CHAPTER **14**

Time-Varying Volatility and ARCH Models

LEARNING OBJECTIVES

Based on the material in this chapter, you should be able to

1. Explain the difference between a constant and a time-varying variance of the error term.
2. Explain the term "conditionally normal."
3. Perform a test for ARCH effects.
4. Estimate an ARCH model.
5. Forecast volatility.
6. Explain the difference between ARCH and GARCH specifications.
7. Explain the distinctive features of a T-GARCH model and a GARCH-in-mean model.

KEYWORDS

ARCH
ARCH-in-mean
conditionally normal
GARCH
GARCH-in-mean
T-ARCH and T-GARCH
time-varying variance

In Chapter 12, our focus was on time-varying mean processes and macroeconomic time series. We were concerned with stationary and nonstationary variables, and, in particular, macroeconomic variables like gross domestic product (GDP), inflation, and interest rates. The nonstationary nature of the variables implied that they had **means that change over time**. In this chapter, we are concerned with stationary series, but with conditional variances that change over time. The model we focus on is called the autoregressive conditional heteroskedastic (**ARCH**) model.

Nobel Prize winner Robert Engle's original work on ARCH was concerned with the volatility of inflation. However, it was applications of the ARCH model to financial time series that established and consolidated the significance of his contribution. For this reason, the examples used in this chapter will be based on financial time series. As we will see, financial time series have characteristics that are well represented by models with dynamic variances. The particular aims of this chapter are to discuss the modeling of dynamic variances using the ARCH class of models of volatility, the estimation of these models, and their use in forecasting.

14.1 The ARCH Model

ARCH stands for **a**uto**r**egressive **c**onditional **h**eteroskedasticity. We have covered the concepts of autoregressive and heteroskedastic errors in Chapters 9 and 8, respectively, so let us begin with a discussion of the concepts of conditional and unconditional means and variances of the error term.

Consider a model with an AR(1) error term

$$y_t = \phi + e_t \tag{14.1a}$$

$$e_t = \rho e_{t-1} + v_t, \qquad |\rho| < 1 \tag{14.1b}$$

$$v_t \sim N\left(0, \sigma_v^2\right) \tag{14.1c}$$

For convenience of exposition, first perform some successive substitution to obtain e_t as the sum of an infinite series of the error term v_t. To do this, note that if $e_t = \rho e_{t-1} + v_t$, then $e_{t-1} = \rho e_{t-2} + v_{t-1}$ and $e_{t-2} = \rho e_{t-3} + v_{t-2}$, and so on. Hence $e_t = v_t + \rho^2 v_{t-2} + \cdots + \rho^t e_0$ where the final term $\rho^t e_0$ is negligible.

The **unconditional mean** of the error is

$$E\left[e_t\right] = E\left[v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \cdots\right] = 0$$

because $E\left[v_{t-j}\right] = 0$ for all j, whereas the **conditional mean** for the error, conditional on information prior to time t, is

$$E\left[e_t | I_{t-1}\right] = E\left[\rho e_{t-1} | I_{t-1}\right] + E\left[v_t\right] = \rho e_{t-1}$$

The information set at time $t-1$, I_{t-1}, includes knowing ρe_{t-1}. Put simply, "unconditional" describes the situation when you have no information, whereas conditional describes the situation when you have information, up to a certain point in time.

The **unconditional variance** of the error is

$$\begin{aligned} E\left[e_t - 0\right]^2 &= E\left[v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \cdots\right]^2 \\ &= E\left[v_t^2 + \rho^2 v_{t-1}^2 + \rho^4 v_{t-2}^2 + \cdots\right] \\ &= \sigma_v{}^2\left[1 + \rho^2 + \rho^4 + \cdots\right] = \frac{\sigma_v{}^2}{1 - \rho^2} \end{aligned}$$

because $E\left[v_{t-j} v_{t-i}\right] = \sigma_v^2$ when $i = j$; $E\left[v_{t-j} v_{t-i}\right] = 0$ when $i \neq j$ and the sum of a geometric series $\left[1 + \rho^2 + \rho^4 + \cdots\right]$ is $1/\left(1 - \rho^2\right)$. The **conditional variance** for the error is

$$E\left[\left(e_t - \rho e_{t-1}\right)^2 \middle| I_{t-1}\right] = E\left[v_t^2 | I_{t-1}\right] = \sigma_v^2$$

Now notice, for this model, that the conditional mean of the error varies over time, while the conditional variance does not. Suppose that instead of a conditional mean that changes over time, we have a conditional variance that changes over time. To introduce this modification, consider a variant of the above model

$$y_t = \beta_0 + e_t \tag{14.2a}$$

$$e_t | I_{t-1} \sim N\left(0, h_t\right) \tag{14.2b}$$

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2, \quad \alpha_0 > 0, \quad 0 \leq \alpha_1 < 1 \tag{14.2c}$$

Equations (14.2b and 14.2c) describe the ARCH class of models. The second equation (14.2b) says that the error term is **conditionally normal** $e_t | I_{t-1} \sim N\left(0, h_t\right)$ where I_{t-1} represents the information available at time $t-1$ with mean 0 and time-varying variance, denoted as h_t, following popular terminology. The third equation (14.2c) models h_t as a function of a constant term and the lagged error squared e_{t-1}^2.

The name ARCH conveys the fact that we are working with time-varying variances (heteroskedasticity) that depend on (are conditional on) lagged effects (autocorrelation). This particular example is an ARCH(1) model since the time-varying variance h_t is a function of a constant term (α_0) plus a term lagged once, the square of the error in the previous period $(\alpha_1 e_{t-1}^2)$. The coefficients, α_0 and α_1, have to be positive to ensure a positive variance. The coefficient α_1 must be less than 1, or h_t will continue to increase over time, eventually exploding. Conditional normality means that the normal distribution is a function of known information at time $t-1$; i.e., when $t=2$, $e_2|I_1 \sim N(0, \alpha_0 + \alpha_1 e_1^2)$ and when $t=3$, $e_3|I_2 \sim N(0, \alpha_0 + \alpha_1 e_2^2)$, and so on. In this particular case, conditioning on I_{t-1} is equivalent to conditioning on the square of the error in the previous period e_{t-1}^2.

Note that while the conditional distribution of the error e_t is assumed to be normal, the unconditional distribution of the error e_t will not be normal. This is not an inconsequential consideration given that a lot of real-world data appear to be drawn from non-normal distributions.

We have noted that, conditional on e_{t-1}^2, the mean and variance of the error term e_t are zero and h_t, respectively. To find the mean and variance of the unconditional distribution of e_t, we note that, conditional on e_{t-1}^2, the standardized errors are standard normal, that is,

$$\left(\frac{e_t}{\sqrt{h_t}}\middle| I_{t-1}\right) = z_t \sim N(0, 1)$$

Because this distribution does not depend on e_{t-1}^2, it follows that the unconditional distribution of $z_t = \left(e_t/\sqrt{h_t}\right)$ is also $N(0, 1)$, and that z_t and e_{t-1}^2 are independent. Thus, we can write

$$E(e_t) = E(z_t)E\left(\sqrt{\alpha_0 + \alpha_1 e_{t-1}^2}\right)$$

and

$$E(e_t^2) = E(z_t^2)E(\alpha_0 + \alpha_1 e_{t-1}^2) = \alpha_0 + \alpha_1 E(e_{t-1}^2)$$

From the first of these equations, we get $E(e_t) = 0$ because $E(z_t) = 0$. From the second of the equations, we get $\text{var}(e_t^2) = E(e_t^2) = \alpha_0/(1-\alpha_1)$ because $E(z_t^2) = 1$ and $E(e_t^2) = E(e_{t-1}^2)$.

The ARCH model has become a very important econometric model because it is able to capture stylized features of real-world volatility. Furthermore, in the context of the ARCH(1) model, knowing the squared error in the previous period e_{t-1}^2 improves our knowledge about the likely magnitude of the variance in period t. This is useful for situations when it is important to understand risk, as measured by the volatility of the variable.

14.2 Time-Varying Volatility

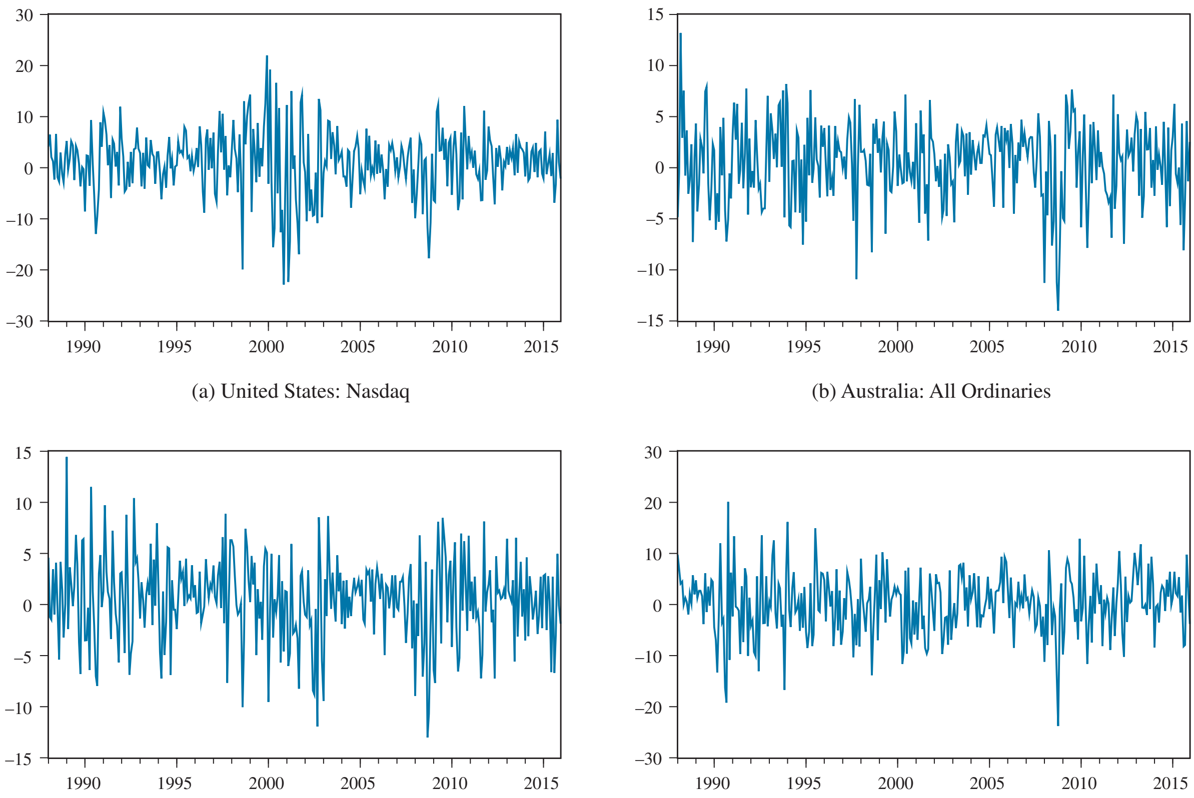
The ARCH model has become a popular one because its variance specification can capture commonly observed features of the time series of financial variables; in particular, it is useful for modeling **volatility** and especially changes in volatility over time. To appreciate what we mean by volatility and time-varying volatility, and how it relates to the ARCH model, let us look at some stylized facts about the behavior of financial variables—for example, the returns to stock price indices (also known as share price indices).

EXAMPLE 14.1 | Characteristics of Financial Variables

Figure 14.1 shows the time series of the monthly returns to a number of stock prices; namely, the U.S. Nasdaq, the Australian All Ordinaries, the Japanese Nikkei, and the UK FTSE over the period 1988M1 to 2015M12 (data file *returns5*). The values of these series change rapidly from period to period in an apparently unpredictable manner; we say the series are volatile. Furthermore, there are periods when large changes are followed by further large changes and periods when small changes are followed by further small changes. In this case the series are said to display time-varying volatility as well as "clustering" of changes.

Figure 14.2 shows the histograms of the returns. All returns display non-normal properties. We can see this more clearly if we draw normal distributions (using the respective sample means and sample variances) on top of these histograms. Note that there are more observations around the mean and in the tails. Distributions with these properties—more peaked around the mean and relatively fat tails—are said to be **leptokurtic**.

Note that the assumption that the conditional distribution for $(y_t|I_{t-1})$ is normal, an assumption that we made in (14.2b), does not necessarily imply that the unconditional distribution for y_t is normal. When we collect empirical observations on y_t into a histogram, we are constructing an estimate of the unconditional distribution for y_t. What we have observed is that the unconditional distribution for y_t is leptokurtic.



(a) United States: Nasdaq

(b) Australia: All Ordinaries

(c) United Kingdom: FTSE

(d) Japan: Nikkei

FIGURE 14.1 **Time series of returns to stock indices.**

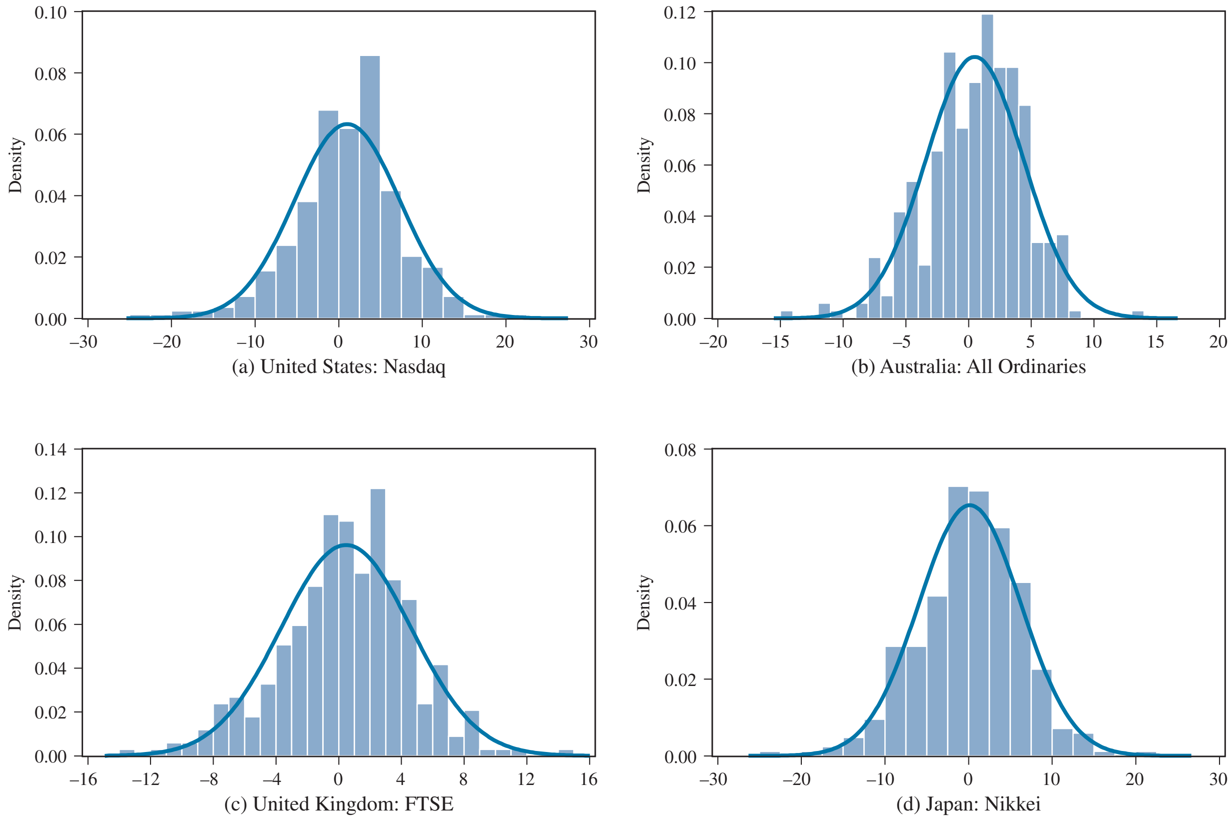


FIGURE 14.2 **Histograms of returns to stock indices.**

EXAMPLE 14.2 | Simulating Time-Varying Volatility

To illustrate how the ARCH model can be used to capture changing volatility and the leptokurtic nature of the distribution for y_t, we generate some simulated data for two models. In both cases we set $\beta_0 = 0$ so that $y_t = e_t$. The left panel in Figure 14.3 illustrates the case when $\alpha_0 = 1$, $\alpha_1 = 0$. These values imply $\text{var}(y_t|I_{t-1}) = h_t = 1$. This variance is constant, and not time varying, because $\alpha_1 = 0$. The right panel in Figure 14.3 illustrates the case when $\alpha_0 = 1$, $\alpha_1 = 0.8$, the case of a time-varying variance given by $\text{var}(y_t|I_{t-1}) = h_t = \alpha_0 + \alpha_1 e_{t-1}^2 = 1 + 0.8e_{t-1}^2$. Note that relative to the series in the left panel, volatility in the right panel is not constant; rather, it changes over time and it clusters—there are periods of small changes (e.g., around observation 100) and periods of big changes (around observation 175).

In Figure 14.4, we present histograms of y_t for the two cases. The top panel is the histogram for the constant variance case where $(y_t|I_{t-1})$ and y_t have the same distribution, namely the noise process $y_t \sim N(0, 1)$ because $h_t = 1$. The bottom panel is the histogram for the time-varying variance case. We know that the conditional distribution for $(y_t|I_{t-1})$

is $N(0, h_t)$. But what about the unconditional distribution for y_t? Again, we can check for normality by superimposing a normal distribution on top of the histogram. In this case, to allow for a meaningful comparison with the histogram in the top panel, we plot the standardized observations of y_t. That is, for each observation we subtract the sample mean and divide by the sample standard deviation. This transformation ensures that the distribution will have a zero mean and variance one, but it preserves the shape of the distribution. Comparing the two panels, we note that the second distribution has higher frequencies around the mean (zero) and higher frequencies in the tails (outside ± 3). This feature of time series with ARCH errors—the unconditional distribution of y_t is non-normal—is consistent with what we observed in the stock return series.

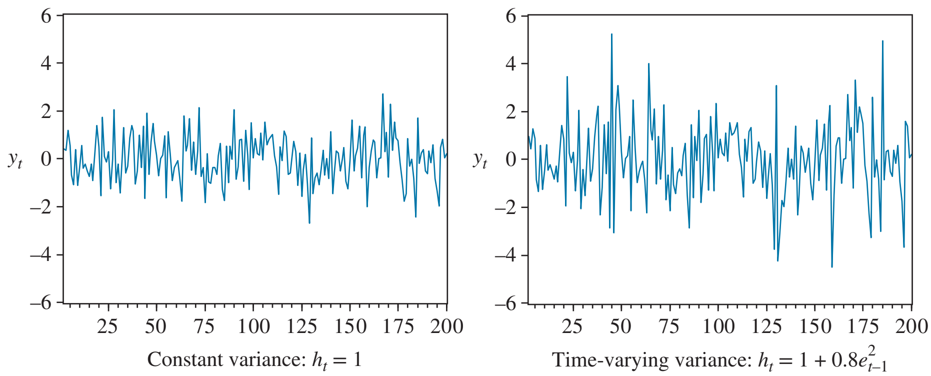


FIGURE 14.3 **Simulated examples of constant and time-varying variances.**

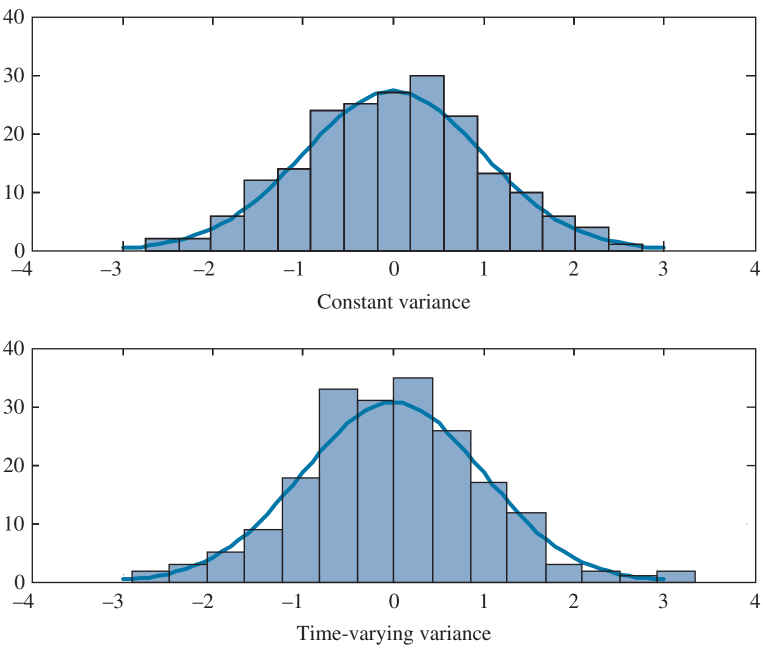


FIGURE 14.4 **Frequency distributions of the simulated models.**

Thus, the ARCH model is intuitively appealing because it seems sensible to explain volatility as a function of the errors e_t. These errors are often called "shocks" or "news" by financial analysts. They represent the unexpected! According to the ARCH model, the larger the shock, the greater the volatility in the series. In addition, this model captures volatility clustering, as big changes in e_t are fed into further big changes in h_t via the lagged effect e_{t-1}. The simulations show how well the ARCH model mimics the behavior of financial time series shown in Figure 14.1, including their non-normal distributions.

14.3 Testing, Estimating, and Forecasting

A Lagrange multiplier (LM) test is often used to test for the presence of ARCH effects. To perform this test, first estimate the **mean equation**, which can be a regression of the variable on a constant (like 14.1) or may include other variables. Then save the estimated residuals $\hat{e}_t$ and obtain their squares $\hat{e}_t^2$. To test for first-order ARCH, regress $\hat{e}_t^2$ on the squared residuals lagged $\hat{e}_{t-1}^2$,

$$\hat{e}_t^2 = \gamma_0 + \gamma_1 \hat{e}_{t-1}^2 + v_t \tag{14.3}$$

where v_t is a random term. The null and alternative hypotheses are

$$H_0: \gamma_1 = 0 \qquad H_1: \gamma_1 \neq 0$$

If there are no ARCH effects, then $\gamma_1 = 0$ and the fit of (14.3) will be poor, and the equation R^2 will be low. If there are ARCH effects, we expect the magnitude of $\hat{e}_t^2$ to depend on its lagged values, and the R^2 will be relatively high. The LM test statistic is $(T-q)R^2$ where T is the sample size, q is the number of $\hat{e}_{t-j}^2$ terms on the right-hand side of (14.3), and R^2 is the coefficient of determination. If the null hypothesis is true, then the test statistic $(T-q)R^2$ is distributed (in large samples) as $\chi^2_{(q)}$, where q is the order of lag, and $T-q$ is the number of complete observations; in this case, $q = 1$. If $(T-q)R^2 \geq \chi^2_{(1-\alpha,\, q)}$, then we reject the null hypothesis that $\gamma_1 = 0$ and conclude that ARCH effects are present.

EXAMPLE 14.3 | Testing for ARCH in BrightenYourDay (BYD) Lighting

To illustrate the test, consider the returns from buying shares in the hypothetical company BYD Lighting. The time series and histogram of the returns are shown in Figure 14.5 (data file *byd*). The time series shows evidence of time-varying volatility and clustering, and the unconditional distribution is non-normal.

To perform the test for ARCH effects, first estimate a mean equation that in this example is $r_t = \beta_0 + e_t$, where r_t is the monthly return on shares of BYD. Second, retrieve the estimated residuals. Third, estimate (14.3). The results for the ARCH test are

$$\hat{e}_t^2 = \underset{(8.409)}{0.908 + 0.353\hat{e}_{t-1}^2} \qquad R^2 = 0.124$$
$$(t)$$

The t-statistic suggests a significant first-order coefficient. The sample size is 500, giving an LM test value of $(T-q)R^2 = 61.876$. Comparing the computed test value to the 5% critical value of a $\chi^2_{(1)}$ distribution $\left(\chi^2_{(0.95,1)} = 3.841\right)$ leads to the rejection of the null hypothesis. In other words, the residuals show the presence of ARCH(1) effects.

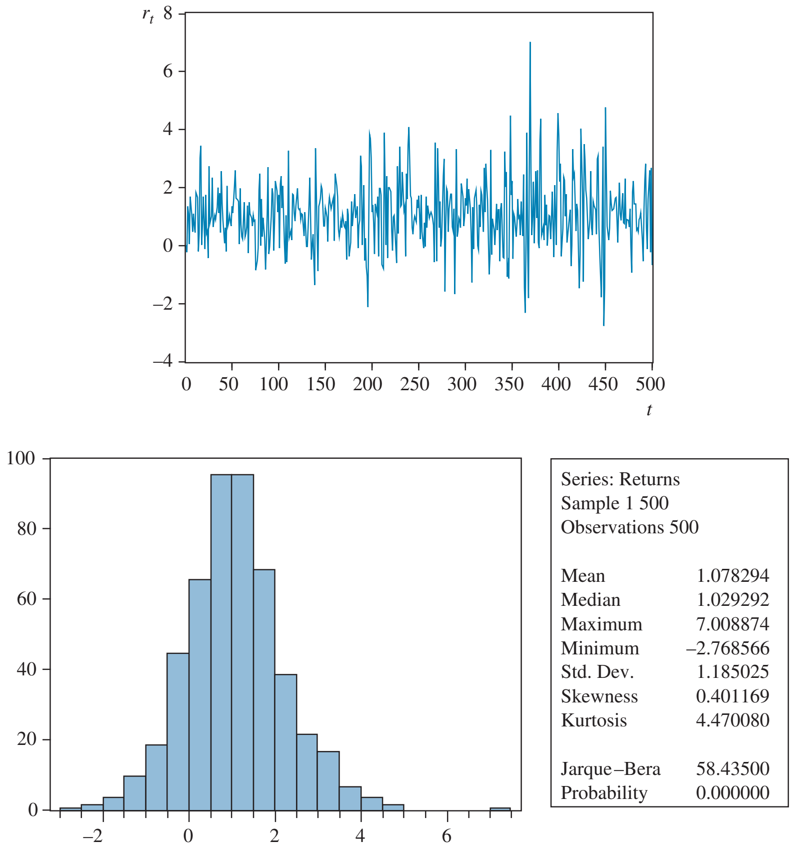


FIGURE 14.5 **Time series and histogram of returns for BYD lighting.**

EXAMPLE 14.4 | ARCH Model Estimates for BrightenYourDay (BYD) Lighting

ARCH models are estimated by the maximum likelihood method. Estimation details are beyond the scope of this book, but the maximum likelihood method (see Appendix C.8) is programmed in most econometric software.

Equation (14.4) shows the results from estimating an ARCH(1) model applied to the monthly returns from buying shares in BrightenYourDay Lighting. The estimated mean of the series is described in (14.4a), while the estimated variance is given in (14.4b).

$$\hat{r}_t = \hat{\beta}_0 = 1.063 \qquad (14.4a)$$

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{e}^2_{t-1} = 0.642 + \underset{(5.536)}{0.569}\hat{e}^2_{t-1} \qquad (14.4b)$$

(t)

The t-statistic of the first-order coefficient (5.536) suggests a significant ARCH(1) coefficient. Recall that one of the requirements of the ARCH model is that $\alpha_0 > 0$ and $\alpha_1 > 0$, so that the implied variances are positive. Note that the estimated coefficients $\hat{\alpha}_0$ and $\hat{\alpha}_1$ satisfy this condition.

EXAMPLE 14.5 | Forecasting BrightenYourDay (BYD) Volatility

Once we have estimated the model, we can use it to forecast next period's return r_{t+1} and the conditional volatility h_{t+1}. When one invests in shares (or stocks), it is important to choose them not just on the basis of their mean returns, but also on the basis of their risk. Volatility gives us a measure of their risk.

For our case study of investing in BYD Lighting, the forecast return and volatility are

$$\hat{r}_{t+1} = \hat{\beta}_0 = 1.063 \tag{14.5a}$$

$$\hat{h}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1\left(r_t - \hat{\beta}_0\right)^2 = 0.642 + 0.569\left(r_t - 1.063\right)^2 \tag{14.5b}$$

Equation (14.5a) gives the estimated return that—because it does not change over time—is both the conditional and unconditional mean return. The estimated error in period t, given by $\hat{e}_t = r_t - \hat{r}_t$, can then be used to obtain the estimated conditional variance (14.5b). The time series of the conditional variance does change over time and is shown in Figure 14.6. Note how the conditional variance around observation 370 coincides with the period of large changes in returns as shown in Figure 14.5.

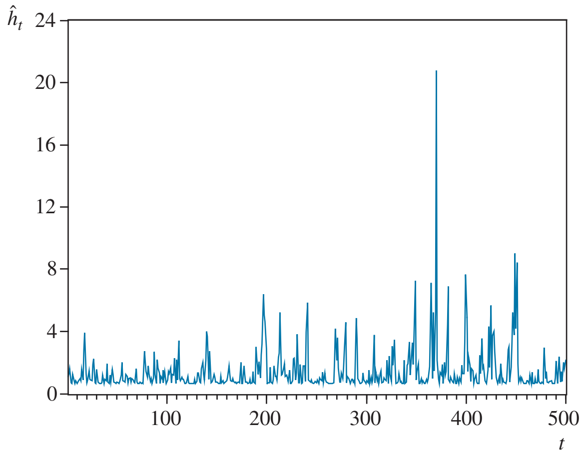


FIGURE 14.6 **Plot of conditional variance.**

14.4 Extensions

The ARCH(1) model can be extended in a number of ways. One obvious extension is to allow for more lags. In general, an ARCH(q) model that includes lags $\hat{e}^2_{t-1}, \ldots, \hat{e}^2_{t-q}$ has a conditional variance function that is given by

$$h_t = \alpha_0 + \alpha_1 e^2_{t-1} + \alpha_2 e^2_{t-2} \cdots + \alpha_q e^2_{t-q} \tag{14.6}$$

In this case the variance or volatility in a given period depends on the magnitudes of the squared errors in the past q periods. Testing, estimating, and forecasting, are natural extensions of the case with one lag.

14.4.1 The GARCH Model—Generalized ARCH

One of the shortcomings of an ARCH(q) model is that there are $q + 1$ parameters to estimate. If q is a large number, we may lose accuracy in the estimation. The generalized ARCH model, or **GARCH**, is an alternative way to capture long lagged effects with fewer parameters. It is a special generalization of the ARCH model and it can be derived as follows. First, consider (14.6) but write it as

$$h_t = \alpha_0 + \alpha_1 e^2_{t-1} + \beta_1\alpha_1 e^2_{t-2} + \beta_1^2\alpha_1 e^2_{t-3} + \cdots$$

In other words, we have imposed a geometric lag structure on the lagged coefficients of the form $\alpha_s = \alpha_1\beta_1^{s-1}$. Next, add and subtract $\beta_1\alpha_0$ and rearrange terms as follows:

$$h_t = \left(\alpha_0 - \beta_1\alpha_0\right) + \alpha_1 e^2_{t-1} + \beta_1\left(\alpha_0 + \alpha_1 e^2_{t-2} + \beta_1\alpha_1 e^2_{t-3} + \cdots\right)$$

Then, since $h_{t-1} = \alpha_0 + \alpha_1 e^2_{t-2} + \beta_1\alpha_1 e^2_{t-3} + \beta_1^2\alpha_1 e^2_{t-4} + \cdots$, we may simplify to

$$h_t = \delta + \alpha_1 e^2_{t-1} + \beta_1 h_{t-1} \tag{14.7}$$

where $\delta = \left(\alpha_0 - \beta_1\alpha_0\right)$. This generalized ARCH model is denoted as GARCH(1, 1). It can be viewed as a special case of the more general GARCH (p, q) model, where p is the number of lagged h terms and q is the number of lagged e^2 terms. We also note that we need $\alpha_1 + \beta_1 < 1$ for stationarity; if $\alpha_1 + \beta_1 \geq 1$ we have a so-called "integrated GARCH" process, or IGARCH.

The GARCH(1, 1) model is a very popular specification because it fits many data series well. It tells us that the volatility changes with lagged shocks $\left(e^2_{t-1}\right)$ but there is also momentum in the system working via h_{t-1}. One reason why this model is so popular is that it can capture long lags in the shocks with only a few parameters. A GARCH(1, 1) model with three parameters $(\delta, \alpha_1,$ and $\beta_1)$ can capture similar effects to an ARCH(q) model requiring the estimation of $(q+1)$ parameters, where q is large, say $q \geq 6$.

EXAMPLE 14.6 | A GARCH Model for BrightenYourDay

To illustrate the GARCH(1, 1) specification, consider again the returns to our shares in BYD Lighting, which we re-estimate (by maximum likelihood) under the new model. The results are

$$\hat{r}_t = 1.049$$
$$\hat{h}_t = 0.401 + 0.492\hat{e}_{t-1}^2 + 0.238\hat{h}_{t-1}$$
$$(t) \qquad (4.834) \qquad (2.136)$$

The significance of the coefficient in front of $\hat{h}_{t-1}$ suggests that the GARCH(1, 1) model is better than the ARCH(1) results shown in (14.4). Plots of the mean equation and the time-varying variance are shown in Figures 14.7(a) and (b), respectively.

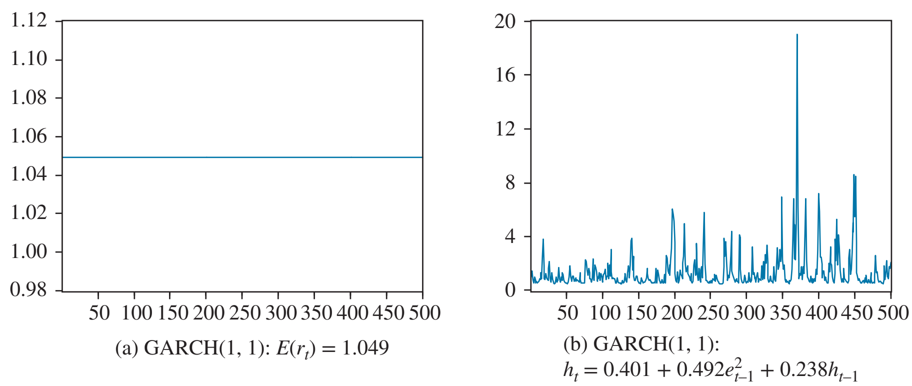


FIGURE 14.7 Estimated mean and variance of GARCH model.

14.4.2 Allowing for an Asymmetric Effect

A standard ARCH model treats bad "news" (negative $e_{t-1} < 0$) and good "news" (positive $e_{t-1} > 0$) symmetrically, that is, the effect on the volatility h_t is the same $(\alpha_1 e_{t-1}^2)$. However, the effects of good and bad news may have asymmetric effects on volatility. In general, when negative news hits a financial market, asset prices tend to enter a turbulent phase and volatility increases, but with positive news volatility tends to be small and the market enters a period of tranquility.

The threshold ARCH model, or **T-ARCH**, is one example where positive and negative news are treated asymmetrically. In the T-GARCH version of the model, the specification of the conditional variance is

$$h_t = \delta + \alpha_1 e_{t-1}^2 + \gamma d_{t-1} e_{t-1}^2 + \beta_1 h_{t-1}$$
$$d_t = \begin{cases} 1 & e_t < 0 \quad \text{(bad news)} \\ 0 & e_t \geq 0 \quad \text{(good news)} \end{cases} \tag{14.8}$$

where γ is known as the asymmetry or leverage term. When $\gamma = 0$, the model collapses to the standard GARCH form. Otherwise, when the shock is positive (i.e., good news) the effect on volatility is α_1, but when the news is negative (i.e., bad news) the effect on volatility is $\alpha_1 + \gamma$. Hence, if γ is significant and positive, negative shocks have a larger effect on h_t than positive shocks.

EXAMPLE 14.7 | A T-GARCH Model for BYD

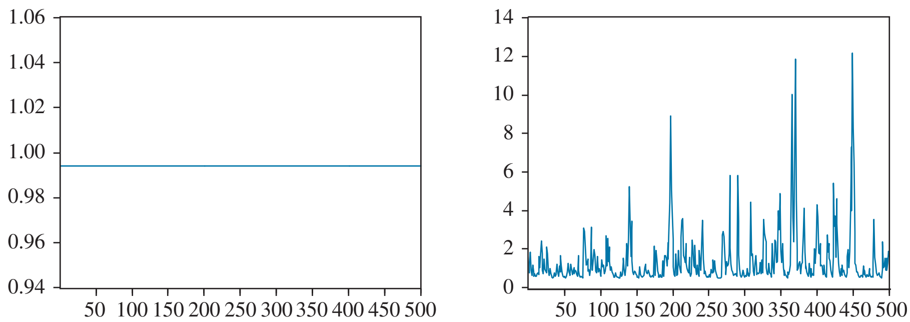
The returns to our shares in BYD Lighting were re-estimated with a T-GARCH(1,1) specification:

$$\hat{r}_t = 0.994$$

$$\hat{h}_t = 0.356 + 0.263\hat{e}^2_{t-1} + 0.492d_{t-1}\hat{e}^2_{t-1} + 0.287\hat{h}_{t-1}$$

$$(t) \qquad (3.267) \qquad (2.405) \qquad (2.488)$$

These results show that when the market observes good news (positive e_t), the contribution of e_t^2 to volatility h_{t+1} is by a factor 0.263, whereas when the market observes bad news (negative e_t), the contribution of e_t^2 to volatility h_{t+1} is by a factor $(0.263 + 0.492)$. Overall, negative shocks create greater volatility in financial markets. The mean and variance are displayed in Figures 14.8(a) and (b). Note that, relative to Figure 14.7(b), the T-GARCH model has highlighted the period around observation 200 as another period of turbulence.



(a) T-GARCH(1, 1): $E(r_t) = 0.994$

(b) T-GARCH(1, 1): $h_t = 0.356 + (0.263 + 0.492d_{t-1})e^2_{t-1} + 0.287h_{t-1}$

FIGURE 14.8 **Estimated mean and variance of T-GARCH model.**

14.4.3 GARCH-in-Mean and Time-Varying Risk Premium

Another popular extension of the GARCH model is the **GARCH-in-mean** model. The positive relationship between risk (often measured by volatility) and return is one of the basic tenets of financial economics. As risk increases, so does the mean return. Intuitively, the return to risky assets tends to be higher than the return to safe assets (low variation in returns) to compensate an investor for taking on the risk of buying the volatile share. However, while we have estimated the mean equation to model returns, and have estimated a GARCH model to capture time-varying volatility, we have not used the risk to explain returns. This is the aim of the GARCH-in-mean models.

The equations of a GARCH-in-mean model are shown below:

$$y_t = \beta_0 + \theta h_t + e_t \tag{14.9a}$$

$$e_t | I_{t-1} \sim N(0, h_t) \tag{14.9b}$$

$$h_t = \delta + \alpha_1 e^2_{t-1} + \beta_1 h_{t-1}, \quad \delta > 0, \quad 0 \le \alpha_1 < 1, \quad 0 \le \beta_1 < 1 \tag{14.9c}$$

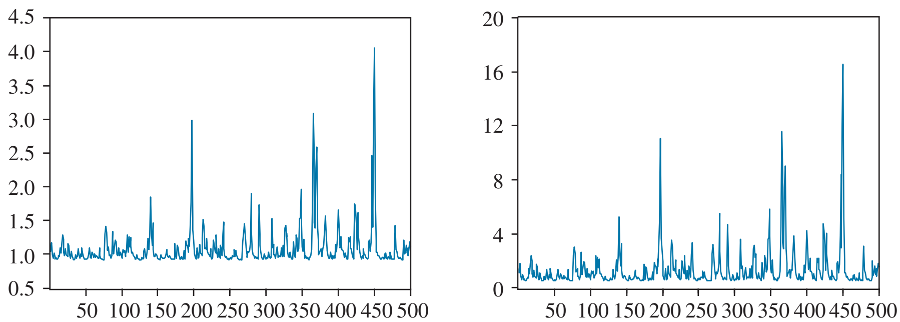
The first equation is the mean equation; it now shows the effect of the conditional variance on the dependent variable. In particular, note that the model postulates that the conditional variance h_t affects y_t by a factor θ. The other two equations are as before.

EXAMPLE 14.8 | A GARCH-in-Mean Model for BYD

The returns to shares in BYD Lighting were reestimated as a GARCH-in-mean model:

$$\hat{r}_t = 0.818 + 0.196h_t$$
$$(t) \qquad (2.915)$$
$$\hat{h}_t = 0.370 + 0.295\hat{e}^2_{t-1} + 0.321d_{t-1}\hat{e}^2_{t-1} + 0.278\hat{h}_{t-1}$$
$$(t) \qquad (3.426) \qquad (1.979) \qquad (2.678)$$

The results show that as volatility increases, the returns correspondingly increase by a factor of 0.196. In other words, this result supports the usual view in financial markets—high risk, high return. The GARCH-in-mean model is shown in Figures 14.9(a) and (b). Note that the expected mean return is no longer a constant value, but rather has high values (e.g., around observation 200) that coincide with higher conditional variances.



(a) GARCH-in-mean: $E(r_t) = 0.818 + 0.196h_t$

(b) GARCH-in-mean: $h_t = 0.370 + (0.295 + 0.321d_{t-1})e^2_{t-1} + 0.278h_{t-1}$

FIGURE 14.9 Estimated mean and variance of GARCH-in-mean model.

One last point before we leave this section. The first equation of the GARCH-in-mean model is sometimes written as a function of the time-varying standard deviation $\sqrt{h_t}$, that is, $y_t = \beta_0 + \theta\sqrt{h_t} + e_t$. This is because both measures—variance and standard deviation—are used by financial analysts to measure risk. There are no hard-and-fast rules about which measure to use. Exercise 14.8 illustrates the case when we use $\sqrt{h_t}$. A standard t-test of significance is often used to decide which is the more suitable measure.

14.4.4 Other Developments

The GARCH, T-GARCH, and GARCH-in-mean models are three important extensions of the original ARCH concept developed by Engle in 1982. There have also been numerous other variations, developed to handle complexities noted in the data, especially in high frequency financial data. One variation, exponential GARCH (EGARCH), has stood the test of time. This model is

$$\ln(h_t) = \delta + \beta_1 \ln(h_{t-1}) + \alpha\left|\frac{e_{t-1}}{\sqrt{h_{t-1}}}\right| + \gamma\left(\frac{e_{t-1}}{\sqrt{h_{t-1}}}\right)$$

where $\left(e_{t-1}/\sqrt{h_{t-1}}\right)$ are the standardized residuals. The model uses a log specification, which ensures the estimated variance remains positive. It also includes two standardized residual terms, with one of them in absolute form to facilitate the testing of the leverage effect. The **leverage effect** refers to the generally observed negative correlation between an asset return and its volatility changes. One potential explanation for this observation is that bad news has a bigger effect on variance than good news. If $\gamma \neq 0$, the effects of good/bad news are asymmetric; if $\gamma < 0$, negative shocks have larger effects.

Another significant development is to allow the conditional distribution of the error term to be non-normal. Because empirical distributions of financial returns generally exhibit fat tails and clustering around zero, the t-distribution has become a popular alternative to the assumption of normality. Also, regressors have been introduced in the variance equation to allow volatility to depend on exogenous or predetermined variables. Shift (dummy) variables are especially popular and have been used to allow for changes in political regimes.

14.5 Exercises

14.5.1 Problems

14.1 The ARCH model is sometimes presented in the following multiplicative form:

$$y_t = \beta_0 + e_t$$
$$e_t = z_t\sqrt{h_t}, \ z_t \sim N(0,1)$$
$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2, \ \ \alpha_0 > 0, \ \ 0 \le \alpha_1 < 1$$

This form describes the distribution of the standardized residuals $e_t/\sqrt{h_t}$ as standard normal z_t. However, the properties of e_t are not altered.

a. Show that the conditional mean $E\left(e_t|I_{t-1}\right) = 0$.

b. Show that the conditional variance $E\left(e_t^2|I_{t-1}\right) = h_t$.

c. Show that $e_t|I_{t-1} \sim N\left(0, h_t\right)$.

14.2 The equations of an **ARCH-in-mean** model are shown below:

$$y_t = \beta_0 + \theta h_t + e_t$$
$$e_t|I_{t-1} \sim N\left(0, h_t\right)$$
$$h_t = \delta + \alpha_1 e_{t-1}^2, \ \ \delta > 0, \ \ 0 \le \alpha_1 < 1$$

Let y_t represent the return from a financial asset and let e_t represent "news" in the financial market. Now use the third equation to substitute out h_t in the first equation, to express the return as

$$y_t = \beta_0 + \theta\left(\delta + \alpha_1 e_{t-1}^2\right) + e_t$$

a. If θ is zero, what is $E_t\left(y_{t+1}\right)$, the conditional mean of y_{t+1}? In other words, what do you expect next period's return to be, given information today?

b. If θ is not zero, what is $E_t\left(y_{t+1}\right)$? What extra information have you used here to forecast the return?

14.3 Consider the following T-ARCH model:

$$h_t = \delta + \alpha_1 e_{t-1}^2 + \gamma d_{t-1} e_{t-1}^2$$
$$d_t = \begin{cases} 1 & e_t < 0 \quad \text{(bad news)} \\ 0 & e_t \ge 0 \quad \text{(good news)} \end{cases}$$

a. If γ is zero, what are the values of h_t when $e_{t-1} = -1$, when $e_{t-1} = 0$, and when $e_{t-1} = 1$?

b. If γ is not zero, what are the values of h_t when $e_{t-1} = -1$, when $e_{t-1} = 0$, and when $e_{t-1} = 1$? What is the key difference between the case $\gamma = 0$ and $\gamma \neq 0$?

14.4 The GARCH(1, 1) model shown below can also be reexpressed as an ARCH(q) model, where q is a large number (in fact, infinity). Derive the ARCH form of a GARCH model using the method of recursive substitution.

$$h_t = \delta + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1}$$

14.5 **a.** Let $I_{t-1} = \{e_{t-1}, e_{t-2}, \ldots\}$. Use the law of iterated iterations to show that $E(e_t|I_{t-1}) = 0$ implies $E(e_t) = 0$.

b. Consider the variance model $h_t = E(e_t^2|I_{t-1}) = \alpha_0 + \alpha_1 e_{t-1}^2$. Use the law of iterated iterations to show that, for $0 < \alpha_1 < 1$, $E(e_t^2) = \alpha_0/(1-\alpha_1)$.

c. Consider the variance model $h_t = E(e_t^2|I_{t-1}) = \delta + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1}$. Use the law of iterated iterations to show that for $0 < \alpha_1 + \beta_1 < 1$, $E(e_t^2) = \delta/(1 - \alpha_1 - \beta_1)$.

14.6 The estimates for the five models in Table 14.1 were obtained using monthly observations on returns to U.S. Nasdaq stock prices from 1985M1 to 2015M12. Use each of the models to estimate the mean and variance of returns for 2016M1.

TABLE 14.1 **Estimates from ARCH Models for U.S. Nasdaq Returns**

Mean function					
Constant	1.4567	1.1789	1.098	1.078	0.931
h_t					0.006
Variance function					
Constant	23.35	19.35	2.076	2.351	2.172
e_{t-1}^2	0.4694	0.3429	0.1329	0.124	0.136
e_{t-2}^2		0.1973			
h_{t-1}			0.8147	0.8006	0.8089
$d_{t-1}e_{t-1}^2$				0.0293	
End-of-sample estimates					
$\hat{e}_{2015M12}$	−3.4388	−3.1610	−3.0803	−3.0605	−3.0760
$\hat{e}_{2015M11}$	−0.3700	−0.0922	−0.0115	0.0083	−0.0296
$\hat{h}_{2015M12}$	23.42	32.64	27.10	27.39	27.27

14.5.2 Computer Exercises

14.7 The data file *share* contains time-series data on the Straits Times share price index of Singapore.

a. Compute the time series of returns using the formula $r_t = 100\ln(y_t/y_{t-1})$, where y_t is the share price index. Generate the correlogram of returns up to at least order 12, since the frequency of the data is monthly. Is there evidence of autocorrelation? If so, it indicates the presence of significant lagged mean effects.

b. Square the returns and generate the correlogram of squared returns. Is there evidence of significant lagged effects? If so, it indicates the presence of significant lagged variance effects.

14.8 The data file *euro* contains 204 monthly observations on the returns to the Euro share price index for the period 1988M1 to 2004M12. A plot of the returns data is shown in Figure 14.10(a), together with its histogram in Figure 14.10(b).

a. What do you notice about the volatility of returns? Identify the periods of big changes and the periods of small changes.

b. Is the distribution of returns normal? Is this the unconditional, or conditional, distribution?

c. Perform a LM test for the presence of first-order ARCH and check that you obtain the following results:

$$\hat{e}_t^2 = 20.509 + \underset{(3.463)}{0.237}\hat{e}_{t-1}^2 \qquad (T-1)R^2 = 11.431$$
$$(t)$$

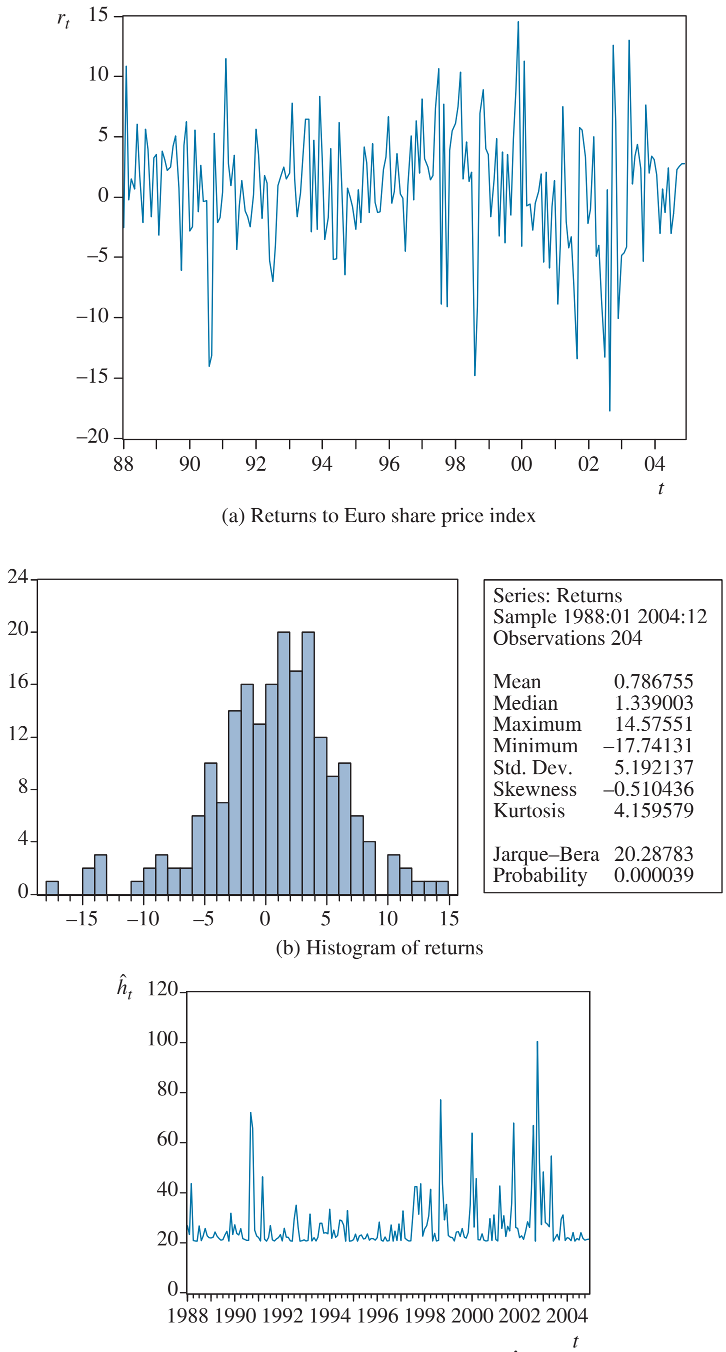
Is there evidence of ARCH effects?

d. Estimate an ARCH(1) model and check that you obtain the following results:

$$\hat{r}_t = 0.879 \qquad \hat{h}_t = 20.604 + 0.230\hat{e}_{t-1}^2$$
$$(t) \quad (2.383) \qquad\qquad (10.968) \;\; (2.198)$$

Interpret the results.

e. A plot of the conditional variance is shown in Figure 14.10(c). Do the periods of high and low conditional variance coincide with the periods of big and small changes in returns?



(a) Returns to Euro share price index

(b) Histogram of returns

(c) Estimated conditional variance $\hat{h}_t$

FIGURE 14.10 **Graphs for Exercise 14.8.**

14.9 Monthly changes in the \$US/\$AUS exchange rate S_t for the period 1985M7 to 2010M6 are stored in the file *exrate5*.

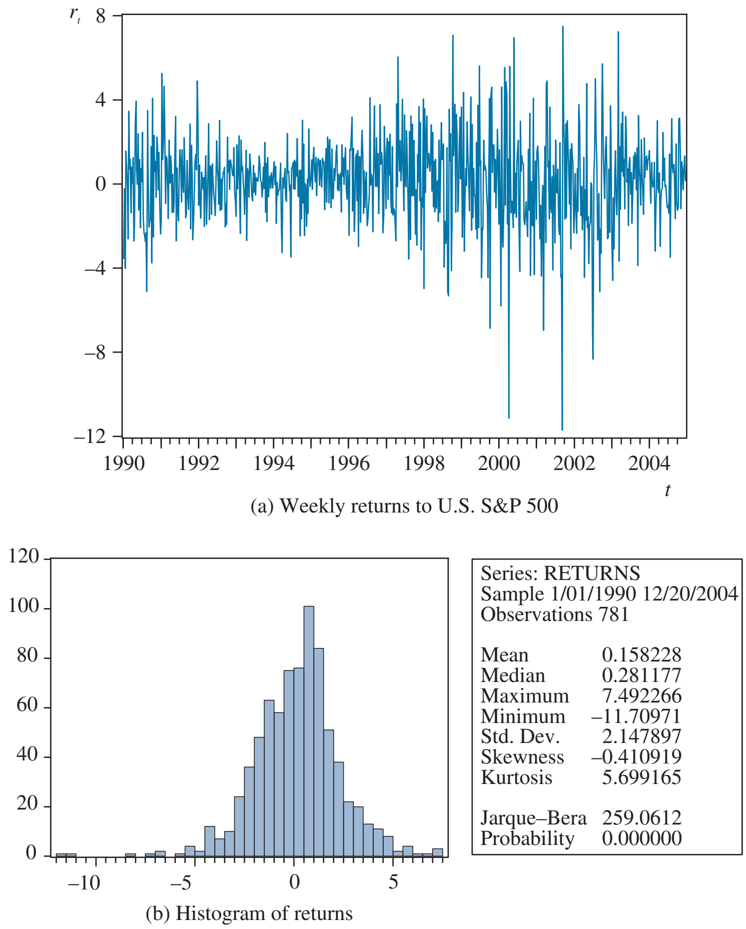
a. Plot the time series of the changes and their histogram. Are there periods of high volatility and periods of low volatility? Does the unconditional distribution of the changes appear to be normally distributed?

b. Estimate the GARCH(1, 1) model $S_t = \beta_0 + e_t, \left(e_t|I_{t-1}\right) \sim N\left(0, h_t\right)$ and $h_t = \delta + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1}$. Comment on the results.

c. Estimate the conditional variance h_t for each observation and create the series $v_t = \hat{e}_t \Big/ \sqrt{\hat{h}_t}$ where $\hat{e}_t$ are the residuals $\hat{e}_t = S_t - \hat{\beta}_0$. Create a histogram for the v_t. Do they appear to be normally distributed?

d. Forecast the conditional mean and variance for 2010M7 and 2010M8.

14.10 Figure 14.11 shows the weekly returns to the U.S. S&P 500 for the sample period January 1990 to December 2004 (data file *sp*).



(a) Weekly returns to U.S. S&P 500

(b) Histogram of returns

FIGURE 14.11 **Graphs for Exercise 14.10.**

a. Estimate an ARCH(1) model and check that you obtain the following results:

$$\hat{r}_t = 0.197 \qquad \hat{h}_t = 3.442 + 0.253\hat{e}_{t-1}^2$$
$$(t) \quad (2.899) \qquad (22.436) \;\; (5.850)$$

What is the value of the conditional variance when the last period's shock was positive, $e_{t-1} = +1$? What about when the last period's shock was negative, $e_{t-1} = -1$?

b. Estimate a T-ARCH model and check that you obtain the following results:

$$\hat{r}_t = 0.147 \qquad \hat{h}_t = 3.437 + (0.123 + 0.268 d_{t-1})\hat{e}_{t-1}^2$$
$$(t) \quad (2.049) \qquad (22.963) \quad (2.330) \ (2.944)$$

c. What is the value of the conditional variance when the last period's shock was positive, $e_{t-1} = +1$? When the last period's shock was negative, $e_{t-1} = -1$?

d. Is the asymmetric T-ARCH model better than the symmetric ARCH model in a financial econometric sense? [*Hint*: Look at the statistical tests for significance.] Is the asymmetric T-ARCH model better than the symmetric ARCH model in a financial economic sense? [*Hint*: Look at the implications of the results.]

14.11 Figure 14.12 shows the daily term premiums between a 180-day bank bill rate and a 90-day bank rate for the period July 1996 to December 1998 (data file *term*). Preliminary unit root tests confirm that the series may be treated as a stationary series, although the value of ρ, the autocorrelation coefficient, is quite high (about 0.9).

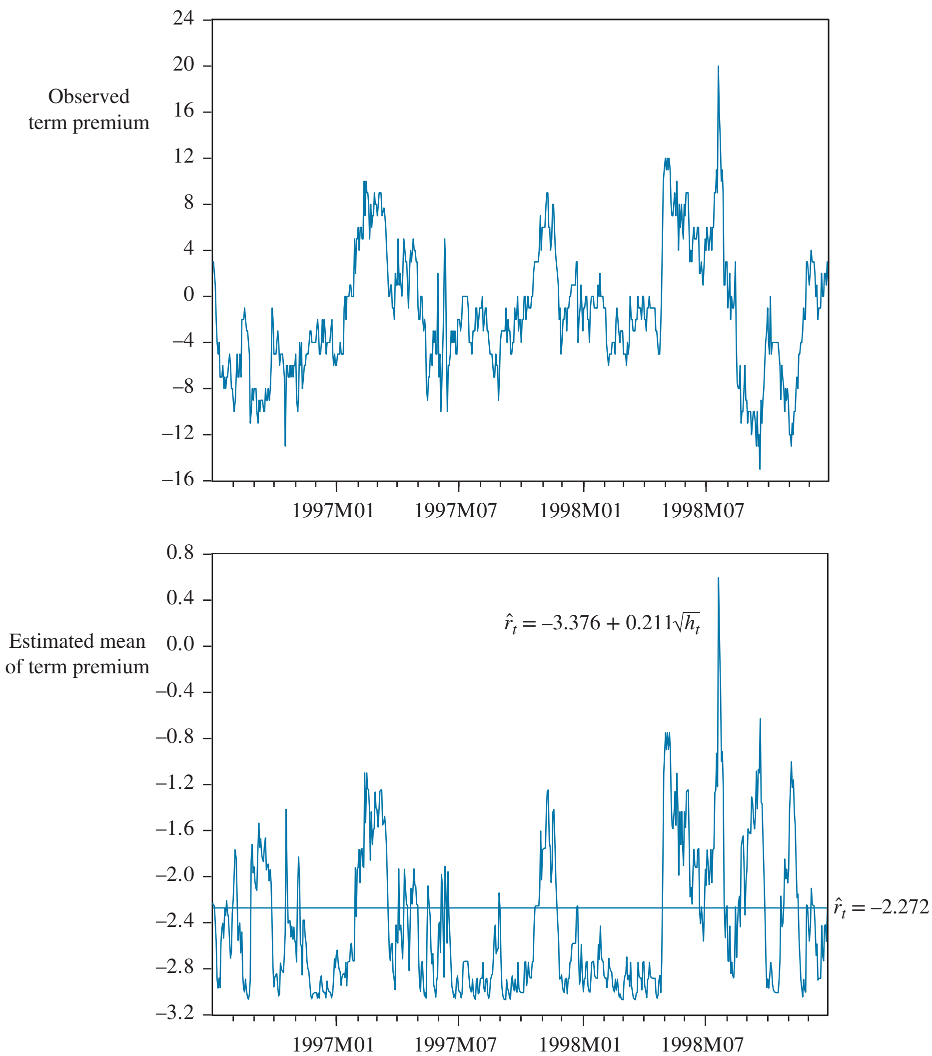


FIGURE 14.12 **Graphs for Exercise 14.11.**

a. Estimate a GARCH model and check that you obtain the following results:

$$\hat{r}_t = -2.272 \qquad \hat{h}_t = \underset{(6.271)}{1.729} + \underset{(6.282)}{0.719}\hat{e}_{t-1}^2 + \underset{(3.993)}{0.224}\hat{h}_{t-1}$$
$$(t)$$

b. Estimate a GARCH-in-mean model and check that you obtain the following results:

$$\hat{r}_t = -3.376 + \underset{(2.807)}{0.211}\sqrt{h_t} \qquad \hat{h}_t = \underset{(5.333)}{1.631} + \underset{(6.327)}{0.730}\hat{e}_{t-1}^2 + \underset{(4.171)}{0.231}\hat{h}_{t-1}$$
$$(t)$$

What is the contribution of volatility to the term premium?

c. Is the GARCH-in-mean model better than the GARCH model in a financial econometric sense? [*Hint*: Look at the statistical tests for significance.] Is the GARCH-in-mean model better than the GARCH model in a financial economic sense? [*Hint*: Look at the implications of the results, in particular the behavior of the term premium.] A plot of the expected term premium estimated for parts (a) and (b) is shown in Figure 14.12.

14.12 The data file *gold* contains 200 daily observations on the returns to shares in a company specializing in gold bullion for the period December 13, 2005, to September 19, 2006.

a. Plot the returns data. What do you notice about the volatility of returns? Identify the periods of big changes and the periods of small changes.

b. Generate the histogram of returns. Is the distribution of returns normal? Is this the unconditional or conditional distribution?

c. Perform a LM test for the presence of first-order ARCH.

d. Estimate a GARCH(1, 1) model. Are the coefficients of the correct sign and magnitude?

e. How would you use the estimated GARCH(1, 1) model to improve your forecasts of returns?

14.13 The seminal paper about ARCH by Robert Engle was concerned with the variance of UK inflation. The data file *uk* contains seasonally adjusted data on the UK consumer price index (*UKCPI*) for the sample period 1957M6 to 2006M6.

a. Compute the monthly rate of inflation (y) for the sample period 1957M7 to 2006M6 using the formula

$$y_t = 100\left[\frac{UKCPI_t - UKCPI_{t-1}}{UKCPI_{t-1}}\right]$$

b. Estimate a T-GARCH-in-mean model and check that you obtain the following results:

$$\hat{y}_t = \underset{(-2.862)}{-0.407} + \underset{(5.243)}{1.983}\sqrt{h_t}$$
$$(t)$$

$$\hat{h}_t = \underset{(4.697)}{0.022} + \left(\underset{(8.952)}{0.211} - \underset{(-8.728)}{0.221}d_{t-1}\right)e_{t-1}^2 + \underset{(27.677)}{0.782}\hat{h}_{t-1}$$
$$(t)$$

c. The negative asymmetric effect (−0.221) suggests that negative shocks (such as falls in prices) reduce volatility in inflation. Is this a sensible result for inflation?

d. What does the positive in-mean effect (1.983) tell you about inflation in the UK and volatility in prices?

14.14 The data file *warner* contains daily returns to holding shares in Time Warner Inc. The sample period is from January 3, 2008 to December 31, 2008 (260 observations), and a graph of the returns appears in Figure 14.13.

a. Estimate a GARCH(1, 1) model and an ARCH(5) model. Which model would you prefer, and why?

b. What is the expected return next period? The expected volatility next period?

c. Use your preferred model to forecast next period's return and next period's volatility.

d. Do good news and bad news have the same effect on return? On volatility?

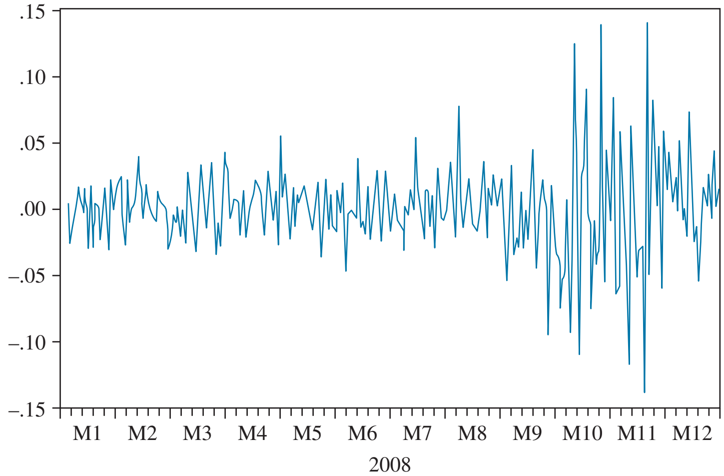


FIGURE 14.13 **Returns to shares in Time Warner.**

14.15 Consider the quarterly rates of growth contained in data file *gfc* used in Exercise 13.14. A researcher in the Euro Area (this is the group of countries in Europe where the Euro currency is the legal tender) is interested in testing the proposition that growth in the Euro region is affected by its own history, growth in the United States, and shocks to economic activity.

- **a.** Specify and estimate an econometric model for the Euro Area based only on its own history and where the expected effect of shocks on the expected quarterly rate of growth is zero.
- **b.** Specify and estimate an econometric model for the Euro Area based only on its own history and where shocks may come from distributions with zero mean, but time-varying variances.
- **c.** Specify and estimate an econometric model for the Euro Area based on its own history, the history of growth in the United States, and where the expected effect of shocks on the expected quarterly rate of growth is zero.
- **d.** Specify and estimate an econometric model for the Euro Area based on its own history and allow shocks in the Euro Area to have an effect of zero on the quarterly rate of growth.
- **e.** Specify and estimate an econometric model for the Euro Area based on its own history, the history of growth in the United States, and where shocks in the Euro Area and in the United States have an effect on the expected quarterly rate of growth.

14.16 The data file *shanghai* contains data on the daily returns to the Shanghai Stock Exchange Composite Index from July 7, 1995 to May 5, 2015.

- **a.** Plot the time series of returns and their histogram. For what observations is volatility the greatest? Describe the shape of the distribution of returns. Does the Jarque–Bera test reject the null hypothesis that returns are normally distributed?
- **b.** Estimate the GARCH model

$$y_t = \beta_0 + e_t \qquad (e_t|I_{t-1}) \sim N(0, h_t) \qquad h_t = \delta + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1}$$

Comment on the results. Plot the within-sample variance estimate $\hat{h}_t$. Have the variance estimates captured the periods of high volatility noted in part (a)?

- **c.** For the model estimated in part (b), compute the series $z_t = \hat{e}_t / \sqrt{\hat{h}_t}$. Does a histogram for the z_t suggest the assumption $z_t \sim N(0, 1)$ is valid? Does the Jarque–Bera test support this assumption?
- **d.** When the normality assumption is violated, the ordinary standard errors are not valid. However, valid robust standard errors can be used.[1] Re-estimate the model in part (b) using the Bollerslev–Wooldridge robust standard errors. Does using these standard errors change any conclusions are about the precision of estimation?

[1]See Bollerslev, T. and Wooldridge, J. (1992), "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances," *Econometric Reviews*, 11, 143–172.

e. Estimate the EGARCH model

$$y_t = \beta_0 + e_t \quad (e_t|I_{t-1}) \sim N(0, h_t) \quad \ln(h_t) = \delta + \beta_1 \ln(h_{t-1}) + \alpha\left|\frac{e_{t-1}}{\sqrt{h_{t-1}}}\right| + \gamma\left(\frac{e_{t-1}}{\sqrt{h_{t-1}}}\right)$$

Comment on the results. Plot the within-sample variance estimate $\hat{h}_t$. Have the variance estimates captured the periods of high volatility noted in part (a)?

f. For the model estimated in part (e), compute the series $z_t = \hat{e}_t / \sqrt{\hat{h}_t}$. Does a histogram for the z_t suggest the assumption $z_t \sim N(0, 1)$ is valid? Does the Jarque–Bera test support this assumption?

g. Reestimate the model in part (e) using the Bollerslev–Wooldridge standard errors. Does using these standard errors change any conclusions about the precision of estimation?

h. Find and compare estimates of $E(y_{T+1}|I_T)$ and $\text{var}(y_{T+1}|I_T)$ from the models in parts (b) and (e).

i. Using the model from part (b), and Bollerslev–Wooldridge variance and covariance estimates, find 95% interval estimates for $E(y_{T+1}|I_T)$ and $\text{var}(y_{T+1}|I_T)$.