3 Modelling Volatility: ARCH

- **Why ARCH processes?**

1) Many economic time series exhibit periods of unusually large volatility followed by periods of relative tranquility (see Figure 2.5 and Figure 3.4). The conditional homoskedasticity assumption for the error term is inappropriate in these cases.

2) Sometimes we are more interested in the conditional variance of a series.

3) Conditional forecasts are superior to unconditional forecasts. For example, study $AR(1)$ model with conditional homoskedasticity assumption for the error term: $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$, where ε_t is a white process satisfying $E(\varepsilon_t | y_{t-1}) = 0$ and $Var(\varepsilon_t | y_{t-1}) = \sigma^2$. Since $E_t y_{t+1} = a_0 + a_1 y_t$, and $Ey_t = a_0/(1-a_1)$, we have

$$\begin{aligned} E_t \left(y_{t+1} - a_0 - a_1 y_t\right)^2 &= E_t(\varepsilon_{t+1}) = \sigma^2 \\ E\left(y_{t+1} - a_0/(1-a_1)\right)^2 &= E\left(\sum_{i=0}^{\infty} a_1^i \varepsilon_{t+1-i}\right)^2 = \frac{\sigma^2}{1-a_1^2} > \sigma^2. \end{aligned}$$

That is, the unconditional forecast has a larger variance than the conditional forecast.

4) Some series appear in volatility clustering, which are different from the series in the random walk process.

- *ARCH* **(Autoregressive Conditional Heteroskedastic) processes:** The main model can be an AR model, an ARMA model, or a standard regression model, i.e.

$$y_t = x_t \gamma + \varepsilon_t,$$

where ε_t is conditionally heteroskedastic in the form of

$$\varepsilon_t = v_t \sqrt{h_t}, \tag{1}$$

where

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2, \quad ARCH(q)$$

and

$$\alpha_0 > 0, \alpha_1 \geq 0, \cdots, \alpha_q \geq 0.$$

Here v_t is a white-noise process with $Var(v_t) = 1$, and v_t and $\varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$ are independent. In ARCH(q), all the shocks from ε_{t-1} to ε_{t-q} have a direct effect on ε_t due to the nonlinear correlation between ε_t and ε_{t-1} through ε_{t-q}: $\varepsilon_t = v_t \sqrt{h_t}$. It is assumed that all the roots of

$$1 - \alpha_1 z - \cdots - \alpha_q z^q = 0$$

lie outside the unit circle. It is easy to verify that

$$\begin{aligned} E\varepsilon_t &= 0, \\ Var(\varepsilon_t) &= \frac{\alpha_0}{1-\alpha_1-\cdots-\alpha_q}, \\ E\varepsilon_t\varepsilon_{t-s} &= 0, \ \forall s \neq 0 \end{aligned}$$

and

$$\begin{aligned} &E[\varepsilon_t|\varepsilon_{t-1},\varepsilon_{t-2},\cdots] = 0, \\ &\sigma_t^2 \equiv E(\varepsilon_t^2|\varepsilon_{t-1},\varepsilon_{t-2},\cdots) = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2. \end{aligned} \tag{2}$$

The series $\{\varepsilon_t\}$ are serially uncorrelated ($E[\varepsilon_t\varepsilon_{t-s}] = 0, s \neq 0$), but they are not independent ($Var(\varepsilon_t|\varepsilon_{t-1},\varepsilon_{t-2},\cdots) \neq 0$). The ARCH model can capture periods of tranquility and volatility in the $\{y_t\}$ series. The conditional variance σ_t^2 has two parts: a constant term α_0 and the linear combination of the information about the squared errors $\varepsilon_{t-1}^2,\cdots,\varepsilon_{t-q}^2$ (i.e. an ARCH term).

Note: (i) Specification (2) can be written in the following form:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 + \eta_t, \tag{3}$$

where η_t is i.i.d with $E\eta_t = 0$, $Var(\eta_t) = \lambda^2$ and $\eta_t \geq -\alpha_0, t = 1,2,\cdots$.The specification (1) is preferred since the multiplicative disturbance of v_t gives the convenient way to simultaneously estimate the parameters in the main model and the conditional variance specification by MLE. (ii) It is proved that $\eta_t = h_t(v_t^2-1)$. Although the error term in (3) is homoskedastic unconditionally, its conditional variance is a function of t: $E(\eta_t^2|\varepsilon_{t-1},\varepsilon_{t-2},\cdots) = h_t^2E(v_t^2-1)^2$. Then

$$\lambda^2 = EE(\eta_t^2|\varepsilon_{t-1},\varepsilon_{t-2},\cdots) = Eh_t^2\cdot E(v_t^2-1)^2.$$

(iii) λ may have non-real solution. For example, assume that $h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2$. Then

$$Eh_t^2 = \frac{\alpha_1^2\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2}$$

and

$$\lambda^2 = \left[\frac{\alpha_1^2\lambda^2}{1-\alpha_1^2} + \frac{\alpha_0^2}{(1-\alpha_1)^2}\right]\cdot E(v_t^2-1)^2.$$

- **Example** In ARCH(1) process

$$\begin{cases} y_t = a_0 + a_1y_{t-1} + \varepsilon_t, \ |a_1| < 1 \\ \varepsilon_t = v_t\sqrt{\alpha_0 + \alpha_1\varepsilon_{t-1}^2}, \ \alpha_0 > 0, \ 0 < \alpha_1 < 1, \end{cases}$$

the series $\{y_t\}$ is stationary since

$$y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i},$$

$$E(y_t) = \frac{a_0}{1-a_1}$$

$$Var(y_t) = \sum_{i=0}^{\infty} a_1^{2i} Var(\varepsilon_{t-i}) = \frac{1}{1-a_1^2}\frac{\alpha_0}{1-\alpha_1},$$

$$Cov(y_t, y_{t-s}) = \sum_{j=0}^{\infty} a_1^{s+2j} Var(\varepsilon_t) = \frac{a_1^s}{1-a_1^2}\frac{\alpha_0}{1-\alpha_1}.$$

The variance of y_t is increasing in α_1 and in the absolute value of a_1. The ARCH error process can be used to model periods of volatility within the univariate framework. However,

$$E_{t-1}y_t = E\left(y_t | y_{t-1}, y_{t-2}, \cdots,\right) = a_0 + a_1 y_{t-1},$$

$$Var\left(y_t | y_{t-1}, y_{t-2}, \cdots,\right) = E_{t-1}[y_t - a_0 - a_1 y_{t-1}]^2 = E_{t-1}\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2.$$

See ex4 for the simulated ARCH processes $y_t = a_1 y_{t-1} + \varepsilon_t$, $\varepsilon_t = \sqrt{v_t(1+0.8\varepsilon_{t-1}^2)}$ with $a_1 = 0.2$ and 0.9.

- **Conditional Maximum Likelihood Estimation of ARCH** Consider the $ARCH(q)$ model

$$y_t = x_t\beta + \varepsilon_t$$

with

$$\varepsilon_t = v_t\sqrt{h_t},$$

where

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2,$$

$v_t \sim$i.i.d.$N(0,1)$, and v_t and $\varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$ are independent. We condition on the first q observations ($t = -q+1, \cdots, 1, 0$) and use observations $t = 1, 2, \cdots, T$ for estimation. Denote

$$\mathbf{Y}_t \equiv \left(y_t, y_{t-1}, \cdots, y_1, y_0, \cdots, y_{-q+1}, x_t, x_{t-1}, \cdots, x_0, \cdots, x_{-q+1}\right)'.$$

Then $(y_t | x_t, \mathbf{Y}_{t-1}) \sim N(x_t\beta, h_t)$. The conditional log likelihood function is

$$\begin{aligned}\ln L &= \ln\left(\prod_{t=1}^{T}\frac{1}{\sqrt{2\pi h_t}}\exp\left(-\frac{(y_t - x_t\beta)^2}{2h_t}\right)\right) \\ &= -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(h_t) - \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - x_t\beta)^2}{h_t}\end{aligned}$$

with $h_t = \alpha_0 + \alpha_1 (y_{t-1} - x_{t-1}\beta)^2 + \cdots + \alpha_q (y_{t-q} - x_{t-q}\beta)^2$.

- **Testing for ARCH(q)** Conduct the standard ACF or Q-statistic test to the squared residuals from the estimated main model of y_t, which can help identify the order of the GARCH process. Also, **Lagrange multiplier test** (Engle(1982)) can be used. The null is $H_0 : \alpha_1 = \cdots = \alpha_q = 0$. That is, the error term ε_t is an white noise process. Use Lagrange multiplier test according to the following steps: (i) LS regress y_t on x_t by using the observations $t = -q+1, -q+2, \cdots, T$ and save the sample residuals $\hat{\varepsilon}_t$ and $h_0 \equiv \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t^2$; (ii) Regress $\hat{\varepsilon}_t^2/h_0 - 1$ on 1, $\hat{\varepsilon}_{t-1}^2, \cdots, \hat{\varepsilon}_{t-q}^2$. Then the sample size T times the R^2 from this regression converges in distribution to $\chi^2(q)$ under the null H_0. Or generate the squared residual sequences $\{\hat{\varepsilon}_t^2\}$, then estimate a regression of the form

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \cdots + \alpha_q \hat{\varepsilon}_{t-q}^2.$$

  Test the null hypothesis using the statistics $TR^2 \sim \chi^2(q)$ or $F \sim F(q, T-q)$.

Note: 1). F test is superior for small sample size T. 2). Unfortunately, there is no available method to test the null of white-noise errors versus the specific alternative of $GARCH(p,q)$ errors.

- $GARCH(p,q)$ **(Generalized Autoregressive Conditional Heteroskedastic)** (Bollerslev (1986)): The error term in the main model $y_t = x_t\beta + \varepsilon_t$ satisfies $\varepsilon_t = v_t\sqrt{h_t}$, where $\{v_t\}$ is a white-noise with $\sigma_v^2 = 1$, independent of h_t, and

$$h_t = \delta_0 + \delta_1 h_{t-1} + \cdots + \delta_p h_{t-p} + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2.$$

  The process $\{h_t\}$ can be seen as an ARMA(∞) process:

$$h_t = \alpha_0 + \pi(L)\varepsilon_t^2$$

  where

$$\pi(L) \equiv \sum_{j=1}^{\infty} \pi_j L^j = \frac{\alpha(L)}{1-\delta(L)} = \frac{\alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q}{1 - \delta_1 L - \delta_2 L^2 - \cdots - \delta_p L^p}$$

  and $(1 - \delta_1 - \delta_2 - \cdots - \delta_p)\alpha_0 = \delta_0$. The GARCH(p,q) process $\{h_t\}$ is stationary if

$$\delta_1 + \delta_2 + \cdots + \delta_p + \alpha_1 + \alpha_2 + \cdots + \alpha_q < 1.$$

  The $GARCH(1,1)$ specification is the most popular form of conditional volatility for financial data. Denote

$$\sigma_t^2 \equiv Var(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots).$$

Since $E[\varepsilon_t|\varepsilon_{t-1},\varepsilon_{t-2},\cdots]=0$ and $E_{t-1}(\varepsilon_t^2)=h_t$, we have $h_t=\sigma_t^2$ and the conditional variance equation

$$\begin{aligned}\sigma_t^2 &= E_{t-1}(\varepsilon_t^2)=h_t\\ &= \delta_0+\delta_1\sigma_{t-1}^2+\cdots+\delta_p\sigma_{t-p}^2+\alpha_1\varepsilon_{t-1}^2+\cdots+\alpha_q\varepsilon_{t-q}^2,\end{aligned}$$

which is not constant. That is, the conditional variance of the disturbances in the model of y_t looks very much like (but is not) an $ARMA(p,q)$ process (derive the process?). For unconditional mean, variance and covariance, we have

$$\begin{aligned}E\varepsilon_t &= 0\\ E\varepsilon_t^2 &= Ev_t^2\cdot Eh_t=Eh_t=E\left(E_{t-1}(\varepsilon_t^2)\right)\\ &= \delta_0+\alpha_1E\varepsilon_{t-1}^2+\cdots+\alpha_qE\varepsilon_{t-q}^2+\beta_1Eh_{t-1}+\cdots+\beta_pEh_{t-p}\\ &= \delta_0+(\alpha_1+\cdots+\alpha_q+\delta_1+\cdots+\delta_p)\,E\varepsilon_t^2\\ &= \delta_0/\left(1-\alpha_1-\cdots-\alpha_q-\delta_1-\cdots-\delta_p\right)\\ &< \infty,\ \text{if } 1-\alpha_1-\cdots-\alpha_q-\delta_1-\cdots-\delta_p>0.\\ E\varepsilon_t\varepsilon_{t-s} &= E\left(v_tv_{t-s}\sqrt{h_th_{t-s}}\right)=0\ \forall s\neq 0.\end{aligned}$$

In a GARCH process, the errors $\{\varepsilon_t\}$ are uncorrelated since $E\varepsilon_t\varepsilon_{t-s}=0$, but the squared errors are dependent since $\varepsilon_t^2=\delta_0+\delta_1\sigma_{t-1}^2+\cdots+\delta_p\sigma_{t-p}^2+\alpha_1\varepsilon_{t-1}^2+\cdots+\alpha_q\varepsilon_{t-q}^2+\xi_t$, where $E_{t-1}(\xi_t)=0$.

- **MLE of GARCH** For the GARCH model

$$\begin{aligned}y_t &= x_t\beta+\varepsilon_t,\quad \varepsilon_t=v_t\sqrt{h_t},\ v_t\sim \text{i.i.d.}N(0,1),\\ h_t &= \delta_0+\delta_1h_{t-1}+\cdots+\delta_ph_{t-p}+\alpha_1\varepsilon_{t-1}^2+\cdots+\alpha_q\varepsilon_{t-q}^2,\end{aligned}$$

the conditional likelihood function is

$$L=\prod_{t=1}^{T}\frac{1}{\sqrt{2\pi h_t}}\exp\left(-\frac{(y_t-x_t\beta)^2}{2h_t}\right).$$

- **Diagnostics for model adequacy:** An estimated GARCH model should capture all dynamic aspects of the model of the mean and the model of the conditional variance. The estimated residuals should be serially uncorrelated ($\hat{\varepsilon}_t$ close to white noise process) and should not display any remaining conditional volatility (the residuals $\hat{w}_t$ in the model of the conditional variance close to white noise process).

1. Use the standardized residuals $\hat{s}_t\equiv\hat{\varepsilon}_t/\hat{h}_t^{1/2}$ and conduct Ljung-Box Q-statistic test to see if the model of the mean is properly specified.

2. Use the squared standardized residuals $\hat{s}_t^2 \equiv \hat{\varepsilon}_t^2/\hat{h}_t = \hat{v}_t^2$ and conduct Ljung-Box Q-statistic test to see if there are remaining GARCH effects in the model of the conditional variance.

- **Assessing the fit of GARCH estimation:** 1) Choose the model with the smallest $RSS' = \sum_{t=1}^{T}\left(\hat{\varepsilon}_t^2 - \hat{h}_t\right)^2$; 2) Select the model with smallest AIC and SBC: $AIC' = -\ln L + 2n$, $SBC = -\ln L + n\ln T$, where $L = -\sum_{t=1}^{T}\left(\ln \hat{h}_t + \hat{\varepsilon}_t^2/\hat{h}_t\right)$.

- **Forecast the mean and the conditional variance:** Consider, for example, the GARCH(1,1) model with $\varepsilon_t = v_t h_t^{1/2}$, where v_t is independent of ε_{t-s} for all $s > 0$ and $h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$ with $\alpha_1 > 0$ and $\beta_1 > 0$. The confidence intervals for the forecast of the mean are

$$E_t y_{t+1} \pm 2h_{t+1}^{1/2},$$
$$E_t y_{t+j} \pm 2h_{t+j}^{1/2}.$$

The forecasts of the conditional variance are

$$\begin{aligned}
E_t h_{t+1} &= \alpha_0 + \alpha_1\varepsilon_t^2 + \beta_1 h_t, \\
E_t h_{t+j} &= \alpha_0 + \alpha_1 E_t\varepsilon_{t+j-1}^2 + \beta_1 E_t h_{t+j-1} \\
&= \alpha_0 + (\alpha_1 + \beta_1) E_t h_{t+j-1} \text{ (since } E_t v_{t+j}^2 = 1, E_t\varepsilon_{t+j-1}^2 = E_t h_{t+j-1}) \\
&= \alpha_0 + \alpha_0(\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 E_t h_{t+j-2} \\
&= \alpha_0\left(1 + (\alpha_1 + \beta_1) + \cdots + (\alpha_1 + \beta_1)^{j-1}\right) + (\alpha_1 + \beta_1)^j E_t h_t \\
&= \alpha_0 \frac{1 - (\alpha_1 + \beta_1)^j}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^j h_t \\
&\rightarrow \alpha_0/(1 - \alpha_1 - \beta_1), \text{ as } j \rightarrow \infty, \text{ if } \alpha_1 + \beta_1 < 1.
\end{aligned}$$

- ARCH-M model (Engle, Lilien and Robins (1987)): Assume that the risk premium is an increasing function of the conditional variance of ε_t.The **ARCH in mean** model of the excess return y_t is

$$y_t = x_t\beta + \delta h_t + \varepsilon_t, \ \varepsilon_t = v_t h_t^{1/2},$$

where h_t is the conditional variance of ε_t and satisfies an $ARCH(q)$ process: $h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2$. Alternatively, the mean equation can also be specified as other forms, e.g.

$$y_t = x_t\beta + \delta h_t^{1/2} + \varepsilon_t$$

or

$$y_t = x_t\beta + \delta \log(h_t) + \varepsilon_t.$$

- Other Models of conditional variance (also see EViews 5 Users Guide for the details): for example,

1. **IGARCH:** For GARCH(1,1), $\alpha_1 + \beta_1 = 1$. The conditional variance is

$$h_t = \alpha_0 + (1-\beta_1)\varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

or

$$h_t = \alpha_0/(1-\beta_1) + (1-\beta_1)\sum_{i=0}^{\infty}\beta_1^i \varepsilon_{t-1-i}^2,$$

which yields a very parsimonious specification of a geometrically decaying conditional variance (in the past realizations of the $\{\varepsilon_t^2\}$).

2. **GARCH with explanatory variables:** Some exogenous factor D_t affects the volatility:

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma D_t, \;\; \gamma > 0.$$

3. **TARCH:** the threshold-GARCH model:

$$h_t = \alpha_0 + (\alpha_1 + \lambda_1 d_{t-1})\,\varepsilon_{t-1}^2 + \beta_1 h_{t-1},$$

where $d_{t-1} = 1$, if $\varepsilon_{t-1} < 0$; 0,otherwise. That is, $\varepsilon_{t-1} = 0$ is a threshold such that shocks greater than the threshold have different effects (on the volatility h_t) from shocks below the threshold. If λ_1 is statistically different from zero, the data contains a threshold effect.

4. **EGARCH:** the exponential-GARCH:

$$\ln h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}/h_{t-1}^{1/2} + \lambda_1\left|\varepsilon_{t-1}/h_{t-1}^{1/2}\right| + \beta_1 \ln h_{t-1}.$$

Note that 1) the volatility h_t can be never negative; 2) the standardized value of ε_{t-1},i.e. $\varepsilon_{t-1}/h_{t-1}^{1/2}$, is used to give a unit-free measure of the volatility; 3) the specification allows for leverage (threshold) effects since

$$\alpha_1\varepsilon_{t-1}/h_{t-1}^{1/2} + \lambda_1\left|\varepsilon_{t-1}/h_{t-1}^{1/2}\right| = \begin{cases} (\alpha_1+\lambda_1)\,\varepsilon_{t-1}/h_{t-1}^{1/2}, & \text{if } \varepsilon_{t-1} > 0 \\ (\alpha_1-\lambda_1)\,\varepsilon_{t-1}/h_{t-1}^{1/2}, & \text{otherwise,} \end{cases}$$

which implies that the effect of the standardized shock on the log of the volatility is $\alpha_1 + \lambda_1$ if ε_{t-1} is positive while the effect is $-\alpha_1 + \lambda_1$ if ε_{t-1} is negative.

5. **Nonparametric Specification:** Corresponding to the linear parametric $ARCH(m)$: $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_m \varepsilon_{t-m}^2$, the conditional variance is specified as

$$h_t = \sum_{\tau=1,\tau\neq t}^{T} w_\tau(t)\varepsilon_\tau^2,$$

where the weights $\{w_\tau(t)\}_{\tau=1,\tau\neq t}^T$ satisfies $\sum_{\tau=1,\tau\neq t}^T w_\tau(t) = 1$. If $\varepsilon_{\tau-1}, \varepsilon_{\tau-2}, \cdots, \varepsilon_{\tau-m}$ are close to $\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-m}$, ε_τ^2 will provide useful information on

$$h_t = E\left(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-m}\right),$$

and we should select a larger weight $w_\tau(t)$. Choose the kernel estimator of h_t:

$$h_t = \frac{\sum_{\tau=1,\tau\neq t}^T \varepsilon_\tau^2 k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right) k\left(\frac{\varepsilon_{\tau-2}-\varepsilon_{t-2}}{h_2}\right)\cdots k\left(\frac{\varepsilon_{\tau-m}-\varepsilon_{t-m}}{h_m}\right)}{\sum_{\tau=1,\tau\neq t}^T k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right) k\left(\frac{\varepsilon_{\tau-2}-\varepsilon_{t-2}}{h_2}\right)\cdots k\left(\frac{\varepsilon_{\tau-m}-\varepsilon_{t-m}}{h_m}\right)},$$

i.e. choose the weights

$$w_\tau(t) = \frac{k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right) k\left(\frac{\varepsilon_{\tau-2}-\varepsilon_{t-2}}{h_2}\right)\cdots k\left(\frac{\varepsilon_{\tau-m}-\varepsilon_{t-m}}{h_m}\right)}{\sum_{\tau=1,\tau\neq t}^T k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right) k\left(\frac{\varepsilon_{\tau-2}-\varepsilon_{t-2}}{h_2}\right)\cdots k\left(\frac{\varepsilon_{\tau-m}-\varepsilon_{t-m}}{h_m}\right)},$$

where $h_1, h_2, \cdots, h_m$ are the bandwidths. Specially, for $ARCH(1)$,

$$h_t = \frac{\sum_{\tau=1,\tau\neq t}^T \varepsilon_\tau^2 k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right)}{\sum_{\tau=1,\tau\neq t}^T k\left(\frac{\varepsilon_{\tau-1}-\varepsilon_{t-1}}{h_1}\right)}.$$

6. **Semiparametric Model:** h_t is specified parametrically while the density of v_t is specified nonparametrically. See Engle R. F. and G-R. Gloria (1991), "Semiparametric ARCH Models", Journal of Business and Economic Statistics 9: 345-359.