CHAPTER 13

Vector Error Correction and Vector Autoregressive Models

LEARNING OBJECTIVES

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

1. Explain why economic variables are dynamically interdependent.
2. Explain the VEC model.
3. Explain the importance of error correction.
4. Explain the VAR model.
5. Explain the relationship between a VEC model and a VAR model.
6. Explain how to estimate the VEC and VAR models for the bivariate case.
7. Explain how to generate impulse response functions and variance decompositions for the simple case when the variables are not contemporaneously interdependent and the shocks are not correlated.

KEYWORDS

dynamic relationships
error correction
forecast error variance decomposition
identification problem
impulse response functions
VAR model
VEC model

In Chapter 12, we studied the time-series properties of data and cointegrating relationships between pairs of nonstationary series. In those examples, we assumed that one of the variables was the dependent variable $\left(\text{let us call it } y_t\right)$ and that the other was the independent variable $\left(\text{say } x_t\right)$, and we treated the relationship between y_t and x_t like a regression model. However, a priori, unless we have good reasons not to, we could just as easily have assumed that y_t is the independent variable and x_t is the dependent variable. Put simply, we are working with two variables $\left\{y_t, x_t\right\}$ and the two possible regression models relating them are

$$y_t = \beta_{10} + \beta_{11}x_t + e_t^y, \quad e_t^y \sim N\left(0, \sigma_y^2\right) \tag{13.1a}$$

$$x_t = \beta_{20} + \beta_{21}y_t + e_t^x, \quad e_t^x \sim N\left(0, \sigma_x^2\right) \tag{13.1b}$$

In this bivariate (two series) system, there can be only one unique relationship between x_t and y_t, and so it must be the case that $\beta_{21} = 1/\beta_{11}$ and $\beta_{20} = -\beta_{10}/\beta_{11}$. A bit of terminology: for (13.1a), we say that we have normalized on y (meaning that the coefficient in front of y is set to 1), whereas for (13.1b), we say that we have normalized on x (meaning that the coefficient in front of x is set to 1).

Is it better to write the relationship as (13.1a) or (13.1b), or is it better to recognize that in many relationships, variables like y and x are simultaneously determined? The aim of this chapter is to explore the causal relationship between pairs of time-series variables. In doing so, we shall be extending our study of time-series data to take account of their dynamic properties and interactions. In particular, we will discuss the **vector error correction (VEC)** and **vector autoregressive (VAR)** models. We will learn how to estimate a **VEC model** when there is cointegration between I(1) variables, and how to estimate a **VAR model** when there is no cointegration. Note that this is an extension of the single-equation models examined in Chapter 12.

Some important terminology emerges here. Univariate analysis examines a single data series. Bivariate analysis examines a pair of series. The term **vector** indicates that we are considering a number of series: two, three, or more. The term "vector" is a generalization of the univariate and bivariate cases.

13.1 VEC and VAR Models

Let us begin with two time-series variables y_t and x_t and generalize the discussion about **dynamic relationships** in Chapter 9 to yield a system of equations:

$$
\begin{aligned}
y_t &= \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} x_{t-1} + v_t^y \\
x_t &= \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} x_{t-1} + v_t^x
\end{aligned}
\tag{13.2}
$$

The equation (13.2) describes a system in which each variable is a function of its own lag and the lag of the other variable in the system. In this case, the system contains two variables y and x. In the first equation y_t is a function of its own lag y_{t-1} and the lag of the other variable in the system x_{t-1}. In the second equation x_t is a function of its own lag x_{t-1} and the lag of the other variable in the system y_{t-1}. Together the equations constitute a system known as a VAR. In this example, since the maximum lag is of order 1, we have a VAR(1).

If y and x are stationary I(0) variables, the above system can be estimated using least squares applied to each equation. If, however, y and x are nonstationary I(1) and not cointegrated, then as discussed in Chapter 12, we work with the first differences. In this case, the VAR model is

$$
\begin{aligned}
\Delta y_t &= \beta_{11} \Delta y_{t-1} + \beta_{12} \Delta x_{t-1} + v_t^{\Delta y} \\
\Delta x_t &= \beta_{21} \Delta y_{t-1} + \beta_{22} \Delta x_{t-1} + v_t^{\Delta x}
\end{aligned}
\tag{13.3}
$$

All variables are now I(0), and the system can again be estimated by least squares. To recap, the VAR model is a general framework to describe the dynamic interrelationship between stationary variables. Thus, if y and x are stationary I(0) variables, the system in (13.2) is used. On the other hand, if y and x are I(1) variables but are not cointegrated, we examine the interrelation between them using a VAR framework in first differences (13.3).

If y and x are I(1) and cointegrated, then we need to modify the system of equations to allow for the cointegrating relationship between the I(1) variables. We do this for two reasons. First, as economists, we like to retain and use valuable information about the cointegrating relationship, and second, as econometricians, we like to ensure that we use the best technique that takes into account the properties of the time-series data. Recall the chapter on simultaneous equations—the cointegrating equation is one way of introducing simultaneous interactions without requiring the data to be stationary. Introducing the cointegrating relationship leads to a model known as the VEC model. We turn now to this model.

Consider two nonstationary variables y_t and x_t that are integrated of order 1: $y_t \sim \mathrm{I}(1)$ and $x_t \sim \mathrm{I}(1)$ and which we have shown to be cointegrated, so that

$$y_t = \beta_0 + \beta_1 x_t + e_t \tag{13.4}$$

and $\hat{e}_t \sim \mathrm{I}(0)$ where $\hat{e}_t$ are the estimated residuals. Note that we could have chosen to normalize on x. Whether we normalize on y or x is often determined from economic theory; the critical point is that there can be at most one fundamental relationship between the two variables.

The VEC model is a special form of the VAR for I(1) variables that are cointegrated. The VEC model is

$$\begin{aligned} \Delta y_t &= \alpha_{10} + \alpha_{11}\left(y_{t-1} - \beta_0 - \beta_1 x_{t-1}\right) + v_t^y \\ \Delta x_t &= \alpha_{20} + \alpha_{21}\left(y_{t-1} - \beta_0 - \beta_1 x_{t-1}\right) + v_t^x \end{aligned} \tag{13.5a}$$

which we can expand as

$$\begin{aligned} y_t &= \alpha_{10} + \left(\alpha_{11} + 1\right)y_{t-1} - \alpha_{11}\beta_0 - \alpha_{11}\beta_1 x_{t-1} + v_t^y \\ x_t &= \alpha_{20} + \alpha_{21}y_{t-1} - \alpha_{21}\beta_0 - \left(\alpha_{21}\beta_1 - 1\right)x_{t-1} + v_t^x \end{aligned} \tag{13.5b}$$

Comparing (13.5b) with (13.2) shows the VEC as a VAR where the I(1) variable y_t is related to other lagged variables $\left(y_{t-1} \text{ and } x_{t-1}\right)$ and where the I(1) variable x_t is also related to the other lagged variables $\left(y_{t-1} \text{ and } x_{t-1}\right)$. Note, however, that the two equations contain the common cointegrating relationship.

The coefficients α_{11}, α_{21} are known as **error correction** coefficients, so named because they show how much Δy_t and Δx_t respond to the cointegrating error $y_{t-1} - \beta_0 - \beta_1 x_{t-1} = e_{t-1}$. The idea that the error leads to a correction comes about because of the conditions put on α_{11}, α_{21} to ensure stability, namely $\left(-1 < \alpha_{11} \le 0\right)$ and $\left(0 \le \alpha_{21} < 1\right)$. To appreciate this idea, consider a positive error $e_{t-1} > 0$ that occurred because $y_{t-1} > \left(\beta_0 + \beta_1 x_{t-1}\right)$. A negative error correction coefficient in the first equation $\left(\alpha_{11}\right)$ ensures that Δy falls, while the positive error correction coefficient in the second equation $\left(\alpha_{21}\right)$ ensures that Δx rises, thereby correcting the error. Having the error correction coefficients less than 1 in absolute value ensures that the system is not explosive. Note that the VEC is a generalization of the error-correction (single-equation) model discussed in Chapter 12. In the VEC (system) model, both y_t and x_t "error-correct."

The error correction model has become an extremely popular model because its interpretation is intuitively appealing. Think about two nonstationary variables, say consumption (let us call it y_t) and income (let us call it x_t), that we expect to be related (cointegrated). Now think about a change in your income Δx_t, say a pay raise! Consumption will most likely increase, but it may take you a while to change your consumption pattern in response to a change in your pay. The VEC model allows us to examine how much consumption will change in response to a change in the explanatory variable $\left(\text{the cointegration part, } y_t = \beta_0 + \beta_1 x_t + e_t\right)$, as well as the speed of the change (the error correction part, $\Delta y_t = \alpha_{10} + \alpha_{11}\left(e_{t-1}\right) + v_t^y$ where e_{t-1} is the cointegrating error).

There is one final point to discuss—the role of the intercept terms. Thus far, we have introduced an intercept term in the cointegrating equation (β_0) as well as in the VEC $\left(\alpha_{10} \text{ and } \alpha_{20}\right)$. However, doing so can create a problem. To see why, we collect all the intercept terms and rewrite (13.5b) as

$$\begin{aligned} y_t &= \left(\alpha_{10} - \alpha_{11}\beta_0\right) + \left(\alpha_{11} + 1\right)y_{t-1} - \alpha_{11}\beta_1 x_{t-1} + v_t^y \\ x_t &= \left(\alpha_{20} - \alpha_{21}\beta_0\right) + \alpha_{21}y_{t-1} - \left(\alpha_{21}\beta_1 - 1\right)x_{t-1} + v_t^x \end{aligned} \tag{13.5c}$$

If we estimate each equation by least squares, we obtain estimates of composite terms $\left(\alpha_{10} - \alpha_{11}\beta_0\right)$ and $\left(\alpha_{20} - \alpha_{21}\beta_0\right)$, and we are not able to disentangle the separate effects of β_0, α_{10}, and α_{20}. In the next section, we discuss a simple two-step least squares procedure that gets around this problem. However, the lesson here is to check whether, and where, an intercept term is needed.

13.2 Estimating a Vector Error Correction Model

There are many econometric methods to estimate the error correction model. Nonlinear (system) least squares is one method, but the most straightforward method is to use a two-step least squares procedure. First, use OLS to estimate the cointegrating relationship $y_t = \beta_0 + \beta_1 x_t + e_t$ and generate the lagged residuals $\hat{e}_{t-1} = y_{t-1} - b_0 - b_1 x_{t-1}$.

Second, use OLS to estimate the equations:

$$\Delta y_t = \alpha_{10} + \alpha_{11}\hat{e}_{t-1} + v_t^y \tag{13.6a}$$

$$\Delta x_t = \alpha_{20} + \alpha_{21}\hat{e}_{t-1} + v_t^x \tag{13.6b}$$

Note that all the variables in (13.6) (Δy, Δx, and $\hat{e}$) are stationary (recall that for y and x to be cointegrated, the residuals $\hat{e}$ must be stationary). Hence, the standard regression analysis studied in earlier chapters may be used to test the significance of the parameters. The usual residual diagnostic tests may be applied.

We need to be careful here about how we combine stationary and nonstationary variables in a regression model. Cointegration is about the relationship between I(1) variables. The cointegrating equation does not contain I(0) variables. The corresponding VEC model, however, relates the change in an I(1) variable (the I(0) variables Δy and Δx) to other I(0) variables, namely, the cointegration residuals $\hat{e}_{t-1}$; if required, other stationary variables may be added. In other words, we should not mix stationary and nonstationary variables: an I(0) dependent variable on the left-hand side of a regression equation should be "explained" by other I(0) variables on the right-hand side and an I(1) dependent variable on the left-hand side of a regression equation should be explained by other I(1) variables on the right-hand side.

EXAMPLE 13.1 | VEC Model for GDP

In Figure 13.1 the quarterly real gross domestic product (GDP) of a small economy (Australia) and a large economy (United States) for the sample period 1970Q1 to 2000Q4 are displayed. Note that the series have been scaled so that

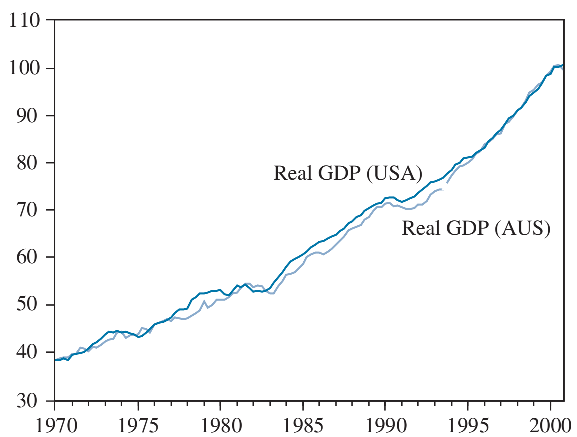

FIGURE 13.1 **Real gross domestic product (GDP = 100 in 2000).**

both economies show a real GDP value of 100 in 2000. They appear in the data file *gdp*. It appears from the figure that both series are nonstationary and possibly cointegrated.

Formal unit root tests of the series confirm that they are indeed nonstationary. To check for cointegration we obtain the fitted equation in (13.7) (the intercept term is omitted because it has no economic meaning):

$$\hat{A}_t = 0.985U_t, \tag{13.7}$$

where A denotes real GDP for Australia and U denotes real GDP for the United States. Note that we have normalized on A because it makes more sense to think of a small economy responding to a large economy. The residuals derived from the cointegrating relationship $\hat{e}_t = A_t - 0.985U_t$ are shown in Figure 13.2. Their first-order autocorrelation is 0.870, and a visual inspection of the time series suggests that the residuals may be stationary.

A formal unit root test is performed, and the estimated unit root test equation is

$$\begin{array}{ll} \widehat{\Delta e_t} = & -0.128\hat{e}_{t-1} \\ (tau) & (-2.889) \end{array} \tag{13.8}$$

Since the cointegrating relationship does not contain an intercept term [see Chapter 12, (12.29a)], the 5% critical value is −2.76. The unit root t-value of −2.889 is less than −2.76. We reject the null of no cointegration and we conclude that the two real GDP series are cointegrated. This result implies that economic activity in the small economy (Australia, A_t) is linked to economic activity in the large economy (United States, U_t). If U_t were to increase by one unit, A_t would increase by 0.985. But the Australian economy may not respond fully by this amount within the quarter. To ascertain how much it will respond within a quarter, we estimate the error correction model by least squares. The estimated VEC model for $\{A_t, U_t\}$ is

$$\begin{aligned}
\widehat{\Delta A_t} &= 0.492 - 0.099\hat{e}_{t-1} \\
(t) & \qquad\qquad (-2.077) \\
\widehat{\Delta U_t} &= 0.510 + 0.030\hat{e}_{t-1} \\
(t) & \qquad\qquad (0.789)
\end{aligned} \tag{13.9}$$

The results show that both error correction coefficients are of the appropriate sign. The negative error correction coefficient in the first equation (−0.099) indicates that ΔA falls (i.e., A_t falls or ΔA_t is negative), while the positive error correction coefficient in the second equation (0.030) indicates that ΔU rises (i.e., U_t rises or ΔU_t is positive), when there is a positive cointegrating error $\left(\hat{e}_{t-1} > 0 \text{ or } A_{t-1} > 0.985U_{t-1}\right)$. This behavior (negative change in A and positive change in U) "corrects" the cointegrating error. The error correction coefficient (−0.099) is significant at the 5% level; it indicates that the quarterly adjustment of A_t will be about 10% of the deviation of A_{t-1} from its cointegrating value $0.985U_{t-1}$. This is a slow rate of adjustment. However, the error correction coefficient in the second equation (0.030) is insignificant; it suggests that ΔU does not react to the cointegrating error. This outcome is consistent with the view that the small economy is likely to react to economic conditions in the large economy, but not vice versa.

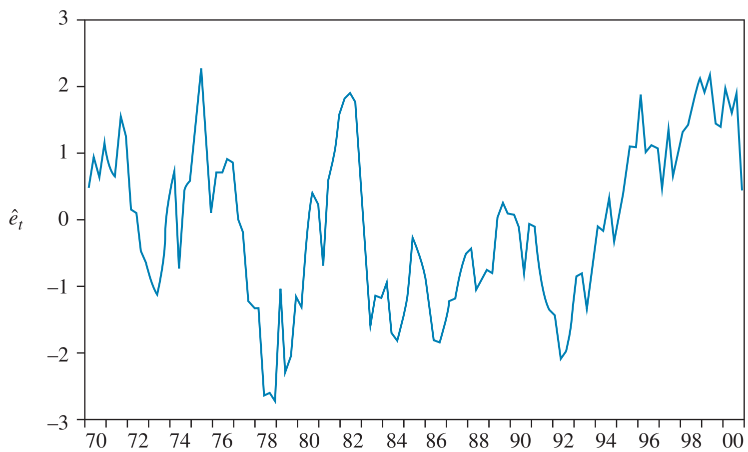

FIGURE 13.2 **Residuals derived from the cointegrating relationship.**

13.3 Estimating a VAR Model

The VEC is a multivariate dynamic model that incorporates a cointegrating equation. It is relevant when, for the bivariate case, we have two variables, say y and x, that are both I(1), but are cointegrated. Now we ask: What should we do if we are interested in the interdependencies between y and x, but they are not cointegrated? In this case, we estimate a VAR model as shown in (13.3).

EXAMPLE 13.2 | VAR Model for Consumption and Income

Consider Figure 13.3 that shows the log of real personal disposable income (RPDI) (denoted as Y) and the log of real personal consumption expenditure (RPCE) (denoted as C) for the U.S. economy over the period 1986Q1 to 2015Q2. Both series appear to be nonstationary, but are they cointegrated? The quarterly data are stored in the data file *fred5*.

The Dickey–Fuller test values for unit roots for C were −0.88 when an intercept only was included and −1.63 when both an intercept and trend term were included. In both cases, there were three augmentation terms. The corresponding values for Y were −1.65 and −0.43. In these cases, one augmentation term was sufficient. The 10% critical values from Table 12.2 are −2.57 without a trend and −3.13 with a trend. Since the test values are greater than the critical values, we cannot conclude that the series are stationary. Using a 10% significance level, unit root tests on the first differences of the series lead to a conclusion that the first differences are stationary, and hence the series are I(1). Testing for cointegration yields the following results:

$$\begin{aligned} \hat{e}_t &= C_t + 0.543 - 1.049Y_t \\ \widehat{\Delta\hat{e}_t} &= -0.203\hat{e}_{t-1} - 0.290\Delta\hat{e}_{t-1} \\ (\tau) &\quad (-3.046) \end{aligned} \tag{13.10}$$

An intercept term has been included to capture the component of (log) consumption that is independent of disposable income. From Table 12.4, the 10% critical value of the test for stationarity in the cointegrating residuals is −3.07. Since the *tau* (unit root *t*-value) of −3.046 is greater than −3.07, it indicates that the errors are not stationary and hence that the relationship between C (i.e., log(RPCE)) and Y (i.e., log(RPDI)) is spurious. That is, we have no cointegration. Thus, we would not apply a VEC model to examine the dynamic relationship between aggregate consumption C and income Y. Instead, we would estimate a VAR model for the set of I(0) variables $\{\Delta C_t, \Delta Y_t\}$.

For illustrative purposes, the order of lag in this example has been restricted to one. In general, one should use significance of the coefficient estimates and serial correlation in the errors to choose a suitable number of lags which may be greater than one. The results are

$$\begin{aligned} \widehat{\Delta C_t} &= 0.00367 + 0.348\Delta C_{t-1} + 0.131\Delta Y_{t-1} \\ (t) &\quad (4.87) \qquad (4.02) \qquad\qquad (2.52) \end{aligned} \tag{13.11a}$$

$$\begin{aligned} \widehat{\Delta Y_t} &= 0.00438 + 0.590\Delta C_{t-1} - 0.291\Delta Y_{t-1} \\ (t) &\quad (3.38) \qquad (3.96) \qquad\quad (-3.25) \end{aligned} \tag{13.11b}$$

The first equation (13.11a) shows that the quarterly growth in consumption (ΔC_t) is significantly related to its own past value (ΔC_{t-1}) and also significantly related to the quarterly growth in last period's income (ΔY_{t-1}). The second equation (13.11b) shows that ΔY_t is significantly negatively related to its own past value but significantly positively related to last period's change in consumption. The constant terms capture the fixed component in the change in log consumption and the change in log income.

Having estimated these models, can we infer anything else? If the system is subjected to an income shock, what is the effect of the shock on the dynamic path of the quarterly growth in consumption and income? Will they rise and by how much? If the system is also subjected to a consumption shock, what is the contribution of an income versus a consumption shock on the variation of income? We turn now to some analysis suited to addressing these questions.

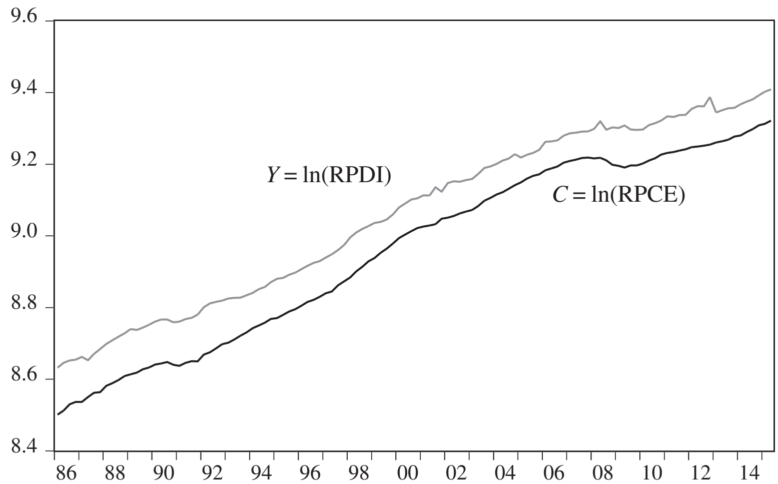

FIGURE 13.3 **The logarithms of real personal disposable income (RPDI) and real personal consumption expenditure (RPCE).**

13.4 Impulse Responses and Variance Decompositions

Impulse response functions and variance decompositions are techniques that are used by macroeconometricians to analyze problems such as the effect of an oil price shock on inflation and GDP growth, and the effect of a change in monetary policy on the economy.

13.4.1 Impulse Response Functions

Impulse response functions show the effects of shocks on the adjustment path of the variables. To help us understand this, we shall first consider a univariate series.

The Univariate Case Consider a univariate series $y_t = \rho y_{t-1} + v_t$ and subject it to a shock of size v in period one. Assume an arbitrary starting value of y at time zero: $y_0 = 0$. (Since we are interested in the dynamic path, the starting point is irrelevant.) At time $t = 1$, following the shock, the value of y will be: $y_1 = \rho y_0 + v_1 = v$. Assume that there are no subsequent shocks in later time periods $\left[v_2 = v_3 = \cdots = 0\right]$, at time $t = 2$, $y_2 = \rho y_1 = \rho v$. At time $t = 3$, $y_3 = \rho y_2 = \rho\left(\rho y_1\right) = \rho^2 v$, and so on. Thus the time-path of y following the shock is $\left\{v, \rho v, \rho^2 v, \ldots\right\}$. The values of the coefficients $\left\{1, \rho, \rho^2, \ldots\right\}$ are known as multipliers, and the time-path of y following the shock is known as the impulse response function.

To illustrate, assume that $\rho = 0.9$ and let the shock be unity: $v = 1$. According to the analysis, y will be $\{1, 0.9, 0.81, \ldots\}$, approaching zero over time. This impulse response function is plotted in Figure 13.4. It shows us what happens to y after a shock. In this case, y initially rises by the full amount of the shock and then it gradually returns to the value before the shock.

The Bivariate Case Now, let us consider an impulse response function analysis with two time series based on a bivariate VAR system of stationary variables:

$$
\begin{aligned}
y_t &= \delta_{10} + \delta_{11} y_{t-1} + \delta_{12} x_{t-1} + v_t^y \\
x_t &= \delta_{20} + \delta_{21} y_{t-1} + \delta_{22} x_{t-1} + v_t^x
\end{aligned}
\tag{13.12}
$$

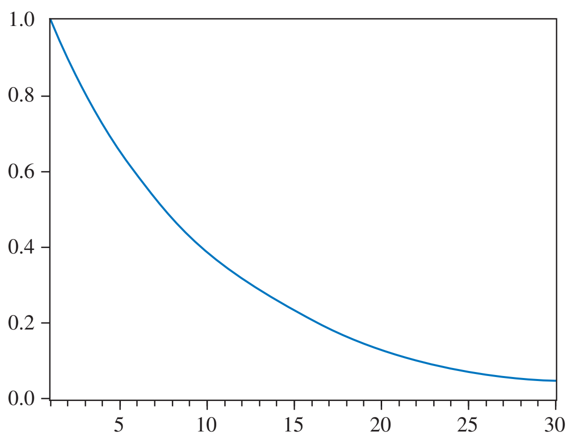

FIGURE 13.4 Impulse responses for an AR(1) model $y_t = 0.9\, y_{t-1} + v_t$ following a unit shock.

In this case, there are two possible shocks to the system—one to y and the other to x. Thus we are interested in four impulse response functions—the effect of a shock to y on the time-paths of y and x and the effect of a shock to x on the time-paths of y and x.

The actual mechanics of generating impulse responses in a system is complicated by the facts that (i) one has to allow for interdependent dynamics (the multivariate analog of generating the multipliers) and (ii) one has to identify the correct shock from unobservable data. Taken together, these two complications lead to what is known as the **identification problem**. In this chapter, we consider a special case where there is no identification problem.[1] This special case occurs when the system that is described in (13.12) is a true representation of the dynamic system—namely, y is related only to lags of y and x, and x is related only to lags of y and x. In other words, y and x are related dynamically, but not contemporaneously. The current value x_t does not appear in the equation for y_t and the current value y_t does not appear in the equation for x_t. Also, we need to assume that the errors v_t^x and v_t^y are contemporaneously uncorrelated.

Consider the case when there is a one standard deviation shock (alternatively called an **innovation**) to y so that at time $t = 1$, $v_1^y = \sigma_y$, and v_t^y is zero thereafter. Assume $v_t^x = 0$ for all t. It is traditional to consider a standard deviation shock (innovation) rather than a unit shock to eliminate units of measurement. Assume $y_0 = x_0 = 0$. Also, since we are focusing on how a shock *changes* the paths of y and x, we can ignore the intercepts. Then

1. When $t = 1$, the effect of a shock of size σ_y on y is $y_1 = v_1^y = \sigma_y$, and the effect on x is $x_1 = v_1^x = 0$.
2. When $t = 2$, the effect of the shock on y is

$$y_2 = \delta_{11}y_1 + \delta_{12}x_1 = \delta_{11}\sigma_y + \delta_{12}0 = \delta_{11}\sigma_y$$

and the effect on x is

$$x_2 = \delta_{21}y_1 + \delta_{22}x_1 = \delta_{21}\sigma_y + \delta_{22}0 = \delta_{21}\sigma_y$$

3. When $t = 3$, the effect of the shock on y is

$$y_3 = \delta_{11}y_2 + \delta_{12}x_2 = \delta_{11}\delta_{11}\sigma_y + \delta_{12}\delta_{21}\sigma_y$$

and the effect on x is

$$x_3 = \delta_{21}y_2 + \delta_{22}x_2 = \delta_{21}\delta_{11}\sigma_y + \delta_{22}\delta_{21}\sigma_y.$$

By repeating the substitutions for $t = 4, 5, \ldots$, we obtain further expressions. The impulse response of the shock (or innovation) to y on y is $\sigma_y\left[1, \delta_{11}, \left(\delta_{11}\delta_{11} + \delta_{12}\delta_{21}\right), \ldots\right]$ and the impulse response of a shock to y on x is $\sigma_y\left[0, \delta_{21}, \left(\delta_{21}\delta_{11} + \delta_{22}\delta_{21}\right), \ldots\right]$.

Now consider what happens when there is a one standard deviation shock to x so that at time $t = 1$, $v_1^x = \sigma_x$, and v_t^x is zero thereafter. Assume $v_t^y = 0$ for all t. In the first period after the shock, the effect of a shock of size σ_x on y is $y_1 = v_1^y = 0$, and the effect of the shock on x is $x_1 = v_1^x = \sigma_x$. Two periods after the shock, when $t = 2$, the effect on y is

$$y_2 = \delta_{11}y_1 + \delta_{12}x_1 = \delta_{11}0 + \delta_{12}\sigma_x = \delta_{12}\sigma_x$$

and the effect on x is

$$x_2 = \delta_{21}y_1 + \delta_{22}x_1 = \delta_{21}0 + \delta_{22}\sigma_x = \delta_{22}\sigma_x$$

Again, by repeated substitutions, we obtain the impulse response of a shock to x on y as $\sigma_x\left[0, \delta_{12}, \left(\delta_{11}\delta_{12} + \delta_{12}\delta_{22}\right), \ldots\right]$, and the impulse response of a shock to x on x as $\sigma_x\left[1, \delta_{22}, \left(\delta_{21}\delta_{12} + \delta_{22}\delta_{22}\right), \ldots\right]$. Figure 13.5 shows the four impulse response functions for numerical values: $\sigma_y = 1$, $\sigma_x = 2$, $\delta_{11} = 0.7$, $\delta_{12} = 0.2$, $\delta_{21} = 0.3$ and $\delta_{22} = 0.6$.

[1] Appendix 13A introduces the general problem.

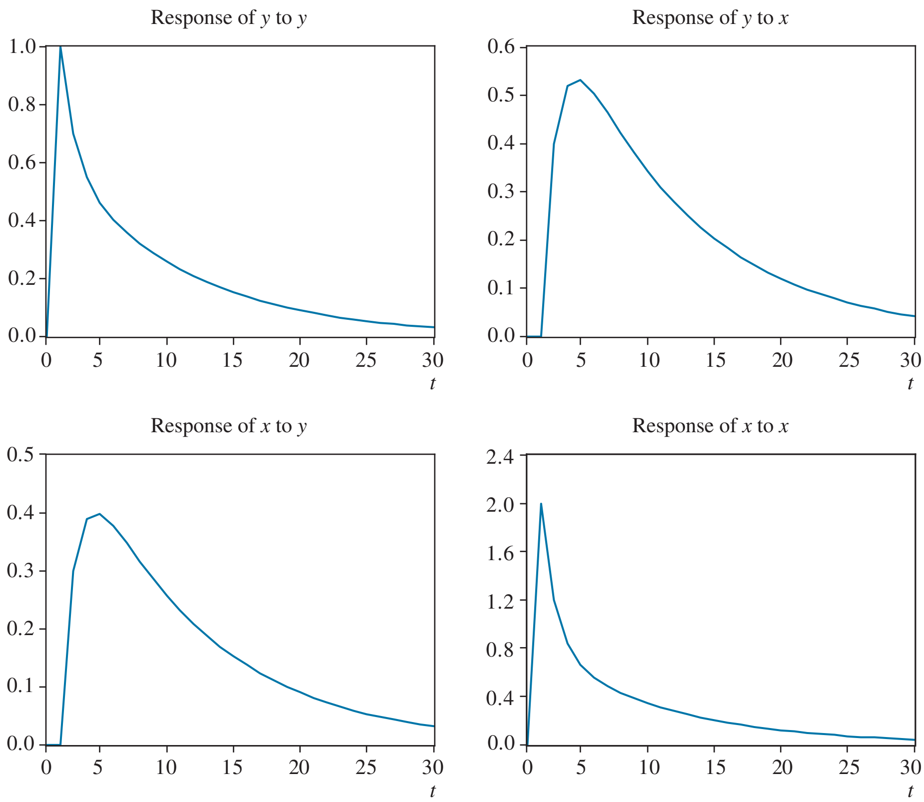

FIGURE 13.5 **Impulse responses to standard deviation shock.**

The advantage of examining impulse response functions (and not just VAR coefficients) is that they show the size of the impact of the shock plus the rate at which the shock dissipates, allowing for interdependencies.

13.4.2 Forecast Error Variance Decompositions

Another way to disentangle the effects of various shocks is to consider the contribution of each type of shock to the forecast error variance.

Univariate Analysis Consider again the univariate series, $y_t = \rho y_{t-1} + v_t$. The best one-step-ahead forecast (alternatively the forecast one period ahead) is

$$y_{t+1}^F = E_t[\rho y_t + v_{t+1}]$$

where E_t is the expected value conditional on information at time t (i.e., we are interested in the mean value of y_{t+1} using what is known at time t). At time t the conditional expectation $E_t[\rho y_t] = \rho y_t$ is known, but the error v_{t+1} is unknown, and so its conditional expectation is zero. Thus the best forecast of y_{t+1} is ρy_t, and the forecast error is

$$y_{t+1} - E_t[y_{t+1}] = y_{t+1} - \rho y_t = v_{t+1}$$

The variance of the one-step forecast error is $\mathrm{var}(v_{t+1}) = \sigma^2$. Suppose we wish to forecast two steps ahead; using the same logic, the two-step forecast becomes

$$y_{t+2}^F = E_t[\rho y_{t+1} + v_{t+2}] = E_t[\rho(\rho y_t + v_{t+1}) + v_{t+2}] = \rho^2 y_t$$

and the two-step forecast error becomes

$$y_{t+2} - E_t[y_{t+2}] = y_{t+2} - \rho^2 y_t = \rho v_{t+1} + v_{t+2}$$

In this case, the variance of the forecast error is $\text{var}(\rho v_{t+1} + v_{t+2}) = \sigma^2(\rho^2 + 1)$, showing that the variance of forecast error increases as we increase the forecast horizon. There is only one shock that leads to a forecast error. Hence the forecast error variance is 100% due to its own shock. The exercise of attributing the source of the variation in the forecast error is known as variance decomposition.

Bivariate Analysis We can perform a **forecast error variance decomposition** for our special bivariate example where there is no identification problem. Ignoring the intercepts (since they are constants), the one-step ahead forecasts are

$$y_{t+1}^F = E_t\left[\delta_{11}y_t + \delta_{12}x_t + v_{t+1}^y\right] = \delta_{11}y_t + \delta_{12}x_t$$

$$x_{t+1}^F = E_t\left[\delta_{21}y_t + \delta_{22}x_t + v_{t+1}^x\right] = \delta_{21}y_t + \delta_{22}x_t$$

The corresponding one-step-ahead forecast errors and variances are

$$FE_1^y = y_{t+1} - E_t\left[y_{t+1}\right] = v_{t+1}^y \qquad \text{var}\left(FE_1^y\right) = \sigma_y^2$$

$$FE_1^x = x_{t+1} - E_t\left[x_{t+1}\right] = v_{t+1}^x \qquad \text{var}\left(FE_1^x\right) = \sigma_x^2$$

Hence in the first period, all variation in the forecast error for y is due to its own shock. Likewise, 100% of the forecast error for x can be explained by its own shock. Using the same technique, the two-step ahead forecast for y is

$$\begin{aligned}
y_{t+2}^F &= E_t\left[\delta_{11}y_{t+1} + \delta_{12}x_{t+1} + v_{t+2}^y\right] \\
&= E_t\left[\delta_{11}\left(\delta_{11}y_t + \delta_{12}x_t + v_{t+1}^y\right) + \delta_{12}\left(\delta_{21}y_t + \delta_{22}x_t + v_{t+1}^x\right) + v_{t+2}^y\right] \\
&= \delta_{11}\left(\delta_{11}y_t + \delta_{12}x_t\right) + \delta_{12}\left(\delta_{21}y_t + \delta_{22}x_t\right)
\end{aligned}$$

and that for x is

$$\begin{aligned}
x_{t+2}^F &= E_t\left[\delta_{21}y_{t+1} + \delta_{22}x_{t+1} + v_{t+2}^x\right] \\
&= E_t\left[\delta_{21}\left(\delta_{11}y_t + \delta_{12}x_t + v_{t+1}^y\right) + \delta_{22}\left(\delta_{21}y_t + \delta_{22}x_t + v_{t+1}^x\right) + v_{t+2}^x\right] \\
&= \delta_{21}\left(\delta_{11}y_t + \delta_{12}x_t\right) + \delta_{22}\left(\delta_{21}y_t + \delta_{22}x_t\right)
\end{aligned}$$

The corresponding two-step-ahead forecast errors and variances are (recall that we are working with the special case of independent errors)

$$FE_2^y = y_{t+2} - E_t\left[y_{t+2}\right] = \left[\delta_{11}v_{t+1}^y + \delta_{12}v_{t+1}^x + v_{t+2}^y\right]$$

$$\text{var}\left(FE_2^y\right) = \delta_{11}^2\sigma_y^2 + \delta_{12}^2\sigma_x^2 + \sigma_y^2$$

$$FE_2^x = x_{t+2} - E_t\left[x_{t+2}\right] = \left[\delta_{21}v_{t+1}^y + \delta_{22}v_{t+1}^x + v_{t+2}^x\right]$$

$$\text{var}\left(FE_2^x\right) = \delta_{21}^2\sigma_y^2 + \delta_{22}^2\sigma_x^2 + \sigma_x^2$$

We can decompose the total variance of the forecast error for y, $\left(\delta_{11}^2\sigma_y^2 + \delta_{12}^2\sigma_x^2 + \sigma_y^2\right)$, into that due to shocks to y, $\left(\delta_{11}^2\sigma_y^2 + \sigma_y^2\right)$, and that due to shocks to x, $\left(\delta_{12}^2\sigma_x^2\right)$. This decomposition is often expressed in proportional terms. The proportion of the two-step forecast error variance of y explained by its "own" shock is

$$\left(\delta_{11}^2\sigma_y^2 + \sigma_y^2\right)\Big/\left(\delta_{11}^2\sigma_y^2 + \delta_{12}^2\sigma_x^2 + \sigma_y^2\right)$$

and the proportion of the two-step forecast error variance of y explained by the "other" shock is

$$\left(\delta_{12}^2\sigma_x^2\right)\Big/\left(\delta_{11}^2\sigma_y^2+\delta_{12}^2\sigma_x^2+\sigma_y^2\right)$$

Similarly, the proportion of the two-step forecast error variance of x explained by its own shock is

$$\left(\delta_{22}^2\sigma_x^2+\sigma_x^2\right)\Big/\left(\delta_{21}^2\sigma_y^2+\delta_{22}^2\sigma_x^2+\sigma_x^2\right)$$

and the proportion of the forecast error of x explained by the other shock is

$$\left(\delta_{21}^2\sigma_y^2\right)\Big/\left(\delta_{21}^2\sigma_y^2+\delta_{22}^2\sigma_x^2+\sigma_x^2\right)$$

For our numerical example with $\sigma_y = 1$, $\sigma_x = 2$, $\delta_{11} = 0.7$, $\delta_{12} = 0.2$, $\delta_{21} = 0.3$, and $\delta_{22} = 0.6$, we find that 90.303% of the two-step forecast error variance of y is due to y, and only 9.697% is due to x.

To sum up, suppose you were interested in the relationship between economic growth and inflation. A VAR model will tell you whether they are significantly related to each other; an impulse response analysis will show how growth and inflation react dynamically to shocks, and a variance decomposition analysis will be informative about the sources of volatility.

The General Case The example above assumes that x and y are not contemporaneously related and that the shocks are uncorrelated. There is no identification problem, and the generation and interpretation of the impulse response functions and decomposition of the forecast error variance are straightforward. In general, this is unlikely to be the case. Contemporaneous interactions and correlated errors complicate the identification of the nature of shocks and hence the interpretation of the impulses and decomposition of the causes of the forecast error variance. This topic is discussed in greater detail in textbooks devoted to time-series analysis.[2] A description of how the identification problem can arise is given in Appendix 13A.

13.5 Exercises

13.5.1 Problems

13.1 Consider the following first-order VAR model of stationary variables:

$$y_t = \delta_{11}y_{t-1} + \delta_{12}x_{t-1} + v_t^y$$
$$x_t = \delta_{21}y_{t-1} + \delta_{22}x_{t-1} + v_t^x$$

Under the assumption that there is no contemporaneous dependence, determine the impulse responses, four periods after a standard deviation shock for

a. y following a shock to y
b. y following a shock to x
c. x following a shock to y
d. x following a shock to x

13.2 Consider the first-order VAR model in Exercise 13.1. Under the assumption that there is no contemporaneous dependence, determine

a. the contribution of a shock to y on the variance of the three-step ahead forecast error for y
b. the contribution of a shock to x on the variance of the three-step ahead forecast error for y
c. the contribution of a shock to y on the variance of the three-step ahead forecast error for x
d. the contribution of a shock to x on the variance of the three-step ahead forecast error for x

[2]One reference you might consider is Lütkepohl, H. (2005) *Introduction to Multiple Time Series Analysis*, Springer, New York, Chapter 9.

13.3 The VEC model is a special form of the VAR for I(1) variables that are cointegrated. Consider the following VEC model:

$$\Delta y_t = \alpha_{10} + \alpha_{11}(y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + v_t^y$$
$$\Delta x_t = \alpha_{20} + \alpha_{21}(y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + v_t^x$$

The VEC model may also be rewritten as a VAR, but the two equations will contain common parameters:

$$y_t = \alpha_{10} + (\alpha_{11} + 1)y_{t-1} - \alpha_{11}\beta_0 - \alpha_{11}\beta_1 x_{t-1} + v_t^y$$
$$x_t = \alpha_{20} + \alpha_{21} y_{t-1} - \alpha_{21}\beta_0 - (\alpha_{21}\beta_1 - 1)x_{t-1} + v_t^x$$

a. Suppose you were given the following results from an estimated VEC model:

$$\widehat{\Delta y_t} = 2 - 0.5(y_{t-1} - 1 - 0.7x_{t-1})$$
$$\widehat{\Delta x_t} = 3 + 0.3(y_{t-1} - 1 - 0.7x_{t-1})$$

Rewrite the model in the VAR form.

b. Now suppose you were given the following results of an estimated VAR model, but you were also told that y and x are cointegrated.

$$\hat{y}_t = 0.7y_{t-1} + 0.3 + 0.24x_{t-1}$$
$$\hat{x}_t = 0.6y_{t-1} - 0.6 + 0.52x_{t-1}$$

Rewrite the model in the VEC form.

13.4 VAR and VEC models are popular forecasting models because they rely on the past history of observed outcomes to predict the expected future values.

a. Consider the following estimated VAR model:

$$y_t = \hat{\delta}_{11} y_{t-1} + \hat{\delta}_{12} x_{t-1} + \hat{v}_{1t}$$
$$x_t = \hat{\delta}_{21} y_{t-1} + \hat{\delta}_{22} x_{t-1} + \hat{v}_{2t}$$

What are the forecasts for y_{t+1} and x_{t+1}?
What are the forecasts for y_{t+2} and x_{t+2}?

b. Consider the following estimated VEC model:

$$\Delta y_t = \hat{\alpha}_{11}\left(y_{t-1} - \hat{\beta}_1 x_{t-1}\right) + \hat{v}_{1t}$$
$$\Delta x_t = \hat{\alpha}_{21}\left(y_{t-1} - \hat{\beta}_1 x_{t-1}\right) + \hat{v}_{2t}$$

What are the forecasts for y_{t+1} and x_{t+1}?
What are the forecasts for y_{t+2} and x_{t+2}?

13.5.2 Computer Exercises

13.5 The data file *gdp* contains quarterly data on the real GDP of Australia (*AUS*) and real GDP of the United States (*USA*) for the sample period 1970Q1 to 2000Q4.

a. Are the series stationary or nonstationary?

b. Test for cointegration allowing for an intercept term. You will find that the intercept is negative. Is this sensible? If not, repeat the test for cointegration excluding the constant term.

c. Save the cointegrating residuals and estimate the VEC model.

13.6 The data file *fred5* contains the log of RPDI (Y) and the log of RPCE (C) for the U.S. economy over the period 1986Q1 to 2015Q2.

a. Are the series stationary, or nonstationary? In particular, test whether the series are trend stationary.

b. Test for cointegration allowing for an intercept term. Are the series cointegrated?

c. Estimate a VAR model for the set of I(0) variables $\{\Delta C_t, \Delta Y_t\}$. Pay particular attention to the order of lags.

13.7 Consider again the data file *fred5* used in Example 13.2 and Exercise 13.6.

a. Estimate a VAR model for $\{\Delta C_t, \Delta Y_t\}$ with three lags of each variable included. Comment on the results. Has serial correlation in the errors been eliminated?

b. The concept of "Granger causality" was introduced in Section 9.3.4. In a VAR involving two variables *x* and *y*, we can ask whether *x* Granger causes *y*, whether *y* Granger causes *x*, and whether there is Granger causality in both directions. Using the model estimated in part (a), test whether ΔY Granger causes ΔC and whether ΔC Granger causes ΔY.

13.8 The data file *vec* contains 100 observations on two generated series of data, *x* and *y*. The variables are nonstationary and cointegrated without a constant term. Save the cointegrating residuals $(\hat{e})$ and estimate the VEC model. As a check, the results for the case normalized on *y* are

$$\begin{aligned}\widehat{\Delta y_t} &= -0.576(\hat{e}_{t-1})\\ (t)\quad &\quad (-6.158)\\ \widehat{\Delta x_t} &= 0.450(\hat{e}_{t-1})\\ (t)\quad &\quad (4.448)\end{aligned}$$

a. The residuals from the error correction model should not be autocorrelated. Are they?

b. Note that one of the error correction terms is negative and the other is positive. Explain why this is necessary.

13.9 The data file *var* contains 100 observations on two generated series of data, *w* and *z*. The variables are nonstationary but not cointegrated. Estimate a VAR model of changes in the variables. As a check, the results are (the intercept terms were not significant):

$$\begin{aligned}\widehat{\Delta w_t} &= 0.743\Delta w_{t-1} + 0.214\Delta z_{t-1}\\ (t)\quad &\quad (11.403)\qquad\quad (2.893)\\ \widehat{\Delta z_t} &= -0.155\Delta w_{t-1} + 0.641\Delta z_{t-1}\\ (t)\quad &\quad (-2.293)\qquad\quad (8.338)\end{aligned}$$

a. The residuals from the VAR model should not be autocorrelated. Is this the case?

b. Determine the impulse responses for the first two periods. (You may assume the special condition that there is no contemporaneous dependence.)

c. Determine the variance decompositions for the first two periods.

13.10 The quantity theory of money says that there is a direct relationship between the quantity of money in the economy and the aggregate price level. Put simply, if the quantity of money doubles, then the price level should also double. Figure 13.6 shows the percentage change in a measure of the quantity of money (*M*) and the percentage change in a measure of aggregate prices (*P*) for the United States between 1961Q1 and 2005Q4 (data file *qtm*). A VEC model was estimated as follows:

$$\begin{aligned}\widehat{\Delta P_t} &= -0.016(P_{t-1} - 1.004M_{t-1} + 0.039) + 0.514\Delta P_{t-1} - 0.005\Delta M_{t-1}\\ (t)\quad &\quad (-2.127)\qquad (-3.696)\qquad (1.714)\quad (7.999)\qquad (-0.215)\\ \widehat{\Delta M_t} &= 0.067(P_{t-1} - 1.004M_{t-1} + 0.039) - 0.336\Delta P_{t-1} - 0.340\Delta M_{t-1}\\ (t)\quad &\quad (3.017)\qquad (-3.696)\qquad (1.714)\ (-1.796)\qquad (-4.802)\end{aligned}$$

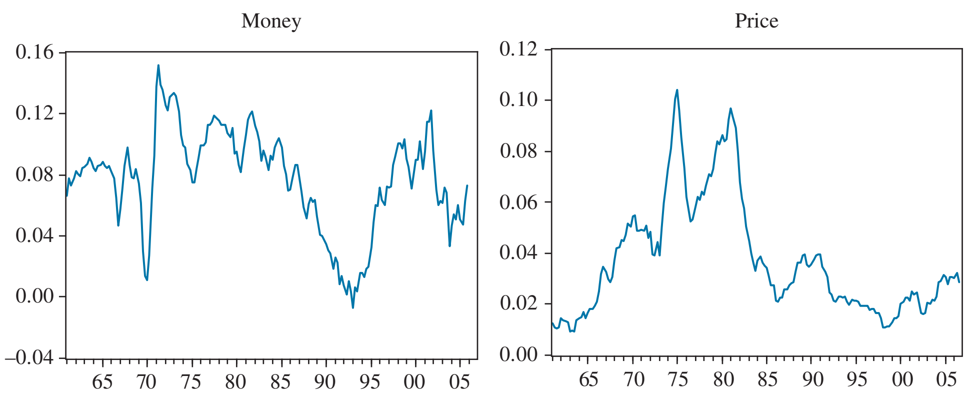

FIGURE 13.6 **Percentage changes in money and price.**

a. Identify the cointegrating relationship between P and M. Is the quantity theory of money supported?
b. Identify the error-correction coefficients. Is the system stable?
c. The above results were estimated using a system approach. Compute the cointegrating residuals and confirm that the series is indeed an I(0) variable.
d. Estimate a VEC model using the cointegrating residuals. (Your results should be the same as above.)

13.11 Research into the Phillips curve is concerned with providing empirical evidence of a tradeoff between inflation and unemployment. Can an economy experience lower unemployment if it is prepared to accept higher inflation? Figure 13.7 plots the changes in a measure of the unemployment rate (DU) and the changes in a measure of inflation (DP) for the United States for the sample period 1970M07 to 2009M06. A VAR model was estimated as follows:

$$\begin{aligned} DU_t &= 0.180 DU_{t-1} - 0.046 DP_{t-1} \\ (t) \quad & \quad (3.905) \qquad (-0.909) \end{aligned}$$

$$\begin{aligned} DP_t &= -0.098 DU_{t-1} + 0.373 DP_{t-1} \\ (t) \quad & \quad (-2.522) \qquad (8.711) \end{aligned}$$

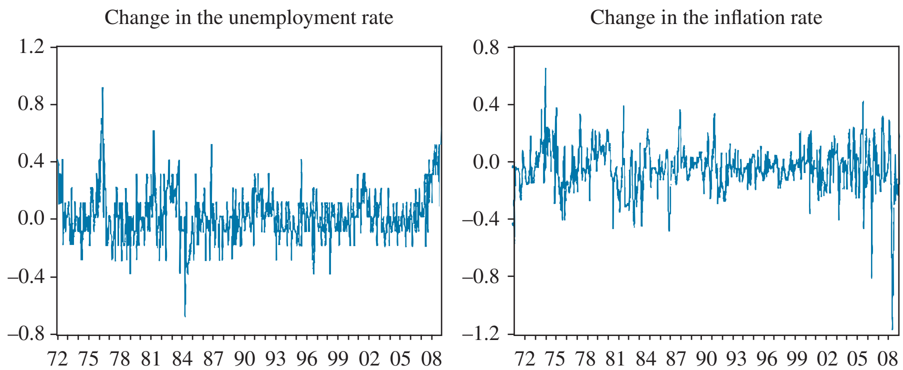

FIGURE 13.7 **Changes in the unemployment and inflation rates.**

a. Is there evidence of an inverse relationship between the change in the unemployment rate (DU) and the change in the inflation rate (DP)?
b. What is the response of DU at time $t+1$ following a unit shock to DU at time t?
c. What is the response of DP at time $t+1$ following a unit shock to DU at time t?
d. What is the response of DU at time $t+2$?
e. What is the response of DP at time $t+2$?

13.12 Figure 13.8 shows the time series for two exchange rates—the *EURO* per \$US and the *STERLING* per \$US (data file *sterling*). Both the levels and the changes in the data are shown.

a. Which set of data would you consider using to estimate a VEC model, and which set to estimate a VAR? Why?
b. Apply the two-step approach suggested in this chapter to estimate a VEC model.
c. Estimate a VAR model paying attention to the order of the lag.

13.13 Financial analysts often debate the role of dividends (DV) in the determination of share prices (SP). Figure 13.9 shows plots of the rate of change in DV and SP computed as

$$DV_t = 100\ln\left(DN_t/DN_{t-1}\right), \quad SP_t = 100\ln\left(PN_t/PN_{t-1}\right)$$

where PN is the Standard and Poor Composite Price Index; DN is the nominal dividend per share (source: Prescott, E. C. and Mehra, R. "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15 March, 1985, pp. 145–161). The data are annual observations over the period 1889–1979.

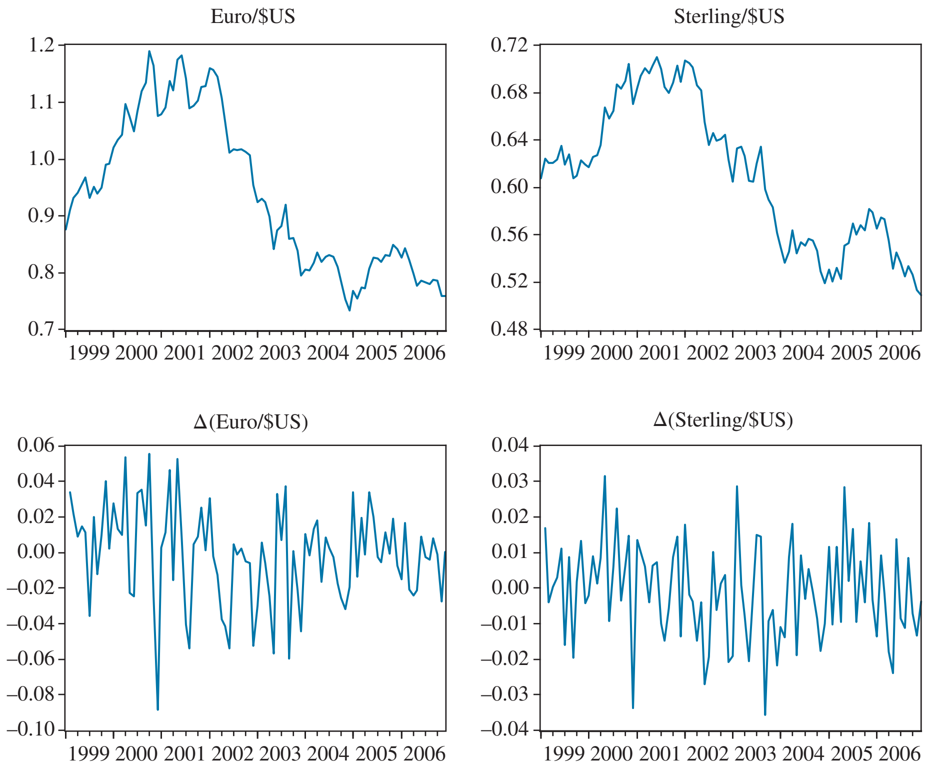

FIGURE 13.8 **Exchange rates.**

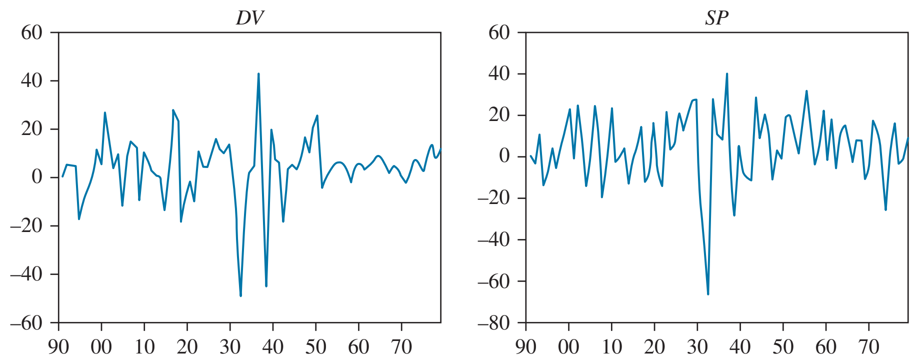

FIGURE 13.9 **Change in dividends (*DV*) and share price (*SP*).**

The data file is called *equity*. Estimate a first-order VAR for *SP* and *DV* by applying least squares to each equation:

$$SP_t = \beta_{10} + \beta_{11} SP_{t-1} + \beta_{12} DV_{t-1} + v_t^s$$
$$DV_t = \beta_{20} + \beta_{21} SP_{t-1} + \beta_{22} DV_{t-1} + v_t^d$$

Estimate an ARDL for each equation:

$$SP_t = \alpha_{10} + \alpha_{11} SP_{t-1} + \alpha_{12} DV_{t-1} + \alpha_{13} DV_t + e_t^s$$
$$DV_t = \alpha_{20} + \alpha_{21} SP_{t-1} + \alpha_{22} DV_{t-1} + \alpha_{23} SP_t + e_t^d$$

Compare the two sets of results and note the importance of the contemporaneous endogenous variable (SP, DV) in each equation.

a. Explain why least squares estimation of the VAR model with lagged variables on the right-hand side yields consistent estimates.

b. Explain why least squares estimation of the model with lagged and contemporaneous variables on the right-hand side yields inconsistent estimates. (You might like to refer to the material in Chapter 11.)

c. What do you infer about the role of dividends in the determination of share prices?

13.14 The file *gfc* contains data about economic activity in two major economies: the United States and the Euro Area (the group of countries in Europe where the Euro currency is the legal tender). Specifically, the data are the logs of their GDP, standardized so that the value of GDP is equal to 100 in 2000. The levels and the change in economic activity are shown in Figure 13.10(a) and (b). The sample period is from 1995Q1 to 2009Q4 and includes the global financial crisis that began in September 2007.

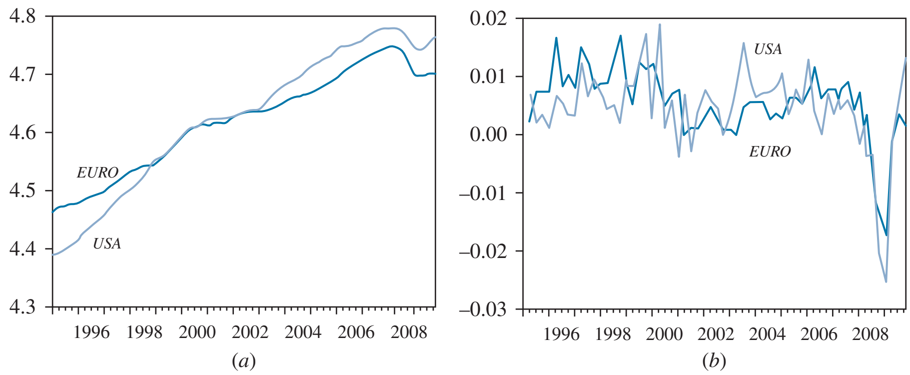

FIGURE 13.10 **Logs of GDP (*a*) and change in logs of GDP (*b*).**

a. Based on a visual inspection of the data, what would you infer about the interactions between the GDPs in the two economies?

b. Do the economies have a long-run relationship? Specify the econometric model and estimate the model. Plot the residuals and comment on their properties.

c. Do the economies have a short-run relationship? Specify the econometric model and estimate the model. Plot the residuals and comment on their properties.

13.15 The file *precious* contains monthly data on the prices of gold and silver (in logs) for the period 1970M1 to 2014M2.

a. Plot the two series and comment on the graph. Do the two prices appear to be moving together?

b. Use a series of hypothesis tests to decide on predictive models for the price of silver and the price of gold.

Appendix 13A The Identification Problem[3]

A bivariate dynamic system with contemporaneous interactions (also known as a structural model) is written as

$$y_t + \beta_1 x_t = \alpha_1 y_{t-1} + \alpha_2 x_{t-1} + e_t^y$$

$$x_t + \beta_2 y_t = \alpha_3 y_{t-1} + \alpha_4 x_{t-1} + e_t^x$$

[3]This appendix requires a basic understanding of matrix notation.

which can be more conveniently expressed in matrix form as

$$\begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} e_t^y \\ e_t^x \end{bmatrix}$$

or rewritten in symbolic form as $BY_t = AY_{t-1} + E_t$, where

$$Y_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} \quad B = \begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix} \quad A = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} \quad E_t = \begin{bmatrix} e_t^y \\ e_t^x \end{bmatrix}$$

A VAR representation (also known as reduced-form model) is written as

$$y_t = \delta_1 y_{t-1} + \delta_2 x_{t-1} + v_t^y$$
$$x_t = \delta_3 y_{t-1} + \delta_4 x_{t-1} + v_t^x$$

or in matrix form as: $Y_t = CY_{t-1} + V_t$, where

$$C = \begin{bmatrix} \delta_1 & \delta_2 \\ \delta_3 & \delta_4 \end{bmatrix} \quad V_t = \begin{bmatrix} v_t^y \\ v_t^x \end{bmatrix}$$

Clearly, there is a relationship between (13.A.1) and (13.A.2): $C = B^{-1}A$ and $V_t = B^{-1}E_t$. The special case considered in the chapter assumes that there are no contemporaneous interactions $(\beta_1 = \beta_2 = 0)$, making B an identity matrix. There is no identification problem in this case because the VAR residuals can be unambiguously "identified" as shocks to y or as shocks to x: $v^y = e^y$, $v^x = e^x$. The generation and interpretation of the impulse responses and variance decompositions are unambiguous.

In general, however, B is not an identity matrix, making v^y and v^x weighted averages of e^y and e^x. In this general case, impulse responses and variance decompositions based on v^y and v^x are not meaningful or useful because we cannot be certain about the source of the shocks. A number of methods exist for "identifying" the structural model from its reduced form.