5.1 Modeling Volatility

In most econometric models the variance of the disturbance term is assumed to be constant (homoscedasticity). However, there is a number of economics and finance series that exhibit periods of unusual large volatility, followed by periods of relative tranquility. Then, in such cases the assumption of homoscedasticity is not appropriate. For example, consider the daily changes in the NYSE International 100 Index, April 5, 2004 - September 20, 2011, shown in Figure 5.1.

One easy way to understand volatility is modeling it explicitly, for example in

\[ y_{t+1} = \epsilon_{t+1} x_t \]  \hspace{1cm} (5.1)

where \( y_{t+1} \) is the variable of interest, \( \epsilon_{t+1} \) is a white-noise disturbance term with variance \( \sigma^2 \), and \( x_t \) is an independent variable that is observed at time \( t \).

If \( x_t = x_{t-1} = x_{t-2} = \cdots = \text{constant} \), then the \( \{ y_t \} \) is a familiar white-noise process with a constant variance. However, if they are not constant, the variance of \( y_{t+1} \) conditional on the observable value of \( x_t \) is

\[ \text{var}(y_{t+1}|x_t) = x_t^2 \sigma^2 \]  \hspace{1cm} (5.2)

If \( \{ x_t \} \) exhibit positive serial correlation, the conditional variance of the \( \{ y_t \} \) sequence will also exhibit positive serial correlation. We can write the model in logarithm form and introduce the coefficients \( a_0 \) and \( a_1 \) to have

\[ \log(y_t) = a_0 + a_1 \log(x_{t-1}) + e_t \]  \hspace{1cm} (5.3)

where \( e_t = \log(\epsilon_t) \).
5.2 ARCH Processes

Engel (1982) shows that it is possible to simultaneously model the mean and the variance of a series. To understand the key idea in Engel’s methodology we need to see how the conditional forecast is superior to the unconditional forecast. Consider the ARMA model

\[ y_t = a_0 + a_1 y_{t-1} + \epsilon_t. \]  

(5.4)

The conditional mean of \( y_{t+1} \) is given by

\[ E_t(y_{t+1}) = a_0 + a_1 y_t \]  

(5.5)

If we use the conditional mean to forecast \( y_{t+1} \), the forecast variance is

\[ E_t[(y_{t+1} - a_0 - a_1 y_t)^2] = E_t(\epsilon_{t+1}^2) = \sigma^2 \]  

(5.6)

However, if we use the unconditional forecast, which is always the long-run mean of the \( \{y_t\} \) sequence, that is, equal to \( a_0/(1 - a_1) \), the unconditional forecast error variance is

\[ E_t\left[(y_{t+1} - \frac{a_0}{1 - a_1})^2\right] = E_t\left[(\epsilon_{t+1} + a_1 \epsilon_t + a_1^2 \epsilon_{t-1} + a_1^3 \epsilon_{t-2} + \cdots)^2\right] \]  

(5.7)
Because $1/(1 - a_1^2) > 1$, the unconditional forecast has a greater variance than the conditional forecast. Hence, conditional forecasts are preferable.

In the same way, if the variance of $\{\varepsilon_t\}$ is not constant, you can estimate any tendency using an ARMA model. For example, if $\{\hat{\varepsilon}_t\}$ are the estimated residuals of the model $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$, the conditional variance of $y_{t+1}$ is

$$\text{var}(y_{t+1}|y_t) = E_t[(y_{t+1} - a_0 - a_1 y_t)^2] \quad (5.8)$$

$$= E_t(\hat{\varepsilon}_{t+1})^2$$

We do not assume that it is constant as in Equation 5.6. Let’s say we can estimate it using the conditional variance as an AR($q$) process using squares of the estimated residuals

$$\hat{\varepsilon}_t^2 = a_0 + a_1 \hat{\varepsilon}_{t-1}^2 + a_2 \hat{\varepsilon}_{t-2}^2 + \cdots + a_q \hat{\varepsilon}_{t-q}^2 + v \quad (5.9)$$

where $v$ is a white-noise process. If $a_1 = a_2 = \cdots = a_q = 0$, the estimated variance is simply a constant $a_0 = \sigma^2$. Otherwise we can use Equation 5.9 to forecast the conditional variance at $t+1$,

$$E_t(\hat{\varepsilon}_{t+1}^2) = a_0 + a_1 \hat{\varepsilon}_{t}^2 + a_2 \hat{\varepsilon}_{t-1}^2 + \cdots + a_q \hat{\varepsilon}_{t+1-q} \quad (5.10)$$

This is why Equation 5.9 is called an autoregressive conditional heteroscedasticity ARCH model. The linear specification in 5.9 is actually not the most convenient. Models for $\{y_t\}$ and the conditional variance can be estimated simultaneously using maximum likelihood. In most cases $v$ enters in a multiplicative fashion.

Engel (1982) proposed the following multiplicative conditional heteroscedasticity model

$$\varepsilon_t = v_t \sqrt{a_0 + a_1 \hat{\varepsilon}_{t-1}^2} \quad (5.11)$$

where $v_t$ is a white-noise process with $\sigma_v^2 = 1$, $v_t$ and $\varepsilon_{t-1}$ are independent of each other, $a_0 > 0$, and $0 \leq a_i \leq 1$.

Consider the properties of $\varepsilon_t$.

1. The unconditional expectation of $\varepsilon_t$ is zero

$$E(\varepsilon_t) = E[v_t(\alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2)^{1/2}] \quad (5.12)$$

$$= E(v_t)E[(\alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2)^{1/2}] = 0$$

2. Because $E(v_t v_{t-1})$,

$$E(\varepsilon_t \varepsilon_{t-1}) = 0 \quad i \neq 0 \quad (5.13)$$
3. The unconditional variance of $\varepsilon_t$ is
\[
E(\varepsilon_t^2) = E[v_t^2(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)] = E(v_t^2)E(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)
\] (5.14)

because $\sigma_t^2 = 1$ and the unconditional variance of $\varepsilon_t$ is equal to the unconditional variance of $\varepsilon_{t-1}$, hence
\[
E(\varepsilon_t^2) = \frac{\alpha_0}{1 - \alpha_1}
\] (5.15)

4. The conditional mean of $\varepsilon_t$ is zero
\[
E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = E_{t-1}(v_t)E_{t-1}(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} = 0
\] (5.16)

5. The conditional variance of $\varepsilon_t$ is
\[
E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2.
\] (5.17)

which just means that the conditional variance of $\varepsilon_t$ depends on the realized value of $\varepsilon_{t-1}^2$.

In Equation 5.17, the conditional variance follows a first-order autoregressive process denoted by ARCH(1). As opposed to the usual autoregression, the coefficients $\alpha_0$ and $\alpha_1$ have to be restricted to make sure that the conditional variance is never negative. Both have to be positive, and in addition, to ensure stability of the process we need to have the restriction $0 \leq \alpha_1 \leq 1$.

The key point in an ARCH process is that even though the process $\{\varepsilon_t\}$ is serially uncorrelated (i.e., $E(\varepsilon_t \varepsilon_{t-s}) = 0, \forall s \neq 0$), the errors are not independent. They are related through their second moment. The heteroscedasticity in $\{\varepsilon_t\}$ results in $\{y_t\}$ being heteroscedastic. Then the ARCH process is able to capture periods of relative tranquility and periods of relative high volatility in the $\{y_t\}$ series.

To understand the intuition behind an ARCH process, consider the simulated white-noise process presented in the upper panel of Figure 5.2. While this is certainly a white noise $\{\varepsilon_t\}$, the lower panel shows the generated heteroscedastic errors $\varepsilon_t = v_t \sqrt{1 + 0.8 \varepsilon_{t-1}^2}$. Notice that when the realized value $\varepsilon_{t-1}$ is far from zero, the variance of $\varepsilon_t$ tends to be large. The Stata code to obtain these graph is (you can try obtaining the graph with different seeds):

```stata
clear
set obs 150
set seed 1001
gen time=_n
tset time
gen white=invnorm(uniform())
twoway line white time, m(o) c(l) scheme(sj) ///
ytitle( "white-noise" ) saving(gwhite, replace)
gen erro = 0
replace erro = white*(sqrt(1+0.8*(l.erro)^2)) if time > 1
twoway line erro time, scheme(sj) ///
```
5.2 ARCH Processes

A white-noise process and the heteroscedastic error $\varepsilon_t = \nu_t \sqrt{1 + 0.8\varepsilon_{t-1}^2}$

The panels in Figure 5.3 show two simulated ARMA processes. The idea is to illustrate how the error structure affect the $\{y_t\}$ sequence. The upper panel shows the simulated path of $\{y_t\}$ when $a_0 = 0.9$, while the lower panel shows the simulated path of $\{y_t\}$ when $a_0 = 0.2$. Notice that when $a_0 = 0$, the $\{y_t\}$ sequence is the same as the $\{v_t\}$ sequence depicted in Figure 5.2. However, the persistence of the series increases with $a_0$. Moreover, notice how the volatility in $\{y_t\}$ is increasing in the value of $a_0$ (it also increase with the value of $\alpha_1$).

gen Y1 = 0
gen Y2 = 0
replace Y1 = +0.90*L.Y1 + erro if time > 1
replace Y2 = +0.20*L.Y2 + erro if time > 1
twoway line Y1 time, scheme(sj) ///
ytitle( "Y1" ) saving(gy1, replace)
twoway line Y2 time, scheme(sj) ///
ytitle( "Y2" ) saving(gy2, replace)
gr combine gy1.gph gy2.gph, col(1) ///
iscale(0.7) fysize(100) ///
title( "Simulated ARCH Processes" )
Formally, the conditional mean and variance of \( \{y_t\} \) can be written as
\[
E_{t-1}(y_t) = a_0 + a_1y_{t-1} \tag{5.18}
\]
and
\[
\text{var}(y_t | y_{t-1}, y_{t-2}, \ldots) = E_{t-1}(a_0 + a_1y_{t-1})^2 = E_{t-1}(\varepsilon)^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1})^2 \tag{5.19}
\]

For a nonzero realization of \( \varepsilon_{t-1} \), the conditional variance is positively related to \( \alpha_1 \). For the unconditional variance, recall that the solution (omitting the constant \( A \)) for the difference equation in 5.4 is
\[
y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \tag{5.20}
\]

Because \( E(\varepsilon) = 0 \), the unconditional expectation is \( E(y_t) = a_0/(1-a_1) \). Moreover, because \( E(\varepsilon_t \varepsilon_{t-i}) = 0, \forall i \neq 0 \), the unconditional variance is
\[
\text{var}(y_t) = \sum_{i=0}^{\infty} a_1^{2i} \text{var}(\varepsilon_{t-i}) \tag{5.21}
\]
5.3 GARCH Processes

\[ \varepsilon_t = v_t \sqrt{h_t} \]

where the last equality follows from the result in Equation 5.15. It is easy to see that the unconditional variance is also increasing in \( \alpha_1 \) (and in the absolute value of \( \alpha_1 \)).

The ARCH process presented in Equation 5.11 can be extended in a number of ways. The most straightforward is considering the higher-order ARCH(\( q \)) process

\[ \varepsilon_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2} \] (5.22)

5.3 GARCH Processes

The ARCH idea was extended in Bollerslev (1986) to allow an ARMA process embedded in the conditional variance. Let the error process be

\[ \varepsilon_t = v_t \sqrt{h_t} \]

where \( \sigma^2 = 1 \), and

\[ h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \]

(5.24)

The conditional and unconditional means of \( \varepsilon_t \) are both zero because \( \{v_t\} \) is a white-noise process. The key point is that the conditional variance of \( \varepsilon_t \) is given by \( E_{t-1}(\varepsilon_t^2) = h_t \), which is the ARMA process given in Equation 5.24.

This heteroscedastic variance that allows autoregressive and moving average components is called GARCH(\( p,q \)), where the G in GARCH denotes generalized. Notice that a GARCH(0,1) is just the ARCH model in Equation 5.11. The important restriction in a GARCH process is that all coefficients in Equation 5.24 must be positive and must ensure that the variance is finite (i.e., its characteristic roots must lie inside the unit circle).

A simple procedure to know if a series \( \{y_t\} \) follows a GARCH process is to estimate the best fitting ARMA process and then obtain the fitted errors \( \{\hat{\varepsilon}_t\} \) and the squares of the fitted errors \( \{\hat{\varepsilon}_t^2\} \). While the ACF and the PACF of the fitted errors should be consistent with a white noise, the squared fitted errors should indicate that they follow an ARMA process. It is also useful that besides the ACF and PACF, we use the Ljung-Box \( Q \)-statistic.

Let’s follow the previous suggested steps in Stata for the previously generated process \( Y_1 \), under the assumption that we know that it follows an ARMA(1,0) process with no constant (otherwise we need to search for the optimal ARMA(p,q)).

```
arima Y1, arima(1,0,0) nocons
predict eserro, res
gen eserro2 = eserro^2
corrgram eserro2, lags(20)
```
The Ljung-Box $Q$-statistics of $\{\hat{\varepsilon}_t^2\}$ show strong evidence that the $\{y_t\}$ follows a GARCH process.

A more formal test is the LM (Lagrange Multiplier) test for ARCH errors developed in McLeod and Li (1983). The idea in this test is to estimate the most appropriate ARMA model and obtain the squared fitted errors $\{\hat{\varepsilon}_t^2\}$. Then, estimate the following equation

\[
\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \cdots + \alpha_q \hat{\varepsilon}_{t-q}^2 + u_t \tag{5.25}
\]

If the estimated coefficients are all jointly equal to zero, $\alpha_0 = \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$. This can be easily checked with an $F$ test with $q$ degrees of freedom. An alternative is to use $T R^2$, which for large samples (large $T$) converges to a $\chi^2_q$. For our $Y_1$ series we have:

\[
\begin{align*}
\text{reg eserro2 1.eserro2 12.eserro2} \\
\text{Source} & | SS df MS Number of obs = 148 \\
\hline
\text{Model} & 173533.28 2 86766.64 Prob > F = 0.0000 \\
\text{Residual} & 159865.35 145 1102.51965 R-squared = 0.5205 \\
\text{Total} & 333398.63 147 2268.01789 Root MSE = 33.204 \\
\hline
\end{align*}
\]

\[
\begin{align*}
\text{eserro2} & | \text{Coef. Std. Err. t P>|t| [95% Conf. Interval]} \\
\hline
\text{L1.} & 0.4021842 .0764732 5.26 0.000 .251038 .5533304 \\
\text{L2.} & 0.3998682 .0764729 5.10 0.000 .2387226 .5410139 \\
\text{_cons} & 3.29114 2.906213 1.13 0.259 -.2452873 9.035152 \\
\hline
\end{align*}
\]

where we clearly reject the null of $H_0 : \alpha_0 = \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$. Hence, there is evidence that $\{y_t\}$ follows a GARCH process. The the $q$ can be selected using any optimal lag selection criteria.

### 5.4 Maximum-Likelihood Estimation of GARCH Models

If we assume that $\{\varepsilon_t\}$ follow a normal distribution with mean zero and constant variance $\sigma^2$, the likelihood of $\varepsilon_t$ is

\[
L_\sigma = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right) \exp \left( -\frac{\varepsilon_t^2}{2\sigma^2} \right) \tag{5.26}
\]
Because \( \{ \varepsilon_t \} \) are independent, the likelihood function of the joint realizations of \( \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T \) is

\[
L = \prod_{t=1}^{T} \left( \frac{1}{\sqrt{2\pi \sigma^2}} \right) \exp \left( -\frac{\varepsilon_t^2}{2\sigma^2} \right). \tag{5.27}
\]

Taking the logs we obtain the log-likelihood function

\[
\log L = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{T} (\varepsilon_t)^2 \tag{5.28}
\]

If we want to estimate the simple model \( y_t = \beta x_t + \varepsilon_t \), the maximum log-likelihood function becomes

\[
\log L = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{T} (y_t - \beta x_t)^2. \tag{5.29}
\]

Then, the maximum likelihood estimates of \( \beta \) and \( \sigma^2 \) can be obtained from the following first-order conditions

\[
\frac{\partial \log L}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^{T} (y_t + \beta x_t)^2 = 0 \tag{5.30}
\]

\[
\frac{\partial \log L}{\partial \beta} = \frac{1}{\sigma^2} \sum_{t=1}^{T} (y_t x_t + \beta x_t^2) = 0 \tag{5.31}
\]

which yield the familiar OLS results \( \hat{\sigma}^2 = \sum \varepsilon_t^2 / T \) and \( \hat{\beta} = \sum x_t y_t / \sum x_t^2 \). A close form solution is possible because the FOC are linear.

How do we introduce heteroscedastic errors in this familiar maximum likelihood procedure? Let’s assume that \( y_t \) follows an ARCH(1) process. Hence, \( \varepsilon_t \) is

\[
\varepsilon_t = \nu_t \sqrt{h_t} \tag{5.32}
\]

where the conditional variance of \( \varepsilon_t \) is \( h_t \). Then the likelihood function becomes

\[
L = \prod_{t=1}^{T} \left( \frac{1}{\sqrt{2\pi h_t}} \right) \exp \left( -\frac{\varepsilon_t^2}{2h_t} \right) \tag{5.33}
\]

and the log-likelihood function is

\[
\log L = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \log h_t - \frac{1}{2} \sum_{t=1}^{T} (\varepsilon_t^2 / h_t). \tag{5.34}
\]

With \( \varepsilon_t = y_t - \beta x_t \) and \( h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \) we have

\[
\log L = -\frac{T-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^{T} \log(\alpha_0 + \alpha_t \varepsilon_{t-1}^2) \tag{5.35}
\]
\[ -\frac{1}{2} \sum_{t=2}^{T} \frac{(y_t - \beta x_t)^2}{(\alpha_0 + \alpha_1 \varepsilon_{t-1})^2} \]

\[ = -\frac{T-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^{T} \log(\alpha_0 + \alpha_1 (y_{t-1} - \beta x_{t-1})^2) \]

\[ -\frac{1}{2} \sum_{t=2}^{T} \frac{(y_t - \beta x_t)^2}{(\alpha_0 + \alpha_1 (y_{t-1} - \beta x_{t-1})^2)}, \]

where the first observation is lost because of the lag. While the FOC yield complicated nonlinear equations that do not have a close-form solution, this log-likelihood function can be maximized using numerical methods.

### 5.4.1 Estimation of GARCH Models in Stata

A simple GARCH process in Stata has the following form

\[ y_t = \text{ARMA}(p,q) + \varepsilon_t \]  

(5.36)

\[ \text{var}(\varepsilon_t) = \sigma_t^2 = \alpha_0 + \alpha_{1,1} \varepsilon_{t-1}^2 + \alpha_{1,2} \varepsilon_{t-2}^2 + \cdots \]  

(5.37)

while both equations (5.36 and 5.37) can be estimated iteratively using the ARMA and OLS methods we already know, Stata has a specific command to obtain the estimates in a single step using maximum likelihood.

For example, if we want to estimate the ARCH(1) process in the simulated series Y1, the command is

```stata
arch Y1, arima(1,0,0) arch(1) nocons
```

(setting optimization to BHHH)

Iteration 0: log likelihood = -360.36487
Iteration 8: log likelihood = -300.84557

ARCH family regression -- AR disturbances

| Sample: 1 - 150 | Number of obs = 150 |
| Distribution: Gaussian | Wald chi2(1) = 3237.07 |
| Log likelihood = -300.8456 | Prob > chi2 = 0.0000 |

| | OPG |
| --- | --- | --- | --- | --- |
| | Y1 | Coef. | Std. Err. | z | P>|z| |
| **ARMA** | | | | | | | [95% Conf. Interval] |
| | **ar** | | | | | | |
| L1. | .9048536 | .0159038 | 56.90 | 0.000 | .8736826 | .9360245 |
| **ARCH** | | | | | | | |
| | **arch** | | | | | | |
| L1. | 1.127475 | .2373718 | 4.75 | 0.000 | .662235 | 1.592715 |
That can be written as

\[ Y_t = 0.90Y_{t-1} + \epsilon_t \]  
(5.38)

\[ \sigma_t^2 = 0.64 + 1.13\epsilon_{t-1}^2 \]  
(5.39)

For an AR(1) with GARCH(1,1) we need

\text{arch Y1, arima(1,0,0) arch(1) garch(1) nocons}

(output omitted)

which yields

\[ Y_t = 0.90Y_{t-1} + \epsilon_t \]  
(5.40)

\[ \sigma_t^2 = 0.65 + 1.13\epsilon_{t-1}^2 - 0.03\sigma_{t-1}^2. \]  
(5.41)

### 5.4.2 Volatility Breaks

A large literature indicates that volatility in macroeconomic variables in industrialized economies decrease in early 1984. For example, Stock and Watson (2002) report that the volatility of the U.S. real GDP growth was smaller after 1984. Figure 5.4 shows the real GDP and the real GDP growth rates from the first quarter of 1947 to the first quarter of 2008. While this figure appears to provide some evidence in favor of higher volatility prior 1984, GARCH models can provide a formal test to verify this claim. Consider the following specification for the variance

\[ \sigma_t^2 = \exp(\lambda_0 + \lambda_1x_t) + \alpha \epsilon_{t-1}^2 \]  
(5.42)

where \( x_t \) is a variable that affects the conditional variance of \( y_t \). For our case, let this variable \( x_t \) a dummy variable \( I_{[t > 1984q1]} \) that is equal to one if after the first quarter of 1984, zero otherwise. Then, to estimate this model in Stata we need to type

```stata
use rgdp.dta, clear
tset date
gen y = log(rgdp/l.rgdp)
gen dum = 0
replace dum = 1 if date >= 149 // 149 is 1984q1
arch y, arima(1,0,0) arch(1) het(dum)
(setting optimization to BHHH)
Iteration 0: log likelihood = 834.7398
Iteration 6: log likelihood = 835.59237
ARCH family regression -- ARMA disturbances and mult. heteroskedasticity
Sample: 2 - 246 Number of obs = 245
Distribution: Gaussian Wald chi2(1) = 22.20
Log likelihood = 835.5924 Prob > chi2 = 0.0000
```

\[ \sigma_t^2 = \exp(\lambda_0 + \lambda_1x_t) + \alpha \epsilon_{t-1}^2 \]  
(5.42)
which can be written as

$$y_t = 0.008 + 0.321y_{t-1} + \epsilon_t$$

$$\sigma_t^2 = \exp(-9.06 - 1.70I_{t>1984q1}) + 0.085\epsilon_{t-1}^2.$$  

The coefficient on the dummy variable $I_{t>1984q1}$ is negative and statistically significant. This means that the volatility is significantly lower after the first quarter of 1984.
The ARCH-in-mean (ARCH-M) models were introduced by Engel, Lilien, and Robins (1987). The idea is to allow for the mean of a sequence to depend on its conditional variance. If the riskiness of an asset can be measured by the variance of its returns, the risk premium will be an increasing function of the conditional variance of returns. This can be formally modeled with ARCH-M models.

Engel, Lilien, and Robins (1987) write the excess return of holding a risky asset as

\[ y_t = \mu_t + \epsilon_t \]  \hspace{1cm} (5.44)

where \( y_t \) is the excess return from holding a long-term asset relative to a one-period treasury bill, \( \mu_t \) is the risk premium necessary to induce the risk-averse agent to hold the long-term asset rather than the one-period bond, and \( \epsilon_t \) is the unforecastable shock to the excess return on the long-term asset.

The expected excess return from holding the long-term asset is equal to the risk premium

\[ E_{t-1}(y_t) = \mu_t. \]  \hspace{1cm} (5.45)

We can model the risk premium as a function of the conditional variance of \( \epsilon_t \). Let \( h_t \) be the conditional variance of \( \epsilon_t \). Then the risk premium can be modeled to depend on the conditional variance

\[ \mu_t = \beta + \delta h_t, \quad \delta > 0 \]  \hspace{1cm} (5.46)

and where the conditional variance \( h_t \) follows an ARCH(\( q \)) process

\[ h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \]  \hspace{1cm} (5.47)

Hence, the ARCH-M model is the combination of Equations 5.44, 5.46, and 5.47. If \( \alpha_1 = \alpha_2 = \cdots = \alpha_q = 0 \), the ARCH-M models becomes the traditional constant risk premium model.

To get an intuition of how volatility affects the mean of a series, we simulate the 150 realizations of a white-noise process \( \{\epsilon_t\} \) and then construct the conditional variance that follows the first-order ARCH process \( h_t = 1 + 0.75 \epsilon_{t-1}^2 \). The upper panel in Figure 5.5 illustrates the white-noise process, while the lower panel shows the constructed conditional variance. From this figure is it easy to see how higher volatility in \( \{\epsilon_t\} \) (large positive or negative realizations) translates into larger conditional variance (a large value of \( h_t \)).

```stata
clear
set obs 150
set seed 1002
gen time=_n
tsset time
gen white=invnorm(uniform())
twoway line white time, m(o) c(l) scheme(sj) ///
ytitle( "white-noise") saving(gwhite, replace)
```
Figure 5.5 shows two simulated ARCH-M processes. The upper panel simulates the series

\[ y_t = -4 + 4h_t + \varepsilon_t \]  \hspace{1cm} (5.48)

that is, \( \beta = -4 \) and \( \delta = 4 \), while the lower panel simulates the series with \( \beta = -1 \) and \( \delta = 1 \),

\[ y_t = -1 + h_t + \varepsilon_t. \]  \hspace{1cm} (5.49)

In the upper panel we can see that during periods of higher volatility are associated with higher values of \( \{y_t\} \). For example, around \( t = 75 \) higher volatility is moving the mean of \( \{y_t\} \) to the positive side. The lower panel has smaller effects for the ARCH-M effects because of the smaller \( \beta \) and \( \delta \) coefficients. Notice that if \( \delta = 0 \), there will be no ARCH-M effects.
Simulated ARCH-M Processes

Fig. 5.6 Upper panel: $Y_1 = -4 + 4h_t + \varepsilon_t$. Lower panel: $Y_2 = -1 + h_t + \varepsilon_t.$

The LM test we described before to test for the presence of an ARCH process can also be used here. However, this test will not help in knowing whether the ARCH process is an ARCH-M.

The Stata estimation of an ARCH-M process is straightforward. The process for $y_t$ can be extended from Equation 5.36 to include the ARCH-M component

$$ y_t = \text{ARMA}(p, q) + \sum_{i=1}^{p} \phi_i g(\sigma_{t-i}^2) + \varepsilon $$

(5.50)

and the conditional variance can follow the same form as before

$$ \text{var}(\varepsilon_t) = \sigma_t^2 = \alpha_0 + \alpha_{1,1} \varepsilon_{t-1}^2 + \alpha_{1,2} \varepsilon_{t-2}^2 + \cdots $$

$$ + \alpha_{2,1} \sigma_{t-1}^2 + \alpha_{2,2} \sigma_{t-2}^2 + \cdots $$

(5.51)

Consider the following example that tests whether demand uncertainty affects mean demand realizations

```
use airlines.dta, clear
```
Autoregressive Conditional Heteroskedasticity Models

```
tset date
arch demand, arima(1,0,0) archm arch(1)

(setting optimization to BHHH)
Iteration 0: log likelihood = 163.58772
Iteration 9: log likelihood = 182.39822

ARCH family regression -- AR disturbances

Sample: 1 - 126 Number of obs = 126
Distribution: Gaussian Wald chi2(2) = 142.56
Log likelihood = 182.3982 Prob > chi2 = 0.0000

------------------------------------------------------------------------------
<table>
<thead>
<tr>
<th>OPG</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>demand</td>
<td>Coef.</td>
<td>Std. Err.</td>
<td>z</td>
</tr>
<tr>
<td>_cons</td>
<td>.7233085</td>
<td>.0227503</td>
<td>31.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.6787187</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.7678982</td>
<td></td>
</tr>
<tr>
<td>ARCHM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sigma2</td>
<td>4.559358</td>
<td>4.980237</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.360</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-5.201728</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.32044</td>
<td></td>
</tr>
<tr>
<td>ARMA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ar</td>
<td>.6990354</td>
<td>.0604844</td>
<td>11.56</td>
</tr>
<tr>
<td>L1.</td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.5804882</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.8175825</td>
<td></td>
</tr>
<tr>
<td>ARCH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>arch</td>
<td>.2915877</td>
<td>.1572736</td>
<td>1.85</td>
</tr>
<tr>
<td>L1.</td>
<td></td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.016663</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.5998383</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>.0024397</td>
<td>.0003445</td>
<td>7.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0017646</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0031149</td>
<td></td>
</tr>
</tbody>
</table>
------------------------------------------------------------------------------

The resulting estimated mean equation is

\[
\text{demand}_t = 0.723 + 0.699 \text{demand}_{t-1} + 4.55 \sigma_t^2 + \epsilon_t
\]  

(5.52)

and the conditional variance equation is

\[
\sigma_t^2 = 0.002 + 0.292 \epsilon_{t-1}^2
\]  

(5.53)

Notice that the \( \phi \) coefficient is not statistically significant. Hence, demand uncertainty as measured by the conditional variance of demand does not affect average demand.

5.6 GARCH Models with Asymmetry

5.6.1 TARCH Models

Glosten, Jaganathan, and Runkle (1994) propose a methodology that allows for an asymmetric effect of shocks on volatility. When modeling stock prices, whether \( \epsilon_t \) is positive (good news) or negative (bad news), the effect on volatility may be different. Consider the following threshold-GARCH (TARCH) process
\[ h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \]  

(5.54)

where \( d_{t-1} \) is a dummy variable that is equal to one if \( \varepsilon_{t-1} < 0 \) and it is equal to zero otherwise.

The intuition in this model is simple. If \( \varepsilon_{t-1} < 0 \), then \( d_{t-1} = 1 \) and the effect of the \( \varepsilon_{t-1} \) shock on volatility \( h_t \) is given by \( (\alpha_1 + \lambda_1) \varepsilon_{t-1}^2 \). On the other hand, if the shock is positive, then \( \varepsilon_{t-1} \geq 0 \) and \( d_{t-1} = 0 \), and the effect of the \( \varepsilon_{t-1} \) shock on volatility \( h_t \) is \( \alpha_1 \varepsilon_{t-1}^2 \). Hence, if \( \lambda_1 > 0 \), negative shocks will have a larger effect on volatility than positive shocks. The statistical significance of the \( \lambda_1 \) coefficient will determine whether there is a threshold effect on the conditional volatility.

The implementation of TARCH models in Stata has the following form

\texttt{use http://www.stata-press.com/data/r11/wpi1, clear}
\texttt{arch D.ln_wpi, ar(1) ma(1 4) arch(1) garch(1) tarch(1)}

(-setting optimization to BHHH)

Iteration 45: log likelihood = 403.871

ARCH family regression -- ARMA disturbances

Sample: 1960q2 - 1990q4
Number of obs = 123
Distribution: Gaussian
Wald chi2(3) = 279.50
Log likelihood = 403.871
Prob > chi2 = 0.0000

| D.ln_wpi | Coef. Std. Err. z P>|z| [95% Conf. Interval] |
|-----------|-----------------|----------|--------|-----------------|
| ln_wpi _cons | .0092552 .0042914 2.16 0.031 .0008442 .0176662 |

<table>
<thead>
<tr>
<th>ARMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ar L1.</td>
</tr>
<tr>
<td>ma L1.</td>
</tr>
<tr>
<td>L4.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>arch L1.</td>
</tr>
<tr>
<td>tarch L1.</td>
</tr>
<tr>
<td>garch L1.</td>
</tr>
<tr>
<td>_cons</td>
</tr>
</tbody>
</table>

Let \( y_t = \delta \log(w_{pi_t}) \), then the estimated equations are written as

\[ y_t = 0.009 + 0.846y_{t-1} - 0.392\varepsilon_{t-1} + 0.219\varepsilon_{t-4} + \varepsilon_t \]  

(5.55)

\[ h_t = 0.000 - 0.066\varepsilon_{t-1}^2 + 0.509d_{t-1}\varepsilon_{t-1}^2 + 0.725h_{t-1} \]  

(5.56)

where the threshold effect is statistically significant at a 10% level.
5.6.2 EGARCH Models

A second model that allows for asymmetries in the effect of $\varepsilon_t$ is the exponential-GARCH (EGARCH) model, as proposed in Nelson (1991). One constraint with the regular GARCH model is that it requires that all estimated coefficients are positive. The EGARCH specify the conditional variance in the following way

$$\log(h_t) = \alpha_0 + \alpha_1 (\varepsilon_{t-1}/\sqrt{h_{t-1}}) + \lambda_1 |\varepsilon_{t-1}/\sqrt{h_{t-1}}| + \beta_1 \log(h_{t-1}).$$  \hspace{1cm} (5.57)

There are three features about this specification:

1. Because of the $\log(h_t)$ form, the implied value of $h_t$ can never be negative (regardless of the values of the coefficients).
2. Instead of using $\varepsilon_{t-1}$, EGARCH uses a standardized value $\varepsilon_{t-1}/\sqrt{h_{t-1}}$ (which is a unit free measure).
3. EGARCH allows for leverage effects. If $\varepsilon_{t-1}/\sqrt{h_{t-1}} > 0$, the effect of the shock on the log of the conditional variance is $\alpha_1 + \lambda_1$. If $\varepsilon_{t-1}/\sqrt{h_{t-1}} < 0$, the effect is $-\alpha_1 + \lambda_1$.

The Stata manual has the following example for the estimation of an EGARCH model.

```
use http://www.stata-press.com/data/r11/wpi1, clear
arch D.ln_wpi, ar(1) ma(1 4) earch(1) egarch(1)
(setting optimization to BHHH)
Iteration 0: log likelihood = 227.5251
Iteration 23: log likelihood = 405.31453

ARCH family regression -- ARMA disturbances
Sample: 1960q2 - 1990q4 Number of obs = 123
Distribution: Gaussian Wald chi2(3) = 156.03
Log likelihood = 405.3145 Prob > chi2 = 0.0000

------------------------------------------------------------------------------
| OPG       D.ln_wpi | Coef.     Std. Err.  z    P>|z|   [95% Conf. Interval]  
-------------+-----------------+-------------------+----------+-----------------+---------------------------+
| OPG       ln_wpi  |                 |                   |          |                |
| _cons     | 0.0087348       | 0.0034005         | 2.570000 | 0.010000       | 0.0020699 - 0.0153997     |

| ARMA      |
| L1.       | 0.7692179       | 0.0968355         | 7.987400 | 0.000000       | 0.5794238 - 0.9590121     |
| L4.       | 0.2414647       | 0.0863819         | 2.800063 | 0.005000       | 0.0721593 - 0.4107701     |

| ARCH      |
| L1.       | 0.406411        | 0.1163508         | 3.490000 | 0.000000       | 0.1783677 - 0.6344543     |
| L1.       | 0.246747        | 0.1233373         | 2.000000 | 0.045000       | 0.0050103 - 0.4884837     |
| L1.       | 0.8417259       | 0.0704077         | 11.960000| 0.000000       | 0.7037294 - 0.9797225     |

------------------------------------------------------------------------------
```
\[ y_t = \delta \log(w_{pi,t}) \], we have
\[ y_t = 0.009 + 0.769 y_{t-1} - 0.355 \varepsilon_{t-1} + 0.241 \varepsilon_{t-4} + \varepsilon_t \quad (5.58) \]
\[ \log(h_t) = -1.488 + 0.406(\varepsilon_{t-1}/\sqrt{h_{t-1}}) \]
\[ + 0.247(\varepsilon_{t-1}/\sqrt{h_{t-1}} - \sqrt{2/\pi}) + 0.842 \log(h_{t-1}) \quad (5.59) \]

Under normally distributed errors \((\varepsilon_t)\), \(\varepsilon_t/\sqrt{h_t}\) follows a standard normal distribution. The results indicate a strong leverage effect. The positive \(\alpha_1\) coefficient implies that positive innovations (unanticipated price increases) are more destabilizing than negative innovations. This effect is larger than the symmetric effect \((\lambda_1)\). Notice that the coefficients can be directly compared because of the standardization of \(\varepsilon_{t-1}/\sqrt{h_{t-1}}\).

### 5.7 Multivariate GARCH Models

If more than one variable appears on the analysis, the conditional volatilities can be modeled jointly. This helps because there may exist contemporaneous shocks that affect the volatility of more than one of the variables. Multivariate GARCH modeling allows for this, in addition to allowing for volatility spillovers where volatility shocks to one variable might affect the volatility of a related variable.

If we want to jointly model \(k\) variables we can define the symmetric variance-covariance matrix of the shocks as

\[
H = \begin{bmatrix}
  h_{11t} & h_{12t} & \ldots & h_{1kt} \\
  h_{21t} & h_{22t} & \ldots & h_{2kt} \\
  \vdots & \vdots & \ddots & \vdots \\
  h_{1kt} & h_{2kt} & \ldots & h_{kkt}
\end{bmatrix}
\]

where \(h_{ijt}\) are just the time \(t\) conditional variance of the shock \(i\) if \(i = j\), and it is the conditional covariance if \(i \neq j\) (i.e., \(h_{ijt} = E_{t-1}(\varepsilon_i \varepsilon_j)\)). Let \(|H|\) be the determinant of \(H\). Then the log-likelihood function to estimate the coefficients is

\[
\log L = -\frac{T}{2} \log(2\pi)^k - \frac{1}{2} \sum_{t=1}^{T} (\log |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t), \quad (5.60)
\]

where \(\varepsilon_t\) in this case is the \(k\) vector \(\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})\).

To keep things simple suppose that there are only two variables, \(y_{1t}\) and \(y_{2t}\). Let the error processes be
We assume that \( \text{var}(v_1) = \text{var}(v_2) = 1 \). The \textit{vech} model is the construction of multivariate GARCH(1,1) process where all volatility terms are allowed to interact with each other. That is,

\[
\begin{align*}
    h_{11t} & = c_{10} + \alpha_{11} \varepsilon_{1t-1}^2 + \alpha_{12} \varepsilon_{1t-1} \varepsilon_{2t-1} + \alpha_{13} \varepsilon_{2t-1}^2 + \beta_{11} h_{11t-1} + \beta_{12} h_{12t-1} + \beta_{13} h_{22t-1} \\
    h_{12t} & = c_{20} + \alpha_{21} \varepsilon_{1t-1}^2 + \alpha_{22} \varepsilon_{1t-1} \varepsilon_{2t-1} + \alpha_{23} \varepsilon_{2t-1}^2 + \beta_{21} h_{11t-1} + \beta_{22} h_{12t-1} + \beta_{23} h_{22t-1} \\
    h_{22t} & = c_{30} + \alpha_{31} \varepsilon_{1t-1}^2 + \alpha_{32} \varepsilon_{1t-1} \varepsilon_{2t-1} + \alpha_{33} \varepsilon_{2t-1}^2 + \beta_{31} h_{11t-1} + \beta_{32} h_{12t-1} + \beta_{33} h_{22t-1}
\end{align*}
\]

These equations show that each conditional variance depends on its own past, the conditional covariance between the two variables, the lagged square errors, and the product of lagged errors. As simple as the model in Equations 5.63, 5.64, and 5.65 appears to be, it is actually difficult to estimate because of the following reasons:

1. There is a large number of parameters that need to be estimated. In the 2 variable case there are 21 parameters plus the parameters in the mean equations.
2. There is no analytical solution for the maximization problem detailed in log-likelihood function of Equation 5.60. Numerical methods do not always find the solution.
3. Because the conditional variance need to be positive, we need to impose restrictions that are more complicated than in the univariate case.

A number of solutions have been proposed to circumvent these problems. A popular solution is to use a diagonal system such that \( h_{ij} \) contains only lags of itself and the cross products of \( \varepsilon_{it} \varepsilon_{jt} \). For example,

\[
\begin{align*}
    h_{11t} & = c_{10} + \alpha_{11} \varepsilon_{1t-1}^2 + \beta_{11} h_{11t-1} \\
    h_{12t} & = c_{20} + \alpha_{22} \varepsilon_{1t-1} \varepsilon_{2t-1} + \beta_{22} h_{12t-1} \\
    h_{22t} & = c_{30} + \alpha_{33} \varepsilon_{2t-1}^2 + \beta_{33} h_{22t-1}
\end{align*}
\]

While this specification is easier to estimate, there are no interactions among the variances.

Another popular solution is the constant-conditional-correlation-GARCH (CCC-GARCH). This model restricts the correlation coefficients to be constant. Hence, for each \( i \neq j \), the model assumes \( h_{ij} = \rho_{ij} \sqrt{h_{ii}h_{jj}} \). While the variance terms are not diagonalized, the covariance terms are always proportional to \( \sqrt{h_{ii}h_{jj}} \). For example, building on the model in Equations 5.63, 5.64, and 5.65,
5.7 Multivariate GARCH Models

\[ h_{12t} = \rho_{12} \sqrt{h_{11t} h_{22t}} \]  

(5.69)

This makes the covariance equation consist of only one parameter instead of seven.\(^1\)

Consider the estimation of the following diagonal \(v\text{ech}\) multivariate GARCH model in Stata. Following the Stata manual, we have data on a secondary market rate of a six-month U.S. Treasury bill, \(tbill\), and on Moody’s seasoned AAA corporate bond yield, \(bond\). We model the first-differences of both, \(tbill\) and \(bond\), in a VAR(1) with an ARCH(1) term,

use http://www.stata-press.com/data/r11/irates4

dvech (D.bond D.tbill = LD.bond LD.tbill), arch(1)

Getting starting values
(setting technique to bhhh)

Iteration 0: log likelihood = 3569.2723
Iteration 8: log likelihood = 4221.6577

Diagonal vech multivariate GARCH model

Sample: 3 - 2456  Number of obs = 2454
Wald chi2(4) = 1183.52
Log likelihood = 4221.6577  Prob > chi2 = 0.0000

------------------------------------------------------------------------------
| Coef.  Std. Err.  z    P>|z|    [95% Conf. Interval] |
-------------+--------------------------------------------------------------------
D.bond       |                      |
    bond      | .2967674   .0247149 12.01  0.000   .2483271   .3452077   |
    LD.       | .0947949   .0098683  9.61  0.000   .0754533   .1141364   |
    _cons     | .0003991   .00143   0.28  0.780   -.0024036   .0032019   |
-------------+--------------------------------------------------------------------
D.tbill      |                      |
    bond      | .0108373   .0301501  0.36  0.719   -.0482558   .0699304   |
    LD.       | .4344747   .0176497 24.62  0.000   .3998819   .4690675   |
    _cons     | .0011611   .0021033  0.55  0.581   -.0029612   .0052835   |
-------------+--------------------------------------------------------------------
Sigma0       |                      |
   1_1        | .004894     .002006  24.40  0.000   .0045008   .0052871   |
   2_1        | .0004986    .0002396 17.10  0.000   .0002628   .0007358   |
   2_2        | .0115149    .0005227 22.03  0.000   .0104904   .0125395   |
-------------+--------------------------------------------------------------------
L.ARCH       |                      |
   1_1        | .4514942    .0456835  9.88  0.000   .3619562   .5410323   |
   2_1        | .2518879    .036736   6.86  0.000   .1798866   .3238893   |
   2_2        | .843368     .0608055 13.87  0.000   .7241914   .9625446   |
-------------+--------------------------------------------------------------------

The estimated equations are

\(^1\) Engel (2002) implemented a generalized CCC-GARCH that allowed the correlations to change over time. This is called the dynamic-conditional-correlation-GARCH (DCC-GARCH). Stata 12 can estimate CCC-GARCH, DCC-GARCH and VCC-GARCH models, the VCC stands for varying conditional correlation.
\[
\Delta \text{bond} = 0.0004 + 0.296 \Delta \text{bond}_{t-1} + 0.095 \Delta \text{tbill}_{t-1} + \varepsilon_{1t} \quad (5.70)
\]
\[
\Delta \text{tbill} = 0.0011 + 0.011 \Delta \text{bond}_{t-1} + 0.434 \Delta \text{tbill}_{t-1} + \varepsilon_{2t} \quad (5.71)
\]
\[
H = \begin{bmatrix}
0.0048\varepsilon_{1t-1}^2 + 0.4515h_{1t-1} & 0.0040\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.2519h_{2t-1} \\
0.0040\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.2519h_{1t-1} & 0.0115\varepsilon_{2t-1}^2 + 0.8434h_{22t-1}
\end{bmatrix}
\]
or
\[
h_{11t} = 0.0048\varepsilon_{1t-1}^2 + 0.4515h_{1t-1} \quad (5.72)
\]
\[
h_{12t} = 0.0040\varepsilon_{1t-1}\varepsilon_{2t-1} + 0.2519h_{12t-1} \quad (5.73)
\]
\[
h_{22t} = 0.0115\varepsilon_{2t-1}^2 + 0.8434h_{22t-1} \quad (5.74)
\]

### 5.8 Supporting .do files

For Figure 5.4:
```stata
use rgdp.dta, clear
tset date
gen y = log(rgdp/l.rgdp)
label variable rgdp "Real GDP"
label variable y "Real GDP Growth Rate"
label variable statadate "Date"
ttitle("Real GDP")
ttitle("Real GDP Growth Rate")
title("Real GDP and Real GDP Growth Rate")
```