercept,  $\alpha_0$ , vary across the models quite markedly. Particularly, introducing a constant risk premium



Bayesian Analysis of a Regime Switching In-Mean Effect ...

Figure 10: Marginal posteriors (bars) and priors (solid line) of past-conditional skewness coefficients within the SV-MS(3)-M model for Agora

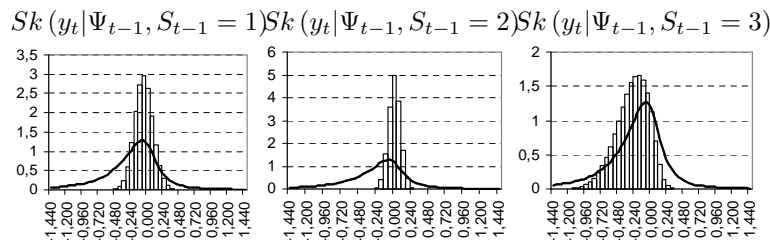
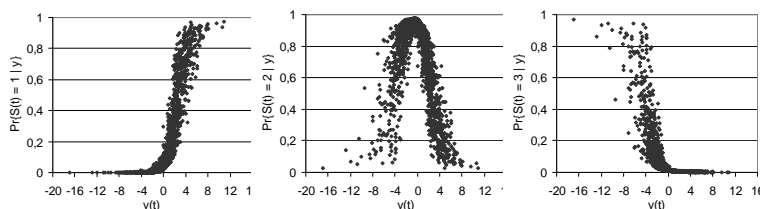


Figure 11: Scatter plots of probabilities  $Pr\{S_t = i|y, M_3\}$  ( $i = 1, 2, 3$ ) against the log-returns on Agora's share prices



parameter to the SV model considerably raises the parameter's posterior standard deviation ( $D(\alpha_0|y, M_1) = 0.113$  as compared with  $D(\alpha_0|y, M_0) = 0.038$ ), with even higher dispersion in the case of the switching models ( $D(\alpha_0|y, M_2) \approx D(\alpha_0|y, M_3) = 0.146$ ); see Table 8. A possible explanation for higher posterior uncertainty about the intercept – upon inclusion and further switching extension of the in-mean structure – is that all the parameters in the observation equation are involved in modelling a single quantity of unconditional mean of the observable process  $\{y_t, t \in \mathbb{Z}\}$ , i.e.  $E(y_t|\theta)$ , the formula for which can easily be derived for the SV-MS-M process given by Definition 1. Taking expectation on both sides of Equation (1),

$$E(y_t|\theta) = \alpha_0 + \alpha_1 E(y_{t-1}|\theta) + E(\gamma_{S_t} h_t^{0.5}|\theta),$$

and noting that

$$E(\gamma_{S_t} h_t^{0.5}|\theta) = E(\gamma_{S_t}|\theta) E(h_t^{0.5}|\theta) = \left( \sum_{j=1}^K \gamma_j \pi_j \right) \exp\left( \frac{\mu}{2(1-\varphi)} + \frac{\sigma^2}{8(1-\sigma^2)} \right),$$

the result – under second-order stationarity conditions (supposedly,  $|\alpha_1| < 1$  and

Łukasz Kwiatkowski

$|\varphi| < 1$ ; see Kwiatkowski, 2010) – is obtained immediately:

$$E(y_t|\theta) = \frac{\alpha_0 + \left(\sum_{j=1}^K \gamma_j \pi_j\right) \exp\left(\frac{\mu}{2(1-\varphi)} + \frac{\sigma^2}{8(1-\sigma^2)}\right)}{1 - \alpha_1}. \quad (20)$$

A corresponding formula for the SV-M model can be deduced from (20), assuming  $\gamma = \gamma_j$  for each  $j = 1, 2, \dots, K$ :

$$E(y_t|\theta) = \frac{\alpha_0 + \gamma \exp\left(\frac{\mu}{2(1-\varphi)} + \frac{\sigma^2}{8(1-\sigma^2)}\right)}{1 - \alpha_1}. \quad (21)$$

According to (21) one can expect a 'tradeoff' between  $\alpha_0$  and  $\gamma$ , or – to be more precise – a negative posterior correlation between the two parameters, which in fact is the case, as  $Corr(\alpha_0, \gamma|y, M_1) = -0.941$ . Posterior correlation coefficients between the observation equation parameters within the two- and the three-state SV-MS-M model usually reveal weaker linear interrelations among the parameters (see Tables 10 and 11).

Table 10: Posterior correlations between the observation equation parameters within the SV-MS(2)-M model for Agora

Parameter	$\alpha_0$	$\alpha_1$	$\gamma_1$	$\gamma_2$
$\alpha_0$	1	-0.354	0.441	-0.360
$\alpha_1$		1	-0.448	0.216
$\gamma_1$			1	-0.112
$\gamma_2$				1

Table 11: Posterior correlations between the observation equation parameters within the SV-MS(3)-M model for Agora

Parameter	$\alpha_0$	$\alpha_1$	$\gamma_1$	$\gamma_2$	$\gamma_3$
$\alpha_0$	1	-0.047	-0.209	-0.804	-0.327
$\alpha_1$		1	-0.221	-0.018	-0.097
$\gamma_1$			1	0.191	-0.266
$\gamma_2$				1	0.563
$\gamma_3$					1

## 5 Concluding remarks

In the paper we extend the methodology of Bayesian stochastic volatility models with Markovian breaks in the in-mean structure developed by Kwiatkowski (2010), by

allowing for three rather than only two states of the underlying chain. The MCMC algorithm is suitably modified and presented for a general  $K$ -state SV-MS-M model. Also, the past-conditional moments' structure of the SV-MS-M process is discussed in some detail, particularly with respect to the model's capability of capturing data skewness.

The methodology is employed to investigate a possibly switching pattern in the risk-return relationship among individual companies' share prices, representing major Polish financial market sectors. Perhaps to one's dismay, the overall results do not provide compelling evidence of either a constant or regime switching market risk premium. Despite superior data fit of the SV-MS-M models (reported for most of the datasets), it seems that introducing Markovian switches into the in-mean parameter does not resolve the issue of its apparent insignificance indicated by the results obtained within simple SV-M models. Moreover, inference on the Markov switching risk premium – via basic posterior characteristics – is handicapped by multimodality and a relatively large dispersion of posterior marginals. Results for Agora, representing a single exception in this regard, are discussed at length. It is found that the best in-sample performance of the three-state SV-MS-M model may stem from its empirical capability of capturing outlying observations and some non-trivial dynamics in their occurrence. Therefore, it is deemed that future research should address a comparison of the SV-MS-M models against common fat-tailed SV structures, with respect to the in-sample fit, as well as in more practical terms of, for instance, market risk analysis - hinged upon modelling the outliers.

## Acknowledgements

Research realised within a project financed by the National Science Centre under decision no. DEC-2011/01/N/HS4/03105.

## References

- [1] Abanto-Valle C.A., Migon H.S., Lachos V.H., (2011), Stochastic volatility in mean models with scale mixtures of normal distributions and correlated errors: A Bayesian approach, *Journal of Statistical Planning and Inference* 141, 1875-1887.
- [2] Backus D.K., Gregory A.W., (1993), Theoretical Relations Between Risk Premiums and Conditional Variances, *Journal of Business & Economic Statistics* 11(2), 177-185.
- [3] Bauwens L., Lubrano M., (1998), Bayesian inference on GARCH models using the Gibbs sampler, *Econometrics Journal* 1, C23-C46.

Łukasz Kwiatkowski

---

- [4] Bollerslev T., Engle R.F., Wooldridge J.M., (1988), A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy* 96(1).
- [5] Bollerslev T., Zhou H., (2006), Volatility puzzles: a simple framework for gauging return-volatility regressions, *Journal of Econometric* 131, 123-150.
- [6] Campbell J.Y., (1987), Stock returns and the term structure, *Journal of Financial Economics* 18(2).
- [7] Carter C.K., Kohn R., (1994), On Gibbs sampling for state space models, *Biometrika* 81, 541-553.
- [8] Chib S., (1996), Calculating posterior distributions and modal estimates in Markov mixture models, *Journal of Econometrics* 75, 79-97.
- [9] Chou R., Engle R.F., Kane A., (1992), Measuring Risk Aversion from Excess Returns on a Stock Index, *Journal of Econometrics* 52, 201-224.
- [10] Clark P.K., (1973), A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica* 41.
- [11] Darrat A.F., Gilley O.W., Li B., Wu Y., (2011), Revisiting the risk/return relations in the Asian Pacific Markets: New evidence from alternative models, *Journal of Business Research* 64, 199-206.
- [12] David A., Veronesi P., (2009), What Ties Return Volatilities to Price Valuations and Fundamentals?, Chicago Booth Research Working Paper No. 10-05, available at: <http://ssrn.com/abstract=1513359>
- [13] Doman M., (2008), Information Impact on Stock Price Dynamics [in:] *Dynamic Econometric Models* 8, [ed.:] Z. Zieliński, Toruń, 13-20.
- [14] Engle R.F., Lilien D.M., Robins R.P., (1987), Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica* 55, 525-542.
- [15] Fiszeder P., Kwiatkowski J., (2005a), Model GARCH-M ze zmiennym parametrem - analiza wybranych spółek i indeksów notowanych na GPW w Warszawie (GARCH-M Model with Time-Varying Parameter – Analysis of Selected Stock and Indices Quoted on the WSE, in Polish), *Statistical Review (Przegląd Statystyczny)* 52 (3), 73-88.
- [16] Fiszeder P., Kwiatkowski J., (2005b), Dynamic Analysis of Relation between Expected Return and Conditional Variance, *Acta Universitatis Nicolai Copernici* 372, 85-98.
- [17] French K.R., Schwert G.W., Stambaugh R.F., (1987), Expected stock returns and volatility, *Journal of Financial Economics* 19(1).

- [18] Gennotte G., Marsh T.A., (1993), Variations in economic uncertainty and risk premiums on capital assets, *European Economic Review* 37(5).
- [19] Glosten L.R., Jagannathan R., Runkle D.E., (1993), On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance* 48, 1791-1801.
- [20] Guo H., Neely C.J., (2008), Investigating the intertemporal risk-return relation in international stock markets with component GARCH model, *Economics Letters* 99.
- [21] Guo H., Whitelaw R., (2006), Uncovering the risk return relation in the stock market, *Journal of Finance* 61(3).
- [22] Hamilton J. D., (1989), A New approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357-384.
- [23] Harrison P., Zhang H., (1999), An investigation of the risk and return relation at long horizons, *Review of Economics and Statistics* 81(3).
- [24] Harvey C.R., (1989), Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24(2).
- [25] Hibbert A.M., Daigler R.T., Dupoyet B., (2008), A behavioral explanation for the negative asymmetric return volatility relation, *Journal of Banking and Finance* vol. 32.
- [26] Jacquier E., Polson N., Rossi P., (1994), Bayesian analysis of stochastic volatility models (with discussion), *Journal of Business & Economic Statistics* 12, 371-417.
- [27] Jacquier E., Polson N., Rossi P., (1999), Stochastic Volatility: Univariate and Multivariate Extensions, Cahiers CIRANO, Centre Interuniversitaire de Recherche en Analyse des Organisations, Montréal.
- [28] Jacquier E., Polson N., Rossi P., (2004), Bayesian Analysis of Stochastic Volatility Models with Fat-tails and Correlated Errors, *Journal of Econometrics* 122, 185-212.
- [29] Kalimipalli M., Susmel R., (2001), Regime-switching stochastic volatility and short-term interest rates, *CEMA Working Papers*, available at: <http://ideas.repec.org/p/cem/doctra/197.html>
- [30] Kim S.-W., Lee B.-S., (2008), Stock returns, asymmetric volatility, risk aversion, and business cycle: some new evidence, *Economic Inquiry* 46(2), 131-148.
- [31] Koopman S.J., Hol Uspensky E., (2002), The Stochastic Volatility In Mean Model: Empirical Evidence from International Stock Markets, *Journal of Applied Econometrics* 17, 667-689.

Łukasz Kwiatkowski

---

- [32] Krolzig H.-M., (1997), Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis, Lecture Notes in Economics and Mathematical Systems, New York/Berlin/Heidelberg: Springer.
- [33] Kwiatkowski Ł., (2010), Markov Switching In-Mean Effect. Bayesian Analysis in Stochastic Volatility Framework, *Central European Journal of Economic Modelling and Econometrics* 2(1), 59-94.
- [34] Li Q., Yang J., Hsiao C., Chang Y.-J., (2005), The relationship between stock returns and volatility in international stock markets, *Journal of Empirical Finance* 12.
- [35] Loudon G.F., (2006), Is the risk-return relation positive? Further evidence from a stochastic volatility in mean approach, *Applied Financial Economics* 16, 981-992.
- [36] Ludvigson S., Ng S., (2007), The empirical risk-return relation: A factor analysis approach, *Journal of Financial Economics* 83(1).
- [37] Nelson D.B., (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.
- [38] Newton M.A., Raftery A.E., (1994), Approximate Bayesian inference by Weighted Likelihood Bootstrap (with discussion), *Journal of the Royal Statistical Society*, series B 56, 3-48.
- [39] Osiewalski J., Pipień M., (2000), GARCH-In-Mean through skewed t conditional distributions: Bayesian inference for exchange rates, [in:] *26-th International Conference MACROMODELS'99*, [ed:] W. Welfe, P. Wdowiński, Łódź, 354-369.
- [40] Pajor A., (2003), Procesy zmienności stochastycznej SV w bayesowskiej analizie finansowych szeregów czasowych (Stochastic Volatility Processes in Bayesian Analysis of Financial Time Series, in Polish), Cracow University of Economics.
- [41] Pipień M., (2007), An approach to measuring the relation between risk and return. Bayesian analysis for WIG data, *Folia Oeconomica Cracoviensia* 48, 95-117
- [42] Pipień M., Osiewalski J., (2001), Model GARCH-M(1,1): specyfikacja i estymacja bayesowska (A GARCH-M(1,1) Model: Specification and Bayesian Estimation, in Polish), *Prace Naukowe Akademii Ekonomicznej we Wrocławiu* (no. 895), 188-197.
- [43] Scruggs J.T., (1998), Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach, *Journal of Finance* 53, 575-603.

---

Bayesian Analysis of a Regime Switching In-Mean Effect ...

---

- [44] Shibata M., Watanabe T., (2005), Bayesian analysis of a Markov switching stochastic volatility model, *Journal of Japan Statistical Society* 35, 205-219.
- [45] Smith D. R., (2007), Risk and Return in Stochastic Volatility Models: Volatility Feedback Matters!, Working Paper, available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=889263](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=889263).
- [46] So M. K. P., Lam K., Li W. K., (1998), A stochastic volatility model with Markov switching, *Journal of Business & Economic Statistics* 16, 244-253.
- [47] Taylor S.J., (1982), *Financial returns modelled by the product of two stochastic processes – a study of daily sugar prices 1961-79* [in:] *Time Series Analysis: Theory and Practice 1*, [ed.:] O.D. Anderson, Amsterdam, North-Holland, 203-226.
- [48] Taylor S.J., (1986), *Modelling Financial Time Series*, Wiley, Chichester.
- [49] Taylor S.J., (2005), *Asset Price Dynamics, Volatility, and Prediction*, Princeton University Press.
- [50] Veronesi P., (2000), How Does Information Quality Affect Stock Returns?, *The Journal of Finance* 55(2).
- [51] Watanabe T., (1999), A non-linear filtering approach to stochastic volatility models with an application to daily stock returns, *Journal of Applied Econometrics* 14.
- [52] Whitelaw R.F., (1994), Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns, *Journal of Finance* 49(2).
- [53] Whitelaw R.F., (2000), Stock Market Risk and Return: An Equilibrium Approach, *The Review of Financial Studies* 13(3).

Łukasz Kwiatkowski

## Appendix

In what follows derivation of the past-conditional moments of the SV-MS-M process (see Theorem 2) is briefly outlined.

Let

$$\bar{\gamma}_{t|t-1}^{(r)} \equiv E(\gamma_{S_t}^r | \Psi_{t-1}, \theta) = E(\gamma_{S_t}^r | S_{t-1}, \theta) = \sum_{j=1}^K \gamma_{S_t=j}^r p_{S_{t-1}, S_t=j} \quad (22)$$

be the past-conditional moment of order  $r \in \mathbb{N}$  for random variable  $\gamma_{S_t}$ . For further considerations note that – due to independence between the Markov chain  $\{S_t\}$  and conditional variance process  $\{h_t\}$  – the following identity holds for each  $k, l \in \mathbb{R}$ :

$$E(\gamma_{S_t}^k h_t^l | \Psi_{t-1}, \theta) = E(\gamma_{S_t}^k | \Psi_{t-1}, \theta) E(h_t^l | \Psi_{t-1}, \theta) = \bar{\gamma}_{t|t-1}^{(r)} E(h_t^l | \Psi_{t-1}, \theta). \quad (23)$$

Conditionally upon  $\Psi_{t-1}$ , the logarithm of conditional variance  $h_t$  is normally distributed with mean  $E(\ln h_t | \Psi_{t-1}, \theta) = \mu + \varphi \ln h_{t-1}$  and variance  $Var(\ln h_t | \Psi_{t-1}, \theta) = \sigma^2$ . Hence, the past-conditional distribution of conditional variance  $h_t$  itself is Log-Normal, with

$$E(h_t^l | \Psi_{t-1}, \theta) = \exp\{(\mu + \varphi \ln h_{t-1})l + 0.5\sigma^2 l^2\}. \quad (24)$$

Past-conditional mean of the SV-MS-M process is obtained immediately:

$$\begin{aligned} E(y_t | \Psi_{t-1}, \theta) &= E(\alpha_0 + \alpha_1 y_{t-1} + \gamma_{S_t} h_t^{0.5} + \varepsilon_t h_t^{0.5} | \Psi_{t-1}, \theta) \\ &= \alpha_0 + \alpha_1 y_{t-1} + \bar{\gamma}_{t|t-1}^{(1)} \exp\left\{\frac{1}{2}(\mu + \varphi \ln h_{t-1}) + \frac{1}{8}\sigma^2\right\}. \end{aligned}$$

To derive the second-order central moment note that:

$$\begin{aligned} Var(y_t | \Psi_{t-1}, \theta) &= Var((\gamma_{S_t} + \varepsilon_t) h_t^{0.5} | \Psi_{t-1}, \theta) \\ &= E((\gamma_{S_t} + \varepsilon_t)^2 | \Psi_{t-1}, \theta) E(h_t | \Psi_{t-1}, \theta) + \\ &\quad - E^2((\gamma_{S_t} + \varepsilon_t) | \Psi_{t-1}, \theta) E^2(h_t^{0.5} | \Psi_{t-1}, \theta) \\ &= (\bar{\gamma}_{t|t-1}^{(2)} + 1) \exp(\mu + \varphi \ln h_{t-1} + \frac{1}{2}\sigma^2) + \\ &\quad - (\bar{\gamma}_{t|t-1}^{(1)})^2 \exp(\mu + \varphi \ln h_{t-1} + \frac{1}{4}\sigma^2) \\ &= \left[ (\bar{\gamma}_{t|t-1}^{(2)} + 1) \exp(\frac{1}{4}\sigma^2) - (\bar{\gamma}_{t|t-1}^{(1)})^2 \right] \exp(\mu + \varphi \ln h_{t-1} + \frac{1}{4}\sigma^2), \end{aligned} \quad (25)$$

which coincides with the result presented in Theorem 2.

Calculation of the conditional skewness coefficient:

$$Sk(y_t | \Psi_{t-1}, \theta) \equiv \frac{E[(y_t - E(y_t | \Psi_{t-1}, \theta))^3 | \Psi_{t-1}, \theta]}{(Var(y_t | \Psi_{t-1}, \theta))^{\frac{3}{2}}},$$



---

 Bayesian Analysis of a Regime Switching In-Mean Effect ...
 

---

requires derivation of a third-order central moment, related with ordinary moments via the identity:

$$\begin{aligned}
 & E \left[ (y_t - E(y_t | \Psi_{t-1}, \theta))^3 | \Psi_{t-1}, \theta \right] \\
 &= E(y_t^3 | \Psi_{t-1}, \theta) - 3E(y_t | \Psi_{t-1}, \theta) E(y_t^2 | \Psi_{t-1}, \theta) + 2E^3(y_t | \Psi_{t-1}, \theta).
 \end{aligned} \tag{26}$$

Some tedious algebra leads to the formula:

$$\begin{aligned}
 & E \left[ (y_t - E(y_t | \Psi_{t-1}, \theta))^3 | \Psi_{t-1}, \theta \right] \\
 &= \left[ \left( \bar{\gamma}_{t|t-1}^{(3)} + 3\bar{\gamma}_{t|t-1}^{(1)} \right) \exp\left(\frac{3}{4}\sigma^2\right) - 3\bar{\gamma}_{t|t-1}^{(1)} \left( \bar{\gamma}_{t|t-1}^{(2)} + 1 \right) \exp\left(\frac{1}{4}\sigma^2\right) + 2 \left( \bar{\gamma}_{t|t-1}^{(1)} \right)^3 \right] \\
 &\times \exp\left(\frac{3(\mu + \varphi \ln h_{t-1})}{2} + \frac{3}{8}\sigma^2\right).
 \end{aligned} \tag{27}$$

Dividing (27) by

$$\begin{aligned}
 & (Var(y_t | \Psi_{t-1}, \theta))^{\frac{3}{2}} = \\
 & \left[ \left( \bar{\gamma}_{t|t-1}^{(2)} + 1 \right) \exp\left(\frac{1}{4}\sigma^2\right) - \left( \bar{\gamma}_{t|t-1}^{(1)} \right)^2 \right]^{\frac{3}{2}} \exp\left(\frac{3(\mu + \varphi \ln h_{t-1})}{2} + \frac{3}{8}\sigma^2\right)
 \end{aligned}$$

reduces the last exponent term in both expressions and yields the formula for conditional skewness coefficient presented in Theorem 2.