$\frac{1}{T}\sum_{t=1}^{T}(\Delta Y_t)^2 \xrightarrow{p} \sigma_u^2$. Under the assumption that $Y_0 = 0$, the first term in Equation (14.22) can be written $Y_T/\sqrt{T} = \sqrt{\frac{1}{T}}\sum_{t=1}^{T}\Delta Y_t = \sqrt{\frac{1}{T}}\sum_{t=1}^{T}u_t$, which in turn obeys the central limit theorem; that is, $Y_T/\sqrt{T} \xrightarrow{d} N(0, \sigma_u^2)$. Thus $(Y_T/\sqrt{T})^2 - \frac{1}{T}\sum_{t=1}^{T}(\Delta Y_t)^2 \xrightarrow{d} \sigma_u^2(Z^2 - 1)$, where Z is a standard normal random variable. Recall, however, that the square of a standard normal distribution has a chi-squared distribution with one degree of freedom. It therefore follows from Equation (14.22) that, under the null hypothesis, the numerator in Equation (14.21) has the limiting distribution

$$\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}\Delta Y_t \xrightarrow{d} \frac{\sigma_u^2}{2}(\chi_1^2 - 1). \tag{14.23}$$

The large-sample distribution in Equation (14.23) is different than the usual large-sample normal distribution when the regressor is stationary. Instead, the numerator of the OLS estimator of the coefficient on Y_t in this Dickey-Fuller regression has a distribution that is proportional to a chi-squared distribution with one degree of freedom, minus one.

This discussion has only considered the numerator of $T\hat{\delta}$. The denominator also behaves unusually under the null hypothesis: because Y_t follows a random walk under the null hypothesis, $\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^2$ does not converge in probability to a constant. Instead, the denominator in Equation (14.21) is a random variable, even in large samples: under the null hypothesis, $\frac{1}{T^2}\sum_{t=1}^{T}Y_{t-1}^2$ converges in distribution jointly with the numerator. The unusual distributions of the numerator and denominator in Equation (14.21) are the source of the nonstandard distribution of the Dickey-Fuller test statistic and the reason that the ADF statistic has its own special table of critical values.

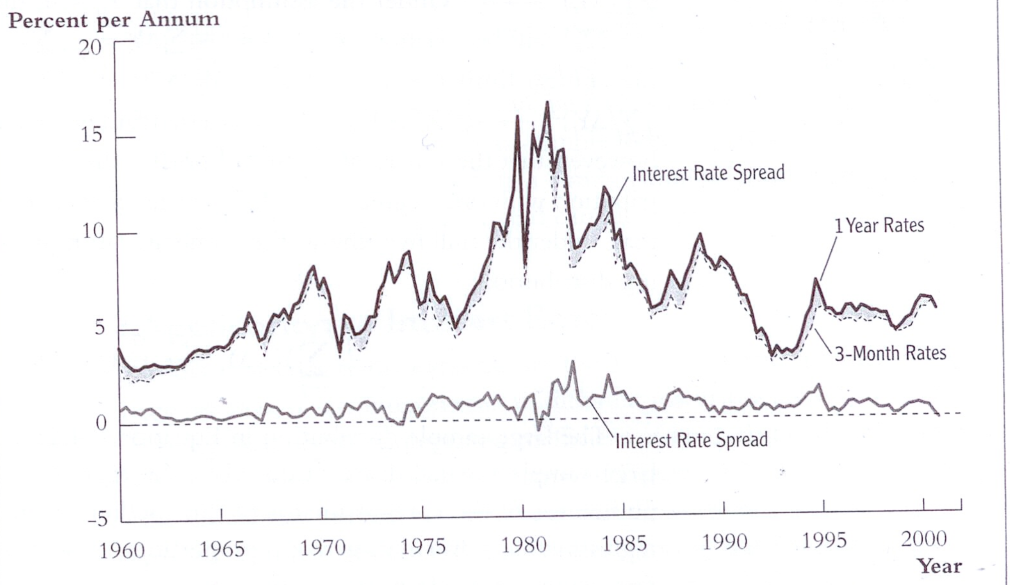
14.4 Cointegration

Sometimes two or more series have the same stochastic trend in common. In this special case, referred to as cointegration, regression analysis can reveal long-run relationships among time series variables, but some new methods are needed.

Cointegration and Error Correction

Two or more time series with stochastic trends can move together so closely over the long run that they appear to have the same trend component, that is, they appear to have a **common trend**. For example, two interest rates on U.S.

FIGURE 14.2 One-Year Interest Rate, Three-Month Interest Rate, and Interest Rate Spread



One-year and three-month interest rates share a common stochastic trend. The spread, or the difference, between the two rates does not exhibit a trend. These two interest rates appear to be cointegrated.

government debt are plotted in Figure 14.2. One of the rates is the interest rate on 90-day U.S. Treasury bills, at an annual rate ($R90_t$); the other is the interest rate on a one-year U.S. Treasury bond ($R1yr_t$); these interest rates are discussed in Appendix 14.1. The interest rates exhibit the same long-run tendencies or trends: both were low in the 1960s, both rose through the 1970s to peaks in the early 1980s, then both fell through the 1990s. Moreover, the difference between the two series, $R1yr_t - R90_t$, which is called the "spread" between the two interest rates and is also plotted in Figure 14.2, does not appear to have a trend. That is, subtracting the 90-day interest rate from the one-year interest rate appears to eliminate the trends in both of the individual rates. Said differently, although the two interest rates differ, they appear to share a common stochastic trend: because the trend in each individual series is eliminated by subtracting one series from the other, the two series must have the same trend, that is, they must have a common stochastic trend.

Two or more series that have a common stochastic trend are said to be **cointegrated**. The formal definition of cointegration (due to Granger, 1983) is given in Key Concept 14.5. In this section, we introduce a test for whether cointegration is present, discuss estimation of the coefficients of regressions relating cointegrated variables, and illustrate the use of the cointegrating relationship for forecasting. The discussion initially focuses on the case that there are only two variables, X_t and Y_t.

Vector error correction model. Until now, we have eliminated the stochastic trend in an $I(1)$ variable Y_t by computing its first difference, ΔY_t; the problems created by stochastic trends were then avoided by using ΔY_t instead of Y_t in time series regressions. If X_t and Y_t are cointegrated, however, another way to eliminate the trend is to compute the difference $Y_t - \theta X_t$. Because the term $Y_t - \theta X_t$ is stationary, it too can be used in regression analysis.

In fact, if X_t and Y_t are cointegrated, the first differences of X_t and Y_t can be modeled using a VAR, augmented by including $Y_{t-1} - \theta X_{t-1}$ as an additional regressor:

$$\Delta Y_t = \beta_{10} + \beta_{11}\Delta Y_{t-1} + \cdots + \beta_{1p}\Delta Y_{t-p} + \gamma_{11}\Delta X_{t-1} + \cdots + \gamma_{1p}\Delta X_{t-p} + \alpha_1(Y_{t-1} - \theta X_{t-1}) + u_{1t} \tag{14.24}$$

$$\Delta X_t = \beta_{20} + \beta_{21}\Delta Y_{t-1} + \cdots + \beta_{2p}\Delta Y_{t-p} + \gamma_{21}\Delta X_{t-1} + \cdots + \gamma_{2p}\Delta X_{t-p} + \alpha_2(Y_{t-1} - \theta X_{t-1}) + u_{2t}. \tag{14.25}$$

The term $Y_t - \theta X_t$ is called the **error correction term**. The combined model in Equations (14.24) and (14.25) is called a **vector error correction model** (VECM). In a VECM, past values of $Y_t - \theta X_t$ help to predict future values of ΔY_t and/or ΔX_t.

How Can You Tell Whether Two Variables Are Cointegrated?

There are three ways to decide whether two variables can plausibly be modeled as cointegrated: use expert knowledge and economic theory, graph the series and see whether they appear to have a common stochastic trend, and perform statistical tests for cointegration. All three methods should be used in practice.

First, you must use your expert knowledge of these variables to decide whether cointegration is in fact plausible. For example, the two interest rates in Figure 14.2 are linked together by the so-called expectations theory of the term structure of interest rates. According to this theory, the interest rate on January 1

Key Concept 14.5

Cointegration

Suppose X_t and Y_t are integrated of order one. If, for some coefficient θ, $Y_t - \theta X_t$ is integrated of order zero, then X_t and Y_t are said to be **cointegrated**. The coefficient θ is called the **cointegrating coefficient**.

If X_t and Y_t are cointegrated, then they have the same, or common, stochastic trend. Computing the difference $Y_t - \theta X_t$ eliminates this common stochastic trend.

on the one-year Treasury bond is the average of the interest rate on a 90-day Treasury bill for the first quarter of the year and the expected interest rates on future 90-day Treasury bills issued in the second, third, and fourth quarters of the year; if not, then investors could expect to make money by holding either the one-year Treasury note or a sequence of four 90-day Treasury bills, and they would bid up prices until the expected returns are equalized. If the 90-day interest rate has a random walk stochastic trend, this theory implies that this stochastic trend is inherited by the one-year interest rate and that the difference between the two rates, that is, the spread, is stationary. Thus, the expectations theory of the term structure implies that, if the interest rates are $I(1)$, then they will be cointegrated with a cointegrating coefficient of $\theta = 1$ (Exercise 14.2).

Second, visual inspection of the series helps to identify cases in which cointegration is plausible. For example, the graph of the two interest rates in Figure 14.2 shows that each of the series appears to be $I(1)$ but that the spread appears to be $I(0)$, so that the two series appear to be cointegrated.

Third, the unit root testing procedures introduced so far can be extended to test for cointegration. The insight on which these tests are based is that if Y_t and X_t are cointegrated with cointegrating coefficient θ, then $Y_t - \theta X_t$ is stationary; otherwise, $Y_t - \theta X_t$ is nonstationary (is $I(1)$). The hypothesis that Y_t and X_t are not cointegrated (that is, that $Y_t - \theta X_t$ is $I(1)$) therefore can be tested by testing the null hypothesis that $Y_t - \theta X_t$ has a unit root; if this hypothesis is rejected, then Y_t and X_t can be modeled as cointegrated. The details of this test depend on whether the cointegrating coefficient θ is known.

Testing for cointegration when θ is known. In some cases expert knowledge or economic theory suggests values of θ. When θ is known, the Dickey-Fuller and

DF-GLS unit root tests can be used to test for cointegration by first constructing the series $z_t = Y_t - \theta X_t$, then testing the null hypothesis that z_t has a unit autoregressive root.

Testing for cointegration when θ is unknown. If the cointegrating coefficient θ is unknown then it must be estimated prior to testing for a unit root in the error correction term. This preliminary step makes it necessary to use different critical values for the subsequent unit root test.

Specifically, in the first step the cointegrating coefficient θ is estimated by OLS estimation of the regression

$$Y_t = \alpha + \theta X_t + z_t. \tag{14.26}$$

In the second step, a Dickey-Fuller t-test (with an intercept but no time trend) is used to test for a unit root in the residual from this regression, $\hat{z}_t$. This two-step procedure is called the Engle-Granger Augmented Dickey-Fuller test for cointegration, or **EG-ADF** (Engle and Granger, 1987).

Critical values of the EG-ADF statistic are given in Table 14.2.[2] The critical values in the first row apply when there is a single regressor in Equation (14.26), so that there are two cointegrated variables (X_t and Y_t). The subsequent rows apply to the case of multiple cointegrated variables, which is discussed at the end of this section.

Estimation of Cointegrating Coefficients

If X_t and Y_t are cointegrated, then the OLS estimator of the coefficient in the cointegrating regression in Equation (14.26) is consistent. However, in general the OLS estimator has a nonnormal distribution, and inferences based on its t-statistics can be misleading whether or not those t-statistics are computed using HAC standard errors. Because of these drawbacks of the OLS estimator of θ, econometricians have developed a number of other estimators of the cointegrating coefficient.

One such estimator of θ that is simple to use in practice is the so-called **dynamic OLS (DOLS)** estimator (Stock and Watson, 1993). The DOLS esti-

[2]The critical values in Table 14.2 are taken from Fuller (1976) and Phillips and Ouliaris (1990). Following a suggestion by Hansen (1992), the critical values in Table 14.2 are chosen so that they apply whether or not X_t and Y_t have drift components.

TABLE 14.2 Critical Values for the Engle-Granger ADF statistic

Number of *X*'s in Equation (14.26)	10%	5%	1%
1	−3.12	−3.41	−3.96
2	−3.52	−3.80	−4.36
3	−3.84	−4.16	−4.73
4	−4.20	−4.49	−5.07

mator is based on a modified version of Equation (14.26) that includes past, present, and future values of the change in X_t:

$$Y_t = \beta_0 + \theta X_t + \sum_{j=-p}^{p} \delta_j \Delta X_{t-j} + u_t. \tag{14.27}$$

Thus, in Equation (14.27), the regressors are X_t, $\Delta X_{t+p}, \ldots, \Delta X_{t-p}$. The DOLS estimator of θ is the OLS estimator of θ in the regression of Equation (14.27).

If X_t and Y_t are cointegrated, then the DOLS estimator is efficient in large samples. Moreover, statistical inferences about θ and the δ's in Equation (14.27) based on HAC standard errors are valid. For example, the *t*-statistic constructed using the DOLS estimator with HAC standard errors has a standard normal distribution in large samples.

One way to interpret Equation (14.27) is to recall from Section 13.3 that cumulative dynamic multipliers can be computed by modifying the distributed lag regression of Y_t on X_t and its lags. Specifically, in Equation (13.7), the cumulative dynamic multipliers were computed by regressing Y_t on ΔX_t, lags of ΔX_t, and X_{t-r}; the coefficient on X_{t-r} in that specification is the long-run cumulative dynamic multiplier. Similarly, if X_t were strictly exogenous, then in Equation (14.27), the coefficient on X_t, θ, would be the long-run cumulative multiplier, that is, the long-run effect on Y of a change in X. If X_t is not strictly exogenous, then the coefficients do not have this interpretation. Nevertheless, because X_t and Y_t have a common stochastic trend if they are cointegrated, the DOLS estimator is consistent even if X_t is endogenous.

The DOLS estimator is not the only efficient estimator of the cointegrating coefficient. The first such estimator was developed by Søren Johansen (Johansen, 1988). For a discussion of Johansen's method and of other ways to estimate the cointegrating coefficient, see Hamilton (1994, Chapter 20).

Even if economic theory does not suggest a specific value of the cointegrating coefficient, it is important to check whether the estimated cointegrating relationship makes sense in practice. Because cointegration tests can be misleading (they can improperly reject the null hypothesis of no cointegration more frequently than they should, and frequently they improperly fail to reject the null), it is especially important to rely on economic theory, institutional knowledge, and common sense when estimating and using cointegrating relationships.

Extension to Multiple Cointegrated Variables

The concepts, tests, and estimators discussed here extend to more than two variables. For example, if there are three variables, Y_t, X_{1t}, and X_{2t}, each of which is $I(1)$, then they are cointegrated with cointegrating coefficients θ_1 and θ_2 if $Y_t - \theta_1 X_{1t} - \theta_2 X_{2t}$ is stationary. When there are three or more variables, there can be multiple cointegrating relationships. For example, consider modeling the relationship among three interest rates: the three-month rate, the one-year rate, and the five-year rate ($R5yr$). If they are $I(1)$, then the expectations theory of the term structure of interest rates suggests that they will all be cointegrated. One cointegrating relationship suggested by the theory is $R90_t - R1yr_t$, and a second relationship is $R90_t - R5yr_t$. (The relationship $R1yr_t - R5yr_t$ is also a cointegrating relationship, but it contains no additional information beyond that in the other relationships because it is perfectly multicollinear with the other two cointegrating relationships.)

The EG-ADF procedure for testing for a single cointegrating relationship among mulitple variables is the same as for the case of two variables, except that the regression in Equation (14.26) is modified so that both X_{1t} and X_{2t} are regressors; the critical values for the EG-ADF test are given in Table 14.2, where the appropriate row depends on the number of regressors in the first-stage OLS cointegrating regression. The DOLS estimator of a single cointegrating relationship among multiple X's involves including the level of each X along with leads and lags of the first difference of each X. Tests for multiple cointegrating relationships can be performed using the system methods, such as Johansen's (1988) method, and the DOLS estimator can be extended to multiple cointegrating relationships by estimating multiple equations, one for each cointegrating relationship. For additional discussion of cointegration methods for multiple variables, see Hamilton (1994).

A cautionary note. If two or more variables are cointegrated then the error correction term can help to forecast these variables and, possibly, other related variables. However, cointegration requires the variables to have the same stochastic trends. Trends in economic variables typically arise from complex interactions of disparate forces, and closely related series can have different trends for subtle reasons. If variables that are not cointegrated are incorrectly modeled using a VECM, then the error correction term will be $I(1)$; this introduces a trend into the forecast that can result in poor out-of-sample forecast performance. Thus forecasting using a VECM must be based on a combination of compelling theoretical arguments in favor of cointegration and careful empirical analysis.

Application to Interest Rates

As discussed above, the expectations theory of the term structure of interest rates implies that, if two interest rates of different maturities are $I(1)$, then they will be cointegrated with a cointegrating coefficient of $\theta = 1$, that is, the spread between the two rates will be stationary. Inspection of Figure 14.2 provides qualitative support for the hypothesis that the one-year and three-month interest rates are cointegrated. We first use unit root and cointegration test statistics to provide more formal evidence on this hypothesis, then estimate a vector error correction model for these two interest rates.

Unit root and cointegration tests. Various unit root and cointegration test statistics for these two series are reported in Table 14.3. The unit root test statistics in the first two rows examine the hypothesis that the two interest rates, the three-month rate ($R90$) and the one-year rate ($R1yr$), individually have a unit root. Two of the four statistics in the first two rows fail to reject this hypothesis at the 10% level, and three of the four fail to reject at the 5% level. The exception is the ADF statistic evaluated for the 90-day Treasury bill rate (−2.96), which rejects the unit root hypothesis at the 5% level. The ADF and DF-GLS statistics lead to

TABLE 14.3 Unit Root and Cointegration Test Statistics for Two Interest Rates

Series	ADF Statistic	DF-GLS Statistic
$R90$	−2.96*	-1.88^{+}
$R1yr$	−2.22	−1.37
$R1yr - R90$	−6.31**	−5.59**
$R1yr - 1.046R90$	−6.97**	—

$R90$ is the interest rate on 90-day U.S. Treasury bills, at an annual rate, and $R1yr$ is the interest rate on one-year U.S. Treasury bonds. Regressions were estimated using quarterly data over the period 1962:I–1999:IV. The number of lags in the unit root test statistic regressions were chosen by AIC (six lags maximum). Unit root test statistics are significant at the $^{+}$10%, *5%, or **1% significance level.