Asymmetric GARCH Models

• The basic GARCH model considered so far assumes that the conditional variance $\sigma^2_t$ depends only on the magnitude and not on the sign of past shocks.

• However, stock market variance tends to react more strongly to bad news than to good news, which is often referred to as the leverage effect.

• To illustrate, we may define the leverage effect at lag $\tau$ as

$$L(\tau) = \text{Corr}(\epsilon_{t-\tau}, |\epsilon_t|).$$

(1)
leverage for the CAC 40
Asymmetric GARCH Models I

- The first asymmetric GARCH model that has been put forward is the Asymmetric GARCH (AGARCH) of Engle (1990), which specifies the conditional variance as

\[
\sigma_t^2 = \omega + \alpha(\epsilon_{t-1} - \theta)^2 + \beta \sigma_{t-1}^2 \\
= \omega + \alpha \theta^2 + \alpha \epsilon_{t-1}^2 - 2\alpha \theta \epsilon_{t-1} + \beta \sigma_{t-1}^2. \tag{3}
\]

- In model (2), the conditional variance, as a function of \(\epsilon_{t-1}\), has its minimum at \(\theta\) rather than at zero.

- Thus, if \(\theta > 0\), negative shocks will have a greater impact on the conditional variance than positive shocks of the same magnitude.

- (3) shows that, if \(\alpha + \beta < 1\), the unconditional variance of this process is

\[
E(\sigma_t^2) = \frac{\omega + \alpha \theta^2}{1 - \alpha - \beta}. \tag{4}
\]
Asymmetric GARCH Models II

• The asymmetric GARCH model proposed by Glosten, Jagannathan and Runkle (1993), referred to as GJR–GARCH, models the conditional variance as

\[ \sigma_t^2 = \omega + (\alpha + \theta S_{t-1})\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2, \]

where

\[ S_{t-1} = \begin{cases} 
1 & \text{if } \epsilon_{t-1} < 0 \\
0 & \text{if } \epsilon_{t-1} \geq 0 
\end{cases} \]

• Clearly \( \theta > 0 \) implies that the change in the next period’s variance is negatively correlated with today's return.

• If the innovation density is symmetric (e.g., normal or Student’s t), the unconditional variance is

\[ \mathbb{E}(\sigma_t^2) = \frac{\omega}{1 - \alpha - \theta/2 - \beta}. \]
Asymmetric GARCH Models III: EGARCH

• The Exponential GARCH process of Nelson (1991)\(^1\) is another popular GARCH specification.

• The EGARCH(1,1) model for the conditional variance of \(\epsilon_t\) is

\[
\log \sigma_t^2 = \omega + g(\eta_{t-1}) + \beta \log \sigma_{t-1}^2,
\]

where \(\eta_t = \epsilon_t / \sigma_t\)

\[
g(\eta_{t-1}) = \theta \eta_{t-1} + \alpha (|\eta_{t-1}| - E(|\eta_{t-1}|)).
\]

For example, if \(\eta_t\) is standard normally distributed, then

\[
E(|\eta_{t-1}|) = \sqrt{\frac{2}{\pi}} \approx 0.7979.
\]

• Note that specification (5)–(6) does not require any parameter constraints to make sure that the conditional variance remains positive.

• Since $\sigma_t^2$ is the exponential of an AR(1) process

• The term $g(\eta_t)$ is iid with zero mean, so that (5) is an AR(1) process for $\log \sigma_t^2$ which is stationary when $|\beta| < 1$.

• Regarding the asymmetric response to shocks,
  
  – for $\eta_t > 0$, $g(\eta_t)$ is linear with slope $\alpha + \theta$
  – for $\eta_t < 0$, $g(\eta_t)$ is linear with slope $\theta - \alpha$

• Thus, a variety of asymmetric response patterns are possible:
  
  – if $\theta = \alpha$, we have a response to only positive shocks, whereas
  – for $\theta = -\alpha$, we have a response to only negative shocks.
Moments of the EGARCH Process

- Calculating the moments of EGARCH models is rather complicated, as compared to standard GARCH processes.

- Just to illustrate, consider process (5)–(6).

- Write the model in MA(∞) form,

\[
\log \sigma_t^2 = \frac{\omega}{1 - \beta} + \sum_{i=1}^{\infty} \beta^{i-1} g(\eta_{t-i})
\]

\[
E(\sigma_t^2) = \exp \left\{ \frac{\omega}{1 - \beta} \right\} \prod_{i=1}^{\infty} E(\exp\{\beta^{i-1} g(\eta_{t-i})\})
\]

- Nelson (1991) provides closed-form expressions for \(E(\exp\{\beta^{i-1} g(\eta_{t-i})\})\) in case that \(\eta_t \sim N(0, 1)\), but the result is still somewhat unwieldy.
• On the other hand,
\[
E(\log \sigma_t^2) = \frac{\omega}{1 - \beta}
\]  
(7)

is easily calculated (cf. the standard AR(1) model) and may be used to initialize the conditional log–likelihood function, as suggested by Nelson (1991).

• In the EGARCH model, due to the exponential transformation, the unconditional variance (and indeed any moments for \( \sigma_t^2 \)) typically do not exist for Student’s \( t \)–type distributions.

• As this may be unappealing from an economic point of view, Nelson (1991) suggests to use the GED instead,
\[
f(\eta_t; p) = \frac{\lambda^p}{2^{1/p+1}\Gamma(1/p)} \exp \left\{ -\frac{|\lambda \eta_t|^p}{2} \right\},
\]
(8)

where \( \lambda = 2^{1/p} \sqrt{\Gamma(3/p)/\Gamma(1/p)} \).
• For the EGARCH model, we need $E(|\eta_t|)$, which for the GED is

$$E(|\eta_t|) = \frac{\Gamma(2/p)}{\sqrt{\Gamma(1/p)\Gamma(3/p)}}.$$ 

• For the GED, we require $p > 1$ for the moments to exist, but this is usually satisfied in practice.
News Impact Curve

• To analyze the asymmetric response of the variance in different GARCH specifications, Engle and Ng (1993) defined the new impact curve (NIC).

• This is defined as the functional relationship

\[ \sigma_t^2 = \sigma_t^2(\epsilon_{t-1}), \]

with all lagged variances evaluated at their unconditional values.

• For example, for the standard symmetric GARCH(1,1) model, we have

\[ \sigma_t^2(\epsilon_{t-1}) = A + \alpha \epsilon_{t-1}^2, \]

where

\[ A = \omega + \beta \sigma^2, \quad \sigma^2 = \frac{\omega}{1 - \alpha - \beta}. \]

• This is a symmetric function of \( \epsilon_{t-1} \).
• Asymmetries may be introduced in various ways: Compared to the standard GARCH, we can change either the position of the slope of the NIC (or both).

• For example, the AGARCH captures asymmetry by allowing its NIC to be centered at a positive $\epsilon_{t-1}$, since

$$
\sigma_t^2(\epsilon_{t-1}) = A + \alpha(\epsilon_{t-1} - \theta)^2,
$$

where

$$
A = \omega + \beta \sigma^2, \quad \sigma^2 = \frac{\omega + \alpha \theta^2}{1 - \alpha - \beta}.
$$
\( \omega = 0.025, \alpha = 0.075, \beta = 0.9 \)
The GJR captures the asymmetry in the impact of news on volatility via a steeper slope for negative than for positive shocks, i.e.,

\[ \sigma_t^2(\epsilon_{t-1}) = A + \begin{cases} 
(\alpha + \theta)\epsilon_{t-1}^2 & \text{if } \epsilon_{t-1} < 0 \\
\alpha\epsilon_{t-1}^2 & \text{if } \epsilon_{t-1} \geq 0, 
\end{cases} \]

but the NIC of the GJR is still centered at zero, i.e., \( \sigma_t^2(\epsilon_{t-1}) \) is minimized for \( \epsilon_{t-1} = 0 \).
Table 1: Asymmetric GARCH(1,1) estimates for various stock return series, January 1990 to October 2009

<table>
<thead>
<tr>
<th>Series</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGARCH (Gaussian)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAC 40</td>
<td>0.0000 (0.0073)</td>
<td>0.0621 (0.0069)</td>
<td>0.9187 (0.0084)</td>
<td>0.7361 (0.0954)</td>
</tr>
<tr>
<td>DAX</td>
<td>0.0087 (0.0069)</td>
<td>0.0709 (0.0073)</td>
<td>0.9081 (0.0088)</td>
<td>0.6524 (0.0829)</td>
</tr>
<tr>
<td>FTSE</td>
<td>0.0000 (0.0036)</td>
<td>0.0673 (0.0071)</td>
<td>0.9189 (0.0079)</td>
<td>0.4693 (0.0664)</td>
</tr>
<tr>
<td>GJR–GARCH (Gaussian)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAC 40</td>
<td>0.0297 (0.0050)</td>
<td>0.0157 (0.0067)</td>
<td>0.9184 (0.0086)</td>
<td>0.0959 (0.0109)</td>
</tr>
<tr>
<td>DAX</td>
<td>0.0364 (0.0053)</td>
<td>0.0220 (0.0072)</td>
<td>0.9042 (0.0093)</td>
<td>0.1049 (0.0126)</td>
</tr>
<tr>
<td>FTSE</td>
<td>0.0119 (0.0021)</td>
<td>0.0187 (0.0064)</td>
<td>0.9227 (0.0073)</td>
<td>0.0943 (0.0104)</td>
</tr>
</tbody>
</table>
### Table 2: Maximized log–likelihood values

<table>
<thead>
<tr>
<th></th>
<th>CAC 40</th>
<th>DAX</th>
<th>FTSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>-8088.5</td>
<td>-8180.9</td>
<td>-6798.8</td>
</tr>
<tr>
<td>AGARCH</td>
<td>-8045.0</td>
<td>-8141.8</td>
<td>-6761.2</td>
</tr>
<tr>
<td>GJR–GARCH</td>
<td>-8043.8</td>
<td>-8138.5</td>
<td>-6755.2</td>
</tr>
</tbody>
</table>

| Differences in log–likelihood | AGARCH – GARCH | 43.5299 | 39.0356 | 37.6334 |
| GJR – GARCH              | 44.6940     | 42.3483 | 43.6754 |

### Table 3: Unconditional variances, $E(\sigma_{t}^{2})$

<table>
<thead>
<tr>
<th></th>
<th>CAC 40</th>
<th>DAX</th>
<th>FTSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>2.0016</td>
<td>2.2133</td>
<td>1.3231</td>
</tr>
<tr>
<td>GARCH</td>
<td>1.9542</td>
<td>2.0610</td>
<td>1.3306</td>
</tr>
<tr>
<td>AGARCH</td>
<td>1.7559</td>
<td>1.8493</td>
<td>1.0772</td>
</tr>
<tr>
<td>GJR–GARCH</td>
<td>1.6649</td>
<td>1.7070</td>
<td>1.0370</td>
</tr>
</tbody>
</table>
NICs for CAC 40

- GARCH
- AGARCH
- GJR-GARCH

\( \sigma_t^2 \) vs. \( \varepsilon_{t-1} \)
Table 4: EGARCH(1,1) estimates for various stock return series, January 1990 to October 2009

<table>
<thead>
<tr>
<th>Series</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC 40</td>
<td>0.0078 (0.0020)</td>
<td>0.1108 (0.0108)</td>
<td>0.9839 (0.0025)</td>
<td>-0.0745 (0.0067)</td>
</tr>
<tr>
<td>DAX</td>
<td>0.0129 (0.0021)</td>
<td>0.1254 (0.0110)</td>
<td>0.9798 (0.0027)</td>
<td>-0.0753 (0.0070)</td>
</tr>
<tr>
<td>FTSE</td>
<td>-0.0006 (0.0015)</td>
<td>0.1205 (0.0119)</td>
<td>0.9883 (0.0019)</td>
<td>-0.0767 (0.0070)</td>
</tr>
</tbody>
</table>

- The Model is

\[
\log \sigma_t^2 = \omega + g(\eta_{t-1}) + \beta \log \sigma_{t-1}^2,
\]  

where $\eta_t = \epsilon_t / \sigma_t$ is assumed to have a standard normal distribution, and

\[
g(\eta_{t-1}) = \theta \eta_{t-1} + \alpha(|\eta_{t-1}| - E(|\eta_{t-1}|)).
\]
<table>
<thead>
<tr>
<th></th>
<th>CAC 40</th>
<th>DAX</th>
<th>FTSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH</td>
<td>-8088.5</td>
<td>-8180.9</td>
<td>-6798.8</td>
</tr>
<tr>
<td>AGARCH</td>
<td>-8045.0</td>
<td>-8141.8</td>
<td>-6761.2</td>
</tr>
<tr>
<td>GJR–GARCH</td>
<td>-8043.8</td>
<td>-8138.5</td>
<td>-6755.2</td>
</tr>
<tr>
<td>EGARCH</td>
<td>-8041.0</td>
<td>-8137.0</td>
<td>-6746.3</td>
</tr>
</tbody>
</table>

**Differences in log–likelihood**

<table>
<thead>
<tr>
<th></th>
<th>AGARCH – GARCH</th>
<th>GJR – GARCH</th>
<th>EGARCH – GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>43.5299</td>
<td>39.0356</td>
<td>37.6334</td>
</tr>
<tr>
<td></td>
<td>44.6940</td>
<td>42.3483</td>
<td>43.6754</td>
</tr>
<tr>
<td></td>
<td>47.5567</td>
<td>43.8836</td>
<td>52.4933</td>
</tr>
</tbody>
</table>
EGARCH Examples

• The NIC is difficult to compute for the EGARCH model due to the unhandy expression for the unconditional variance.

• To visualize the degree of asymmetry, however, we may plot the function $g$ in the model

$$\log \sigma_t^2 = \omega + g(\eta_{t-1}) + \beta \log \sigma_{t-1}^2,$$

where $\eta_t = \epsilon_t / \sigma_t$ is assumed to have a standard normal distribution, and

$$g(\eta_{t-1}) = \theta \eta_{t-1} + \alpha(|\eta_{t-1}| - \mathbb{E}(|\eta_{t-1}|)).$$
EGARCH function $g(\eta_t)$ for CAC 40

EGARCH function $g(\eta_t)$ for DAX 30

EGARCH function $g(\eta_t)$ for FTSE 100
Implied Conditional Volatilities

conditional volatilities, symmetric GARCH, CAC 40

conditional volatilities, EGARCH, CAC 40
conditional volatilities, symmetric GARCH, DAX 30

conditional volatilities, EGARCH, DAX 30
conditional volatilities, symmetric GARCH, FTSE 100

conditional volatilities, EGARCH, FTSE 100
Implied Conditional Risk Measures: Value–at–Risk (VaR)

- For a given model, the VaR at level $\xi$ for period $t$, denoted by $\text{VaR}_t(\xi)$, is implicitly defined by

$$
\hat{F}(\text{VaR}_t(\xi)|I_{t-1}) = \xi,
$$

where $\hat{F}(\cdot|I_{t-1})$ is the conditional (based on information $I_{t-1}$) cumulative distribution function (cdf) implied by an estimated model.

- Thus, $\text{VaR}_t(\xi)$ is just the $\xi$–quantile of the conditional return distribution.

- Economically, this means that with probability $1 - \xi$, our loss will not exceed the $\text{VaR}_t(\xi)$.

- Under conditional normality, we have

$$
\text{VaR}_t(\xi) = \mu_t + z_{\xi}\sigma_t,
$$

where $\mu_t$ is the conditional mean of the return (often just a constant), $z_{\xi}$ is the $\xi$–quantile of the standard normal distribution (e.g., $z_{0.01} = -2.3263$), and $\sigma_t$ is the conditional standard deviation.
conditional 1% Value–at–Risk, symmetric GARCH, CAC 40

conditional 1% Value–at–Risk, EGARCH, CAC 40
conditional 1% Value-at-Risk, symmetric GARCH, FTSE 100

conditional 1% Value-at-Risk, EGARCH, FTSE 100
More Flexible Error Distributions

• In practice, we would consider the asymmetric GARCH models with more flexible error distributions allowing for fat tails and/or skewness where appropriate.
APARCH Model (Asymmetric Power ARCH)

- A rather general GARCH process was suggested by Ding, Granger, and Engle (1993).²

- They considered a very long time series of the S&P 500 returns from 1928 to 1991 and studied the properties of autocorrelations of the form

\[ \varrho(\tau; \delta) = \text{Corr}(|r_t|^\delta, |r_{t-\tau}|^\delta), \quad \delta > 0. \]  

(11)

- Let us consider the same time series as Ding, Granger, and Engle (1993), but from 1928 to 2001, so we have 19600 daily observations.

---

Autocorrelations of power–transformed absolute returns

• Let us consider the sample autocorrelations of the absolute returns, $|r_t|$, and those of the squared returns, $r_t^2$.

• We may also fix $\tau$ and consider the autocorrelations as functions of $\delta$.

• The observation that the autocorrelations are often strongest for $\delta \approx 1$ is sometimes referred to as the *Taylor effect*.

• It may suggest to use a model of the form

$$\sigma_t = \omega + \alpha |\epsilon_{t-1}| + \beta \sigma_{t-1}$$

rather than

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$
Sample autocorrelations of absolute and squared returns

lag, $\tau$

absolute returns
squared returns
A general model is the asymmetric power GARCH(1,1)\(^3\)

\[
\sigma_t^\delta = \omega + \alpha(|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta + \beta \sigma_{t-1}^\delta, \quad \gamma \in (-1, 1),
\]

where \(\delta > 0\) is a parameter to be estimated.

The leverage effect is captured by parameter \(\gamma\):

- For \(\epsilon_{t-1} < 0\), the reaction coefficient is \(\alpha(1 + \gamma)^2\).
- For \(\epsilon_{t-1} > 0\), the reaction coefficient is \(\alpha(1 - \gamma)^2\),

so the slope of the NIC is different for positive and negative shocks.

\(^3\)Generalization to asymmetric power GARCH\((p, q)\) is straightforward.
<table>
<thead>
<tr>
<th>Series</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\delta}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC 40</td>
<td>0.0245 (0.0040)</td>
<td>0.0606 (0.0073)</td>
<td>0.9300 (0.0074)</td>
<td>0.6278 (0.0937)</td>
<td>1.2466 (0.1513)</td>
<td>$\infty$</td>
</tr>
<tr>
<td>DAX</td>
<td>0.0302 (0.0041)</td>
<td>0.0702 (0.0071)</td>
<td>0.9196 (0.0076)</td>
<td>0.5814 (0.0734)</td>
<td>1.2216 (0.1111)</td>
<td>$\infty$</td>
</tr>
<tr>
<td>FTSE</td>
<td>0.0121 (0.0020)</td>
<td>0.0640 (0.0067)</td>
<td>0.9355 (0.0065)</td>
<td>0.6333 (0.0918)</td>
<td>1.1182 (0.1527)</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Series</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\delta}$</th>
<th>$\hat{\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC 40</td>
<td>0.0198 (0.0036)</td>
<td>0.0636 (0.0073)</td>
<td>0.9318 (0.0068)</td>
<td>0.6728 (0.0912)</td>
<td>1.1697 (0.1380)</td>
<td>12.2296 (1.7009)</td>
</tr>
<tr>
<td>DAX</td>
<td>0.0187 (0.0034)</td>
<td>0.0806 (0.0079)</td>
<td>0.9210 (0.0074)</td>
<td>0.5176 (0.0695)</td>
<td>1.1192 (0.1138)</td>
<td>9.6036 (1.0677)</td>
</tr>
<tr>
<td>FTSE</td>
<td>0.0117 (0.0021)</td>
<td>0.0613 (0.0073)</td>
<td>0.9359 (0.0068)</td>
<td>0.6647 (0.1045)</td>
<td>1.2014 (0.1705)</td>
<td>14.8298 (2.4877)</td>
</tr>
</tbody>
</table>
Table 7: APARCH Maximized log–likelihood values; for \( \delta = 1 \) and \( \delta = 2 \), the difference to “\( \delta \) free” is reported

<table>
<thead>
<tr>
<th></th>
<th>Gaussian</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CAC 40</td>
<td>DAX</td>
<td>FTSE</td>
<td></td>
<td>CAC 40</td>
<td>DAX</td>
</tr>
<tr>
<td>( \delta ) free</td>
<td>−8035.6</td>
<td>−8123.2</td>
<td>−6745.0</td>
<td>( \delta ) free</td>
<td>−7977.3</td>
<td>−8001.2</td>
</tr>
<tr>
<td>( \delta = 1 )</td>
<td>1.5259</td>
<td>2.2772</td>
<td>0.3251</td>
<td>( \delta = 1 )</td>
<td>0.8287</td>
<td>0.5854</td>
</tr>
<tr>
<td>( \delta = 2 )</td>
<td>8.2124</td>
<td>15.2846</td>
<td>10.1215</td>
<td>( \delta = 2 )</td>
<td>11.4483</td>
<td>18.3154</td>
</tr>
</tbody>
</table>

- For normal innovations, a likelihood ratio test would reject \( H_0 : \delta = 1 \) at the 5% level for the DAX, and at the 10% level for the CAC 40.\(^4\)

\(^4\)5% critical value with 1 degree of freedom is 3.8415.
• For Student's $t$ innovations, $H_0 : \delta = 1$ is never rejected at the 10% level.

• $H_0 : \delta = 2$ is always rejected against the more general model with freely estimated $\delta$. 
Long Memory in Volatility?

- Go back to the ACFs of the S&P 500.

- It might appear as if there is something like “long memory”.

- That is, the autocorrelations are significantly positive over very long lags.

- Moreover, the autocorrelations decrease fast in the first few weeks and then decrease very slowly.

- This is typical for long memory processes, where the autocorrelations exhibit an (asymptotically) hyperbolic decay of the autocorrelation function, where

\[ \rho(\tau) \approx c\tau^{-h}, \]

where \( c > 0 \) is constant and \( 0 < h \leq 1 \).
$0.5 \times \tau^{-0.4}$
• We have seen that the ACF of GARCH models mimics that of a usual ARMA, i.e., it tends to zero at an exponential rate, i.e., much faster.

• It might seem at first sight that the GARCH model may thus not be appropriate to capture the dependence structure in volatility.

• Long memory GARCH models have been developed (e.g., FIGARCH).

• Also, Ding and Granger (1996) proposed the Component GARCH Model, which is able to better approximate the aforementioned autocorrelation pattern.
The Component GARCH Model

- The underlying idea is that the dependence structure of returns is rooted in the presence of different volatility components.

- With two components, the component GARCH(1,1) process (in absolute values) is defined as

  \[
  \sigma_t = \sigma_{1,t} + \sigma_{2,t} \quad (12)
  \]

  \[
  \sigma_{1,t} = \omega + \alpha_1 |\epsilon_{t-1}| + \beta_1 \sigma_{1,t-1}, \quad \omega > 0, \quad \alpha_1, \beta_1 \geq 0 \quad (13)
  \]

  \[
  \sigma_{2,t} = \alpha_2 |\epsilon_{t-1}| + \beta_2 \sigma_{2,t-1}, \quad \alpha_2, \beta_2 \geq 0. \quad (14)
  \]

- One volatility component may have a big short-run effect, but dies out very quickly (fast decay at the beginning), and one component may have a relatively small short-run effect, but it lasts for a long time period (slow decay at higher lags).

- We can write

  \[
  \sigma_{1,t} = \frac{\omega}{1 - \beta_1} + \frac{\alpha_1 |\epsilon_{t-1}|}{1 - \beta_1 L}, \quad \sigma_{2,t} = \frac{\alpha_2 |\epsilon_{t-1}|}{1 - \beta_2 L},
  \]
so

\[
\sigma_t = \frac{\omega}{1 - \beta_1} + \frac{\alpha_1 |\epsilon_{t-1}|}{1 - \beta_1 L} + \frac{\alpha_2 |\epsilon_{t-1}|}{1 - \beta_2 L},
\]

and multiplying through by \((1 - \beta_1 L)(1 - \beta_2 L)\) gives

\[
\sigma_t = \omega(1 - \beta_2) + (\alpha_1 + \alpha_2)|\epsilon_{t-1}| - (\alpha_1 \beta_2 + \alpha_2 \beta_1)|\epsilon_{t-2}|
+ (\beta_1 + \beta_2)\sigma_{t-1} - \beta_1 \beta_2 \sigma_{t-2},
\]

which shows that the process has a GARCH(2,2) representation with negative coefficients of \(|\epsilon_{t-2}|\) and \(\sigma_{t-2}\).

- Consider a general GARCH(2,2) process,

\[
\sigma_t = \omega + \alpha_1 |\epsilon_{t-1}| + \alpha_2 |\epsilon_{t-2}| + \beta_1 \sigma_{t-1} + \beta_2 \sigma_{t-2},
\]

with ACF of absolute returns (from the ARMA representation of the GARCH process)

\[
\rho(\tau; 1) = c_1 \lambda_1^\tau + c_2 \lambda_2^\tau,
\]
where \( \lambda_1 \) and \( \lambda_2 \) are the roots of

\[
\lambda^2 - (\alpha_1 + \beta_1)\lambda - (\alpha_2 + \beta_2) = 0
\]

i.e.,

\[
\lambda_{1/2} = \frac{\alpha_1 + \beta_1 \pm \sqrt{(\alpha_1 + \beta_1)^2 + 4(\alpha_2 + \beta_2)}}{2}.
\]

- Clearly \( \lambda_1 \) and \( \lambda_2 \) can both be positive only if the sum of the lag–two coefficients is negative.

- Then one of the roots (the smaller one) captures the fast decay at the beginning, and the other root captures the slow decay at higher lags.

- This approximates long memory (hyperbolic decay).

- Thus the ability of the component model to mimic a long memory–like shape of the ACF is a result that the lag–two coefficients in the GARCH(2,2) representation are negative.
• When we estimate the model for the S&P 500 returns, we get

\[
\begin{align*}
\sigma_{1t} &= 0.018 + 0.127|\epsilon_{t-1}| + 0.835\sigma_{1,t-1} \\
\sigma_{2t} &= 0.004|\epsilon_{t-1}| + 0.991\sigma_{2,t-1},
\end{align*}
\]

where the first component has greater short–run impact, but decays very fast, and the second has smaller short–run impact, but decays slowly.

• The ACF of the absolute values is

\[
\rho(\tau; 1) = 0.139(0.932)^\tau + 0.139(0.999)^\tau.
\]

• The first term decays fast, and the second decays slowly.

Empirical and theoretical autocorrelations of absolute S&P 500 returns

- Empirical
- PGARCH\(^+(2,2;1)\)
- ComPGARCH(1)
Empirical and theoretical autocorrelations of absolute NYSE returns

- **empirical**
- **PGARCH**$^+(2,2;1)$
- **ComPGARCH(1)**

The graph shows the autocorrelation function for lag $\tau$ with data points and fitted models for empirical and theoretical autocorrelations.
Empirical and theoretical autocorrelations of absolute DJIA returns

- **Empirical**
- **PGARCH**$^+(2,2;1)$
- **ComPGARCH(1)**

The graph shows the autocorrelation coefficients as a function of lag, with theoretical models compared to empirical data. The x-axis represents the lag, while the y-axis shows the autocorrelation coefficient, $\rho(\tau)$. The empirical data is depicted by a solid line, whereas the theoretical models are represented by dashed lines.
Long Memory?

• After all, the apparent long memory effects in the volatility of the S&P 500 (and other long time series) may also be due to structural breaks in the volatility process over this long time period.

• Intuitively, suppose volatility is driven by a GARCH process up to time $t_1$,

$$\sigma_t^2 = \omega_1 + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t = 1, \ldots, t_1. \quad (16)$$

• At time $t_1 + 1$, there is a structural break, that is,

$$\sigma_t^2 = \omega_2 + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad t \geq t_1 + 1, \quad (17)$$

where $\omega_1 \neq \omega_2$.

• If this structural break in the volatility level is not accounted for, it may appear as a shock to volatility that never dies out (⇒ apparent long memory).
ARCH–M

- In the finance literature, a link is often made between the expected return and the risk of an asset.

- Investors are willing to hold risky assets only if their expected return compensate for the risk.

- A model that incorporates this link is the GARCH–in–mean or GARCH–M model, which can be written as

\[ r_t = \delta g(\sigma_t^2) + \epsilon_t, \]

where \( \epsilon_t \) is a GARCH error process, and \( g \) is a known function such as \( g(\sigma_t^2) = \sigma_t^2 \), \( g(\sigma_t^2) = \sigma_t \), or \( g(\sigma_t^2) = \log(\sigma_t^2) \).

- If \( \delta > 0 \) and \( g \) is monotonically increasing, then the term \( \delta g(\sigma_t^2) \) can be interpreted as a risk premium that increases expected returns if conditional volatility \( \sigma_t^2 \) is high.

- In practice \( g(\sigma_t^2) = \sigma_t \) appears to be the preferred specification.