The Continuous Limit of GARCH Processes

ISMA Centre Discussion Papers in Finance 2004-10

First Version: July 2004
This Version: February 2005

Carol Alexander
ISMA Centre, University of Reading

Emese Lazar
ISMA Centre, University of Reading

Copyright 2004 Alexander and Lazar. All rights reserved.
Abstract

Contrary to popular belief, the diffusion limit of a GARCH variance process is not a diffusion model unless one makes a very specific assumption that cannot be generalized. In fact, the GARCH price of European call and puts are identical to the Black-Scholes prices based on the average of a deterministic variance process. In the case of GARCH models with several normal components – and these are more realistic representations of option prices and returns behaviour – we show that the continuous limit is a stochastic model with uncertainty over which deterministic local volatility governs the return. We also extend the normal mixture GARCH framework to a model with time-varying mixing parameter and show the direct relationship of these models with Markov Switching GARCH models in discrete and continuous time. An interesting area to be considered for application of these multi-state GARCH models is path dependent option pricing and hedging. Since the transition price densities are lognormal mixtures although the marginal densities are normal, the mixture GARCH option pricing model is not equivalent to the mixture option pricing models that have previously been discussed by several authors.

JEL Classification Codes: C32, G13.
Keywords: GARCH diffusion, normal mixture, stochastic volatility, time aggregation

Author Details:
Professor Carol Alexander
Chair of Risk Management and Director of Research,
ISMA Centre, Business School,
The University of Reading, PO Box 242
Reading RG6 6BA, United Kingdom
Email: c.alexander@ismacentre.rdg.ac.uk
Tel: +44 (0)1183 786431 (ISMA Centre) Fax: +44 (0)1189 314741

Emese Lazar
PhD Student,
ISMA Centre, Business School,
The University of Reading, PO Box 242
Reading RG6 6BA, United Kingdom
Email: e.lazar@ismacentre.rdg.ac.uk
Tel: +44 (0)1183 786675 (ISMA Centre) Fax: +44 (0)1189 314741

§ We are grateful for useful comments on an earlier draft from Leonardo Nogueira of the ISMA Centre, University of Reading, Hyungsok Ahn, of Commerz Bank, London and Valentina Corradi of Queen Mary, University of London.
I Introduction

It has long been supposed that GARCH variance processes have continuous time limits that are stochastic volatility models (see for instance, Wilmott, 2000, p. 379 and Lewis, 2000, p.191). The seminal paper by Nelson (1990) derived general results on the weak convergence of stochastic difference equations to stochastic differential equations that are based on finite moment conditions of both the returns and the variance processes. His application of these results to the normal GARCH(1,1) process assumed that the GARCH error coefficient decreases at a rate that is proportional to the square root of the observation interval \( h \). In other words, \( \lim_{h \to 0} (b^{-1} \alpha_h^2) \) has a finite positive value. Then this value is the diffusion coefficient for the continuous-time model. Obviously, if one assumes that \( \lim_{h \to 0} (b^{-1} \alpha_h^2) = 0 \), the GARCH limit is not a diffusion at all.

Recent research has placed considerable doubt over the ‘GARCH’ mean reverting stochastic volatility models that are now commonly used. Corradi (2000) proved not only that the normal GARCH (1,1) limit is a variance diffusion only under the specific limiting assumption above, but also that any (non-degenerate) diffusion model has a discrete time equivalent that is not a GARCH model – in fact it remains stochastic in discrete time. Wang (2002) used the asymptotic non-equivalence of the likelihood functions to demonstrate that the continuous limit of normal GARCH(1,1) must have a deterministic variance, i.e. it cannot be a diffusion model. Brown, Wang and Zhao (2002) consider stronger convergence conditions and again show that there can be no diffusion term in the continuous limit of normal GARCH models. In fact, a common theme is all these papers is that the transition from continuous variance diffusion to discrete time variance model yields a discrete time stochastic volatility model, such as the autoregressive volatility model that was introduced by Taylor (1986) and not a GARCH model.

This paper will add further weight to this research as our main theorem forces us to focus on the assumption that \( \lim_{h \to 0} (b^{-1} \alpha_h^2) \) has a finite positive value. We show that this assumption is impossible to generalize to GARCH processes that we know to be better representations of financial asset returns. We also consider the limit of asymmetric GARCH specifications as well as the normal GARCH(1,1) process. We extend the normal mixture GARCH model to one with time-varying mixing parameter and discuss its limit. Markov Switching GARCH models – which are closely related to GARCH processes with normal mixture distributed errors – are treated as well in a continuous-time framework.

Our starting point is the acknowledged fact that the time series properties of financial asset returns are not well captured by the normal GARCH(1,1) process. To better explain the unconditional skewness and leptokurtosis in financial data Bollerslev (1987) introduced the Student’s \( \xi \)-GARCH(1,1) model and Fernandez and Steel (1998) extended this to the skewed \( \xi \)-distribution. The models of Nelson (1991), Engle and Ng (1993) and Glosten, Jagannathan and Runkle (1993) amongst others were introduced to capture the asymmetry in volatility.
that arises, for example, from leverage effects in equity markets. However, none of these models are able to explain the time variation that is observed in the skewness and kurtosis of conditional returns densities in virtually all financial markets. This type of variation has been added exogenously, augmenting the GARCH process as in Hansen (1994) and Harvey and Siddique (1999), but the only general class of GARCH processes that can endogenously explain the time variability in conditional higher moments are those that specify the existence of two or more market regimes, or states in the model, each state being governed by a different GARCH variance processes. In this case the conditional densities of returns will follow a normal mixture process.

Normal mixture GARCH(1,1) models (which we denote NM-GARCH) with restricted conditional densities have been extensively studied in the works of Vlaar and Palm (1993), Ding and Granger (1996), Bauwens, Bos and van Dijk (1999), Bai, Russell and Tiao (2001, 2003), Roberts (2001) and others. In addition to the GARCH variance process parameters, an estimation of the NM-GARCH model provides estimates of the unconditional regime, or state, probabilities. Although none of these models capture time variation in conditional skewness and few have time-varying conditional kurtosis, recently Haas, Mittnik and Paolella (2004a) and Alexander and Lazar (2004, 2005) have studied general unrestricted NM-GARCH models that have very flexible individual variance processes and time-varying conditional higher moments. Alexander and Lazar (2004, 2005) provide strong evidence that unrestricted models with just two GARCH variance components provide closer fits to the conditional and unconditional physical returns densities than normal and symmetric or skewed t-GARCH models with only one variance component. Both symmetric and asymmetric components were considered for all models and all major exchange rates and stock index returns were examined.

Markov switching (from now on denoted by MS) models were introduced by Hamilton (1989) and they were first applied to ARCH models by Cai (1994) and Hamilton and Susmel (1994). The Markov switching framework was extended by Gray (1996) and subsequently by Klaassen (2002) to GARCH models, but these models did not become popular due to their estimation problems. In their specification the individual variances do not only depend on their lagged values, but on a mixture of the previous individual variances. A recent extension is that of Haas et al. (2004b) in which the individual variances depend on their own lagged values. The only difference between this new Markov switching model and the NM-GARCH model with time-varying mixing parameter is in the state dynamics. Whilst in the NM-GARCH model at each point in time the selection of the state is random and does not depend on the previous state (ex post it depends only on the observations), in a Markov switching specification the prevailing state depends on the previous state as well as the observations.

It should come as no surprise to market practitioners that a mean-reverting volatility model with more than one state fits market data much better than models with only one possible mean-reversion mechanism. Different
types of shocks lead to quite different speeds of mean-reversion. For instance, a rumour may have an enormous effect but die out very quickly whereas an announcement of important changes to economic policy may persist and even raise the general level of volatility for some time. At least two mean-reversion mechanisms should be specified in any volatility model in order to capture such effects. Normal GARCH models have only one mean-reversion mechanism so the one they detect will depend on the weighted average of the mean reversion under different circumstances as well as on the data frequency. Rapid mean-reversion will dominate the high frequency data and low frequency data can only contain the slower mean-reverting effects.

Whilst the above remarks are intuitively obvious, up to now there has been little research on option pricing with more than one mean-reversion mechanism. And pricing European options with normal mixture GARCH and MS-GARCH models is a virtually undiscovered area. For this we require a specification of the continuous time limit of the normal mixture and Markov switching GARCH processes, and this is the task that we have set ourselves here. Since these processes offer an improved fit in the physical measure and have features such as time variation in higher conditional moments and multiple mean-reverting mechanisms that we know to be essential for any model that pertains to be consistent with the behaviour of implied volatility (see Garcia, Ghysels and Renault, 2005), this task seems eminently worthwhile.

We show that the properties of the process in discrete time carry over into continuous time. That is, in both discrete and continuous time there is uncertainty over a finite set of deterministic mean-reverting variance processes. The continuous limit of normal mixture GARCH is therefore a stochastic volatility model, but not in the traditional sense where there is a correlated diffusion for the variance process. Similar conclusions can be drawn for the Markov switching GARCH model, too.

Perhaps the result that will have the most immediate impact comes as a corollary to the main theorem. We already know from Corradi (2000) and others that the limiting assumption made by Nelson (1990) is arbitrary and special. Here we show that it cannot be generalised to the case where more than one GARCH variance component governs returns, depending on the market regime. When assumptions that can be generalized to mixture GARCH processes are made, then both symmetric and asymmetric normal GARCH models will converge to deterministic local volatility models. Hence their marginal price densities will be lognormal, as in the Black-Scholes (1973) model and the GARCH prices of European calls and puts will be equivalent to Black-Scholes prices.

The remainder of this paper is organized as follows: Section II derives the results for symmetric and different asymmetric normal mixture GARCH variance processes; Section III applies these results to symmetric and asymmetric normal GARCH models. We show that the continuous time limits of the normal GARCH,
AGARCH and GJR-GARCH models are geometric Brownian motions with deterministic local volatility. Section IV defines a weak normal mixture GARCH process that has the time-aggregation property and shows that this process also has a deterministic variance process in the limit. Section V extends this model to one that has a time-dependent mixing parameter, this model being the connection between normal mixture and Markov switching GARCH models. Section VI derives the continuous-time limit of Markov Switching GARCH models. Section VII presents a discussion of the previous results and Section VIII summarizes and concludes.

II The Continuous Limit of Normal Mixture GARCH Processes

The following takes as starting point the normal mixture GARCH(1,1) model, denoted 'NM(K)-GARCH(1, 1)' as specified by Alexander and Lazar (2005). We extend it by including a constant in the mean equation and express the model as function of the returns $y_t$. Hence the discrete-time specification is

\begin{equation}
    y_t = \mu + \varepsilon_t,
\end{equation}

where

\begin{equation}
    y_t = \frac{S_t - S_{t-1}}{S_{t-1}} \approx \ln(S_t / S_{t-1})
\end{equation}

\begin{equation}
    \sigma^2_t = \omega + \alpha_i \varepsilon^2_{t-1} + \beta \sigma^2_{t-1} \quad i = 1, \ldots, K
\end{equation}

\begin{equation}
    \sigma^2_t = \sum_{i=1}^{K} p_i \sigma^2_{i} + \sum_{i=1}^{K} p_i \mu^2_{i}
\end{equation}

\begin{equation}
    \sum_{i=1}^{K} p_i = 1 \quad \text{and} \quad \sum_{i=1}^{K} p_i \mu_{i} = 0
\end{equation}

The error term is assumed to follow a conditional normal mixture distribution:

\begin{equation}
    \varepsilon_t \mid \varepsilon_{t-1} \sim \text{NM}\left(p_1, \ldots, p_K; \mu_1, \ldots, \mu_K; \sigma^2_1, \ldots, \sigma^2_K\right)
\end{equation}

We would like to mention that in equation (2) the individual variance depends on its previous realized value, and not on the previous overall variance. We consider that this is a better approach since it allows the individual variances to drift away from the overall variance, thus giving a better fit.

This model is characterized by time varying conditional skewness and excess kurtosis given by the following formulae:

\begin{equation}
    \tau_t = \frac{E_{t-1}(\varepsilon^3_t)}{\left(\sigma^2_t\right)^{3/2}} = \frac{3 \sum_{i=1}^{K} p_i \mu_i \sigma^2_i + \sum_{i=1}^{K} p_i \mu^2_i}{\left(\sigma^2_t\right)^{3/2}}
\end{equation}

\begin{equation}
    \kappa_t = \frac{E_{t-1}(\varepsilon^4_t)}{\left(\sigma^2_t\right)^2} - 3 = \frac{3 \sum_{i=1}^{K} p_i \sigma^4_i + 6 \sum_{i=1}^{K} p_i \mu^2_i \sigma^2_i + \sum_{i=1}^{K} p_i \mu^4_i}{\left(\sigma^2_t\right)^2} - 3
\end{equation}

That is, the conditional density functions are

\[ f_i(\varepsilon_t) = \sum_{i=1}^{K} p_i \varphi_i(\varepsilon_t) \] where

\[ \varphi_i(\varepsilon_t) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(\varepsilon_t - \mu_i)^2}{2\sigma^2_i}\right) \]
and non-zero unconditional skewness and excess kurtosis. These properties will still hold in the continuous time limit: we show in Theorem 3 below that the log price density in the physical measure is not normal, but a mixture of normal densities with the mixing law, means and variances given above.

We examine the behaviour of the data generation process (1) – (5) as the time interval between realisations changes by introducing a ‘step-length’ \( h \) between realisations and a notation for a time series with step-length \( h \) that uses a pre-subscript \( b \) and indexes time as \( k \) with \( k = 1, 2, \ldots \) Figure 1 illustrates the error process for \( h = 2 \), denoted \( \varepsilon_{2k} \) and compares this with the error process for \( h = 1 \), denoted \( \varepsilon_{kr} \) for \( k = 1, 2, \ldots \). Note that we drop the pre-suffix when \( b = 1 \). At time \( t = 2k \) we have \( \varepsilon_{2k} = \varepsilon_{2k} + \varepsilon_{2k-1} \) so the variance of the two-step error process is twice the magnitude of the variance of the one-step error process.

Figure 1: Error Processes for \( h = 2 \) and \( h = 1 \).

\[
\begin{align*}
\text{\( h = 2 \)} & \quad \varepsilon_{2k} \quad \varepsilon_{2(k+1)} \\
\text{\( h = 1 \)} & \quad \varepsilon_{2k-1} \quad \varepsilon_{2k} \quad \varepsilon_{2k+1} \quad \varepsilon_{2k+2} \\
& \text{time}
\end{align*}
\]

In the following we need to compare the individual variances (2) for processes of different time-steps, so we shall divide the \( b \)-step squared error process by \( b \) to express the individual variance series. Also, we adopt the interpretation for the normal mixture distribution according to which at each point in time we have one of the states ruling, and in each state the returns follow a normal distribution. We thus re-write (1) – (5) for the annualised variance processes, now allowing the series and the parameters to depend on \( h \) as follows:

(8) \( b_y(kh)h = b_y + b_y \varepsilon_{kh} \) where \( b_y(kh)h = \frac{S_{kh} - S_{(k-1)h}}{S_{(k-1)h}} \approx bh \left( \frac{S_{kh}}{S_{(k-1)h}} \right) \)

(9) \( b_y(kh) = b_y + b_y \varepsilon_{kh}^2 / h + b_y b_y \varepsilon_{kh}^2 \)

(10) \( b_y(kh) = \sum_{i=1}^{K} P_i b_y \sigma_i^2 + b_y \sum_{i=1}^{K} P_i \mu_i^2 \)

This last equation gives the overall variance. The time-invariant state-probabilities are given by:

(11) \( b_y(kh) = P(b_y(kh) = i) \) with \( \sum_{i=1}^{K} P_i = 1 \) and \( \sum_{i=1}^{K} P_i \mu_i = 0 \)

Conditionally, in each state we assume a normal distribution:

(12) \( b_y(kh) \mid \{ I_{(k-1)h}, b_y \} = i \sim N(b_y(kh), b_y \sigma_i^2) \) where \( \{ I_{(k-1)h}, b_y \} = \{ b_y(kh), b_y(k(2)h), \ldots \} \)
With these definitions, the continuous limit of the price and variance components can now be defined as:

\[(13)\quad S(t) := \lim_{h \to 0} \frac{1}{h} S_t, \quad \text{where} \quad \frac{1}{h} S_t := S_{t+h} \quad \text{for} \quad kh \leq t < (k+1)b\]

\[(14)\quad \sigma_i^2(t) := \lim_{h \to 0} \frac{1}{h} \sigma_{i,t}^2, \quad \text{where} \quad \frac{1}{h} \sigma_{i,t}^2 := \sigma_{i,t+h}^2 \quad \text{for} \quad kh \leq t < (k+1)b\]

and

\[(15)\quad \sigma^2(t) := \lim_{h \to 0} \frac{1}{h} \sigma^2_t, \quad \text{where} \quad \frac{1}{h} \sigma_t^2 := \sigma_{t+h}^2 \quad \text{for} \quad kh \leq t < (k+1)b\]

Similarly, for the state variable we have:

\[(16)\quad s(t) := \lim_{h \to 0} \frac{1}{h} s_t, \quad \text{where} \quad \frac{1}{h} s_t := s_{t+h} \quad \text{for} \quad kh \leq t < (k+1)b\]

We employ the weak convergence results of Nelson (1990) to derive the continuous time limit of the process (8) – (12), and for this we must make some assumptions about the limiting behaviour of the parameters as the step length tends to zero. For the continuous limit of normal mixture GARCH(1,1) process to exist we must assume that the following limits exist and are finite:

\[(17)\quad \omega_i := \lim_{h \to 0} \left( \frac{h \omega_i}{b} \right), \quad \psi_i := \lim_{h \to 0} \left( \frac{h \psi_i}{b} \right), \quad \psi_i := \lim_{h \to 0} \left( \frac{1-\psi_i}{b} \right), \quad p_i := \lim_{h \to 0} (\frac{\psi_i}{b}), \quad \mu_i := \lim_{h \to 0} (\frac{\mu_i}{b})\]

with \(\psi_i \geq 0, \psi_i > 0, 0 < p_i < 1\) and \(\sum_{i=1}^{K} p_i = 1\).

What differentiates this from the previous literature is that we find it necessary to assume the existence of separate limits for the \(\alpha\) and \(\beta\) parameters whereas previously a limit was only defined on \(\alpha + \beta\). Note that under (17), \(\lim_{h \to 0} \left( b^{-1} \sigma^2_i \right) = 0\).

The first result in this section proves that the continuous time limit of the normal mixture GARCH(1,1) variance process is a mixture model with uncertainty over a latent state variable governing a set of deterministic mean-reverting component variance processes. This is not stochastic volatility model in the traditional sense because there can be no diffusion term in the variance processes. In fact, each GARCH variance component converges to a deterministic variance process and all the uncertainty in the model is captured by the mixing law \((p_1, \ldots, p_K)\).

**Theorem 1:** The continuous time limit of the normal mixture GARCH process (8) – (12) has, with probability \(p_i\):

\[(18)\quad \frac{dS(t)}{S(t)} = (\mu + \mu_i) \, dt + \sigma_i(t) \, dB(t)\]

where

\[(19)\quad d\sigma_i^2(t) = \left( \omega_i + \psi_i \sigma^2(t) - \psi_i \sigma_i^2(t) \right) dt, \quad \sigma_i^2(t) = \sum_{i=1}^{K} p_i \sigma_i^2(t)\]
and \( B(t) \) is a Brownian motion under the physical measure. For this limit to exist, condition (17) must be satisfied.

**Proof:** In the following, to ease notation, we shall drop the pre-subscript \( h \) whenever it is clear from the context. We shall apply the weak convergence result of Nelson (1990) that states that a sequence of discrete processes (based on step-length \( h \)) converges in distribution as \( h \downarrow 0 \) to a continuous Ito process (i.e. defined by an appropriate stochastic integral) if the moments per unit time, conditional on information up to time \((k-1)h\), of the limit process converge to a well-behaved limit.

Consider first the returns process. For the drift per unit time we have:

\[
E\left( b^{-1} \left( \frac{S_{k,h} - S_{I_{(k-1)h}}}{S_{I_{(k-1)h}}} \right) \right| I_{(k-1)h}, I_{(k-1)h} = i) = b^{-1} \left( I_{(k-1)h} - h \mu_i \right) = \left( \mu + \mu_i \right)
\]

Hence, ignoring terms of second and higher order, in state \( i \) the percentage change in \( S_k \) can be approximated by \( (\mu + \mu_i) \) and the drift per unit time converges to \( (\mu + \mu_i) \) under our assumptions (17). For the second moment of the returns process per unit time we have:

\[
E\left( b^{-1} \left( \frac{S_{k,h} - S_{I_{(k-1)h}}}{S_{I_{(k-1)h}}} \right)^2 \right| I_{(k-1)h}, I_{(k-1)h} = i) = \sigma_{1,k,h}^2
\]

so the conditional second moment per unit time converges to \( \sigma_i^2(t) \) as \( b \downarrow 0 \). For the variance process we have:

\[
E\left( b^{-1} \left( \sigma_{j,k+1,h}^2 - \sigma_{1,k,h}^2 \right) \right| I_{(k-1)h}) = \frac{\omega_i}{b} + \frac{h \alpha_j}{b} E\left( \sigma_{k,h}^2 \right| I_{(k-1)h}) / b + \left( \beta_j - \frac{1}{b} \right) \sigma_{1,k,h}^2
\]

We need the limit to exist and be finite when \( b \to 0 \) for \( i = 1, \ldots, K \) and it is this that forces us to have separate limits for the \( \alpha \) and \( \beta \) parameters. The variances and covariances of the individual variance components are not significant, considering (12):

\[
E\left( b^{-1} \left( \sigma_{j,k+1,h}^2 - \sigma_{1,k,h}^2 \right) \right) = b \left( \frac{\alpha_j}{b} \right) \sigma_{1,k,h}^2 = \sigma_j^4 + 6b \left( \frac{\alpha_j}{b} \right) \sigma_{1,k,h}^2 + b^2 \left( \frac{\alpha_j}{b} \right)^2 \sigma_{1,k,h}^4 = o(1)
\]
The covariances between the percentage returns and the changes in the variances are also small:

\[
E\left(\frac{S_{t+h} - S_{t}}{S_{t}} - b(\mu + \mu_{f})\right)\left(\sigma_{t+h}^{2} - \sigma_{t}^{2} - \left(\delta_{t+h}^{2} - \delta_{t}^{2}\right)\right)
\]

(24)

\[
E\left(\frac{\epsilon_{t+h}^{2}}{b} - b^{2}\mu_{f}\right)\left(\epsilon_{t+h}^{2} / b - \sigma_{t}^{2}\right)\left(I_{(t+h),t}\right) = i
\]

\[
= E\left(\frac{\epsilon_{t+h}^{2}}{b} - b^{2}\mu_{f}\right)\left(\epsilon_{t+h}^{2} / b - \sigma_{t}^{2}\right)\left(I_{(t+h),t}\right) - \sqrt{b}\frac{\alpha_{f}}{b} \left(b_{f}\mu_{f}^{3} - \mu_{f}\sigma_{t}^{2} - \sigma_{t}^{2}\right)
\]

\[
= b\frac{\alpha_{f}}{b} \left(b_{f}\mu_{f}^{3} + 3b_{f}\mu_{f}\sigma_{t}^{2} - b_{f}\mu_{f}\sigma_{t}^{2} - \mu_{f}\sigma_{t}^{2}\right) - \sqrt{b}\frac{\alpha_{f}}{b} \left(b_{f}\mu_{f}^{3} + \sigma_{t}^{2} - \sigma_{t}^{2}\right) = o(1)
\]

We also verified that the third and fourth moments of the NM(K)-GARCH(1,1) process exist and are finite when \( b \to 0 \) (see Appendix A). Now the proof follows from the main convergence results of stochastic difference equations to differential equations, derived in section 2.1 of Nelson (1990), or from the theorems for weak convergence of discrete time Markov chains to continuous processes applied as in Corradi (2000). □

From Theorem 1 we know that each of the deterministic variance components in (19) converges to a constant steady state forward variance, given by:

\[
\sigma_{i}^{2} = E(\sigma_{i}^{2}) = \frac{\omega_{j} + \psi_{i}\bar{\sigma}}{\theta_{i}}
\]

(25)

where \( \theta_{i} \) is the speed of mean reversion. This follows on setting \( \sigma_{i}^{2} := \lim_{b \to 0} \gamma_{i} \) with

\[
\gamma_{i} := \lim_{b \to 0} \sigma_{i}^{2} = \frac{b^{2}\omega_{j} + b\alpha_{f}\bar{\sigma}}{1 - b^{2}\beta_{f}} = \frac{b^{2}\omega_{j} + b\alpha_{f}\bar{\sigma}}{1 - b^{2}\beta_{f}}
\]

(26)

The overall steady state forward variance is given by:

\[
\bar{\sigma}^{2} = E(\sigma^{2}) = \sum_{i=1}^{K} \frac{p_{i}\omega_{j}}{\theta_{i}} \left(1 - \sum_{i=1}^{K} \frac{p_{i}\phi_{i}}{\theta_{i}} \right)
\]

(27)

To see this, define

\[
\sigma_{i}^{2} := \lim_{b \to \infty} \sigma_{i}^{2} = \frac{\sum_{i=1}^{K} p_{i}b_{i}\mu_{i}^{2} + \sum_{i=1}^{K} p_{i}b_{i}\bar{\sigma}}{1 - \sum_{i=1}^{K} (1 - b^{2}\beta_{i})} = \frac{\sum_{i=1}^{K} p_{i}b_{i}\mu_{i}^{2} + \sum_{i=1}^{K} p_{i}b_{i}\bar{\sigma}}{1 - \sum_{i=1}^{K} (1 - b^{2}\beta_{i})} = \frac{\sum_{i=1}^{K} p_{i}b_{i}\mu_{i}^{2} + \sum_{i=1}^{K} p_{i}b_{i}\bar{\sigma}}{1 - \sum_{i=1}^{K} (1 - b^{2}\beta_{i})}
\]

(28)

Then
The relationship between the overall and individual unconditional variances is:

\[ \sigma^2 = \sum_{i=1}^{K} \sigma_i^2 \]

Also, as already noted, the skewness and excess kurtosis of the discrete model converge towards a finite value (see Appendix A).

The \( T \)-period variances of each component are given by:

\[ \sigma_i^2(T) = \omega_i T + \varphi_i \int_{t=0}^{T} \sigma_i^2(t) - \theta_i \int_{t=0}^{T} \sigma_i^2(t) \]

and the overall \( T \)-period variance in the model is:

\[ \sigma^2(T) = \sum_{i=1}^{K} p_i \sigma_i^2(T) \]

**Corollary 1:** In the mixture GARCH option pricing model the marginal price density is lognormal.

The proof can be found in Appendix B.

Theorem 1 and its corollary can be extended to normal mixture GARCH models with asymmetric variance components given by either the AGARCH model of Engle and Ng (1993), where now

\[ \sigma_i^2 = \omega_i + \alpha_i (z_{i,t-1} - \lambda_i)^2 + \beta_i \sigma_{i,t-1}^2 \]

or the GJR-GARCH model of Glosten et al (1993) where the variance components are specified as

\[ \sigma_i^2 = \omega_i + \alpha_i z_{i,t-1}^2 + \lambda_i d_{i,t-1} z_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2 \]

where \( d_i = 1 \) if \( z_i < 0 \), and 0 otherwise. Interestingly, only the AGARCH components have leverage effects in the limit. The continuous limit of the normal mixture GARCH with GJR variance components is similar to the model given in Theorem 1, as shown by the following theorem:

**Theorem 2:** The continuous time limit of the normal mixture AGARCH process (8) – (12) with (9) replaced by

\[ \alpha_{i,k,k+1} = \omega_i + \alpha_i \left( \frac{a_{ik}}{b} - \lambda_i \right)^2 + \beta_i \sigma_{i,k,k}^2 \]

has, with probability \( p_i \):
(36) \[
\frac{dS(t)}{S(t)} = (\mu + \mu_i) \, dt + \sigma_i(t) \, dB(t)
\]

where

(37) \[
d\sigma_i^2(t) = \left( \omega_i + \psi_i \sigma_i^2(t) + \psi_i \sigma_i^2(t) - 0_i \sigma_i^2(t) \right) \, dt,
\]

(38) \[
\sigma^2(t) = \sum_{i=1}^{K} p_i \sigma_i^2(t)
\]

and \( B(t) \) is a Brownian motion under the physical measure. For this limit to exist, condition (17) and the additional condition \( \psi_i \to \psi_i \) must be satisfied.

**Proof:** Similar to the proof of Theorem 1.

The speed of mean-reversion remains unchanged but the unconditional variance to which the component variances mean-revert to is:

(39) \[
\bar{\sigma}^2 = \frac{\omega_i + \psi_i \sigma_i^2 + \psi_i \sigma_i^2}{0_i}
\]

The overall steady state forward variance changes to:

(40) \[
\bar{\sigma}^2 = \frac{\sum_{i=1}^{K} p_i \left( \omega_i + \psi_i \sigma_i^2 \right)}{1 - \sum_{i=1}^{K} p_i \psi_i}
\]

Note that in the case of GJR-GARCH variance components, where we write

(41) \[
_b \sigma^2_{i, k \neq i, b} = \omega_i + \alpha_i \varepsilon_t^2 / b + \lambda_i \varepsilon_{t-k}^2 / b + \beta_i \sigma^2_{i, k \neq i, b} \quad i = 1, \ldots, K
\]

with \( \varepsilon_t = 1 \) if \( \varepsilon_t < 0 \), and 0 otherwise, \( \varepsilon_t \) is finite but does not converge when \( b \downarrow 0 \). Hence the following needs to be added to the condition set (17):

(42) \[
0 = \lim_{b \downarrow 0} \left( \frac{\lambda_i}{b} \right)
\]

With this extended condition, Theorem 1 remains unchanged.

### III The continuous limit of normal GARCH processes

The limiting assumptions that can be made in the normal GARCH case are somewhat arbitrary. As noted by Corradi (2000) there are few parameter constraints, and hence some flexibility about the limiting assumptions one can choose. But when one attempts to widen the family of GARCH processes, so that instead of a single
normal GARCH variance component there are several normal GARCH variance components, the possibilities for our limiting assumptions become more restricted. In fact, the limiting assumptions made by Nelson cannot be generalised to the normal mixture case. In particular it is no longer possible to assume:

\[
\lim_{b \downarrow 0} \left( b^{-1} \sigma^2_b \right) \text{ exists and is finite and strictly positive.}
\]

However, it is exactly this assumption that introduces a non-zero diffusion coefficient in the ‘GARCH diffusion’.

Consider the case of Theorem 1 with \( K = 1 \), i.e. there is only one normal variance process

\[
\sigma^2_t = \omega + \alpha \sigma^2_{t-1} + \beta \sigma^2_{t-1}
\]

Under the assumptions (17), which may be written as:

\[
\lim_{b \downarrow 0} \frac{\sigma^2_b}{b} = \omega \quad \lim_{b \downarrow 0} \frac{\alpha}{b} = \psi \quad \lim_{b \downarrow 0} \frac{\beta}{b} = \theta
\]

our main theorem shows that the continuous time counterpart of the normal GARCH(1,1) process is the deterministic local volatility model.

\[
\frac{dS(t)}{S(t)} = \mu \, dt + \sigma(t) dB(t)
\]

\[
d\sigma^2(t) = \left( \omega - (\theta - \psi) \sigma^2(t) \right) dt
\]

Clearly, \( T \)-maturity European option prices under (48) are equal to the Black-Scholes prices based on

\[
\sigma^2(T) = \omega T - (\theta - \psi) \int_0^T \sigma^2(t) \, dt
\]

From above we also know that the same limit applies to the GJR-GARCH model, although the A-GARCH model has continuous time variance process

\[
\frac{d\sigma^2(t)}{\sigma^2(t)} = \left( \omega + \psi \sigma^2 - (\theta - \psi) \sigma^2(t) \right) dt
\]

assuming \( \psi = \lim_{b \downarrow 0} \lambda \).

---

2 Note that with only one variance component to consider, Nelson uses the notation \( \alpha_i \) for the time-step dependent parameter – equivalent to \( \alpha_i \) with \( i = 1 \) in our notation.

3 Local volatility models were introduced by Dupire (1994) and have been developed many others since. The assumption of these models is that the forward variance is a deterministic function of time and the underlying asset price.
IV Time Aggregation of Normal Mixture GARCH Models

When the continuous time limit of the model is under discussion, time aggregation of a time variant model naturally comes into question: is the model keeping its properties when analyzed under a different frequency? Drost and Nijman (1993) have studied this property extensively, and showed that the normal GARCH process is not time aggregating. They subsequently defined a ‘weak’ normal GARCH process that is invariant under changing the length of the time interval of observations. Their results have no straightforward generalization to the normal mixture GARCH process.

A trick is needed to take care of the dependence of the conditional variance on all previous disturbances. Taking expectations in (3) shows that the total unconditional variance in the model can be expressed as a sum of two terms: the mean of the unconditional variances plus the variance of the conditional means of the components (under the discrete mixing law distribution):

\[
E\left(\sigma_i^2\right) = \sum_{i=1}^{K} p_i E\left(\sigma_i^2\right) + \sum_{i=1}^{K} p_i \mu_i^2
\]

In this decomposition of the total variance we are interested in the relative contribution of each component variance, and to this end we define:

\[
q_i = E\left(\sigma_i^2\right) / E\left(\sigma_i^2\right)
\]

\[
\eta_i = q_i \sigma_i^2 - \sigma^2
\]

It is easy to see that \(E(\eta_i) = 0\). Re-writing (53) in the form

\[
\eta_i = q_i \left(\varepsilon_i^2 - \sigma_i^2\right) + \left(q_i \sigma_i^2 - \sigma^2\right)
\]

we see that \(\eta_i\) captures two time-varying effects for each variance component: the deviation of the squared return from the overall conditional variance, and the difference between the conditional variance of the component and a proportion \(q_i\) of the overall conditional variance.

**Definition:** A random variable \(Y_t\) is said to follow a weak NM(K)-GARCH process if, in addition to (1), (2), (3), (4), (52) and (53) the following properties hold:

\[
P(s_i = i) = p_i
\]

\[
E(\varepsilon_i | I_{t-1}, s_i = i) = \mu_i
\]

Note that the weak NM(K)-GARCH(1,1) process is not an exact generalization of the strong NM(K)-GARCH(1,1) process. It can be shown that under (5) some, but not all the correlations defined in (56) and (57) are exactly zero. However, those that are not zero are extremely small (less than 0.01 and 0.0005 respectively) and we believe they tend to zero as the sample size tends to infinity, although this has yet to be proved.
Based on time steps of length $h$, the weak NM(K)-GARCH process definition can be rewritten in terms of annualized variances, as:

\[ q_i = E\left( \sigma^2_{h,k,i} \right)/E\left( \sigma^2_{h,k,i} \right) \]

\[ \eta_{h,k,i} = q_i \cdot h^{-1} \cdot \eta^2_{h,k} - h^{-1} \cdot \sigma^2_{h,k} \]

\[ E\left( \eta^2_{h,k} | I_{(k-1)h} \right) = h \cdot k \cdot \mu_i \]

\[ Corr\left( \eta_{h,k,i}, \eta_{h,k-(j-1)h} \right) = 0, \quad j > 0 \]

\[ Corr\left( \eta_{h,k,i}, \eta_{h,k-(j-1)h} \right) = 0, \quad j > 0 \]

\[ \left| E\left( \epsilon^3_{(k-1)h} | I_{(k-1)h} \right) \right| < \infty \quad \text{and} \quad \left| E\left( \epsilon^1_{(k-1)h} | I_{(k-1)h} \right) \right| < \infty \]

in addition to (8), (9), (10) and (11).

**Theorem 3:** The class of weak NM(K)–GARCH processes is closed under temporal aggregation. After annualisation, the parameters of $\sigma^2_{h,k+1,h}$ will be given by:

\[ \beta_i = \frac{1 - \left( \frac{\alpha_i}{q_i} + \beta_i \right)^h}{1 - \left( \frac{\alpha_i}{q_i} + \beta_i \right)} \]

\[ \beta_i = \frac{1 - \left( \frac{\alpha_i}{q_i} + \beta_i \right)^h}{1 - \left( \frac{\alpha_i}{q_i} + \beta_i \right)} \]

and $\beta_i \in (0,1)$ is the solution to

\[ \frac{\beta_i}{1 + \beta_i^2} = -Corr\left( v_{i,k+1,h}, v_{i,(k-1)h} \right) \]

where

\[ v_{i,k+1,h} = \frac{1}{2} \sum_{p=0}^{h-1} \min(p,h) \cdot \eta_{i,k+1,h-p} + \sum_{0 \leq r < h} \eta_{i,k+1,h-r} - \frac{\alpha_i}{q_i} - \beta_i \sum_{0 \leq r < h} \left( \sum_{0 \leq r < h} \left( \eta_{i,k+1,h-r} - \eta_{i,(k-1)h-r} \right) \right) \]
and 

$$
\epsilon_k = \begin{cases} 
1 & \text{if } k = 0 \\
\frac{\alpha_i}{q_i} \left( \frac{\beta_i}{q_i} + \right)^{k-1} & \text{if } 1 \leq k \leq b - 1 \\
-\beta_i \left( \frac{\alpha_i}{q_i} + \right)^{k-1} & \text{if } k = b 
\end{cases}
$$

Also,

$$\epsilon_i \mu_i = \mu_i \quad \text{and} \quad \epsilon_i \rho_i = \rho_i$$

As the proof is rather long we put this in an Appendix. Our next result shows that the weak NM(K)-GARCH process has a continuous time limit with a single deterministic volatility component:

**Theorem 4.** The continuous-time limit of the weak normal mixture GARCH process under the physical measure $P$ has, with probability $p_i$:

\[
\frac{dS(t)}{S(t)} = \left( \mu_i + \mu_j \right) dt + q(t) dB(t)
\]

where

\[
dq^2(t) = \left( \sum_{i=1}^{K} p_i \left( \omega_i + \psi_i q_i^2(t) + 0, \sigma_i^2(t) \right) + c(t) \right) dt
\]

\[
c(t) = \lim_{b \to 0} b \epsilon_i
\]

exists and is finite, given $b \epsilon_i = \epsilon_{kb}$ for $kb \leq t < (k+1)b$,

\[
\epsilon_{kb} = E \left( \frac{b_{k+1} - b_k}{b} \right)_{I_{k+1}b}^{I_{k+1}b}
\]

\[
b_{kb} = q_{kb}^2 - \sigma_{kb}^2 \quad \text{where} \quad E(b_{kb}) = 0, \quad b_{kb} > -\sigma_{kb}^2 \quad \text{for } k > 0
\]

and $B(t)$ is a Brownian motion. For this limit to exist, condition (17) must be satisfied.

Again the proof is left to an Appendix. Theorem 4 proves the interesting result that in the limit of the weak normal mixture GARCH process there is no uncertainty over the variance, only over the drift.
V The continuous limit of normal mixture GARCH processes with time-varying probabilities

This model is an extension of NM-GARCH models so that the state probabilities are time-varying. This way, in addition to (8), (9), (10) and (12) we define the time-varying probabilities and their expectations as:

\[ b_i \pi_i := E \left( b_i \pi_i \right), \quad \sum_{i=1}^{K} b_i \pi_i = 1; \quad b_i \pi_i := \left( b_i \pi_i \right) \]

The information-based expectation of the time-dependent state-probabilities can be written as:

\[ \varphi_{ij}(b_k, \varepsilon_k) := \frac{1}{b_i \sigma_{i,k} \sqrt{2\pi}} \exp \left( -\frac{(\varepsilon_k - \mu_i)^2}{2 b_i \sigma_{i,k}^2} \right) \]

In addition to the continuous limits given in (13), (14), (15) and (16) we define the limit of state-probability and its conditional expectation:

\[ \pi_i(t) := \lim_{b \to 0} b_i \pi_i, \quad \hat{\pi}_i(t) := \lim_{b \to 0} b_i \hat{\pi}_i \]

The set of conditions that the parameters must satisfy is:

\[ \omega_i := \lim_{b \to 0} \left( \frac{\omega_i}{b} \right); \quad \psi_i := \lim_{b \to 0} \left( \frac{\sigma_i}{b} \right); \quad \theta_i := \lim_{b \to 0} \left( \frac{1 - \beta_i}{b} \right); \quad \pi_i := \lim_{b \to 0} \left( b_i \pi_i \right); \quad \mu_i := \lim_{b \to 0} \left( b_i \mu_i \right) \]

with \( \psi_i > 0, \theta_i > 0, \mu_i > 0, \) and \( \sum_{i=1}^{K} \pi_i = 1. \)

**Theorem 5:** If condition (75) is satisfied, the continuous time limit of the normal mixture GARCH process with time-varying probabilities given by equations (8), (9), (10), (12), (70), (71) and (72) has, with probability \( \pi_i(t) : \)

\[ \frac{dS(t)}{S(t)} = (\mu + \mu_i) \ dt + \sigma_i(t) dB(t) \]

where

\[ d\sigma^2_i(t) = \left( \omega_i + \psi_i \sigma^2_i(t) - \theta_i \sigma^2_i(t) \right) dt, \quad \sigma^2_i(t) = \sum_{i=1}^{K} \hat{\pi}_i(t) \sigma^2_i(t) \]

\[ \hat{\pi}_i(t) = \frac{\pi_i \varphi_i(x(t))}{\sum_{i=1}^{K} \pi_j \varphi_j(x(t))} \quad \text{with} \quad x(t) = \mu_i dt + \sigma_i(t) dB(t) \]

\[ \varphi_i(x(t)) := \frac{1}{\sigma_i(t) \sqrt{2\pi}} \exp \left( -\frac{(x(t) - \mu_i dt)^2}{2 \sigma^2_i(t)} \right) \]
Proof: Equations (20), (21), (22) and (24) can be used, and the only change in equation (23) is the substitution of \( \hat{p}_m \) with \( \hat{p}_m(t) \).

It can be seen that this model is a restricted form of Markov Switching GARCH process (discussed in the next section) where the transition matrix can be expressed as:

\[ \hat{Q} = \pi \cdot 1' \]

This way, the results from section VI can be applied to the normal mixture GARCH model with time-dependent state probabilities. Notably, Corollary 1 from section II applies in this case as well.

VI Markov Switching GARCH Processes

The Markov Switching GARCH is an extension of the NM-GARCH model with time-varying parameters, where the transition matrix is not singular. The model is given by the following equations (in addition to (8), (9), (10) and (12)):

\[ q_{i,j} = P(h_s = i | h_s = j, h_s = k-1) = P(h_s = i | h_s = j, h_s = k-2) = I, \ldots) \]

\[ \hat{Q} := \left( q_{i,j} \right)_{i,j} \text{ with } \sum_{i=1}^{K} q_{i,j} = 1 \quad \text{and} \quad q_{i,j} \geq 0 \]

where \( \hat{Q} \) is the transition matrix and we define \( \pi \) the vector of unconditional probabilities so that:

\[ \pi := E(h_s | \hat{p}_h) \]

It can be shown that \( \pi \) does not depend on the time step. It is also true that:

\[ \lim_{h \to \infty} Q = \pi \cdot 1' \quad \text{where} \ 1 \ \text{is a vector of ones} \]

Furthermore, we have:

\[ \lim_{n \to \infty} Q^n = \pi \cdot 1' \]

\[ \hat{Q} \cdot \pi = \pi \]

We use the following notations:

\[ \hat{p}_{hh} := E(h_s | I_{k-1} | h_s = i, I_{k-1} | h_s = j) = E(h_s | I_{k-1} | h_s = i, I_{k-1} | h_s = j) = E(h_s | I_{k-1} | h_s = i, I_{k-1} | h_s = j) \]

Based on (80) we write the following updating formulae:
This second equation gives the time-varying series of state-probabilities that is of most interest. Here $\otimes$ denotes element-by-element multiplication and $\varphi$ is the vector of normal density functions (different rows specify different states).

We assume that the limits given in equations (13) – (16), (73) and (74) exist; additionally, we define:

$$\tilde{p}_i(t) := \lim_{b \to 0} \tilde{p}_{i,b} \quad \text{where} \quad \tilde{p}_{i,b} := \frac{1}{b} \tilde{P}_{i,b} \quad \text{for} \quad kb \leq t < (k+1)b; \quad \tilde{p}(t) = (\tilde{p}_i(t))_i$$

In this case condition (17) can be rewritten as:

$$\omega_i := \lim_{b \to 0} \left( \frac{b \omega_i}{b} \right); \quad \psi_j := \lim_{b \to 0} \left( \frac{b \psi_j}{b} \right); \quad \theta_j := \lim_{b \to 0} \left( \frac{1-b \theta_j}{b} \right); \quad Q := \lim_{b \to 0} \left( \frac{b Q - I}{b} \right); \quad \mu_i := \lim_{b \to 0} \left( b \mu_i \right)$$

with $\omega_i \geq 0$, $\theta_j > 0$, $q_{i,j} > 0$ for $i \neq j$ and $\sum_{j=1}^{K} q_{i,j} = 0$.

$Q$ is called the transition rate matrix or generator matrix and it can also be expressed as:

$$Q = \frac{d}{db} Q \big|_{b=0}$$

Also, we have that:

$$\frac{d}{db} Q = Q - Q \cdot Q$$

and:

$$Q \cdot \pi = 0 \quad \text{where} \quad 0 \text{ is a vector of zeros}$$

**Theorem 6:** The continuous time limit of the Markov switching GARCH process given by (8), (9), (10), (12) and (80) – (89) is a model where with probability $\hat{p}_i(t)$ we have that:

$$\frac{dS(t)}{S(t)} = \left( \mu_i + \mu_j \right) dt + \sigma_i(t) dB(t)$$

$$\frac{d\sigma_i^2(t)}{\sigma_i^2(t)} = \left( \omega_i + \psi_j \sigma_j^2(t) - \theta_j \sigma_i^2(t) \right) dt, \quad \sigma_i^2(t) = \sum_{j=1}^{K} p_{i,j}(t) \sigma_j^2(t)$$

where

$$d\tilde{p}(t) = Q \cdot \tilde{p}(t) dt$$
\[ \dot{p}_i(t) = \frac{\pi_i \varphi_i(\alpha(t))}{\sum_{j=1}^n \pi_j \varphi_j(\alpha(t))} ; \quad \text{with} \quad \alpha(t) = \mu_i \, dt + \sigma_i(t) dB(t) \quad \text{and} \]
\[ \varphi_i(\alpha(t)) := \frac{1}{\sigma_i(t) \sqrt{2\pi}} \exp \left( -\frac{(\alpha(t) - \mu_i dt)^2}{2\sigma_i^2(t)} \right) \]

For this limit to exist, condition (91) must be satisfied. In this case, \( s(t) \) is a continuous time Markov chain.

**Proof:** For the drift part, we use equations (17) and (18). For the variance we can write (dropping the pre-subscript \( h \)):

\[
E(\dot{\beta}^{-1}(\sigma_{j,k}^{2} - \sigma_{j,k}^{2}) \mid I_{(k-1),j}) = \beta \sigma_{j,k}^{2} + \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2} \\
= \beta \sigma_{j,k}^{2} + \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2}
\]

The variances and covariances of the individual variances are again not significant:

\[
E \left( \sigma_{j,k}^{2} - \sigma_{j,k}^{2} \mid I_{(k-1),j} \right) = \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2} + \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2}
\]

Similarly to equation (21), we obtain that the covariances between the returns and the changes in the individual variances are small as well.

To derive the process for the probability parameter, we write based on (88) that:

\[
E(\dot{\beta}^{-1}(\beta_{j,k}^{2} - \beta_{j,k}^{2}) \mid I_{(k-1),j}) = \beta \sigma_{j,k}^{2} + \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2} = o(1)
\]

Similarly to equation (21), we obtain that the covariances between the returns and the changes in the individual variances are small as well.

To derive the process for the probability parameter, we write based on (88) that:

\[
E(\dot{\beta}^{-1}(p_{j,k} - p_{j,k}^{2}) \mid I_{(k-1),j}) = \beta \sigma_{j,k}^{2} + \frac{1}{b} \frac{\beta_{j} - 1}{b} \sigma_{j,k}^{2} = o(1)
\]

where \( e_i \) is the \( i^{th} \) row of the identity matrix \( I \) and the subscript \( i \) for \( Q \) symbolizes the \( i^{th} \) row of the matrix. □
**Properties:** If the continuous time Markov chain \( \{s(t)\} \) has holding times \( \{H_k\}_k \) (i.e. the duration of the time periods spent in a state) and jump times \( \{J_k\}_k \), (i.e. the points in time at which the chain switches to a different state) where \( J_k = \sum_{i=1}^{k} H_i \) and \( J_0 = 0 \), then we have:

(a) The holding times \( \{H_k\}_k \) are independent and follow an exponential distribution:

\[
P(\text{the chain jumps in an interval of length } t) = P(H_k \leq t) = 1 - \exp(-\lambda t)
\]

(b) The jump times \( \{J_k\}_k \) follow a Poisson distribution:

\[
P(k \text{ jumps in an interval of length } t) = P(J_k \leq t) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t) \quad \text{where } \lambda = \sum_{i \neq i} q_{ij}
\]

**VII Discussion**

A considerable body of research in mathematical finance concerns the use of mixture terminal price densities for pricing European options. For example, Ritchey (1990), Mellick and Thomas (1997), Brigo and Mercurio (2001) and others adopt a deterministic approach to the normal mixture model, corresponding to a behavioural view of financial markets that was introduced by Epps and Epps (1976). Here traders in options on a given underlying hold different expectations about the future volatility of the underlying. As they are certain of their own views there is no additional source of randomness in the model and the market price merely represents an equilibrium price based on the traders’ heterogeneous price expectations.

This behavioural view of demand and supply can explain the characteristics of option prices and the implied volatility skew in particular. Buyers of out-of-the-money put options on an equity index are risk-averse investors that fear a general market decline; as such options provide an attractive form of insurance. Sellers of these options are the market makers that hold less pessimistic views about volatility. The market markets are typically in relatively short supply, hence the market prices of out-of-the-money put options are often far higher than the Black-Scholes (1973) model prices based on the at-the-money volatility. This may be why we observe a more pronounced implied volatility skew when investor’s fear of a general market decline increases following crash periods.

In these deterministic ‘local volatility’ mixture models the mixing law \((p_1, \ldots, p_K)\) represents the relative frequencies of the different types of traders and hence are simply regarded as parameters of the price density, just as the means and the variances of the components. Hence European option prices and hedge ratios are constructed in a complete market as simple weighted averages of Black-Scholes prices and hedge ratios, based

---

6 Brigo and Mercurio (2001, 2002) showed that if the price density is a lognormal mixture with deterministic variance components that satisfy certain continuity and regularity conditions, then the price process is a geometric Brownian with a local volatility function. Equivalently, one can assume the log price density is a normal mixture as the mixing law is invariant under a monotonic transformation of the random variable. See Alexander and Nogueira (2004) for more details.
on the average volatilities in each component and with weights defined by the mixing law. However, the mixture GARCH option pricing model is not deterministic.

Instead, the mixture GARCH models derived in Section II are in the general class of mixture models for option pricing described in Garcia, Ghysels and Renault (2005). The latent variable $s_t \in [1, 2, \ldots, K]$ denotes the ‘state of nature’ at time $t$ with each regime corresponding to a different GARCH(1,1) variance process for generating the returns. Thus the behavioural justification of the mixture GARCH model rests on heterogeneous volatility regimes with, for instance, different types of mean-reversion rather than heterogeneous traders. The price transition densities are normal mixtures where the mixing law gives the weights of these states. The market is incomplete due to the uncertainty about the regime that governs future returns and option prices may reflect investor’s preferences, in which case a non-zero volatility risk premium will enter the option price.

Note that mixture GARCH models are not equivalent to the well-known uncertain volatility mixture models of Brigo (2002), Mercurio (2002) and Brigo, Mercurio and Rapisada (2004). There an uncertain volatility representation of the complete markets lognormal mixture diffusion is employed, where all uncertainty is resolved after an infinitesimally small time interval. In a similar vein, Garcia, Luger and Renault (2003) provide an empirical analysis of a mixture model in which the state variable is also observed after one period. In these models the transition densities after the first time period are normal, as after the first period the volatility regime is revealed, thereafter the agents have no uncertainty about the future regime. This assumption has been criticised (for instance, by Pieterbarg) as an artificial construct with no rational foundation.

It is important to specify both the state of nature and the state of knowledge for the mixture GARCH model. If the state of nature is that only one volatility regime can occur but agents just don’t know which one it is at time zero, then agents must be assumed to learn nothing about this state. That is, the state of knowledge at any time $t$ is determined by the mixing law $(p_1, \ldots, p_K)$ and this does not evolve at all. This is entirely unrealistic – and we have jumped to completely the opposite extreme now in terms of agents learning. Obviously at time zero an agent will know that he can observe prices at $t > 0$. If only one state can occur then he also knows that observing prices will give information about that state. Hence we cannot, realistically, assume that only one state can occur, because the model then implies that there is no learning at all about that state.

Instead, the state of nature at time $t$ is that any one of the volatility regimes can occur. In this case the mixing law represents the relative frequency of each volatility regime over a long period of time. Whilst the state of knowledge at time zero is represented by the mixing law, as the underlying asset price evolves in the future there will be a non-zero probability of switching from one regime to another. In fact the conditional regime probabilities are time-varying and it is these that represent the state of knowledge at time $t > 0$. 

Copyright © 2004. Alexander and Lazar. All Rights Reserved.
The Markov Switching GARCH model of Haas, Mittnik and Paolella (2004b) extends the normal mixture GARCH model to include (constant) state transition probabilities. However, in the case of just two volatility regimes the ratio of the two unconditional regime probabilities that are estimated from the normal mixture GARCH will determine the ratio of the switching transition probabilities, as a standard equilibrium argument shows. Since an option pricing model does not (usually) require that we specify agent’s state of knowledge, we do not actually need to know the time-varying conditional regime probabilities, which is good because these are notoriously difficult to estimate. This simplification should help our progress in a fruitful but difficult area for future research.

**VIII Summary and Conclusion**

It has been known for some time that normal GARCH models are not as good for fitting the conditional dynamics of returns in most financial markets as GARCH variance processes with non-normal conditional distributions. One of the important properties to capture is time variation in the higher conditional moments of financial returns and for this we require the GARCH process to be based on at least two volatility states. A large body of empirical research on normal mixture GARCH models now spans more than a decade, and one of the conclusions is that normal mixture GARCH processes provide a more intuitive model of, and a closer fit to, the unconditional physical returns density than the normal and the symmetric and skewed \( t \)-GARCH models even when leverage effects are included in the variance process.

Our main result is that the variance dynamics in the continuous limit of the symmetric and asymmetric normal mixture GARCH variance processes are not governed by standard diffusion processes. The same result applies to normal mixture GARCH models with time-varying regime probabilities and to Markov switching GARCH models. It was not possible to generalize the limiting assumption for the GARCH alpha parameter that is necessary for the normal GARCH(1,1) process to have a mean-reverting diffusion process as its continuous time limit. When there are two or more variance components, the limiting conditions on the parameters become much less flexible and one is forced to assume that the diffusion coefficient is zero for each of the variance components in order for a continuous limit to exist. The continuous limit of normal mixture GARCH is a stochastic volatility model, but not in the traditional sense.

We also addressed the question of time aggregation. Since the strong NM(\( K \))-GARCH(1,1) model is not aggregating in time, we gave a weak definition of this model that is not sensitive to the choice of step length. The weak definition leads to single deterministic variance process but there is uncertainty about the drift.

Option pricing with normal mixture and Markov Switching GARCH processes is likely to be an important area for future research. These processes can be consistent with a volatility surface that has smile and skew effects.
that change with maturity but still persist longer than the central limit theorem would imply, as the conditional
skewness and kurtosis are time varying and the unconditional skewness and kurtosis is non-zero. Very few
other volatility models are able to capture such an effect. Hence the explanation of the observed behaviour of
implied volatility does not necessarily rest on the existence of a time-varying volatility risk premium, as it does
in other GARCH models. An important result in this paper is that there is no justification to introduce a
volatility risk premium to normal GARCH option pricing models anyway. Their continuous limit is more
naturally a deterministic volatility model.
Appendix A. The convergence of the Higher Moments of the NM(K)-GARCH(1,1) Model

The higher moments of the NM(K)-GARCH(1,1) model defined in (9) – (13) are given in Alexander and Lazar (2005): The unconditional skewness is:

\[
\sum_{i=1}^{K} p_i \left( 3 b \omega_i + 3 b \omega_i \right) / b^{3/2}.
\]

and the unconditional kurtosis may be expressed as:

\[
\left( \frac{3 b p' B^{-1} f + b \beta}{1 - 3 b p' B^{-1} f g} \right) / b^{1/2} \quad \text{where } b \mathbf{p} = (\beta_1, \ldots, \beta_K)'
\]

where\[ b \mathbf{B} = \begin{bmatrix}
1 - b \beta_1^2 - 2 b \alpha_1 b \beta_1 \epsilon_{11} & -2 b \alpha_1 b \beta_1 \epsilon_{12} & \cdots & -2 b \alpha_1 b \beta_1 \epsilon_{1K} \\
-2 b \alpha_2 b \beta_2 \epsilon_{21} & 1 - b \beta_2^2 - 2 b \alpha_2 b \beta_2 \epsilon_{22} & \cdots & -2 b \alpha_2 b \beta_2 \epsilon_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
-2 b \alpha_K b \beta_K \epsilon_{K1} & -2 b \alpha_K b \beta_K \epsilon_{K2} & \cdots & 1 - b \beta_K^2 - 2 b \alpha_K b \beta_K \epsilon_{KK}
\end{bmatrix}
\]

\[ b \mathbf{f} = \begin{bmatrix}
 b \omega_1 + 2 b \alpha_1 b \beta_1 \epsilon_{11} \\
\vdots \\
b \omega_K + 2 b \alpha_K b \beta_K \epsilon_{K1}
\end{bmatrix}
\]

\[ b \mathbf{g} = \begin{bmatrix}
 b \alpha_1^2 + 2 b \alpha_1 b \beta_1 d_1 \\
\vdots \\
b \alpha_K^2 + 2 b \alpha_K b \beta_K d_K
\end{bmatrix}
\]

\[ b \mathbf{A} = \begin{bmatrix}
1 - \sum_{k=1}^{K} \frac{b \delta_k b \beta_{1, k} \alpha_k}{1 - b \beta_{1, k} b \beta_k} & -b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{1, k} \beta_2}{1 - b \beta_{1, k} b \beta_k} & \cdots & -b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{1, k} \beta_K}{1 - b \beta_{1, k} b \beta_K} \\
-b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{2, k} \alpha_k}{1 - b \beta_{2, k} b \beta_1} & 1 - \sum_{k=1}^{K} \frac{b \delta_k b \beta_{2, k} \beta_2}{1 - b \beta_{2, k} b \beta_k} & \cdots & -b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{2, k} \beta_K}{1 - b \beta_{2, k} b \beta_K} \\
\vdots & \vdots & \ddots & \vdots \\
-b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{K, k} \alpha_k}{1 - b \beta_{K, k} b \beta_1} & -b \sum_{k=1}^{K} \frac{b \delta_k b \beta_{K, k} \beta_2}{1 - b \beta_{K, k} b \beta_k} & \cdots & 1 - \sum_{k=1}^{K} \frac{b \delta_k b \beta_{K, k} \beta_K}{1 - b \beta_{K, k} b \beta_K}
\end{bmatrix}
\]

\[ b \mathbf{q} = \begin{bmatrix}
 b \omega_j + b \omega_k + b \omega_j b \omega_k + b \omega_j b \omega_k \\
\vdots \\
b \omega_j + b \omega_k + b \omega_j b \omega_k + b \omega_j b \omega_k
\end{bmatrix}
\]

We can write the following limits:

\[ \lim_{b \rightarrow 0} b \beta_j = 1 \]

Copyright © 2004. Alexander and Lazar. All Rights Reserved.
\[
\lim_{b \to 0} \frac{1 - \beta_{i \cdot k} \beta_{j \cdot k}}{b} = 0, + 0 \]

\[ P = \lim_{b \to 0} P = (p_1, \ldots, p_K)'
\]

\[ \lim_{b \to 0} d_i = 0 \]

\[ \lim_{b \to 0} g / b = 0 \]

\[ w_j = \lim_{b \to 0} w_j / b = 2\omega_j \sigma_j^2 \]

\[ q = \lim_{b \to 0} q = \sum_{i=1}^{K} p_i \mu_i^2 \]

\[ r_{ik} = \lim_{b \to 0} r_{ik} / b = \sigma_j^2 \omega_k + \sigma_k^2 \omega_i \]

\[ A = (a_{ij}) = \lim_{b \to 0} A = \left[ \begin{array}{cccc}
1 - \sum_{k=0}^{K} p_k \psi_k & -p_2 \psi_1 & \cdots & -p_K \psi_1 \\
-\frac{p_1 \psi_2}{1 + \mu_2} & -\sum_{k=2}^{K} p_k \psi_k & \cdots & -\frac{p_K \psi_2}{1 + \mu_2} \\
\vdots & \vdots & \ddots & \vdots \\
-\frac{p_1 \psi_K}{1 + \mu_K} & -\frac{p_2 \psi_K}{1 + \mu_K} & \cdots & -\sum_{k=K}^{K} p_k \psi_k
\end{array} \right] \]

\[ \epsilon_i = \lim_{b \to 0} \epsilon_i = \sum_{j=1}^{K} a_{ij} \left[ \sum_{k \neq j} \frac{p_k \psi_{jk}}{1 + \mu_k} + \bar{\sigma}_j \right] \]

\[ f = \lim_{b \to 0} f / b = \left( \begin{array}{c}
w_1 + 2\psi_1 \epsilon_1 \\
\vdots \\
w_K + 2\psi_K \epsilon_K
\end{array} \right) \]

\[ e_j = \lim_{b \to 0} e_j = a_{ij} p_j \]

\[ B = \lim_{b \to 0} B / b = \left[ \begin{array}{cccc}
2\omega_1 - 2\psi_1 \epsilon_{11} & -2\psi_1 \epsilon_{12} & \cdots & -2\psi_1 \epsilon_{1K} \\
-2\psi_2 \epsilon_{21} & 2\omega_2 - 2\psi_2 \epsilon_{22} & \cdots & -2\psi_2 \epsilon_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
-2\psi_K \epsilon_{K1} & -2\psi_K \epsilon_{K2} & \cdots & 2\omega_K - 2\psi_K \epsilon_{KK}
\end{array} \right] \]

\[ s = \lim_{b \to 0} s = \sum_{i=1}^{K} p_i \left( \mu_i^2 \bar{\sigma}_j^2 + \mu_i^4 \right) \]

This way, the limit of the unconditional kurtosis is:

\[
\left( 3p' B^{-1} f + s \right) / \bar{\sigma}_j^2
\]
Appendix B. Proof of Corollary 1

For this, we show that the unconditional distribution of \( Y(t) = \ln S(t) \) is normal. Let \( s = t/b \) and \( w = \lfloor t/b \rfloor \) be the integer part. We shall use that, by approximation,

\[
Y(t) = s \sum_{a=0}^{s} \Delta Y_{ab} = \sum_{a=0}^{s} Y_{ab}
\]

The moment generating function of \( Y(t) \) is:

\[
E(e^{Y(t)}) = \lim_{b \to 0} E(e^{\sum_{a=0}^{s} \Delta Y_{ab}}) = \lim_{b \to 0} E\left(e^{\sum_{a=0}^{s} \epsilon \gamma_{a}}\right) = \lim_{b \to 0} \prod_{a=0}^{s} E(e^{\gamma_{a}})
\]

\[
= \lim_{b \to 0} \prod_{a=0}^{s} \left(1 + \left(\sum_{i=1}^{K} \beta_{i} e^{\left(\mu_{i} + \beta_{i} \frac{\epsilon}{2}\gamma_{i}\right)} - 1\right)\right) = \lim_{b \to 0} \prod_{a=0}^{s} \left(1 + \frac{A_{s}}{s}\right)
\]

where \( A_{s} = \sum_{i=1}^{K} \beta_{i} e^{\left(\mu_{i} + \beta_{i} \frac{\epsilon}{2}\gamma_{i}\right)} - s, \ b = t/s. \) Since \( \mu_{s} = \lim_{b \to 0} \left(\mu_{s}\right) < \infty, \ p_{s} = \lim_{b \to 0} \left(p_{s}\right) \) and \( \sigma_{s}^{2} = \lim_{b \to 0} \left(\gamma_{s}\right), \) we have:

\[
\lim_{b \to 0} \left(\frac{e^{t} - 1}{t}\right) = a,
\]

and since \( \lim_{t \to \infty} \left(\frac{e^{t} - 1}{t}\right) = a, \) after some algebra it follows that \( \lim_{b \to 0} \left(\frac{A_{s}}{s}\right) = \sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right) \)

To show this, for any \( u > 0 \) and any \( t > 0 \) we define \( v = u/\left(p_{s} a\right) \). According to the definition of convergence, we can find \( s^{*} \) such that for any \( s > s^{*} \) we have \( \mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2} - v < \mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2} y_{i} < \mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2} + v \). Based on this, we have:

\[
\sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2} - v\right) < \lim_{t \to \infty} \left(\frac{A_{s}}{s}\right) < \sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2} + v\right)
\]

\[
\sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right) - u < \lim_{t \to \infty} \left(\frac{A_{s}}{s}\right) < \sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right) + u
\]

This gives the limit of \( A_{s} \). Similar arguments show that:

\[
\sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right) - u < \lim_{t \to \infty} \left(\frac{1 + \frac{A_{s}}{s}}{s}\right)^{t} < \sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right) + u
\]

as \( \lim_{t \to \infty} \left(\frac{1 + \frac{a}{s}}{s}\right)^{t} = e^{a}. \) Hence \( E(e^{Y(t)}) = e^{\sum_{i=1}^{K} p_{i} a \left(\mu_{i} + \frac{\epsilon}{2} \sigma_{i}^{2}\right)} \)

and this is the mgf of a normal distribution with mean \( tp \) and variance \( \sigma^{2} \). \( \square \)
Appendix C. Proof of Theorem 3

We present the proof for \( h = 2 \). The proof for general \( h \)-period time intervals follows by induction.

From (9) we have:

\[
\sigma_{t, 2k+2}^2 = \omega_t + \left( \frac{\alpha_i}{q_i} + \beta_i \right) q_i \sigma_{t, 2k+1}^2 - \beta_i \eta_{i, 2k+1}
\]

This can be rearranged as:

\[
q_i \sigma_{t, 2k+2}^2 = \omega_t \left( 1 + \frac{\alpha_i}{q_i} + \beta_i \right) + \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 q_i \sigma_{t, 2k}^2 + \eta_{i, 2k+2} + \left( \frac{\alpha_i}{q_i} \right) \eta_{i, 2k+1} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \eta_{i, 2k}
\]

Repeating this equation for time step \( 2k+1 \) and summing yields:

\[
q_i \sigma_{t, 2k+1}^2 / 2 = \omega_t \left( 1 + \frac{\alpha_i}{q_i} + \beta_i \right) + \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 q_i \sigma_{t, 2k}^2 / 2 + q_i \sigma_{t, 2k}^2 \sigma_{t, 2k+1} - q_i \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 \sigma_{t, 2k}^2 \sigma_{t, 2k+1}
\]

\[
+ \frac{1}{2} \left( \eta_{i, 2k+2} + \left( 1 + \frac{\alpha_i}{q_i} \right) \eta_{i, 2k+1} + \left( \frac{\alpha_i}{q_i} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \eta_{i, 2k} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \eta_{i, 2k-1} \right)
\]

where we have used the following relationship for returns:

(C1) \( \sigma_{t, 2k+1}^2 = \left( \sigma_{t, 2k}^2 + \sigma_{t, 2k+1}^2 \right)^2 = \sigma_{t, 2k}^2 + \sigma_{t, 2k+1}^2 + 2 \sigma_{t, 2k} \sigma_{t, 2k+1} \)

Letting

\[
v_{t, 2k+1} = q_i \sigma_{t, 2k} + \sigma_{t, 2k+1} - q_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \sigma_{t, 2k-1}
\]

\[
+ \frac{1}{2} \left( \eta_{i, 2k+2} + \left( 1 + \frac{\alpha_i}{q_i} \right) \eta_{i, 2k+1} + \left( \frac{\alpha_i}{q_i} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \eta_{i, 2k} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \eta_{i, 2k-1} \right)
\]

we can now write

\[
q_i \sigma_{t, 2k+1}^2 / 2 = \omega_t \left( 1 + \frac{\alpha_i}{q_i} + \beta_i \right) + \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 q_i \sigma_{t, 2k}^2 / 2 + v_{t, 2k+1}
\]

and it is easy to see that:

\[
E(v_{t, 2k}) = 0
\]

(C2) \( E(v_{t, 2k+1} v_{t, 2(k+1)}) = 0 \) for \( l > 1 \)

\[
Corr(v_{t, 2k+1}, v_{t, 2k}) = \frac{-\alpha + d}{\epsilon + \epsilon^2 + f}
\]

where:
\[ \epsilon = \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 \]

\[ d = \frac{\alpha_i}{q_i} - \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) - \beta_i \left( 1 + \frac{\alpha_i}{q_i} \right) \left( \frac{\alpha_i}{q_i} + \beta_i \right) \]

\[ e = q_i^2 \frac{E(\varepsilon_i^2 | \varepsilon_{i-1}^2)}{E(\varepsilon_i^2)} \]

and

\[ f = 1 + \left( 1 + \frac{\alpha_i}{q_i} \right)^2 \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 + \left( \beta_i \left( \frac{\alpha_i}{q_i} + \beta_i \right) \right)^2 \]

Denoting the above correlation by \( a \), we have \( a \in (-1/2, 0) \). Consider the function \( f(x) = x^2 + a^{-1}x + 1 \) for \( a \in (-1/2, 0) \). Since \( f \) is continuous, \( f(0) = 1 \) and \( f(1) < 0 \), \( f \) must have a root \( \lambda \) between 0 and 1. And if \( f(\lambda) = 0 \), then \( \frac{\lambda}{1 + \lambda^2} = -a \). Hence there exists \( 0 < \lambda_i < 1 \) such that:

(C3) \[ -\text{Corr}(v_{i,2(k+1)}, v_{i,2k}) = \frac{\lambda_i}{1 + \lambda_i^2} \]

We now define \( w_{i,2(k+1)} \) in the following way:

(C4) \[ w_{i,0} = v_{i,0} \text{ and } w_{i,2(k+1)} = v_{i,2(k+1)} + \lambda_i w_{i,2k} \]

Obviously \( E(w_{i,2k}) = 0 \). Also, it can be shown that:

\[ \text{Corr}(w_{i,2(k+1)}, w_{i,2(k+1-l)}) = 0 \text{ for } l \geq 1. \]

Rewriting (c4) as

(C5) \[ v_{i,2(k+1)} = w_{i,2(k+1)} - \lambda_i w_{i,2k} \]

we have:

\[ q_{i,2}^2 \varepsilon_{2(k+1)}^2 / 2 = \omega_i \left( 1 + \frac{\alpha_i}{q_i} + \beta_i \right)^2 \left( \frac{\alpha_i}{q_i} + \beta_i \right)^2 q_{i,2}^2 \varepsilon_{2k}^2 / 2 + w_{i,2(k+1)} - \lambda_i w_{i,2k} \]

that is:

\[ 7 \text{ First, it can be shown that } a \text{ is negative. Also: } -1/2 < a \Leftrightarrow \epsilon + \epsilon^2 + f > 2\epsilon - 2d \text{ which is true because } \epsilon \text{ is positive, } 1 + \epsilon^2 > 2\epsilon \text{ and } f + 2d > 0, \text{ being a sum of squares.} \]

\[ 8 \text{ Let } \epsilon_{2l} = \text{Corr}(w_{i,2(k+1)}, w_{i,2(k+1-l)}) \text{. From (C3) and (C4) we have that } \epsilon_{4} = \epsilon_2 \left( 1 + \lambda_i^2 \right)/\left( 1 + \lambda_i^2 \right). \text{ Assuming } \epsilon_2 \geq 0 \text{ (the case for } \epsilon_2 \leq 0 \text{ follows similarly) we have } \epsilon_4 \geq \epsilon_2. \text{ Applying (B2) similarly we obtain that } \epsilon_{2(l+1)} \geq \epsilon_{2l}, \text{ with equality only if } \epsilon_2 = \epsilon_4 = \ldots = 0. \text{ Summing (15) for } l = 2, 3, \ldots \text{ gives } \left( 1 - \lambda_i \right)^2 \sum_{l=2}^{n} \epsilon_{2l} + \lambda_i \left( \epsilon_4 - \epsilon_2 \right) = 0 \text{ which holds only if } \epsilon_{2l} = 0 \text{ for } l \geq 1. \]
Now, similarly to (33), we define

\begin{equation}
\sigma_{i,2k}^2 = \frac{q_i \sigma_{2k}^2}{2 - w_{i,2k}}
\end{equation}

Hence (C6) becomes:

\begin{equation}
\sigma_{i,2(k+1)}^2 = \omega_i \left(1 + \frac{\alpha_i}{q_i} + \beta_i\right) + \left(\frac{\alpha_i}{q_i} + \beta_i\right)^2 - \lambda_i q_i \sigma_{2k}^2 / 2 + \lambda_i \sigma_{i,2k}^2
\end{equation}

and this is the updating formula for the new two-period conditional variances of each component.

Equation (33) needs to be satisfied for the new two-step returns series. Since \(2\sigma_{2k}^2 = \sigma_{2k} + \sigma_{2k-1}\) we have:

\begin{equation}
E(\sigma_{2k}^2 | I_{2(k-1)}, s_{2k} = i, s_{2k-1} = i) = 2\mu_i
\end{equation}

so \(2\mu_i = \mu_i\). The overall variance is given by:

\begin{equation}
2\sigma_{2k}^2 = \sum_{i=1}^{K} p_i 2\sigma_{i,2k}^2 + 2 \sum_{i=1}^{K} \sum_{j=1}^{K} \Delta_i^2
\end{equation}

Now equation (34) for the new two-step NM(K)-GARCH(1,1) model (replacing \(\eta\) by \(w\) and using steps of length 2) is obviously satisfied. Equations (35) and (36) can also be proved easily using the results derived above. Hence when \(b = 2\) the new parameters are

\begin{align*}
2\omega_i &= \omega_i \left(1 + \frac{\alpha_i}{q_i} + \beta_i\right) \\
2\alpha_i &= q_i \left(\frac{\alpha_i}{q_i} + \beta_i\right)^2 - \lambda_i \beta_i
\end{align*}

and \(2\beta_i \in (0,1)\) is the solution to \(\frac{\beta_i}{1 + 2\beta_i^2} = \frac{-ce + d}{e + e' + f}\), with \(c, e, d\) and \(f\) defined above. Furthermore, \(2\mu_i = \mu_i\) and \(2\beta_i = \beta_i\). □
Appendix D: Proof of Theorem 4.

We have that:

\[ (D1) \quad E \left( \frac{S_{kh} - S_{(k-1)b}}{S_{(k-1)b}} \right) I_{(k-1)b} \bigg| s_{kh} = i \bigg) = E(y_{kh} \bigg| I_{(k-1)b}, s_{kh} = i) = \delta(\mu_{ih}) \]

and

\[ (D2) \quad E \left( \frac{S_{kh} - S_{(k-1)b}}{S_{(k-1)b}} - \delta(\mu_{ih}) \right)^2 I_{(k-1)b} \bigg| s_{kh} = i \bigg) = E\left( (y_{kh} - \delta(\mu_{ih}))^2 \bigg| I_{(k-1)b}, s_{kh} = i \right) = E\left( (y_{kh} - \delta(\mu_{ih}))^2 \bigg| I_{(k-1)b}, s_{kh} = i \right) = \delta(\mu_{ih})^2 \]

The new variance process is:

\[ q_{kh}^2 = E(\delta^{-2} y_{kh}^2 \bigg| I_{(k-1)b}) = \sigma_{kh}^2 + b_{kh} \]

so that

\[ E(\delta b_{kh}) = 0 \quad \text{and} \quad q_{kh}^2 > 0 \quad \text{for} \quad 0 \leq k \]

We define

\[ b(t) = \lim_{b \to 0} b_t ; \quad b_t = \delta b_{kh} \quad \text{for} \quad kh \leq t < (k+1)b \]

For the variance process we can write:

\[ (D3) \quad E\left( q_{(k+1)b}^2 - q_{kh}^2 \bigg| I_{(k-1)b} \right) = E\left( \sum_{j=1}^{k} \delta \beta_j \left( \sigma_{kh}^2 + \sigma_{kh}^2 - \right) b_{(k+1)b} - b_{kh} \right) I_{(k-1)b} \]

where

\[ \epsilon_{kh} = E\left( \frac{b_{(k+1)b} - b_{kh}}{b} \bigg| I_{(k-1)b} \right) = E\left( \frac{b_{(k+1)b} - b_{kh}}{b} \right) I_{(k-1)b} \]

We need the existence of the limit \( \epsilon(t) = \lim_{b \to 0} \epsilon_t \) where \( b \epsilon_t = \epsilon_{kh} \) for \( k \leq t < (k+1)b \).

The variance of the variance process is not significant:
The covariances between the percentage returns and the changes in the variances are also very small:

\[
E \left( \left( \frac{S_{b} - S_{(k-1)b}}{S_{(k-1)b}} \right) \left( q_{(k+1)b} - q_{kb} - \left( \sum_{i=1}^{K} p_{i} \left( \omega_{i} + \frac{\alpha_{i} q_{e_{kb} / b - q_{kb}}}{2} \right) \right) \right) \right) I_{(k-1)b}, s_{kb} = i
= E \left( \left( \frac{S_{b} - S_{(k-1)b}}{S_{(k-1)b}} \right) \left( q_{(k+1)b} - q_{kb} - \left( \sum_{i=1}^{K} p_{i} \left( \omega_{i} + \frac{\alpha_{i} q_{e_{kb} / b - q_{kb}}}{2} \right) \right) \right) \right) I_{(k-1)b}, s_{kb} = i
= b \sqrt{b} \left( E \left( \left( \frac{S_{b} - S_{(k-1)b}}{S_{(k-1)b}} \right) \left( \sum_{i=1}^{K} p_{i} \left( \alpha_{i} q_{e_{kb} / b - q_{kb}} \right) \right) \right) I_{(k-1)b}, s_{kb} = i \right)
= b \sqrt{b} \left( E \left( \left( \frac{S_{b} - S_{(k-1)b}}{S_{(k-1)b}} \right) \left( \sum_{i=1}^{K} p_{i} \left( \alpha_{i} q_{e_{kb} / b - q_{kb}} \right) \right) \right) I_{(k-1)b}, s_{kb} = i \right)
= a(3/2)
\]

Now the result follows from Nelson’s (1990) convergence results of stochastic difference equations to differential equations. □
References:


Pieterbarg


Copyright © 2004. Alexander and Lazar. All Rights Reserved.