3 Empirical Regularities of Asset Returns

1. Thick tails

   (a) Excess kurtosis decreases with aggregation

2. Volatility clustering.

   (a) Large changes followed by large changes; small changes followed by small changes

3. Leverage effects

   (a) changes in prices often negatively correlated with changes in volatility

4. Non trading periods
   (a) Volatility is smaller over periods when markets are closed than when they are open

5. Forecastable events
   (a) Forecastable releases of information are associated with high ex ante volatility

6. Volatility and serial correlation
   (a) Inverse relationship between volatility and serial correlation of stock indices

7. Volatility co-movements

   (a) Evidence of common factors to explain volatility in multiple series

4 Engle's ARCH(p) Model

The ARCH(p) model for $y_t = \ln P_t - \ln P_{t-1}$ is

$$
\begin{aligned}
y_t &= E_{t-1}[y_t] + \epsilon_t,\ \epsilon_t | I_{t-1} \sim iid\ (0, \sigma_t^2) \\
\sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 + \cdots + a_p \epsilon_{t-p}^2,\ a_0 > 0, a_i \geq 0 \\
&= a_0 + a(L)\epsilon_t^2,\ a(L) = \sum_{j=1}^{p} a_j L^j
\end{aligned}
$$

Alternative error specification

$$
\begin{aligned}
\epsilon_t &= z_t \sigma_t \\
& \quad z_t \sim iid\ (0, 1) \\
\sigma_t^2 &= a_0 + a(L)\epsilon_t^2
\end{aligned}
$$

Remark: The random variable z_t doesn't have to be normal. It can have a fat-tailed distribution; e.g. Student's-t

4.1 Properties of ARCH Errors

Note: Derivations utilize heavily the law of iterated expectations (note: $E_{t-1}[\varepsilon_t] = E[\varepsilon_t | I_{t-1}]$)

- $\{\epsilon_t, I_{t-1}\}$ is a MDS with conditionally heteroskedastic errors

$$
\begin{aligned}
E[\epsilon_t | I_{t-1}] &= E[z_t \sigma_t | I_{t-1}] = \sigma_t E[z_t | I_{t-1}] = 0 \\
\text{var}(\epsilon_t | I_{t-1}) &= E[\epsilon_t^2 | I_{t-1}] = \sigma_t^2 E[z_t^2 | I_{t-1}] = \sigma_t^2 \\
E[\epsilon_t^m | I_{t-1}] &= 0 \text{ for } m \text{ odd.}
\end{aligned}
$$

Since $\epsilon_t \sim$ MDS it is an uncorrelated process: $E[\epsilon_t \epsilon_{t-j}] = 0$ for $j = 1, 2, \ldots$.

- The error ϵ_t is stationary with mean zero and constant unconditional variance

$$
\begin{aligned}
E[\epsilon_t] &= E[E[z_t\sigma_t|I_{t-1}]] \\
&= E[\sigma_t E[z_t|I_{t-1}]] = 0 \\
\text{var}(\epsilon_t) &= E[\epsilon_t^2] = E[E[z_t^2\sigma_t^2|I_{t-1}]] \\
&= E[\sigma_t^2 E[z_t^2|I_{t-1}]] = E[\sigma_t^2]
\end{aligned}
$$

Assuming stationarity

$$
\begin{aligned}
E[\sigma_t^2] &= E[a_0 + a(L)\epsilon_t^2] \\
&= a_0 + a_1 E[\epsilon_{t-1}^2] + \cdots + a_p E[\epsilon_{t-p}^2] \\
&= a_0 + a_1 E[\sigma_t^2] + \cdots + a_p E[\sigma_t^2]
\end{aligned}
$$

which implies that

$$
E[\sigma_t^2] = \bar{\sigma}^2 = \frac{a_0}{1 - a_1 - \cdots - a_p} = \frac{a_0}{a(1)}, a(1) > 0
$$

- ϵ_t is *leptokurtic*

$$\begin{aligned}
E[\epsilon_t^4] &= E[\sigma_t^4 E[z_t^4 | I_{t-1}]] = E[\sigma_t^4] \cdot 3 \\
&\geq \left(E[\sigma_t^2]\right)^2 \cdot 3 \text{ by Jensen's inequality} \\
&= \left(E[\epsilon_t^2]\right)^2 \cdot 3 \\
&\Rightarrow \frac{E[\epsilon_t^4]}{\left(E[\epsilon_t^2]\right)^2} > 3
\end{aligned}$$

That, is

$$\text{kurt}(\epsilon_t) > 3 = \text{kurt}(\text{normal})$$

- σ_t^2 is a serially correlated random variable

$$\begin{aligned} \sigma_t^2 &= a_0 + a(L)\epsilon_t^2, \\ E[\sigma_t^2] &= \frac{a_0}{1 - a(1)} = \bar{\sigma}^2 \end{aligned}$$

Using $a_0 = (1 - a(1))\bar{\sigma}^2$, σ_t^2 may be expressed as

$$\sigma_t^2 - \bar{\sigma}^2 = a(L)(\epsilon_t^2 - \bar{\sigma}^2)$$

- ϵ_t^2 has a stationary AR(p) representation.

$$\begin{aligned} \sigma_t^2 + \epsilon_t^2 &= a_0 + a(L)\epsilon_t^2 + \epsilon_t^2 \\ &\Rightarrow \epsilon_t^2 = a_0 + a(L)\epsilon_t^2 + (\epsilon_t^2 - \sigma_t^2) \end{aligned}$$

where $(\epsilon_t^2 - \sigma_t^2) = v_t$ is a conditionally heteroskedastic MDS.

- ϵ_t^2 exhibits *volatility mean reversion.*

Example: Consider ARCH(1) with $0 < a < 1$

$$\begin{aligned}
\sigma_t^2 &= a_0 + a\epsilon_{t-1}^2 \\
E[\epsilon_t^2] &= E[\sigma_t^2] = \bar{\sigma}^2 = a_0/(1-a) \Rightarrow \\
(\epsilon_t^2 - \bar{\sigma}^2) &= a(\epsilon_{t-1}^2 - \bar{\sigma}^2) + v_t \Rightarrow \\
E[\epsilon_{t+k}^2 | I_{t-1}] - \bar{\sigma}^2 &= a^k(\epsilon_{t-1}^2 - \bar{\sigma}^2) \to 0 \text{ as } k \to \infty
\end{aligned}$$

5 Bollerslev's GARCH Model

Idea: ARCH is like an AR model for volatility. GARCH is like an ARMA model for volatility.

The GARCH(p, q) model is

$$\begin{aligned}
\epsilon_t &= z_t \sigma_t, \ z_t \sim iid\,(0,1) \\
\sigma_t^2 &= a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2, \ a_0 > 0 \\
a(L) &= a_1 L + \cdots + a_p L^p, \ a_i \geq 0 \\
b(L) &= b_1 L + \cdots + b_q L^q, \ b_j \geq 0
\end{aligned}$$

Note: for identification of β_j, must have at least one ARCH coefficient $a_i > 0$

5.1 Properties of GARCH model

- GARCH(p,q) is equivalent to ARCH(∞). If $1-b(z)=0$ has all roots outside unit circle then

$$\begin{aligned}\sigma_t^2 &= \frac{a_0}{1-b(1)}+\frac{a(L)}{1-b(L)}\epsilon_t^2 \\ &= a_0^*+\delta(L)\epsilon_t^2,\ \delta(L)=\sum_{k=0}^{\infty}\delta_k L^k\end{aligned}$$

- ϵ_t is a stationary and ergodic MDS with finite variance provided $a(1)+b(1)<1$

$$\begin{aligned}E[\epsilon_t] &= 0 \\ \mathrm{var}(\epsilon_t) &= E[\epsilon_t^2]=\frac{a_0}{1-a(1)-b(1)} \\ \epsilon_t^2\sim \mathrm{ARMA}(m,q),\ m &= \max(p,q)\end{aligned}$$

5.2 GARCH(1,1)

The most commonly used GARCH(p,q) model is the GARCH(1,1)

$$\sigma_t^2 = a_0 + a_1\epsilon_t^2 + b_1\sigma_t^2$$

Properties:

$$\begin{aligned}
\text{stationarity condition:} \quad & a_1 + b_1 < 1 \\
\text{ARCH}(\infty) \quad : \quad & a_i = a_1 b_1^{i-1} \\
\text{ARMA(1,1):} \quad & \epsilon_t^2 = a_0 + (a_1 + b_1)\epsilon_{t-1}^2 + u_t - b_1 u_{t-1}, \\
u_t \quad = \quad & \epsilon_t^2 - E_{t-1}(\epsilon_t^2) \\
\text{unconditional variance} \quad : \quad & \bar{\sigma}^2 = a_0/(1 - a_1 - b_1)
\end{aligned}$$

5.3 Conditional Mean Specification

- $E_{t-1}[y_t]$ is typically specified as a constant or possibly a low order ARMA process to capture autocorrelation caused by market microstructure effects (e.g., bid-ask bounce) or non-trading effects.

- If extreme or unusual market events have happened during sample period, then dummy variables associated with these events are often added to the conditional mean specification to remove these effects. The typical conditional mean specification is of the form

$$E_{t-1}[y_t] = c + \sum_{i=1}^{r} \phi_i y_{t-i} + \sum_{j=1}^{s} \theta_j \epsilon_{t-j} + \sum_{l=0}^{L} \boldsymbol{\beta}_l' \mathbf{x}_{t-l} + \epsilon_t,$$

where $\mathbf{x}_t$ is a $k \times 1$ vector of exogenous explanatory variables.

5.4 Explanatory Variables in the Conditional Variance Equation

- Exogenous explanatory variables may also be added to the conditional variance formula

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2 + \sum_{k=1}^{K} \boldsymbol{\delta}_k' \mathbf{z}_{t-k},$$

  where $\mathbf{z}_t$ is a $m \times 1$ vector of variables, and $\boldsymbol{\delta}$ is a $m \times 1$ vector of positive coefficients.

- Variables that have been shown to help predict volatility are trading volume, interest rates, macroeconomic news announcements, implied volatility from option prices and realized volatility, overnight returns, and after hours realized volatility

5.5 GARCH-in-Mean (GARCH-M)

Idea: Modern finance theory suggests that volatility may be related to risk premia on assets

The GARCH-M model allows time-varying volatility to be realted to expected returns

$$\begin{aligned} y_t &= c + \alpha g(\sigma_t) + \epsilon_t \\ & \epsilon_t \sim \text{ GARCH} \\ g(\sigma_t) &= \begin{cases} \sigma_t \\ \sigma_t^2 \\ \ln(\sigma_t^2) \end{cases} \end{aligned}$$

5.6 Temporal Aggregation

- Volatility clustering and non-Gaussian behavior in financial returns is typically seen in weekly, daily or intraday data. The persistence of conditional volatility tends to increase with the sampling frequency.

- For GARCH models there is no simple aggregation principle that links the parameters of the model at one sampling frequency to the parameters at another frequency. This occurs because GARCH models imply that the squared residual process follows an ARMA type process with MDS innovations which is not closed under temporal aggregation.

- The practical result is that GARCH models tend to be fit to the frequency at hand. This strategy, however, may not provide the best out-of-sample volatility forecasts. For example,Martens (2002) showed that a GARCH model fit to S&P 500 daily returns produces better forecasts of weekly and monthly volatility than GARCH models fit to weekly or monthly returns, respectively.

6 Testing for ARCH Effects

Consider testing the hypotheses

$$\begin{aligned} H_0 &: \text{(No ARCH) } a_1 = a_2 = \cdots = a_p = 0 \\ H_1 &: \text{(ARCH) at least one } a_i \neq 0 \end{aligned}$$

Engle derived a simple LM test

- Step 1: Compute squared residuals ϵ_t from mean equation regression

- Step 2: Estimate auxiliary regression

$$\hat{\epsilon}_t^2 = a_0 + a_1 \hat{\epsilon}_{t-1}^2 + \cdots + a_p^2 \hat{\epsilon}_{t-p}^2 + error_t$$

- Step 3. Form the LM test statistic

$$LM_{ARCH} = T \cdot R^2_{AUX}$$

where $T =$ sample size from auxiliary regression. Under H_0 :(No ARCH) $LM_{ARCH} \overset{A}{\sim} \chi^2(p)$

Remark:

- Test has power against GARCH(p, q) alternatives

7 Estimating GARCH by MLE

Consider estimating the model

$$\begin{aligned} y_t &= E_t[y_{t-1}] + \epsilon_t = \mathbf{x}_t'\boldsymbol{\beta} + \epsilon_t \\ \epsilon_t &= z_t\sigma_t,\ z_t \sim iid\ N(0,1) \\ \sigma_t^2 &= a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2 \end{aligned}$$

Result: The regression parameters $\boldsymbol{\beta}$ and GARCH parameters

$\boldsymbol{\gamma} = (a_0, a_1, \ldots, a_p, b_1, \ldots, b_q)'$ can be estimated separately because the information matrix for $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')'$ is block diagonal.

- Step 1: Estimate $\boldsymbol{\beta}$ by OLS ignoring ARCH errors and form residuals $\hat{\epsilon}_t = y_t - \mathbf{x}_t'\hat{\boldsymbol{\beta}}$

- Step 2: Estimate ARCH process for residuals $\hat{\epsilon}_t$ by mle.

Warning: Block diagonality of information matrix fails if

- pdf of z_t is not a symmetric density

- $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are not variation free; e.g. GARCH-M model

7.1 GARCH Likelihood Function Under Normality

Assume $E_t[y_{t-1}] = 0$. Let $\boldsymbol{\theta} = (a_0, a_1, \ldots, a_p, b_1, \ldots, b_q)'$ denote the parameters to be estimated. Since $\epsilon_t = z_t \sigma_t$

$$
\begin{aligned}
f(\epsilon_t | I_{t-1}; \boldsymbol{\theta}) &= f(z_t) \left| \frac{dz_t}{d\epsilon_t} \right| = f\left(\frac{\epsilon_t}{\sigma_t}\right) \left| \frac{1}{\sigma_t} \right| \\
&= (2\pi\sigma_t^2)^{-1/2} \exp\left\{ \frac{-1}{2\sigma_t^2} \epsilon_t^2 \right\}
\end{aligned}
$$

For a sample of size T, the prediction error decomposition gives

$$\begin{aligned}
& f(\epsilon_T, \epsilon_{T-1}, \ldots, \epsilon_1; \boldsymbol{\theta}) \\
= & \left(\prod_{t=m+1}^{T} f(\epsilon_t | I_{t-1}; \boldsymbol{\theta}) \right) \cdot f(\epsilon_1, \ldots, \epsilon_m; \boldsymbol{\theta}) \\
= & \left(\prod_{t=m+1}^{T} (2\pi\sigma_t^2)^{-1/2} \exp\left\{ \frac{-1}{2\sigma_t^2} \epsilon_t^2 \right\} \right) \cdot f(\epsilon_1, \ldots, \epsilon_m; \boldsymbol{\theta})
\end{aligned}$$

where $\sigma_t^2 = a_0 + a(L)\epsilon_t^2 + b(L)\sigma_t^2$ may be evaluated recursively given starting values for σ_t^2. The log-likelihood function is

$$\begin{aligned}
& \frac{-(T-m+1)}{2} \ln(2\pi) - \sum_{t=m+1}^{T} \left[\frac{1}{2} \ln(\sigma_t^2) + \frac{1}{2} \frac{\epsilon_t^2}{\sigma_t^2} \right] \\
& + \ln(f(\epsilon_1, \ldots, \epsilon_m; \boldsymbol{\gamma})
\end{aligned}$$

Problem: the marginal density for the initial values $f(\varepsilon_1, \ldots, \varepsilon_m; \boldsymbol{\theta})$ does not have a closed form expression so exact mle is not possible. In practice, the initial values $\varepsilon_1, \ldots, \varepsilon_m$ are set equal to zero and the marginal density $f(\varepsilon_1, \ldots, \varepsilon_m; \boldsymbol{\theta})$ is ignored. This is conditional mle.

7.1.1 Practical issues

- Starting values for the model parameters a_i $(i = 0, \cdots, p)$ and b_j $(j = 1, \cdots, q)$ need to be chosen and an initialization of ϵ_t^2 and σ_t^2 must be supplied.

- Zero values are often given for the conditional variance parameters other than a_0 and a_1, and a_0 is set equal to the unconditional variance of y_t. For the initial values of σ_t^2, a popular choice is

$$\sigma_t^2 = \epsilon_t^2 = \frac{1}{T} \sum_{s=m+1}^{T} y_s^2, \ t \leq m,$$

- Once the log-likelihood is initialized, it can be maximized using numerical optimization techniques. The most common method is based on a Newton-Raphson iteration of the form

$$\hat{\boldsymbol{\theta}}_{n+1} = \hat{\boldsymbol{\theta}}_n - \lambda_n \mathbf{H}(\hat{\boldsymbol{\theta}}_n)^{-1} \mathbf{s}(\hat{\boldsymbol{\theta}}_n),$$

- For GARCH models, the BHHH algorithm is often used. This algorithm approximates the Hessian matrix using only first derivative information

$$-\mathbf{H}(\boldsymbol{\theta}) \approx \mathbf{B}(\boldsymbol{\theta}) = \sum_{t=1}^{T} \frac{\partial l_t}{\partial \boldsymbol{\theta}} \frac{\partial l_t}{\partial \boldsymbol{\theta}'}.$$

- Under suitable regularity conditions, the ML estimates are consistent and asymptotically normally distributed and an estimate of the asymptotic covariance matrix of the ML estimates is constructed from an estimate of the final Hessian matrix from the optimization algorithm used.

7.2 Numerical Accuracy of GARCH Estimates

- GARCH estimation is widely available in a number of commercial software packages (e.g. EVIEWS, GAUSS, MATLAB, Ox, RATS, S-PLUS, TSP) and there are also a few free open source implementations. (Even Excel!)

- Starting values, optimization algorithm choice, and use of analytic or numerical derivatives, and convergence criteria all influence the resulting numerical estimates of the GARCH parameters.

- The GARCH log-likelihood function is not always well behaved, especially in complicated models with many parameters, and reaching a global maximum of the log-likelihood function is not guaranteed using standard optimization techniques. Poor choice of starting values can lead to an ill-behaved log-likelihood and cause convergence problems.

- In many empirical applications of the GARCH(1,1) model, the estimate of a_1 is close to zero and the estimate of b_1 is close to unity. This situation is of some concern since the GARCH parameter b_1 becomes unidentified if $a_1 = 0$, and it is well known that the distribution of ML estimates can become ill-behaved in models with nearly unidentified parameters.

- Ma, Nelson and Startz (2007) studied the accuracy of ML estimates of the GARCH parameters a_0, a_1 and b_1 when a_1 is close to zero. They found that the estimated standard error for b_1 is spuriously small and that the t-statistics for testing hypotheses about the true value of b_1 are severely size distorted. They also showed that the concentrated loglikelihood as a function of b_1 exhibits multiple maxima.

- To guard against spurious inference they recommended comparing estimates from pure ARCH(p) models, which do not suffer from the identification problem, with estimates from the GARCH(1,1). If the volatility dynamics from these models are similar then the spurious inference problem is not likely to be present.

7.3 Quasi-Maximum Likelihood Estimation

- The assumption of conditional normality is not always appropriate.

- However, even when normality is inappropriately assumed, maximizing the Gaussian log-likelihood results in quasi-maximum likelihood estimates (QMLEs) that are consistent and asymptotically normally distributed provided the conditional mean and variance functions of the GARCH model are correctly specified.

- An asymptotic covariance matrix for the QMLEs that is robust to conditional non-normality is estimated using
$$\mathbf{H}(\hat{\boldsymbol{\theta}}_{QML})^{-1}\mathbf{B}(\hat{\boldsymbol{\theta}}_{QML})\mathbf{H}(\hat{\boldsymbol{\theta}}_{QML})^{-1},$$
where $\hat{\boldsymbol{\theta}}_{QML}$ denotes the QMLE of $\boldsymbol{\theta}$, and is often called the "sandwich" estimator.

7.3.1 Determining lag length

- Use model selection criteria (AIC or BIC)

- For GARCH(p, q) models, those with $p, q \leq 2$ are typically selected by AIC and BIC.

- Low order GARCH(p,q) models are generally preferred to a high order ARCH(p) for reasons of parsimony and better numerical stability of estimation (high order GARCH(p, q) processes often have many local maxima and minima).

- For many applications, it is hard to beat the simple GARCH(1,1) model.

7.4 Model Diagnostics

Correct model specification implies

$$\frac{\hat{\epsilon}_t}{\hat{\sigma}_t} \sim iid\ N(0,1)$$

- Test for normality - Jarque-Bera, QQ-plot

- Test for serial correlation - Ljung-box, SACF, SPACF

- Test for ARCH effects - serial correlation in squared standardized residuals, LM test for ARCH

7.5 GARCH and Forecasts for the Conditional Mean

- Suppose one is interested in forecasting future values of y_T in the standard GARCH model. For simplicity assume that $E_T[y_{T+1}] = c$. Then the minimum mean squared error $h-$ step ahead forecast of y_{T+h} is just c, which does not depend on the GARCH parameters, and the corresponding forecast error is

$$\epsilon_{T+h} = y_{T+h} - E_T[y_{T+h}].$$

- The conditional variance of this forecast error is then

$$\mathrm{var}_T(\epsilon_{T+h}) = E_T[\sigma^2_{T+h}],$$

which does depend on the GARCH parameters. Therefore, in order to produce confidence bands for the $h-$step ahead forecast the $h-$step ahead volatility forecast $E_T[\sigma^2_{T+h}]$ is needed.

7.6 Forecasting From GARCH Models

Consider the basic GARCH(1,1) model

$$\sigma_t^2 = a_0 + a_1\epsilon_{t-1}^2 + b_1\sigma_{t-1}^2$$

from $t = 1, \ldots, T$. The best linear predictor of σ_{T+1}^2 using information at time T is

$$\begin{aligned} E[\sigma_{T+1}^2|I_T] &= a_0 + a_1E[\epsilon_T^2|I_T] + b_1E[\sigma_T^2|I_T] \\ &= a_0 + a_1\epsilon_T^2 + b_1\sigma_T^2 \end{aligned}$$

Using the chain-rule of forecasting and $E[\epsilon_{T+1}^2|I_T] = E[\sigma_{T+1}^2|I_T]$

$$\begin{aligned} E[\sigma_{T+2}^2|I_T] &= a_0 + a_1E[\epsilon_{T+1}^2|I_T] + b_1E[\sigma_{T+1}^2|I_T] \\ &= a_0 + (a_1 + b_1)E[\sigma_{T+1}^2|I_T] \end{aligned}$$

In general, for $k \geq 2$

$$
\begin{aligned}
E[\sigma^2_{T+k}|I_T] &= a_0 + (a_1 + b_1)E[\sigma^2_{T+k-1}|I_T] \\
&= a_0 \sum_{i=0}^{k-1} (a_1 + b_1)^i + (a_1 + b_1)^{k-1}(a_1\epsilon^2_T + b_1\sigma^2_T).
\end{aligned}
$$

Note: as $k \longrightarrow \infty$

$$
E[\sigma^2_{T+k}|I_T] \longrightarrow E[\sigma^2_T] = \frac{a_0}{1 - a_1 - b_1}
$$

An alternative representation of the forecasting equation starts with the mean-adjusted form

$$
\sigma^2_{T+1} - \bar{\sigma}^2 = a_1(\epsilon^2_T - \bar{\sigma}^2) + b_1(\sigma^2_T - \bar{\sigma}^2),
$$

where $\bar{\sigma}^2 = a_0/(1 - a_1 - b_1)$ is the unconditional variance. Then by recursive substitution

$$
E_T[\sigma^2_{T+k}] - \bar{\sigma}^2 = (a_1 + b_1)^{k-1}(E[\sigma^2_{T+1}] - \bar{\sigma}^2).
$$

7.8 Asymmetric Leverage Effects and News Impact

- In the basic GARCH model, since only squared residuals ϵ_{t-i}^2 enter the conditional variance equation, the signs of the residuals or shocks have no effect on conditional volatility.

- A stylized fact of financial volatility is that bad news (negative shocks) tends to have a larger impact on volatility than good news (positive shocks). That is, volatility tends to be higher in a falling market than in a rising market. Black (1976) attributed this effect to the fact that bad news tends to drive down the stock price, thus increasing the leverage (i.e., the debt-equity ratio) of the stock and causing the stock to be more volatile. Based on this conjecture, the asymmetric news impact on volatility is commonly referred to as the leverage effect.

7.8.1 Testing for Asymmetric Effects on Conditional Volatility

- A simple diagnostic for uncovering possible asymmetric leverage effects is the sample correlation between r_t^2 and r_{t-1}. A negative value of this correlation provides some evidence for potential leverage effects.

- Other simple diagnostics, result from estimating the following test regression

$$\hat{\epsilon}_t^2 = \beta_0 + \beta_1 \hat{w}_{t-1} + \xi_t,$$

where $\hat{w}_{t-1}$ is a variable constructed from $\hat{\epsilon}_{t-1}$ and the sign of $\hat{\epsilon}_{t-1}$. A significant value of β_1 indicates evidence for asymmetric effects on conditional volatility.

- Let S_{t-1}^{-} denote a dummy variable equal to unity when $\hat{\epsilon}_{t-1}$ is negative, and zero otherwise. Engle and Ng consider three tests for asymmetry. Setting $\hat{w}_{t-1} = S_{t-1}^{-}$ gives the Sign Bias test; setting $\hat{w}_{t-1} = S_{t-1}^{-}\hat{\epsilon}_{t-1}$ gives the Negative Size Bias test; and setting $\hat{w}_{t-1} = S_{t-1}^{+}\hat{\epsilon}_{t-1}$ gives the Positive Size Bias test.

7.9 EGARCH Model

Define $h_t = \ln(\sigma_t^2)$. Nelson's *exponential* GARCH model is then

$$h_t = a_0 + \sum_{i=1}^{p} a_i \frac{|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{q} b_j h_{t-j}$$

- Variance is always positive because $\sigma_t^2 = \exp(h_t)$

- Total effect of positive shocks (good news) to ϵ_{t-i}

$$(1 + \gamma_i)|\epsilon_{t-i}|$$

- Total effect of negative shocks (bad news) to ϵ_{t-i}

$$(1 - \gamma_i)|\epsilon_{t-i}|$$

- *Leverage effect* implies that $\gamma_i < 0$

- EGARCH is covariance stationary provided $\sum_{j=1}^{q} b_j < 1$.

7.10 TGARCH Model

Zakoian's *threshold* GARCH model is

$$
\begin{aligned}
\sigma_t^2 &= a_0 + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \gamma_i S_{t-i} \epsilon_{t-i}^2 + \sum_{j=1}^{q} b_j \sigma_{t-j}^2 \\
S_{t-i} &= \begin{cases} 1 & \text{if } \epsilon_{t-i} < 0 \\ 0 & \text{if } \epsilon_{t-i} \geq 0 \end{cases}
\end{aligned}
$$

- When ϵ_{t-i} is positive, the total effects are $a_i \epsilon_{t-i}^2$

- when ϵ_{t-i} is negative, the total effects are $(a_i + \gamma_i)\epsilon_{t-i}^2$

- *Leverage effect* implies that $\gamma_i > 0$

7.11 PGARCH Model

Ding, Granger and Engle's power GARCH model for $d > 0$

$$\sigma_t^d = a_0 + \sum_{i=1}^{p} a_i(|\epsilon_{t-i}| + \gamma_i \epsilon_{t-i})^d + \sum_{j=1}^{q} b_j \sigma_{t-j}^d$$

- Leverage effect implies that $\gamma_i < 0$
- $d = 2$ gives a regular GARCH model with leverage effects
- $d = 1$ gives a model for σ_t and is more robust to outliers than when $d = 2$
- d can be fixed at a particular value or estimated by mle

7.13 Forecasts from Asymmetric GARCH(1,1) Models

- Consider the TGARCH(1,1) model at time T

$$\sigma_T^2 = a_0 + a_1\epsilon_{T-1}^2 + \gamma_1 S_{T-1}\epsilon_{T-1}^2 + b_1\sigma_{T-1}^2.$$

- Assume that ϵ_t has a symmetric distribution about zero. The forecast for $T+1$ based on information at time T is

$$E_T[\sigma_{T+1}^2] = a_0 + a_1\epsilon_T^2 + \gamma_1 S_T\epsilon_T^2 + b_1\sigma_T^2,$$

  where it assumed that ϵ_T^2, S_T and σ_T^2 are known. Hence, the TGARCH(1,1) forecast for $T+1$ will be different than the GARCH(1,1) forecast if $S_T = 1$ ($\epsilon_T < 0$).

- The forecast at $T+2$ is

$$\begin{aligned} E_T[\sigma^2_{T+2}] &= a_0 + a_1 E_T[\epsilon^2_{T+1}] + \gamma_1 E_T[S_{T+1}\epsilon^2_{T+1}] + b_1 E_T[\sigma^2_{T+1}] \\ &= a_0 + \left(\frac{\gamma_1}{2} + a_1 + b_1\right) E_T[\sigma^2_{T+1}], \end{aligned}$$

which follows since

$$E_T[S_{T+1}\epsilon^2_{T+1}] = E_T[S_{T+1}]E_T[\epsilon^2_{T+1}] = \frac{1}{2}E_T[\sigma^2_{T+1}]$$

Notice that the asymmetric impact of leverage is present even if $S_T = 0$.

- By recursive substitution for the forecast at $T+h$ is

$$E_T[\sigma^2_{T+h}] = a_0 + \left(\frac{\gamma_1}{2} + a_1 + b_1\right)^{h-1} E_T[\sigma^2_{T+1}],$$

which is similar to the GARCH(1,1) forecast.

- The mean reverting form is

$$E_T[\sigma^2_{T+h}] - \bar{\sigma}^2 = \left(\frac{\gamma_1}{2} + a_1 + b_1\right)^{h-1} \left(E_T[\sigma^2_{T+h}] - \bar{\sigma}^2\right)$$

where $\bar{\sigma}^2 = a_0/(1 - \frac{\gamma_1}{2} - a_1 - b_1)$ is the long run variance.

- Forecasting algorithms for σ^d_{T+h} in the PGARCH$(1, d, 1)$ and for $\ln \sigma^2_{T+h}$ in the EGARCH(1,1) follow in a similar manner.