Econometric Theory, **23**, 2007, 485–500. Printed in the United States of America.
DOI: 10.1017/S0266466607070211

ON THE STATIONARITY OF MARKOV-SWITCHING GARCH PROCESSES

Ari Abramson and Israel Cohen
Technion—Israel Institute of Technology

Generalized autoregressive conditional heteroskedasticity (GARCH) models with Markov-switching regimes are often used for volatility analysis of financial time series. Such models imply less persistence in the conditional variance than the standard GARCH model and potentially provide a significant improvement in volatility forecast. Nevertheless, conditions for asymptotic wide-sense stationarity have been derived only for some degenerated models. In this paper, we introduce a comprehensive approach for stationarity analysis of Markov-switching GARCH models, which manipulates a backward recursion of the model's second-order moment. A recursive formulation of the state-dependent conditional variances is developed, and the corresponding conditions for stationarity are obtained. In particular, we derive necessary and sufficient conditions for the asymptotic wide-sense stationarity of two different variants of Markov-switching GARCH processes and obtain expressions for their asymptotic variances in the general case of m-state Markov chains and (p,q)-order GARCH processes.

1. INTRODUCTION

Volatility analysis of financial time series is of major importance in many financial applications. The generalized autoregressive conditional heteroskedasticity (GARCH) model, first introduced by Bollerslev (1986), has been applied quite extensively in the field of econometrics, both by practitioners and by researchers. It has been shown useful for the analysis of the volatility of time-varying processes such as those pertaining to financial markets. Incorporating GARCH models with a hidden Markov chain, where each state of the chain (regime) allows a different GARCH behavior and thus a different volatility structure, extends the dynamic formulation of the model and potentially enables improved forecasts of the volatility; e.g., see Gray (1996), Klaassen (2002), Haas, Mittnik, and Paolella (2004b), Marcucci (2005), Dueker (1997), and Frömmel (2004).

The authors thank Professor Rami Atar for helpful discussions. The authors thank the co-editor Bruce Hansen and the three anonymous referees for their helpful comments and suggestions and in particular the referee who proposed a generalization of the proof in Appendix B. This research was supported by the Israel Science Foundation (grant 1085/05). Address correspondence to Israel Cohen, Department of Electrical Engineering, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel; e-mail: icohen@ee.technion.ac.il.

 0266-4666/07 $15.00

Unfortunately, the volatility of a GARCH process with switching regimes depends on the entire history of the process, including the regime path, which makes the derivation of a volatility estimator impractical.

Cai (1994) and Hamilton and Susmel (1994) applied the idea of regime-switching parameters to autoregressive conditional heteroskedasticity (ARCH) specification. The conditional variance of an ARCH model depends only on past observations, and accordingly the restriction to ARCH models avoids problems of infinite path dependency. Gray (1996), Klaassen (2002), and Haas et al. (2004b) proposed different variants of Markov-switching GARCH models, which also avoid the problem of dependency on the regime's path. Gray introduced a Markov-switching GARCH model relying on the assumption that the conditional variance at any regime depends on the *expectation* of previous conditional variances, rather than their values. Accordingly, the conditional variance depends only on some finite set of past state-dependent expected values via their conditional state probabilities and thus can be constructed from past observations. Klaassen proposed modifying Gray's model by conditioning the expectation of previous conditional variances on all available observations and also on the current regime. A different concept of Markov-switching GARCH model has recently been introduced by Haas et al. (2004b). Accordingly, a finite state-space Markov chain is assumed to govern the ARCH parameters, whereas the autoregressive behavior of the conditional variance is subject to the assumption that past conditional variances are in the same regime as that of the current one.

Markov-switching GARCH processes, and also the standard GARCH process, are nonstationary as their second-order moments change recursively over time. However, if these processes are asymptotically wide-sense stationary then their variances are guaranteed to be finite. A necessary and sufficient condition for the stationarity of a (single-regime) GARCH(p,q) process has been developed by Bollerslev (1986). A condition for the stationarity of a *natural* path-dependent Markov-switching GARCH(p,q) model, has been developed by Francq, Roussignol, and Zakoïan (2001), and a thorough analysis of the probabilistic structure of that model, with conditions for the existence of moments of any order, was presented by Francq and Zakoïan (2005). Wong and Li (2001), Alexander and Lazar (2004), and Haas, Mittnik, and Paolella (2004a) derived stationarity analysis for some mixing models of conditional heteroskedasticity. Yang (2000), Francq and Zakoïan (2002), Francq and Zakoïan (2001), Yao (2001), and Timmermann (2000) derived conditions for the asymptotic stationarity of some autoregressive (AR) and autoregressive moving average (ARMA) models with Markov-regimes. However, for the Markov-switching GARCH models described previously, which avoid the dependency of the conditional variance on the chain's history, stationarity conditions are known in the literature only for some special cases. Klaassen (2002) developed necessary (but not necessarily sufficient) conditions for stationarity of his model in the special case of two regimes and GARCH modeling of order (1,1). A necessary and sufficient stationarity condition has been developed by Haas et al. (2004b) for their Markov-switching GARCH model, but only in the case of GARCH(1,1) behavior in each regime.

In this paper, we develop a comprehensive approach for stationarity analysis of Markov-switching GARCH models in the general case of m-state Markov chains and (p,q)-order GARCH processes. We specify the unconditional variance of the process using the expectation of the regime-dependent conditional variances, and we assume no historical knowledge of the process except for the model parameters. The expectation of the conditional variance at a given regime is then recursively constructed from the conditional expectation of both previous conditional and unconditional variances. Consequently, we obtain a complete recursion for the expected vector of state-dependent conditional variances. The recursive vector form is constructed by means of a representative matrix that is built from the model parameters. We show that constraining the largest absolute eigenvalue of the representative matrix to be less than one is necessary and sufficient for the convergence of the unconditional variance and, therefore, for the asymptotic stationarity of the process. We derive stationarity conditions for the general formulation of the two variants of Markov-switching GARCH models. We show that our results reduce in some degenerated cases to the stationarity conditions developed by Bollerslev (1986), Klaassen (2002), and Haas et al. (2004b). Furthermore, we show that the stationarity conditions developed by Klaassen are not only necessary but also sufficient for asymptotic stationarity of his model.

This paper is organized as follows. In Section 2, we review the variants proposed by Klaassen (2002) and Haas et al. (2004b) for Markov-switching GARCH models and develop comprehensive necessary and sufficient conditions for asymptotic stationarity appropriate for the general formulation of the models. In Section 3, we derive relations between our results and previous works. Section 4 concludes.

2. STATIONARITY OF MARKOV-SWITCHING GARCH MODELS

Let $S_t \in \{1,\ldots,m\}$ denote the (unobserved) regime at a discrete time t and let s_t be a realization of S_t, assuming that $\{S_t\}$ is a first-order stationary Markov chain with transition probabilities $a_{ij} \triangleq p(S_t = j | S_{t-1} = i)$, a transition probabilities matrix A, $\{A\}_{ij} = a_{ij}$, and stationary probabilities $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_m]'$, $\pi_i \triangleq p(S_t = i)$, where $'$ denotes the transpose operation. Let $\mathcal{I}_t$ denote the observation set up to time t and let $\{v_t\}$ be a zero-mean unit-variance random process with independent and identically distributed elements. Given that $S_t = s_t$, a Markov-switching GARCH model of order (p,q) can be formulated as

$$\varepsilon_t = \sigma_{t,s_t} v_t, \tag{1}$$

where the conditional variance of the process $\sigma^2_{t,s_t} = E\{\varepsilon_t^2 | S_t = s_t, \mathcal{I}_{t-1}\}$ is a function of p previous conditional variances and q previous squared observations.

Klaassen (2002) and Haas et al. (2004b) proposed different variants of Markov-switching GARCH models. The former is a modification of the model proposed by Gray (1996). Each of these overcomes the problem of dependency on the regime's path encountered when naturally integrating the GARCH model with switching regimes. However, conditions for these models to be asymptotically wide-sense stationary, and therefore to guarantee finite second-order moments, are known only for some special cases. Klaassen developed necessary conditions for the stationarity of his model in the case of two-state Markov chain and GARCH of order (1,1). Haas et al. (2004b) gave a necessary and sufficient stationarity condition for their model, but this condition is restricted to a first-order GARCH model in each of the regimes (i.e., $p = q = 1$). We first review these variants of Markov-switching GARCH models, which we call MSG-I and MSG-II, respectively. Then we develop necessary and sufficient conditions for their asymptotic wide-sense stationarity and derive their stationary variances.

2.1. MSG-I Model

Gray (1996) proposed modeling the conditional variance of a Markov-switching GARCH model as dependent on the *expectation* of its past values over the entire set of states, rather than dependent on past states and the corresponding conditional variances. Accordingly, the state-dependent conditional variance follows:

$$\begin{aligned}\sigma_{t,s_t}^2 &= \xi_{s_t} + \sum_{i=1}^{q} \alpha_{i,s_t} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{j,s_t} E(\varepsilon_{t-j}^2 | \mathcal{I}_{t-j-1}) \\ &= \xi_{s_t} + \sum_{i=1}^{q} \alpha_{i,s_t} \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_{j,s_t} \sum_{s_{t-j}=1}^{m} p(S_{t-j} = s_{t-j} | \mathcal{I}_{t-j-1}) \sigma_{t-j,s_{t-j}}^2, \end{aligned} \tag{2}$$

and the following constraints:

$$\begin{aligned}&\xi_{s_t} > 0, \qquad \alpha_{i,s_t} \geq 0, \qquad \beta_{j,s_t} \geq 0, \\ &\quad i = 1, \ldots, q, \qquad j = 1, \ldots, p, \qquad s_t = 1, \ldots, m\end{aligned} \tag{3}$$

are sufficient for the positivity of the conditional variance.

Gray's model integrates out the unobserved regime path so that the conditional variance can be constructed from previous observations only. As a consequence, there is no path dependency problem although GARCH effects are still allowed. Empirical analysis of modeling financial time series demonstrates that this Markov-switching GARCH model implies less persistence in the conditional variance than the standard GARCH model, and in addition, its one-step-ahead volatility forecast significantly outperforms the single-regime GARCH model (see, e.g., Gray, 1996; Marcucci, 2005; Frömmel, 2004).

Klaassen (2002) proposed modifying Gray's model by replacing $p(S_{t-j} = s_{t-j} | \mathcal{I}_{t-j-1})$ in (2) by $p(S_{t-j} = s_{t-j} | \mathcal{I}_{t-1}, S_t = s_t)$ while evaluating σ_{t,s_t}^2. Consequently, all available observations are used, in addition to the given regime in which the conditional variance is calculated. The conditional variance according to Klaassen's model (denoted here as MSG-I) is given by

$$\sigma^2_{t,s_t} = \xi_{s_t} + \sum_{i=1}^{q} \alpha_{i,s_t} \varepsilon^2_{t-i} + \sum_{j=1}^{p} \beta_{j,s_t} \sum_{s_{t-j}=1}^{m} p(S_{t-j} = s_{t-j} | \mathcal{I}_{t-1}, S_t = s_t) \sigma^2_{t-j,s_{t-j}}, \tag{4}$$

and the same constraints in (3) are sufficient for the positivity of the conditional variance.

Both models integrate out the unobserved regimes for evaluating the conditional variance. However, Klaassen's model employs all the available information, whereas Gray's model employs only part of it because it does not utilize all the available observations and the assumed regime in which the conditional variance is being calculated. Specifically, if process regimes are highly persistent, then both the current state s_t and the previous innovation ε_{t-1} give much information about previous states, and thus the conditional probability of s_{t-1} given all the observations up to time $t-1$ and the next state is substantially different from the probability of s_{t-1}, which is conditioned only on observations up to time $t-2$. In contrast to Gray, Klaassen does manipulate this information in his model while evaluating the expectation of previous conditional variances. Furthermore, the formulation (4) better exploits the available information, and its structure yields straightforward expressions for the multi-step-ahead volatility forecasts; see Klaassen (2002) and Marcucci (2005).

The unconditional variance of the MSG-I process, defined in (1) and (4), can be calculated as follows:

$$E[\varepsilon_t^2] = E_{\mathcal{I}_{t-1},S_t}[E(\varepsilon_t^2 | \mathcal{I}_{t-1}, s_t)] = E_{S_t}[E_{\mathcal{I}_{t-1}}(\sigma^2_{t,s_t} | s_t)] = \sum_{s_t=1}^{m} \pi_{s_t} E_{\mathcal{I}_{t-1}}(\sigma^2_{t,s_t} | s_t). \tag{5}$$

For notation simplification, we shall use $E(\cdot | s_t)$ and $p(\cdot | s_t)$ to represent $E(\cdot | S_t = s_t)$ and $p(\cdot | S_t = s_t)$, respectively, where s_t represents the regime realization at time t. Furthermore, we shall use $E_t(\cdot)$ to denote the expectation over the information up to time t, i.e., $E_{\mathcal{I}_t}(\cdot)$. The expectation of the regime-dependent conditional variance follows:

$$E_{t-1}[\sigma^2_{t,s_t} | s_t] = \xi_{s_t} + \sum_{i=1}^{q} \alpha_{i,s_t} E_{t-1}[\varepsilon^2_{t-i} | s_t] + \sum_{j=1}^{p} \beta_{j,s_t} \sum_{s_{t-j}=1}^{m} E_{t-1}[p(s_{t-j} | \mathcal{I}_{t-1}, s_t) \sigma^2_{t-j,s_{t-j}} | s_t], \tag{6}$$

where the expectation over ε_{t-i}^2 can be obtained by

$$E_{t-1}[\varepsilon_{t-i}^2|s_t] = \sum_{s_{t-i}=1}^{m} \int_{\mathcal{I}_{t-1}} \varepsilon_{t-i}^2 p(\mathcal{I}_{t-1}|s_t, s_{t-i}) p(s_{t-i}|s_t)\, d\mathcal{I}_{t-1}$$

$$= \sum_{s_{t-i}=1}^{m} p(s_{t-i}|s_t) E_{t-1}[\varepsilon_{t-i}^2|s_{t-i}, s_t]. \quad \textbf{(7)}$$

Note that given the current active state, the expected absolute value is independent of any future states. Therefore,

$$\begin{aligned} &E_{t-1}[\varepsilon_{t-i}^2|s_{t-i}, s_t] \\ &\quad = E_{t-1}[\varepsilon_{t-i}^2|s_{t-i}] \\ &\quad = \int_{\mathcal{I}_{t-i-1}} \int_{\varepsilon_{t-i}} \varepsilon_{t-i}^2 p(\varepsilon_{t-i}|\mathcal{I}_{t-i-1}, s_{t-i}) p(\mathcal{I}_{t-i-1}|s_{t-i})\, d\varepsilon_{t-i}\, d\mathcal{I}_{t-i-1} \\ &\quad = E_{t-i-1}[E(\varepsilon_{t-i}^2|\mathcal{I}_{t-i-1}, s_{t-i})|s_{t-i}] \\ &\quad = E_{t-i-1}[\sigma_{t-i,s_{t-i}}^2|s_{t-i}]. \end{aligned} \quad \textbf{(8)}$$

Furthermore, the conditional expectation over the conditional variance in (6), weighted by the current state probability, can be obtained by

$$\begin{aligned} &E_{t-1}[p(s_{t-j}|\mathcal{I}_{t-1}, s_t)\sigma_{t-j,s_{t-j}}^2|s_t] \\ &\quad = \int_{\mathcal{I}_{t-1}} \sigma_{t-j,s_{t-j}}^2 p(s_{t-j}|\mathcal{I}_{t-1}, s_t) p(\mathcal{I}_{t-1}|s_t)\, d\mathcal{I}_{t-1} \\ &\quad = \int_{\mathcal{I}_{t-1}} \sigma_{t-j,s_{t-j}}^2 p(\mathcal{I}_{t-1}|s_{t-j}, s_t) p(s_{t-j}|s_t)\, d\mathcal{I}_{t-1} \\ &\quad = p(s_{t-j}|s_t) E_{t-j-1}[\sigma_{t-j,s_{t-j}}^2|s_{t-j}]. \end{aligned} \quad \textbf{(9)}$$

Consequently, the expectation of the conditional variance at a given regime s_t can be recursively constructed, according to the model definitions, from both expectation of previous conditional variances and expected squared values given the current regime s_t. Let $r = \max\{p, q\}$ and define $\alpha_{i,s} \triangleq 0$ for all $i > q$ and $\beta_{i,s} \triangleq 0$ for all $i > p$. Then, by substituting (7)–(9) into (6) we obtain

$$E_{t-1}[\sigma_{t,s_t}^2|s_t] = \xi_{s_t} + \sum_{i=1}^{r} (\alpha_{i,s_t} + \beta_{i,s_t}) \sum_{s_{t-i}=1}^{m} p(s_{t-i}|s_t) E_{t-i-1}[\sigma_{t-i,s_{t-i}}^2|s_{t-i}], \quad \textbf{(10)}$$

and applying Bayes' rule we have

$$p(s_{t-i}|s_t) = \frac{\pi_{s_{t-i}}}{\pi_{s_t}} p(s_t|s_{t-i}) = \frac{\pi_{s_{t-i}}}{\pi_{s_t}} \{A^i\}_{s_{t-i},s_t}. \tag{11}$$

The expected state-dependent conditional variance (10) is recursively generated from a weighted sum of its previous expected values through their conditioned probabilities and the model parameters. Let $\boldsymbol{\xi} \triangleq [\xi_1,\ldots,\xi_m]'$, let $\mathcal{K}^{(i)}$ be an m-by-m matrix with elements

$$\{\mathcal{K}^{(i)}\}_{s,\tilde{s}} \triangleq (\alpha_{i,s} + \beta_{i,s}) \frac{\pi_{\tilde{s}}}{\pi_s} \{A^i\}_{\tilde{s},s}, \qquad s,\tilde{s} = 1,\ldots,m, \tag{12}$$

and let $\mathbf{h}_t \triangleq [E_{t-1}(\sigma^2_{t,1}|S_t = 1),\ldots,E_{t-1}(\sigma^2_{t,m}|S_t = m)]'$ be an m-by-1 vector of the expected state-dependent conditional variances. Then, we have

$$\mathbf{h}_t = \boldsymbol{\xi} + \sum_{i=1}^{r} \mathcal{K}^{(i)} \mathbf{h}_{t-i}. \tag{13}$$

Define the rm-by-1 vectors $\tilde{\mathbf{h}}_t \triangleq [\mathbf{h}'_t, \mathbf{h}'_{t-1},\ldots,\mathbf{h}'_{t-r+1}]'$ and $\tilde{\boldsymbol{\xi}} \triangleq [\boldsymbol{\xi}',0,\ldots,0]'$ and let

$$\Psi_I \triangleq \begin{bmatrix} \mathcal{K}^{(1)} & \mathcal{K}^{(2)} & & \cdots & & \mathcal{K}^{(r)} \\ I_m & 0_m & & \cdots & & 0_m \\ 0_m & I_m & & & & \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ & & & & & \\ 0_m & \cdots & & 0 & I_m & 0_m \end{bmatrix} \tag{14}$$

be an mr-by-mr matrix where I_m represents the identity matrix of size m-by-m and 0_m is an m-by-m matrix of zeros. Then a recursive vector form of the expected conditional variance (13) can be written as

$$\tilde{\mathbf{h}}_t = \tilde{\boldsymbol{\xi}} + \Psi_I \tilde{\mathbf{h}}_{t-1}, \qquad t \geq 0, \tag{15}$$

with some initial conditions $\tilde{\mathbf{h}}_{-1}$.

Let $\rho(\cdot)$ denote the spectral radius of a matrix, i.e., its largest eigenvalue in modulus, and let Λ_I be an m-by-m square matrix built from the mr-by-mr matrix $(I - \Psi_I)^{-1}$ such that $\{\Lambda_I\}_{ij} = \{(I - \Psi_I)^{-1}\}_{ij}$, $i,j = 1,\ldots,m$. Then we have the following theorem.

THEOREM 1. *An MSG-I process as defined by (1) and (4) is asymptotically wide-sense stationary with variance* $\lim_{t\to\infty} E(\varepsilon_t^2) = \boldsymbol{\pi}'\Lambda_I\boldsymbol{\xi}$, *if and only if* $\rho(\Psi_I) < 1$.[1]

Proof. The recursive equation (15) can be written as

$$\tilde{\mathbf{h}}_t = \Psi_I^t \tilde{\mathbf{h}}_0 + \sum_{i=0}^{t-1} \Psi_I^t \tilde{\boldsymbol{\xi}}, \qquad t \geq 0. \tag{16}$$

According to the matrix convergence theorem (e.g., Lancaster and Tismenetsky, 1985, pp. 327–329), a necessary and sufficient condition for the convergence of (16) for $t \to \infty$ is $\rho(\Psi_I) < 1$. Under this condition, Ψ_I^t converges to zero as t goes to infinity and $\sum_{i=0}^{t-1} \Psi_I^i$ converges to $(I - \Psi_I)^{-1}$, where the matrix $(I - \Psi_I)$ is then guaranteed to be invertible. Therefore, if $\rho(\Psi_I) < 1$, equation (16) yields

$$\lim_{t\to\infty} \tilde{\mathbf{h}}_t = (I - \Psi_I)^{-1} \tilde{\boldsymbol{\xi}}. \tag{17}$$

By definition, the first m elements of $\tilde{\mathbf{h}}_t$ constitute the vector $\mathbf{h}_t$, the first m elements of $\tilde{\boldsymbol{\xi}}$ constitute the vector $\boldsymbol{\xi}$, and the remaining elements of $\tilde{\boldsymbol{\xi}}$ are zeros. Consequently,

$$\lim_{t\to\infty} \mathbf{h}_t = \Lambda_I \boldsymbol{\xi}, \tag{18}$$

and using (5) we have

$$\lim_{t\to\infty} E(\varepsilon_t^2) = \boldsymbol{\pi}' \Lambda_I \boldsymbol{\xi}. \tag{19}$$

Otherwise, if $\rho(\Psi_I) \geq 1$, the expected variance goes to infinity with the growth of the time index. ■

2.2. MSG-II Model

Another variant of Markov-switching GARCH model has recently been proposed by Haas et al. (2004b). This model assumes that a Markov chain controls the ARCH parameters at each regime (i.e., ξ_s and $\alpha_{i,s}$), whereas the *autoregressive* behavior in each regime is subject to the assumption that past conditional variances are in the same regime as that of the current conditional variance. Specifically, the vector of conditional variances $\boldsymbol{\sigma}_t^2 \triangleq [\sigma_{t,1}^2, \sigma_{t,2}^2, \ldots, \sigma_{t,m}^2]'$ is given by

$$\boldsymbol{\sigma}_t^2 = \boldsymbol{\xi} + \sum_{i=1}^{q} \boldsymbol{\alpha}_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} B^{(j)} \boldsymbol{\sigma}_{t-j}^2, \tag{20}$$

where $\boldsymbol{\alpha}_i \triangleq [\alpha_{i,1}, \ldots, \alpha_{i,m}]'$, $i = 1, \ldots, q$, and $\boldsymbol{\beta}_j \triangleq [\beta_{j,1}, \ldots, \beta_{j,m}]'$, $j = 1, \ldots, p$, are vectors of state-dependent GARCH parameters and $B^{(j)} \triangleq \text{diag}\{\boldsymbol{\beta}_j\}$ is a diagonal matrix with elements $\boldsymbol{\beta}_j$ on its diagonal. The same constraints that are sufficient to ensure a positive conditional variance in the MSG-I model (3) are also applied here to guarantee the positivity of the conditional variance.

Note that the conditional variance at a specific regime depends on previous conditional variances of the same regime through the diagonal matrices $B^{(j)}$. Consequently, this model allows derivation of the conditional variance at a given time from past observations only. Furthermore, Haas et al. (2004b) showed that the MSG-II model is analytically more tractable than the MSG-I model and its conditional variance can be straightforwardly constructed because the conditional variance at a specific time does not depend on previous state probabilities but only on previous observations and previous conditional variances.

Let $\boldsymbol{\alpha}_i$ be an m-by-1 vector of zeros for $i > q$ and let $B^{(j)} = 0_m$ for $j > p$. Let $\Omega^{(i)}$ denote an m^2-by-m^2 block matrix of basic dimension m-by-m

$$\Omega^{(i)} \triangleq \begin{bmatrix} \Omega_{11}^{(i)} & \Omega_{21}^{(i)} & \dots & \Omega_{m1}^{(i)} \\ \Omega_{12}^{(i)} & \Omega_{22}^{(i)} & \dots & \Omega_{m2}^{(i)} \\ \vdots & & & \vdots \\ \Omega_{1m}^{(i)} & \Omega_{2m}^{(i)} & \dots & \Omega_{mm}^{(i)} \end{bmatrix}, \tag{21}$$

with each block given by

$$\Omega_{s\tilde{s}}^{(i)} \triangleq p(S_{t-i} = s | S_t = \tilde{s})(\boldsymbol{\alpha}_i \mathbf{e}_s' + \mathbf{B}^{(i)}), \qquad s, \tilde{s} = 1, \dots, m, \tag{22}$$

where $\mathbf{e}_s$ is an m-by-1 vector of all zeros, except its sth element, which is one. We define an rm^2-by-rm^2 matrix by

$$\Psi_{II} \triangleq \begin{bmatrix} \Omega^{(1)} & \Omega^{(2)} & & \dots & & \Omega^{(r)} \\ I_{m^2} & 0_{m^2} & & \dots & & 0_{m^2} \\ 0_{m^2} & I_{m^2} & & & & \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ & & & & & \\ 0_{m^2} & \dots & & 0_{m^2} & I_{m^2} & 0_{m^2} \end{bmatrix}. \tag{23}$$

Let Λ_{II} be an m^2-by-m^2 matrix that is built from the rm^2-by-rm^2 matrix $(I - \Psi_{II})^{-1}$ such that $\{\Lambda_{II}\}_{ij} = \{(I - \Psi_{II})^{-1}\}_{ij}$, $i, j = 1, \dots, m^2$. Let $\bar{\boldsymbol{\pi}} \triangleq [\pi_1 \mathbf{e}_1', \pi_2 \mathbf{e}_2', \dots, \pi_m \mathbf{e}_m']'$; then we get the following theorem for the stationarity condition of an MSG-II process.

THEOREM 2. *An MSG-II process as defined by (1) and (20) is asymptotically wide-sense stationary with variance* $\lim_{t\to\infty} E(\varepsilon_t^2) = \bar{\boldsymbol{\pi}}' \Lambda_{II} \boldsymbol{\xi}$, *if and only if* $\rho(\Psi_{II}) < 1$.

The proof of Theorem 2 is given in Appendix A.

2.3. Comparison of Stationarity Conditions

It has been pointed out by Haas et al. (2004b) that stationarity of the MSG-II model with $p = q = 1$ requires that the regression parameters $\beta_{1,s} < 1$ for all s. It follows from (20) that for general order (p,q), it is necessary that $\sum_{i=1}^{m} \beta_{i,s} < 1$. However, the reaction parameters $\alpha_{i,s}$ may become rather large with correspondence to the regime probabilities. For the MSG-I model, the reaction parameters $\alpha_{i,s}$ and also the regression parameters $\beta_{i,s}$ may be larger than one, provided that the corresponding regime probabilities are sufficiently small. Furthermore, in the representative matrix Ψ_I in (14) the reaction parameters and the regression parameters are weighted by the same weights $p(S_{t-i} = s \mid S_t = \tilde{s})$. Consequently, for a given state s, the values of $\alpha_{i,s}$ and $\beta_{i,s}$ in the MSG-I model have the same contribution to the model stationarity,[2] but for the MSG-II model, each of them affects the heteroskedasticity evolution differently. Figure 1 illustrates the stationarity regions for the MSG-I model (solid line) and the MSG-II model (dashed-dotted line), in the case of two-state Markov chains and GARCH of order (1,1). In Figure 1a, the regime transition probabilities are $a_{1,1} = 0.6$ and $a_{2,2} = 0.7$, and the reaction parameters are $\alpha_{1,1} = 0.4$ and $\alpha_{1,2} = 0.5$. The stationarity region is the interior intersection of each curve and the two axes. In Figure 1b, $a_{1,1} = 0.2$, $a_{2,2} = 0.3$ are considered with reaction parameters $\alpha_{1,1} = 0.8$ and $\alpha_{1,2} = 0.2$. For the MSG-I model, stationarity is allowed with regression parameters larger than one, whereas for the MSG-II model, $\beta_{1,1}$ and $\beta_{1,2}$ must both be smaller than one for stationarity. In both cases, $\pi_2 > \pi_1$; however, in Figure 1a the stationarity region of the MSG-II model is contained in the stationarity region of the MSG-I model whereas in Figure 1b, in which case $\alpha_{1,1} \gg \alpha_{1,2}$, for $\beta_{1,1} \in [0.2, 0.55]$ stationarity is achieved with a larger $\beta_{1,2}$ for the MSG-II model than for the MSG-I model.

3. RELATION TO OTHER WORKS

Klaassen (2002) developed conditions that are necessary, but not necessarily sufficient, for asymptotic stationarity of a two-state MSG-I model of order (1,1). Consider the 2-by-2 matrix C with elements

$$c_{ij} = a_{ji}(\alpha_{1,i} + \beta_{1,i})\pi_j / \pi_i, \qquad i,j = 1,2. \tag{24}$$

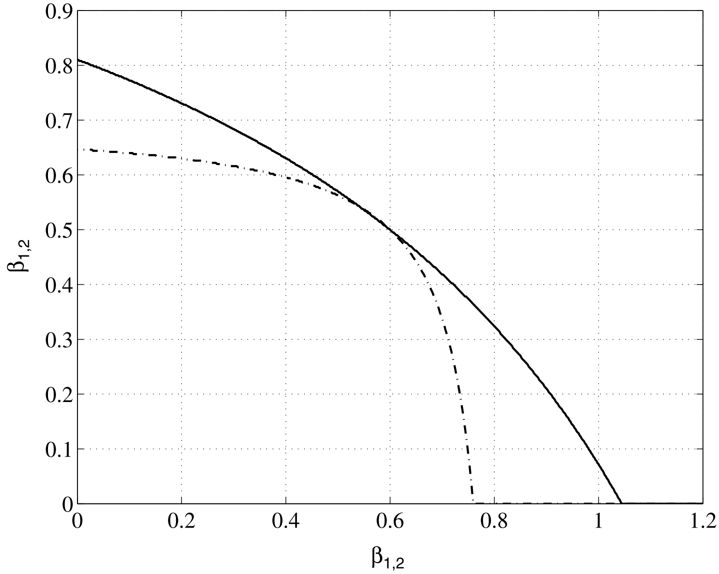
Klaassen showed that the stationary variance of the process is given by

$$\sigma^2 = \boldsymbol{\pi}'(I - C)^{-1}\boldsymbol{\xi} \tag{25}$$

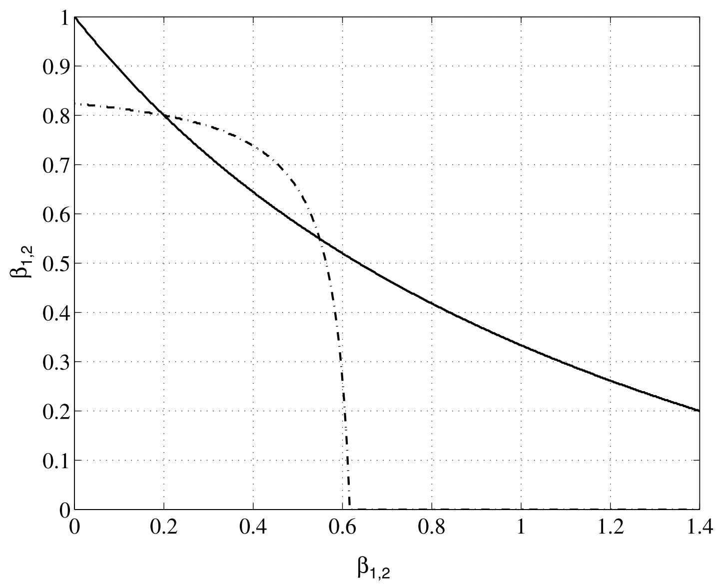
and that the conditions

$$c_{11}, c_{22} < 1 \quad \text{and} \quad \det(I - C) > 0 \tag{26}$$

are necessary to ensure that the stationary variance is finite and positive.



(a)



(b)

FIGURE 1. Stationarity regions for two-state Markov-chains with GARCH of order (1,1) corresponding to MSG-I (solid line) and MSG-II (dashed-dotted line). The regime transition probabilities and the reaction parameters are (a) $a_{1,1} = 0.6$, $a_{2,2} = 0.7$ and $\alpha_{1,1} = 0.4$, $\alpha_{1,2} = 0.5$; (b) $a_{1,1} = 0.2$, $a_{2,2} = 0.3$ and $\alpha_{1,1} = 0.8$, $\alpha_{1,2} = 0.2$.

For the special case of our analysis for GARCH orders of (1,1) and the MSG-I model with two states, the representative matrix Ψ_I reduces to matrix C, and the stationary variance reduces to the expression given in (25). Metzler (1950) showed that for a nonnegative matrix C (i.e., $c_{ij} \geq 0$), $\rho(C) < 1$ if and only if all of the principal minors of $(I - C)$ are positive. Furthermore, together with the Hawkins–Simon condition (Hawkins and Simon, 1949), $\rho(C)$ is less than one if and only if $(I - C)^{-1}$ has no negative elements. Therefore, for the nonnegative matrix C, the condition $c_{11}, c_{22} < 1$ implies $\det(I - C) > 0$, and it is equivalent to $\rho(C) < 1$. Accordingly, the conditions of Klaassen are not only necessary but also sufficient for asymptotic stationarity.

A necessary and sufficient condition for asymptotic stationarity of an MSG-II model of order (1,1) has been developed by Haas et al. (2004b). Accordingly, the largest eigenvalue in modulus of an m^2-by-m^2 block matrix D is constrained to be less than one, where

$$D = \begin{bmatrix} D_{11} & D_{21} & \dots & D_{m1} \\ D_{12} & D_{22} & \dots & D_{m2} \\ \vdots & \vdots & & \vdots \\ D_{1m} & D_{2m} & \dots & D_{mm} \end{bmatrix} \tag{27}$$

is built from matrices D_{ij} of size m-by-m that are obtained by

$$D_{ij} = a_{ij}(B^{(1)} + \boldsymbol{\alpha}_1 \mathbf{e}'_j). \tag{28}$$

The stationarity analysis of Haas et al. (2004b) for an MSG-II process employs a forward recursive calculation of the expected conditional variance, assuming some initial conditions. As a result, the probabilities of state transitions, a_{ij}, are used for evaluating the expectation of the one-step-ahead conditional variance. Our analysis manipulates a backward recursion of the conditional variance expectation, and thus it uses the stationary probabilities of the Markov chain, along with the transition probabilities, to generate previous conditional state probabilities $p(s_{t-i}|s_t)$. Therefore, when we degenerate the MSG-II model to order $p = q = 1$ (which is the case analyzed by Haas et al., 2004b) the block matrices D in (27) and Ψ_{II} in (23) are not identical, and specifically, for that order of model we have $\Psi_{II} = \Omega^{(1)}$ and

$$\Omega_{ij}^{(1)} = \frac{\pi_i a_{ij}}{\pi_j a_{ji}} D_{ji}. \tag{29}$$

Although our representative matrix Ψ_{II} and that developed by Haas et al. (2004b) do not share the same elements, we show in Appendix B that their eigenvalues are identical and therefore both conditions are equivalent for that order of MSG-II.

A special case of any of the MSG models is a degenerated case of having a single regime of order (p,q) (the models reduce to a standard GARCH(p,q) model). In that case, the representative matrices are equal, $\Psi_I = \Psi_{II}$. Francq et al. (2001) developed a stationarity condition for the *natural* case of Markov-switching GARCH model, in which case the conditional variance depends on the active regime path. For the special case of a single-regime model they got the transition matrix of a standard GARCH(p,q) model, which is equal to that which is derived, e.g., by substituting $\mathcal{K}^{(i)} = \alpha_{i,1} + \beta_{i,1}$ and $I_1 = 1$ in (14). They showed that having the spectral radius of that matrix less than one is equivalent to Bollerslev's condition for the asymptotic wide-sense stationarity of a GARCH(p,q) model, $\sum_{i=1}^{r}(\alpha_{i,1} + \beta_{i,1}) < 1$ (Bollerslev, 1986).

4. CONCLUSIONS

Conditions for asymptotic wide-sense stationarity of random processes with time-variant distributions are useful for ensuring the existence of a finite asymptotic volatility of the process. We developed a comprehensive approach for stationarity analysis of Markov-switching GARCH processes where finite state-space Markov chains control the switching between regimes and GARCH models of order (p,q) are active in each regime. Necessary and sufficient conditions for the asymptotic stationarity are obtained by constraining the spectral radius of representative matrices, which are built from the model parameters. These matrices also enable derivation of compact expressions for the stationary variance of the processes.

NOTES

1. Note that for a matrix with nonnegative elements, there exists a real eigenvalue that is equal to the spectral radius (Horn and Johnson, 1985, p. 288).
2. This also holds for the natural extension of GARCH(p,q) to Markov switching, which has been analyzed by Francq et al. (2001).

REFERENCES

Alexander, C. & E. Lazar (2004) Normal Mixture GARCH(1,1) Applications to Exchange Modelling. ISMA Centre Discussion Papers in Finance 2004–06, the Business School for Financial Markets at University of Reading.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.

Cai, J. (1994) A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics* 12, 309–316.

Dueker, M.J. (1997) Markov switching in GARCH processes and mean reverting stock market volatility. *Journal of Business & Economic Statistics* 15, 26–34.

Francq, C., M. Roussignol, & J.-M. Zakoïan (2001) Conditional heteroskedasticity driven by hidden Markov chains. *Journal of Time Series Analysis* 22, 197–220.

Francq, C. & J.-M. Zakoïan (2001) Stationarity of multivariate Markov-switching ARMA models. *Journal of Econometrics* 102, 339–364.

Francq, C. & J.-M. Zakoïan (2002) Comments on the paper by Minxian Yang: "Some properties of vector autoregressive processes with Markov-switching coefficients." *Econometric Theory* 18, 815–818.

Francq, C. & J.-M. Zakoïan (2005) The l^2-structure of standard and switching-regime GARCH models. *Stochastic Processes and Their Applications* 115, 1557–1582.

Frömmel, M. (2004) Modelling Exchange Rate Volatility in the Run-up to EMU Using Markov Switching GARCH Model. Discussion paper, Hannover University.

Gray, S.F. (1996) Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* 42, 27–62.

Haas, M., S. Mittnik, & M.S. Paolella (2004a) Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics* 2, 211–250.

Haas, M., S. Mittnik, & M.S. Paolella (2004b) A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics* 2, 493–530.

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

Hawkins, D. & H. Simon (1949) Note: Some conditions of macroeconomics stability. *Econometrica* 17, 245–248.

Horn, R.A. & C.R. Johnson (1985) *Matrix Analysis*. Cambridge University Press.

Klaassen, F. (2002) Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics* 27, 363–394.

Lancaster, P. & M. Tismenetsky (1985) *The Theory of Matrices*, 2nd ed. Academic Press.

Marcucci, J. (2005) Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics* 9, 1145.

Metzler, L.A. (1950) A multiple-region theory of income and trade. *Econometrica* 18, 329–354.

Timmermann, A. (2000) Moments of Markov switching models. *Journal of Econometrics* 96, 75–111.

Wong, C.S. & W.K. Li (2001) On a mixture autoregressive conditional heteroscedastic model. *Journal of the American Statistical Association* 96, 982–985.

Yang, M. (2000) Some properties of vector autoregressive processes with Markov-switching coefficients. *Econometric Theory* 16, 23–43.

Yao, J. (2001) On square-integrability of an AR process with Markov switching. *Statistics and Probability Letters* 52, 265–270.

APPENDIX A

Proof of Theorem 2. In this Appendix we prove Theorem 2, which gives necessary and sufficient conditions for the asymptotic wide-sense stationarity of the MSG-II model and also its stationary variance.

Following (20) and (5), the expectation of the MSG-II conditional variance under a chain state, s, follows:

$$E_{t-1}(\sigma_{t,s}^2|s_t) = \xi_s + \sum_{i=1}^{q} \alpha_{i,s} E_{t-1}(\varepsilon_{t-i}^2|s_t) + \sum_{j=1}^{p} \beta_{j,s} E_{t-1}(\sigma_{t-j,s}^2|s_t), \tag{A.1}$$

where using (7) and (8)

$$E_{t-1}(\varepsilon_{t-i}^2|s_t) = \sum_{s_{t-i=1}}^{m} p(s_{t-i}|s_t) E_{t-i-1}(\sigma_{t-i,s_{t-i}}^2|s_{t-i}) \tag{A.2}$$

and

$$E_{t-1}(\sigma^2_{t-j,s}|s_t) = E_{S_{t-j}}[E_{t-1}(\sigma^2_{t-j,s}|s_{t-j},s_t)]$$
$$= \sum_{s_{t-j}=1}^{m} p(s_{t-j}|s_t)E_{t-i-1}(\sigma^2_{t-j,s}|s_{t-j}). \quad \textbf{(A.3)}$$

The main difference between an MSG-II model and an MSG-I model is that the conditional variance depends on previous conditional variances of the same regime, regardless of the past regime path. By contrast, for the MSG-I model, the conditional variance is a linear combination of past state-dependent conditional variances, where for each one the state is conditioned to be the active one. Consequently, the computation of the unconditional variance for an MSG-II model requires the terms $E_{t-j-1}(\sigma^2_{t-j,s}|s_{t-j})$ for all $s = 1,\ldots,m$, whereas in the case of the MSG-I model, only $E_{t-j-1}(\sigma^2_{t-j,s_{t-j}}|s_{t-j})$ is relevant to calculate the expectation of the unconditional variance. Accordingly, an m^2-by-1 vector is necessary to represent $E_{t-1}(\sigma^2_{t,s}|s_t)$ elements, and an rm^2-by-rm^2 matrix is employed for the recursive formulation.

By substituting (A.3) and (A.2) into (A.1) we have

$$E_{t-1}(\sigma^2_{t,s}|s_t) = \xi_s + \sum_{i=1}^{r}\sum_{s_{t-i}=1}^{m} p(s_{t-i}|s_t)$$
$$\times [\alpha_{i,s}E_{t-i-1}(\sigma^2_{t-i,s_{t-i}}|s_{t-i}) + \beta_{i,s}E_{t-i-1}(\sigma^2_{t-i,s}|s_{t-i})]. \quad \textbf{(A.4)}$$

Let $g_t(s,s_t) \triangleq E_{t-1}(\sigma^2_{t,s}|s_t)$ and let $\mathbf{g}_t \triangleq [g_t(1,1), g_t(2,1),\ldots,g_t(m,1), g_t(1,2),\ldots, g_t(m,m)]'$ be a vector of expected, state-dependent, conditional variances. Then, a recursive formulation of the conditional variance is given by

$$\tilde{\mathbf{g}}_t = \tilde{\boldsymbol{\xi}} + \Psi_{II}\tilde{\mathbf{g}}_{t-1}, \qquad t \geq 0, \quad \textbf{(A.5)}$$

where $\tilde{\mathbf{g}}_t \triangleq [\mathbf{g}'_t, \mathbf{g}'_{t-1},\ldots,\mathbf{g}'_{t-r+1}]'$. The completion of this proof follows the proof of Theorem 1. ■

APPENDIX B

Equivalence with Haas Condition. In this Appendix we show that the eigenvalues of matrices D in (27) and $\Psi_{II} = \Omega^{(1)}$ in (23) are equal for the case of an m-state MSG-II model of order (1,1).

Let $\tilde{D}$ denote an m^2-by-m^2 matrix that is given by

$$\tilde{D} \triangleq \begin{bmatrix} B^{(1)} + \boldsymbol{\alpha}_1\mathbf{e}'_1 & 0_m & \cdots & 0_m \\ 0_m & B^{(1)} + \boldsymbol{\alpha}_1\mathbf{e}'_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_m \\ 0_m & \cdots & 0_m & B^{(1)} + \boldsymbol{\alpha}_1\mathbf{e}'_m \end{bmatrix} \quad \textbf{(B.1)}$$

and let $\otimes$ denote the Kronecker product. Then $D = \tilde{D}(A' \otimes I_m)$. Let $\mathbf{1}_m$ denote an m-by-1 vector of ones and let $P \triangleq \mathrm{diag}(\boldsymbol{\pi} \otimes \mathbf{1}_m)$. By substituting (28) into (29), we have

$$\Omega_{ij}^{(1)} = \frac{\pi_i}{\pi_j} a_{ij} (B^{(1)} + \boldsymbol{\alpha}_1 \mathbf{e}_i') \tag{B.2}$$

and

$$\Omega^{(1)} = P^{-1} (A' \otimes I_m) \tilde{D} P. \tag{B.3}$$

Therefore, $\Omega^{(1)}$ and $(A' \otimes I_m)\tilde{D}$ are similar matrices, and the spectrum of $\Omega^{(1)}$, $\mathrm{eig}\{\Omega^{(1)}\}$, satisfies

$$\mathrm{eig}\{\Omega^{(1)}\} = \mathrm{eig}\{(A' \otimes I_m)\tilde{D}\} = \mathrm{eig}\{\tilde{D}(A' \otimes I_m)\} = \mathrm{eig}\{D\}. \tag{B.4}$$