Modelling good and bad volatility

Matteo M. Pelagatti*

Department of Statistics, Università degli Studi di Milano-Bicocca, Milan, Italy

Abstract

The returns of many financial assets show significant skewness, but in the literature this issue is only marginally dealt with. Our conjecture is that this distributional asymmetry may be due to two different dynamics in positive and negative returns. In this paper we propose a process that allows the simultaneous modelling of skewed conditional returns and different dynamics in their conditional second moments. The main stochastic properties of the model are analyzed and necessary and sufficient conditions for weak and strict stationarity are derived.

An application to the daily returns on the principal index of the London Stock Exchange supports our model when compared to other frequently used GARCH-type models, which are nested into ours.

Keywords: Volatility, Skewness, GARCH, Asymmetric Dynamics, Stationarity.

JEL codes: C22, C53, G10.

1 Introduction

It is quite common for financial assets returns to show conditional heteroskedasticity, leptokurtosis and skewness. The first two properties are usually dealt with GARCH or Stochastic Volatility type models possibly with fat-tailed distributions. However, skewness has received much less attention in literature. In fact, skewness is usually obtained as a byproduct of models for leverage effects such as the Exponential GARCH (Nelson 1991) and the GJR-GARCH (Glosten et al. 1993). In most cases, though, the standardized residuals of these models still show significant skewness. Table 1 reports the sample skewness and its significance for the daily percentage returns of the FTSE100 index and for the standardized residuals with respect to asymmetric GARCH models with GED (conditional) distribution: the significance of the skewness coefficient appears even enhanced in the standardized residuals. Similarly, the absolute values of the minima are circa twice as large as the maxima, evidencing a significantly fatter negative tail. On some stock returns like ENEL1 the skewness coefficient gets even larger after passing the EGARCH and GJR-GARCH filters.

Some authors have modelled skewness by using GARCH-type processes with conditional skewed distributions (Harris et al. 2004, Lambert and Laurent 2001, Lambert and Laurent 2002, Lanne and Saikkonen 2004, Miettinen 2005), and in the OxMetrics package G@rch

* Full address: Department of Statistics, Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, I-20126 Milano. E-mail: matteo.pelagatti@unimib.it
1 ENEL is Italy’s largest power company (the privatized former public monopolist) and Europe’s third largest listed utility.
Table 1: Daily returns on FTSE100 from 2nd Jan 1984 to 11th Oct 2007.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>St.Dev.</th>
<th>Skew</th>
<th>Ex.Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns (%)</td>
<td>0.013</td>
<td>3.299</td>
<td>-5.658</td>
<td>0.439</td>
<td>-0.734*</td>
<td>10.426</td>
</tr>
<tr>
<td>St.Resid EGARCH</td>
<td>-0.011</td>
<td>5.830</td>
<td>-13.047</td>
<td>1.002</td>
<td>-0.530**</td>
<td>5.285</td>
</tr>
<tr>
<td>St.Resid GJR-GARCH</td>
<td>-0.012</td>
<td>6.144</td>
<td>-13.326</td>
<td>1.004</td>
<td>-0.554**</td>
<td>5.666</td>
</tr>
</tbody>
</table>

* and ** denote significance at, respectively, 5% and 1% levels, and are based on the bootstrap test of Lisi (2005).

4 by Laurent and Peters the skewed Student’s t distribution may be used for modelling conditional returns. Some other authors have built GARCH type models for the conditional skewness (Hansen 1994, Harvey and Siddique 1999).

In this paper we pursue a different approach. We argue that the dynamics of negative returns may differ from that of positive returns and propose a class of processes that may be used to model these different behaviours. More precisely, we allow the conditional second moments of positive returns (good volatility) to evolve differently from the conditional second moments of negative returns (bad volatility).

The plan of the rest of the article is as follows. In Section 2 we define the process in terms of its conditional distribution. The main stochastic properties of the process, such as the conditional moments and the conditions for the existence of weakly and strictly stationary solutions are analyzed in Section 3. Section 4 contains an application of the process to the main index of the London Stock Exchange (FTSE100), which demonstrates that the dynamics of real data may be captured by our model better than by other frequently used GARCH type processes. Finally, we give some concluding remarks in Section 5. Two appendices at the very end of the article report the density and the moments of the half generalized error distribution used throughout the paper and some eigenvalues and matrix norms necessary for assessing the stationarity of the process.

2 The Model

Let \( r_t \) be the time series of returns. Define the series of positive and negative returns as

\[ r_t^+ := r_t \cdot I(r_t \geq 0), \quad r_t^- := r_t \cdot I(r_t < 0) \]

with \( I(\cdot) \) indicator function. The returns are modelled as

\[ r_t | \{ r_t > 0, F_{t-1} \} \sim D^+(h_t^+), \quad r_t | \{ r_t < 0, F_{t-1} \} \sim D^-(h_t^-) \]

where \( F_t \) is the \( \sigma \)-field generated by \( r_1, r_{t-1}, \ldots, D^+ \) a family of distributions with positive support and second non-central moment \( h_{t+} \) and \( D^- \) a family of distributions with negative support and second non-central moment \( h_{t-} \). The two distributions may, of course, depend on other parameters, such as location and tail parameters.

^2As far as we know, this is the only mainstream software package that allows for skewed conditional distributions in GARCH modelling. We used this package for estimating the EGARCH and GJR-GARCH models of Table 1.

^3While presenting this work at the 5th OxMetrics User Conference, Sébastien Laurent acquainted me with the article by El Babsiri and Zakoian (2001), whose basic idea of treating positive and negative returns asymmetrically is the same as the one we pursue in this paper. However, the model we propose is different from theirs and has the advantage of nesting other frequently used models such as the GARCH and the GJR-GARCH, both with possibly fat-tailed and skewed conditional distributions. This allows the use of maximum-likelihood based statistics for testing against these subset models.
It is useful for later developments to use conditional distributions with unitary second moments, such as the Half-GED, whose main properties are reported in Appendix A. By doing so, we can write
\[ r_t|\mathcal{F}_{t-1} = \xi_t \mathbb{1}(\xi_t \geq 0) \sqrt{h^+_t} + \xi_t \mathbb{1}(\xi_t < 0) \sqrt{h^-_t} \] (2.1)
with \( \xi_t \) i.i.d. process with
\[ \mathbb{E}[\xi^2_t | \xi_t \geq 0] = \mathbb{E}[\xi^2_t | \xi_t < 0] = 1, \]
\[ p := \mathbb{P}(\xi_t \geq 0), \quad q := \mathbb{P}(\xi_t < 0) = 1 - p. \]
Thus, \( p \) and \( q \) are the probabilities of, respectively, positive and negative returns, and will be assumed constant throughout the paper.

The conditional second moments are let evolve according to the following bivariate difference equation
\[ \begin{bmatrix} h^+_t \\ h^-_t \end{bmatrix} = \omega + A \begin{bmatrix} r^+_t \\ r^-_{t-1} \end{bmatrix} + B \begin{bmatrix} h^+_t \\ h^-_{t-1} \end{bmatrix} \] (2.2)
with \( \omega \) (2 \times 1) non-negative vector, \( A \) and \( B \) (2 \times 2) matrices with non-negative elements. The non-negativity conditions are sufficient and necessary for the non-negativity of the conditional second moments. The matrix \( B \) will be taken diagonal, so that each element of \( h_t \) depends only on the same element in \( h_{t-1} \). This will avoid collinearity problems, since we expect the two components of \( h_t \) to be highly correlated in real world applications.

3 Stochastic Properties of the Model

Form equations (2.1) and (2.2) we have the following representation of the square of positive and negative returns:
\[ r^2_t|\mathcal{F}_{t-1} = X_t h_t \] (3.3)
\[ h_t = \omega + (AX_{t-1} + B)h_{t-1} \] (3.4)
with
\[ r_t := \begin{bmatrix} r^+_t \\ r^-_t \end{bmatrix} \quad h_t := \begin{bmatrix} h^+_t \\ h^-_t \end{bmatrix} \quad X_t := \begin{bmatrix} \xi^2_t \mathbb{1}(\xi_t \geq 0) & 0 \\ 0 & \xi^2_t \mathbb{1}(\xi_t < 0) \end{bmatrix}. \]
Notice that \( X_t \) is an i.i.d. sequence of nonnegative random matrices.

3.1 Conditional moments

It is easy to see that the return process \( r_t \) is not a martingale difference (MD) sequence. In fact, if we model \( r^+_t \) and \( r^-_t \) using two independent distributions, the conditional expectation may be easily obtained as
\[ \mu_{t|t-1} := \mathbb{E}[r_t|\mathcal{F}_{t-1}] \]
\[ = \mathbb{E}\left[ \xi_t \mathbb{1}(\xi_t \geq 0) \sqrt{h^+_t} | \mathcal{F}_{t-1} \right] + \mathbb{E}\left[ \xi_t \mathbb{1}(\xi_t < 0) \sqrt{h^-_t} | \mathcal{F}_{t-1} \right] \]
\[ = \mu^+ \cdot p \cdot \sqrt{h^+_t} + \mu^- \cdot q \cdot \sqrt{h^-_t}, \]
where \( \mu^+ := \mathbb{E}(\xi_t | \xi_t \geq 0) \) and \( \mu^- := \mathbb{E}(\xi_t | \xi_t < 0) \).
Using the Half-GED described in Appendix A, we have
\[
\mu^+ = \Gamma(2/\nu^+) \left[ \Gamma(1/\nu^+) \Gamma(3/\nu^+) \right]^{-1/2}
\]
\[
\mu^- = -\Gamma(2/\nu^-) \left[ \Gamma(1/\nu^-) \Gamma(3/\nu^-) \right]^{-1/2}.
\]

If the MD property is considered an essential feature, it may be obtained by subtracting the conditional mean from the process \( r_t \):
\[
\tilde{r}_t =: r_t - \mu_t|_{t-1}.
\]

We do not use the process \( \tilde{r}_t \) in this paper directly, since a financial economist would anyway change the conditional expectation of the process for embedding risk-preferences or some form of absence of arbitrage.

The \( k \)-step ahead prediction (conditional expectation) of \( r_t \) is nontrivial since it involves conditional expectations of the process \( \sqrt{h_t} \). Probably the easiest way to get approximations is by means of the following Taylor expansion. Let \( y \) be a nonnegative random variable with expectation \( \mu \), then, if the needed moments exist, we have
\[
E[\sqrt{y}] \approx \mu^{1/2} - \frac{1}{8} \mu^{-3/2} E[(y - \mu)^2] + \frac{1}{16} \mu^{-5/2} E[(y - \mu)^3].
\]

Of course, the above approximation may be stopped at any desired order.

The equations for forecasting the volatilities are easily obtained by the law of iterated expectations:
\[
\begin{align*}
E[h_{t+1}|F_t] &= h_{t+1} \\
E[h_{t+k}|F_t] &= \sum_{i=0}^{k-2} (AX + B)^i \omega + (AX + B)^{k-1} h_{t+1}, \quad k = 2, 3, \ldots
\end{align*}
\]

As for the conditional variance, since \( h_t \) is \( F_{t-1} \)-measurable, we have \( E[r_t^2|F_{t-1}] = h_t \), and after some easy but tedious algebra we obtain
\[
m_{2t-1} := E[\tilde{r}_t^2|F_{t-1}]
= p \left( 1 - p\mu^+ \right) h_t^+ - 2pq\mu^+ \mu^- \sqrt{h_t^+ h_t^-} + q \left( 1 - q\mu^- \right) h_t^-.
\]

The third conditional central moment, necessary for computing the conditional skewness, is given by
\[
m_{3t-1} := p \left( \mu_3^+ - 2p\mu^+ + p^2\mu^+^3 \right) (h_t^+)^{3/2} + \\
- 2pq\mu^- \left( 1 - 2p\mu^+ \right) h_t^+ \sqrt{h_t^-} + \\
- 2pq\mu^+ \left( 1 - 2q\mu^- \right) \sqrt{h_t^+ h_t^-} + \\
+ q \left( \mu_3^- - 2q\mu^- + q^2\mu^-^3 \right) (h_t^-)^{3/2},
\]

where \( \mu_3^+ \) and \( \mu_3^- \) are the third moments of the distributions of \( \xi_t|\xi \geq 0 \) and \( \xi_t|\xi < 0 \), respectively (refer to Appendix A for the moments of the HGED).
Dividing equation (3.8) by equation (3.7) raised to the power of 3/2, we obtain the formula for the conditional skewness:

$$
\gamma_{1t|t-1} := \frac{m_{3t|t-1}}{m_{2t|t-1}^{3/2}}.
$$

### 3.2 Weak stationarity and second moments

By taking expectations of both sides of equation (3.4), we get

$$
\bar{h}_t = \omega + (A\bar{X} + B)\bar{h}_{t-1},
$$

where we have set $\bar{h}_t := \mathbb{E}[h_t]$ and

$$
\bar{X} := \mathbb{E}(X_t) = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}.
$$

It is well-known that a necessary and sufficient condition for $\bar{h}_t$ to converge to an equilibrium, say $\bar{h}$, which does not depend on initial conditions is

$$
\rho(A\bar{X} + B) < 1, \tag{3.9}
$$

where $\rho(\cdot)$ denotes the spectral radius. The eigenvalues of $A\bar{X} + B$ are reported in Appendix B. When this condition holds the long run volatilities are given by

$$
\bar{h} = (I - A\bar{X} - B)^{-1}\omega,
$$

provided that $(I - A\bar{X} - B)$ is invertible, and

$$
\mathbb{E}[r_{2t}] = \bar{X}\bar{h}. \tag{3.10}
$$

The same results may be obtained from the VARMA representation of the squared positive and negative returns process. Let’s define the process

$$
n_t := r_{2t}^2 - \bar{X}h_t. \tag{3.11}
$$

It easy to show that $n_t$ is a MD sequence:

$$
\mathbb{E}[n_t|\mathcal{F}_{t-1}] = \mathbb{E}[r_{2t}^2|\mathcal{F}_{t-1}] - \bar{X} \mathbb{E}[h_t|\mathcal{F}_{t-1}] = \bar{X}h_t - \bar{X}h_t = 0.
$$

From equation (3.11) we have

$$
h_t = \bar{X}^{-1}r_{2t}^2 - \bar{X}^{-1}n_t,
$$

and substituting in equation (2.2), we obtain the VARMA representation of $r_{2t}^2$,

$$
r_{2t}^2 = \bar{X}\omega + (\bar{X}A + \bar{X}B\bar{X}^{-1})r_{2t-1}^2 + n_t + \bar{X}B\bar{X}^{-1}n_{t-1},
$$

which is well known to have causal stationary solutions iff

$$
\rho(\bar{X}A + \bar{X}B\bar{X}^{-1}) < 1. \tag{3.12}
$$
Since the spectrum of a matrix is invariant under similarity transformations, conditions (3.9) and (3.12) are equivalent, in fact:

\[ \bar{X}^{-1}(\bar{X}A + \bar{X}B\bar{X}^{-1})\bar{X} = A\bar{X} + B. \]

The expectation of \( r_t^2 \) may be derived thorough the VARMA representation as well:

\[ \mathbb{E}(r_t^2) = (I_2 - \bar{X}A - \bar{X}B\bar{X}^{-1})^{-1}\bar{X}\omega. \] (3.13)

The equivalence of this expressions with (3.10) is easily proved applying the result (Lütkepohl 1996, p.29)

\[ (I + XY)^{-1}X = X(I + YX)^{-1} \]

that holds if \( I + XY \) is non-singular and \( X \) and \( Y \) are square matrices.

We summarize these results in the following proposition. For notational brevity’s sake put

\[ G_t = AX_t + B. \]

**Proposition 1.** The process \( r_t \) is weakly stationary if and only if \( \rho(\mathbb{E}[G_0]) < 1 \). Under this condition, we have

\[ h_t := \mathbb{E}(h_t) = (I - \mathbb{E}[G_0])^{-1}\omega, \]

\[ \mathbb{E}(r_t^2) = \bar{X}h, \]

\[ \mathbb{E}(r_t) = [\mu^+, \mu^-] \mathbb{E}\left(\sqrt{h_t}\right). \]

An approximation to the unconditional expectation of \( r_t \) may be obtained using the Taylor expansion (3.6).

### 3.3 Strict stationarity and ergodicity

In order to generalize the results of Nelson (1990) we have to rely on the literature on strong law of large numbers and weak convergence for products of i.i.d. random matrices originated by the articles of Bellman (1954), Furstenberg and Kesten (1960) and Kesten and Spitzer (1984). In fact, by iterating equation (3.4) for the times 1, 2, ..., \( t \) we can write

\[ h_t = \left(I + G_{t-1} + G_{t-1}G_{t-2} + G_{t-1}G_{t-2}G_{t-3} + \ldots + G_{t-1} \cdots G_1\right)\omega \]

\[ + G_{t-1} \cdots G_1G_0h_0, \] (3.14)

where \( G_t = AX_t + B \) and \( h_0 \) is strictly positive and finite with probability one and independent of \( G_0, G_1, \ldots \). For notational convenience put \( H_t = G_{t-1}G_{t-2} \cdots G_0 \) for \( t \in \mathbb{N} \).

In order to make this paper as much self-contained as possible, we summarize theorems 1 and 2 by Furstenberg and Kesten (1960), theorem 5 by Kingman (1973) and theorem 2 by Hennion (1997) in the following Theorem 1.

In the rest of the paper we use the following notation: let \( x \) be a \( q \)-vector and \( X \) a \( q \times q \)
matrix,
\[
\|x\|_1 = \sum_{i=1}^{q} |x_i|
\]
\[
\|X\|_1 = \max_j \sum_{i=1}^{q} |x_{ij}|
\]
\[
v(X) = \min_j \sum_{i=1}^{q} |x_{ij}|
\]
and the relational operators \(<, >, =\) applied to vectors and matrices are to be intended elementwise (i.e. \(X > Y\) is true iff every element of \(X\) is greater than the corresponding element of \(Y\)).

**Theorem 1** (Furstenberg, Kesten, Kingman and Hennion). Let \(X_1, X_2, X_3, \ldots\) form a stationary and ergodic stochastic process with values in the set of \(k \times k\) nonnegative real matrices. If
\[
\mathbb{E} \left| \log \|X'_1\|_1 \right| + \mathbb{E} |v(X'_1)| < +\infty,
\]
and for the products \(Y_n = X_n \cdots X_1\)
\[
\Pr\{Y_n > 0\} > 0 \quad \text{for some } n,
\]
holds, then
\[
a) \quad \lim_{n \to \infty} n^{-1} \mathbb{E} |\log \|Y_n\|_1| = \lambda
\]
\[
b) \quad \lim_{n \to \infty} n^{-1} \log \|Y_n\|_1 = \lambda \quad \text{a.s.,}
\]
\[
c) \quad \lim_{n \to \infty} n^{-1} \log \rho(Y_n) = \lambda \quad \text{a.s.,}
\]
\[
d) \quad \lim_{n \to \infty} n^{-1} \log |Y_n|_{ij} = \lambda \quad \text{a.s.,}
\]
where \(\lambda\), usually referred to as greatest Lyapunov exponent or greatest characteristic exponent, is a nonrandom point in \(\mathbb{R}\), \(\rho(Y_n)\) is the spectral radius of the matrix \(Y_n\) and \(|Y_n|_{ij}\) is the \((i, j)\)-th element of \(Y_n\).


Now, paralleling Nelson (1990), we are in the condition to prove the following propositions. Let \(\lambda\) be the greatest Lyapunov exponent relative to the matrix product \(H_t\) and suppose that the assumptions of Theorem 1 hold for the sequence of matrices \(G_i\).

**Proposition 2.** Let \(\omega = 0\), as \(t \to \infty\):
\[
a) \quad \text{if } \lambda > 0, \text{ then } h_t \to +\infty \quad \text{a.s.;}
\]
\[
b) \quad \text{if } \lambda = 0, \text{ then } h_t \text{ is tight in } \mathbb{R}^{+2} \text{ and has a limit distribution which depends on } h_0 \text{ and for some } h_0 \text{ is concentrated on strictly positive vectors;}
\]
\[
c) \quad \text{if } \lambda < 0, \text{ then } h_t \to 0 \quad \text{a.s.;}
\]

\(^{\text{4}}\)The easier to verify condition \(\max_{i,j} \mathbb{E} |\log |X'_1|_{ij}| < +\infty\) implies the condition stated in the proposition.
Proof. Let \((\Omega, F, \mu)\) be the probability space on which \(G_i, \) for \(i = 1, 2, \ldots, \) are defined. By Theorem \([\text{point d)}\) and the definition of almost sure convergence, there is a \(F\)-measurable \(M, \) such that for \(n > M \) and \(\lambda \neq 0\)

\[-\frac{|\lambda|}{2} < \frac{1}{t} \log \{|H|_{ij} - \lambda < \frac{|\lambda|}{2} \text{ a.s.} \}

from which

\[e^{-\frac{|\lambda|}{2} + t\lambda} < |H|_{ij} < e^{\frac{|\lambda|}{2} + t\lambda} \text{ a.s.} \]

Now, for \(\lambda > 0\) the (almost sure) limit of the above expression is \(+\infty\), while for \(\lambda < 0\) the limit is 0. When \(\lambda = 0\), we have to relay on the results of Kesten and Spitzer (1984). Since \(G_0\) is non-negative and, when \(\beta_{ii} > 0\) for \(i = 1, 2, \) \(\Pr\{G_0\text{ has a zero row or column}\} = 0,\) then assumptions \(H_1\) and \(H_2\) of Kesten and Spitzer (1984) are fulfilled and by their Theorem 1

\[
C_I : \{H_t \text{ is tight with limit distribution concentrated on positive matrices}\} \\
\downarrow \\
H_3 : \{E[G_0]_{ij} < \infty \text{ and log } \rho(E[G_0]) = 0\} \text{ and } C_{IV} : \{\lambda = 0 \text{ a.s.}\}.
\]

By Theorem 2 of Kesten and Spitzer (1984), under assumptions \(H_1, H_2\) and \(H_3\)

\[C_{IV} \Leftrightarrow C_I. \]

Therefore, if \(H_1, H_2\) and \(E[G_0]_{ij} < \infty\) hold, the condition \(\lambda = 0\) is necessary and sufficient for \(H_t\) to be tight and converge in probability to a nonzero random matrix. This implies that for arbitrary positive \(h_0 < \infty\) the random vector \(h_t = H_t h_0\) is tight and its limiting distribution depends on \(h_0,\) and, apart from possible subspaces of \(\mathbb{R}^{+2}\) where \(h_t\) may converge to 0 almost surely, \(h_t\) converges weakly to a distribution with positive support. \(\square\)

**Proposition 3.** Let \(\omega > 0,\) as \(t \to \infty:\)

a) if \(\lambda \geq 0,\) then \(h_t \to +\infty \text{ a.s.};\)

b) if \(\lambda < 0,\) then \(h_t\) is asymptotically strictly stationary and ergodic.

**Proof.** Part a) is a trivial consequence of Proposition 3a). As for part b), we have already proven that the second addend of equation \([3.14]\) converges almost surely to zero when \(\lambda < 0.\) In the proof of Proposition \(k\) we have shown that the rate of almost sure convergence to zero of all the elements of \(H_t\) is \(O(e^{t\lambda/2})\), which assures the convergence of the partial sums of matrix products in \([3.14]\).

Now consider the process

\[h_t := (I + G_{t-1} + G_{t-2} + G_{t-1}G_{t-3} + \ldots)\omega \]

By Stout (1974, Th. 3.5.8) \(h_t\) is asymptotically stationary and ergodic if

\[R_{t}^\infty := I + G_{t-1} + G_{t-1}G_{t-2} + G_{t-1}G_{t-2}G_{t-3} + \ldots \]

is measurable. Now, take the process

\[R_{t}^{k} := I + G_{t-1} + G_{t-1}G_{t-2} + G_{t-1}G_{t-2}G_{t-3} + \ldots + G_{t-1}G_{t-2}G_{t-3} \cdots G_{t-k}. \]
$R^k_t$ involves only sums and products of measurable functions and therefore is measurable for any finite $k$. Since $R^k_t$ is increasing in $k$ we have $\sup_k R^k_t = R^\infty_t$, which is measurable. Thus, $\tilde{h}_t = R^\infty_t \omega$ is stationary and ergodic.

Now we have to show that $\tilde{h}_t$ converges almost surely to $\hat{h}_t$ as $t$ increases. Consider the difference

$$h_t - \hat{h}_t = G_{t-1}G_{t-2} \cdots G_0h_0$$

$$- \left( G_{t-1}G_{t-2} \cdots G_0 + G_{t-1}G_{t-2} \cdots G_{-1} + G_{t-1}G_{t-2} \cdots G_{-2} + \cdots \right) \omega.$$

We have proven in Proposition 2 that the first addend on the right hand side converges almost surely to zero as $t$ increases if $\lambda < 0$. Under the same condition, also the second addend converges almost surely to zero, since each element of the infinite sum converges almost surely to zero at exponential rate as $t$ increases. This proves the asymptotic stationarity and ergodicity of the process $h_t$ when $\lambda < 0$.

Remarks

i) Lyapunov exponents are usually very hard to compute, but under the conditions of Theorem 1 it is easy to see that $\rho(E(G_0)) < 1$ implies $\lambda < 0$ (i.e. weak stationarity is sufficient for strict stationarity), in fact by equation (1.4) in Kesten and Spitzer (1984)

$$\lambda = \lim_{t \to \infty} \frac{1}{t} \log \|H_t\| < \log \rho(E(G_0)).$$

ii) Exploiting the submultiplicativity of any operator norm $\|\cdot\|$, another sufficient condition for strict stationarity could be derived using the inequality

$$\frac{1}{t} \log \|H_t\| \leq \frac{1}{t} \sum_{i=1}^{t} ||G_{i-1}||,$$

that for $t \to \infty$ becomes

$$\lambda \leq E[\log \|G_0\||].$$

If the right hand side is smaller then zero, the strict stationarity condition holds. But since by a well known result of matrix algebra $\rho(E[G_0]) \leq \|E[G_0]\|$, the bound of the previous remark is tighter.

4 Application

The model, with $p = q = 0.5$, has been applied to the daily returns on the FTSE100 index, and the the estimated good and bad volatilities are depicted in log-scale in Figure 1. As expected the positive and negative volatilities are highly correlated, but negative volatility seems more reactive to shocks than positive volatility.

By observing Table 2 it emerges that negative returns tend to raise volatilities much more than positive shocks. This feature, known as leverage effect, is well documented in the literature and can be captured using EGARCH and or GJR-GARCH processes. What these processes are not able to model is the case in which negative volatility is not driven
Figure 1: Positive and negative volatilities (log scale) estimated on FTSE100 returns in the period 2nd Jan 1984 – 11th Oct 2007.

<table>
<thead>
<tr>
<th>$h^+_t$</th>
<th>Estimate</th>
<th>St. Error</th>
<th>t-ratio</th>
<th>$h^-_t$</th>
<th>Estimate</th>
<th>St. Error</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Const</td>
<td>0.016</td>
<td>0.030</td>
<td>0.531</td>
<td></td>
<td>0.023</td>
<td>0.037</td>
<td>0.603</td>
</tr>
<tr>
<td>$r_{t-1}^+$</td>
<td>0.057</td>
<td>0.011</td>
<td>4.986</td>
<td>$r_{t-1}^-$</td>
<td>0.014</td>
<td>0.009</td>
<td>1.486</td>
</tr>
<tr>
<td>$r_{t-1}^-$</td>
<td>0.090</td>
<td>0.012</td>
<td>7.721</td>
<td>$r_{t-1}^-$</td>
<td>0.156</td>
<td>0.023</td>
<td>6.683</td>
</tr>
<tr>
<td>$h^+_t$</td>
<td>0.908</td>
<td>0.012</td>
<td>73.654</td>
<td>$h^-_t$</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h^-_t$</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>$h^-_t$</td>
<td>0.904</td>
<td>0.015</td>
<td>58.338</td>
</tr>
<tr>
<td>Tail</td>
<td>1.479</td>
<td>0.040</td>
<td>36.909</td>
<td></td>
<td>1.967</td>
<td>0.049</td>
<td>40.133</td>
</tr>
<tr>
<td>$\rho[E[G_0]]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.987</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restrictions</td>
<td>Wald Stat.</td>
<td>DoF</td>
<td>p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewed. GARCH</td>
<td>108.76</td>
<td>5</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewed. GJR-GARCH</td>
<td>20.26</td>
<td>4</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: ML estimates of the model for daily FTSE100 returns from 2nd Jan 1984 to 11th Oct 2007.

by positive shocks, which however have significant effects on positive volatility. This fact is a clear feature of our data, as it appears from the second line of Table 2.

Figure 2 depicts the news impact curves, both for positive and negative volatilities. The curves are computed as

$$h^+_t = \omega_1 + a_{11} r_{t-1}^2 I(r_{t-1} \geq 0) + a_{12} r_{t-1}^2 I(r_{t-1} < 0) + b_{11} \bar{h}^+$$

$$h^-_t = \omega_2 + a_{21} r_{t-1}^2 I(r_{t-1} \geq 0) + a_{22} r_{t-1}^2 I(r_{t-1} < 0) + b_{22} \bar{h}^-$$

where $\bar{h}^+$ and $\bar{h}^-$ are the unconditional second moments and $r_{t-1}$ ranges between the minimum and maximum observed on the FTSE100 daily percentage returns series.
Since our model nests the simple GARCH and the GJR-GARCH both skewed and symmetric, it is possible to carry out one of the classic ML-based tests (Wald, LR, LM). We report Wald type statistics in the lower panel of Table 2 and both hypotheses are to be rejected at any usual level.

5 Conclusion

We have proposed a novel approach to attack the issue of skewness in financial asset returns. The relevant literature dealing with this problem is still scarce and mostly relies on GARCH type processes with conditional skewed distributions (Harris et al. 2004, Lambert and Laurent 2001, Lambert and Laurent 2002, Lanne and Saikkonen 2004, Miettinen 2005) or with time-varying conditional skewness (Hansen 1994, Harvey and Siddique 1999).

Our approach consists in the asymmetrical modelling of second moments for positive and negative returns. The model we have developed nests the widespread GARCH and GJR-GARCH processes as special cases and this allows the use of ML-based tests for assessing the relevance of our model when compared to these simpler ones.

The application of our model to the main index of the London Stock Exchange reveals that the second moments of positive and negative returns do show different dynamics. In fact, the restrictions corresponding to the skewed GARCH and GJR-GARCH models are strongly rejected by the data.

Practitioners should get benefits from the application of our model, both when evaluating the risk of holding an asset or by exploiting the asymmetric dynamics of positive and negative volatilities for implementing option strategies.\footnote{This was suggested to me by professor David Hendry, whom I greatly acknowledge, at the 5th OxMetrics User Conference, London, 20th-21st September 2007.}
References


Appendices

A  Half Generalized Error Distribution

The half generalized error distribution (HGED) has positive support and in $\mathbb{R}^+$ is proportional to the well known GED. Its density is

$$f(x) = \frac{\nu \exp\left\{-\frac{1}{2} \left(\frac{x}{\omega}\right)^\nu\right\}}{\omega 2^{\frac{\nu}{2}} \Gamma\left(\frac{1}{\nu}\right)}$$

with

$$\omega = 2^{-\frac{1}{\nu}} \left(\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{1}{\nu}\right)}\right)^{\frac{1}{2}}, \quad x \in \mathbb{R}^+ \cup \{0\}, \quad \nu \in \mathbb{R}^+.$$

Moments:

$$\mu_k = \mathbb{E}(X^k) = \Gamma\left(\frac{1+k}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right)^{\frac{k}{2}} \Gamma\left(\frac{3}{\nu}\right)^{-\frac{k}{2}};$$

particularly

$$\mu = \mathbb{E}(X) = \Gamma\left(\frac{2}{\nu}\right) \left[\Gamma\left(\frac{1}{\nu}\right) \Gamma\left(\frac{3}{\nu}\right)\right]^{-\frac{1}{2}}$$

and

$$\mu_2 = \mathbb{E}(X^2) = 1.$$

B  Relevant norms and eigenvalues

Eigenvalues of $A\bar{X} + B$:

$$\rho_{1,2} = \frac{1}{2} \left\{ b_{11} + b_{22} + a_{11}p + a_{22}q \pm \left[ (b_{11} + b_{22} + a_{11}p + a_{22}q)^2 ight. ight.$$  
$$
- 4(b_{11}b_{22} + a_{11}b_{22}p + a_{22}b_{11}q - a_{12}a_{21}pq + a_{11}a_{22}pq) \right\}^{1/2} \}.
$$

Norms of $G'_t$:

$$\|G'_t\|_1 = \left[ (a_{11} + a_{12})\xi_t I(\xi_t \geq 0) + b_{11} \right] \lor \left[ (a_{21} + a_{22})\xi_t I(\xi_t < 0) + b_{22} \right],$$

$$v(G'_t) = \left[ (a_{11} + a_{12})\xi_t I(\xi_t \geq 0) + b_{11} \right] \land \left[ (a_{21} + a_{22})\xi_t I(\xi_t < 0) + b_{22} \right].$$