TOPIC 4

MODELING LONG-RUN RELATIONSHIPS IN FINANCE

1. Nonstationarity and spurious regression

The series y_t is a random walk with drift if $y_t = \mu + y_{t-1} + \varepsilon_t$ where ε_t is a white noise process, often referred to as shocks or innovations. This specification can be thought of as an AR(1) process $y_t = \mu + \rho y_{t-1} + \varepsilon_t$, with AR parameter $\rho = 1$.[1] By recursive substitution,

$$y_{t+j} = j \cdot \mu + y_t + \varepsilon_{t+j} + \varepsilon_{t+j-1} + \ldots + \varepsilon_{t+1}$$

It then follows that

$$E(y_{t+j} \mid y_t) = j \cdot \mu + y_t$$
$$\mathrm{var}(y_{t+j} \mid y_t) = j \cdot \sigma^2$$

As $j \to \infty$, the mean and variance are unbounded. Consequently, a random walk is not a covariance stationary process. It is also called as stochastically non-stationary process. To induce non-stationarity we need to difference the time series $y_t = \mu + y_{t-1} + \varepsilon_t$ using difference operator $\Delta y_t = y_t - y_{t-1}$:

$$y_t - y_{t-1} = \mu + y_{t-1} - y_{t-1} + \varepsilon_t$$
$$\Delta y_t = \mu + \varepsilon_t$$

Obviously, differences series Δy_t are white noise now. In this case y_t is called integrated of order one, I(1) and the differenced series Δy_t is integrated of order zero, I(0).

It may be the case that differencing only one time does not make the series stationary. E.g., to induce stationary in $y_t = \mu + 2y_{t-1} - y_{t-2} + \varepsilon_t$ we need to difference twice:

$$y_t - y_{t-1} = \mu + 2y_{t-1} - y_{t-2} - y_{t-1} + \varepsilon_t \qquad \text{or} \qquad \Delta y_t = \mu + \Delta y_{t-1} + \varepsilon_t$$

Δy_t is still unit root process. Second differencing

$$(y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = \mu + (y_{t-1} - y_{t-2}) - (y_{t-1} - y_{t-2}) + \varepsilon_t \text{ or}$$
$$\Delta(\Delta y_t) = \mu + \Delta(\Delta y_{t-1}) + \varepsilon_t = \mu + \varepsilon_t$$

induces stationarity. In this case y_t is called integrated of order two, I(2).

---

[1] Note, when $\rho = -1$, we also have unit root process where shocks are alternating in sign. This case is similar to when $\rho = 1$ and it rarely observed in finance.

We may have I(k) processes in general, but processes with k>2 are not typical in finance and economics. Most of financial and economic time series are I(1), some prices series may be I(2).

Another potential source of nonstationarity is presence of time trend in the data. Time series generated by

$$y_t = \mu + \beta t + \varepsilon_t$$

are said to contain a deterministic trend, or time trend.

To induce stationarity we need to estimate the model and filter out time trend (use residuals).

Using nonstationary variables in linear regression may (but does not always!) produce the phenomenon of spurious regression. Spurious regression may have high R^2, and large test statistics (small *p*-values) suggesting relationship between dependent and independent variables, where in fact may be not related. Therefore before using potentially non-stationary variables in the regression one needs to test for unit root.

2. Dickey-Fuller test and other tests for nonstationarity.

The AR(1) model: $y_t = \rho y_{t-1} + \varepsilon_t$ can be reparameterised as

$$\Delta y_t = \psi y_{t-1} + \varepsilon_t \text{ where } \psi = \rho - 1$$

The Dickey-Fuller test is a test of the hypothesis that

H_0 : $\psi = 0$ (i.e. $\rho = 1$) "unit root" against the alternative
H_1 : $\psi < 0$ (i.e. $\rho < 1$) "stationary process"

The t-ratio or t-statistic associated with the estimated value of ψ is used to test this hypothesis by comparison with the appropriate critical value. Often the regression above is augmented with lags of Δy_t appearing on the right-hand side to account for serial correlation. In this case, the test is referred to as an augmented Dickey-Fuller test and is performed in exactly the same way as the Dickey-Fuller test.

Note that under the null the test statistics is not distributed according to Student's t distributions, because if presence of a non-stationary term. Special Dickey-Fuller test critical values need to be used. Moreover, if you assume that your equation contains either a drift term, or a drift term and a time trend different critical values need to be used.

Augmented Dickey-Fullers (ADF) test ties to incorporate any autocorrelations in residuals ε_t and take the form

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \varepsilon_t .$$

Phillips and Perron (PP) have developed a more comprehensive theory of unit root nonstationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow for autocorrelated residuals.
The tests usually give the same conclusions as the ADF tests, and the calculation of the test statistics is complex.
PP and ADF tests perform poor when we are in the situation close to unit root, that is they are poor at deciding if $\rho = 1$ or $\rho = 0.95$, especially with small sample sizes.

KPSS test (Kwaitowski, Phillips, Schmidt and Shin, 1992) tries to address this problem by using the reversed null hypothesis:

H_0 : "stationary process" against the alternative
H_1 : "unit root"

Both sets of test may be performed for robustness.

2. Cointegration and Common Trends

Often we observe the tendency for financial time series, for example, interest rates, to move together over time, even though individually each time series is characterized as an I(1) process. When two I(1) series move together over time, it is possible that there is a long-run relationship between them. Further, we discuss how a long-run relationship among two time series may be detected and, if such a relationship is uncovered, how the long-run relationship is restored when the series deviate in the short-run from the long-run relationship.

Suppose we have data on two nominal interest rate series, one a short-term interest rate denoted r_t^s and the other a long-term interest rate denoted r_t^l. Typically, we find that both interest rate series are I(1) processes on the basis of, say, ADF tests. One possible explanation for this finding is provided by the following model for interest rates:

$$r_t^s = a_1 + i_t + \varepsilon_t^s \qquad (1)$$
$$r_t^l = a_2 + i_t + \varepsilon_t^l \qquad (2)$$
$$i_t = \alpha + i_{t-1} + \upsilon_t \qquad (3)$$

where i_t is the rate of inflation from *t-1* to *t* and the error term in each equation is a white noise process. Both nominal interest rates are postulated to vary one-for-one with the rate of inflation as first suggested by Irving Fisher. Both the interest rate series are I(1) because the rate of inflation i_t is I(1); specifically the rate of inflation is a random walk process with drift. Although both nominal interest rates are I(1) processes, they will move together because they share a common unit root component, namely, the unit root in inflation. We say that r_t^s and r_t^l share a *common stochastic trend*, which is the unit root component of inflation.

In this setup, the spread, defined as $r_t^l - r_t^s$, is stationary; it is an I(0) process. To see this, subtract equation (1) from equation (2) to get:

$$r_t^l - r_t^s = a_2 - a_1 + \varepsilon_t^l - \varepsilon_t^s \tag{4}$$

The spread is an I(0) process, that is, it is stationary, because the difference of two white noise error terms is stationary (that is, $\varepsilon_t^l - \varepsilon_t^s$, is stationary). The spread can be thought of as a linear combination of the long and short nominal interest rate. In particular, the stationary linear combination of the interest rate series is:

$$\begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} r_t^l \\ r_t^s \end{pmatrix} \tag{5}$$

We say that the spread is *cointegrated* with cointegrating vector, denoted β, as

$$\beta = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{6}$$

To summarize, because there is a linear combination of the interest rate series which is stationary, the two interest rates are cointegrated. The specific linear combination is shown by the cointegrating vector, which in this case is $\beta' = (1 \quad -1)$.

There is cointegration between the two interest rate series because they share a common stochastic trend or a common unit root component. If the random walk component in r_t^l were independent of the random walk component in r_t^s, the two interest rate series would not share a common stochastic trend and thus would not be cointegrated, in which case there will *not* exist a linear combination of the two interest rates which is stationary. However, that does not happen in our model because the two interest rates share the same unit root component, which is the unit root in inflation.

Without loss of generality, we can define $a = a_2 - a_1$ and $\varepsilon_t = \varepsilon_t^l - \varepsilon_t^s$ so that equation (4) can be written as:

$$r_t^l - r_t^s = a + \varepsilon_t$$

Because ε_t is a white noise process and can be thought of as a stationary deviation, the long-run relationship between r_t^l and r_t^s is given as:

$$r_t^l - r_t^s = a$$

The deviation from the long-run or cointegrating relationship is given by ε_t where $\varepsilon_t = r_t^l - r_t^s - a$.

Let $\beta = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $X_t = \begin{pmatrix} r_t^l \\ r_t^s \end{pmatrix}$. Then the long-run equilibrium relationship can be expressed as:

$$\beta' X_t = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} r_t^l \\ r_t^s \end{pmatrix} = a$$

and the deviations from the long-run equilibrium relationship are $\varepsilon_t = \beta' X_t - a$.

Consider a more general setup. Consider two I(1) variables Y_t and X_t that are generated by two independent random walks:

$$Y_t = Y_{t-1} + \varepsilon_{1t}$$
$$X_t = X_{t-1} + \varepsilon_{2t}$$

where the ε_{1t} and ε_{2t} are white noise processes. Suppose we estimate a regression of the form:

$$Y_t = \alpha_0 + \beta X_t + \varepsilon_t \tag{7}$$

which can be written as

$$Y_t - \beta X_t = \alpha_0 + \varepsilon_t$$

We can interpret $Y_t - \beta X_t = \alpha_0$ as the long-run relationship between Y_t and X_t as long as ε_t is a stationary process. However, in this case, ε_t will be non-stationary because both Y_t and X_t are independent non-stationary processes so that there will be *no* linear combination of them which will be stationary.

Let us change the set-up and assume that:

$$Y_t = \beta X_{t-1} + \varepsilon_{1t} \tag{7a}$$
$$X_t = X_{t-1} + \varepsilon_{2t} \tag{7b}$$

In this case, both variables are I(1) processes. The variable X_t is a random walk process and Y_t is non-stationary because it depends on the random walk process X_t. Again, suppose we estimate a regression of the form:

$$Y_t = \alpha_0 + \beta X_t + \varepsilon_t \tag{8}$$

which can be written as:

$$Y_t - \beta X_t = \alpha_0 + \varepsilon_t \tag{9}$$

In this case, we can interpret $Y_t - \beta X_t = \alpha_0$ as the long-run relationship between Y_t and X_t because ε_t is a stationary process. To see this, substitute (7a) and (7b) into $Y_t - \beta X_t$ to get:

$$\begin{aligned} Y_t - \beta X_t &= \beta X_{t-1} + \varepsilon_{1t} - \beta(X_{t-1} + \varepsilon_{2t}) \\ &= \varepsilon_{1t} - \beta\varepsilon_{2t} \end{aligned}$$

Hence $Y_t - \beta X_t$ is a stationary series since it is given by $\varepsilon_{1t} - \beta\varepsilon_{2t}$. Thus ε_t in equation (8) or (9) is a stationary error term. In this case, the series are cointegrated because the unit root in Y_t is due to the unit root in X_t so that there is a common unit root in both of the series. In other words, it is legitimate to estimate a regression of the form of equation (8) where an I(1) dependent variable is regressed on an I(1) independent variable as long as the variables are cointegrated so that the regression residual is stationary.

3. The Engle-Granger Test for Cointegration

In view of the discussion so far, a natural test for cointegration between two I(1) variables, Y_t and X_t is to estimate the regression

$$Y_t = \alpha_0 + \beta X_t + \varepsilon_t \tag{9a}$$

by OLS and test whether the estimated regression residual is stationary. Specifically, denote the OLS residuals from this regression as e_t. Estimate the Augmented Dickey Fuller regression:

$$\Delta e_t = \gamma e_{t-1} + \sum_{i=1}^{n} \delta_i \Delta e_{t-i} + \upsilon_t$$

Since the e_t's are the residuals from a regression equation, the mean of the e_t's is zero so there is no need to include an intercept term in the ADF regression. The choice of *n* is determined on the basis of some information criteria, for example, the AIC criterion. If the estimated residuals exhibit serial correlation, then *n* is chosen to be larger then zero and the augmented form of the Dickey-Fuller regression is used. If we can reject the null hypothesis of a unit root in the residuals (that is, we can reject the null of $\gamma = 0$), we conclude that the two variables are cointegrated and share a common stochastic trend. If we cannot reject the null hypothesis of a unit root in the residuals (that is, we cannot reject the null of $\gamma = 0$), we conclude that the two variables are not cointegrated and that the random walk in one variable is independent of the random walk in the other variable.

The regression of Y_t on a constant and X_t and the subsequent testing for a unit root in the estimated regression residual is referred to as the Engle-Granger procedure.

In the application of the ADF test to the regression residuals, it is not appropriate to use the standard Dickey-Fuller critical values in testing for a unit root. The reason is that the OLS estimates of α_0 and β in equation (9a) are chosen to minimize the sum of squared residuals (that is, to make the residual variance as small as possible). Thus, the procedure is biased toward finding a stationary error process, since unit root processes have large variance in finite samples. The critical values to test the null hypothesis of a unit root in the residuals must take account of this bias. Appropriate critical values for the ADF test when it is used as a test for cointegration have been derived by Engle and Granger and refined by MacKinnon. These critical values depend on the number of variables used in the analysis and on the sample size. For the case of two variables, the critical values are shown in the table below.

Table 1: Critical Values for the Engle-Granger Test for Two Variables

Sample Size (T)	1%	5%	10%
50	-4.123	-3.461	-3.130
100	-4.008	-3.398	-3.087
200	-3.954	-3.368	-3.067
500	-3.921	-3.350	-3.054

Note: The critical values are for bivariate cointegrating relations (with a constant in the cointegrating vector) estimated using the Engle-Granger procedure.

By comparison, for example, the standard ADF critical value for an ADF regression with no constant term is -2.60 for T=100. The critical values in Table 1 are larger to take account of the fact that the testing procedure is biased in favour of cointegration.

4. Cointegration and Error Correction

Engle and Granger have shown that if Y_t and X_t are cointegrated, then there exists an error correction model which describes how Y_t and X_t adjust in the short-run following a deviation from long-run equilibrium. (This is known as the Granger Representation Theorem). In an error-correction model, the short-term dynamics of the variables in the system are influenced by the deviation from equilibrium.

In the interest rate model, a simple error correction model that could apply is:

$$\Delta r_t^s = \alpha_s (r_{t-1}^l - r_{t-1}^s - a) + \eta_{st}$$
$$\Delta r_t^l = -\alpha_l (r_{t-1}^l - r_{t-1}^s - a) + \eta_{lt}$$

where $\alpha_s > 0, \alpha_l > 0$ and η_{st}, η_{lt} are white-noise disturbance terms. The short and long term interest rates change in response to stochastic shocks (represented by η_{st} and η_{lt})

and in response to the previous period's deviation from long-run equilibrium. Everything else being equal, if this deviation happened to be positive ($r_{t-1}^l - r_{t-1}^s - a$ >0), the short-term interest rate would rise and the long-term rate would fall. Long-run equilibrium is attained when $r_t^l = a + r_t^s$. In practice, we would estimate the cointegrating regression

$$r_t^l = a + \beta r_t^s + \varepsilon_t$$

and test whether the residuals are stationary. Suppose we find that our OLS estimates of a and β are, respectively, $\hat{a}$ and $\hat{\beta}$ and that the estimated OLS residuals (e_t) are stationary. This would agree with our theory provided $\hat{\beta}$ is very close to one. Then the estimated deviation from the long-run equilibrium relationship is given by:

$$e_t = r_t^l - \hat{a} - \hat{\beta} r_t^s$$

and the error-correction model can be written as

$$\Delta r_t^s = \alpha_s (r_{t-1}^l - \hat{\beta} r_{t-1}^s - \hat{a}) + \eta_{st} \quad (10)$$
$$\Delta r_t^l = -\alpha_l (r_{t-1}^l - \hat{\beta} r_{t-1}^s - \hat{a}) + \eta_{lt} \quad (11)$$

or equivalently as

$$\Delta r_t^s = \alpha_s (e_{t-1}) + \eta_{st} \quad (12)$$
$$\Delta r_t^l = -\alpha_l (e_{t-1}) + \eta_{lt} \quad (13)$$

We can estimate equations (10) and (11), or equivalently equations (12) and (13), by OLS since the independent variable $e_{t-1} = r_{t-1}^l - \hat{\beta} r_{t-1}^s - \hat{a}$ is stationary by construction. The OLS estimates of α_s and α_l (denoted $\hat{\alpha}_s$ and $\hat{\alpha}_l$) are often referred to as the estimates of the speed of adjustment parameters and measure how quickly r_t^s and r_t^l adjust to bring about long-run equilibrium. For example, the larger $\hat{\alpha}_s$ the larger the response of the short-term interest rate to the previous period's deviation from long-run equilibrium. Suppose for argument sake that the OLS estimate of $\hat{\alpha}_l = 0$ and $\hat{\alpha}_s > 0$. In this case, all the adjustment to long-run equilibrium is through movements in the short-term interest rate because the long-term interest rate does not adjust to last period's equilibrium error. We say that the long-term interest rate is weakly exogenous in this case. (It cannot be the case that both $\hat{\alpha}_l = 0$ and $\hat{\alpha}_s = 0$ since that would imply both variables are governed by independent random walks as they are not adjusting to any equilibrium relation and thus cannot be cointegrated).

Finally, we can formulate a more general error-correction model by including lagged changes of each interest rate into equations (12) and (13) or, for that matter, equations (10) and (11):

$$\Delta r_t^s = \alpha_s(e_{t-1}) + \sum_{i=1}^{p} a_{11,i}\Delta r_{t-i}^s + \sum_{i=1}^{q} a_{12,i}\Delta r_{t-i}^l + \eta_{st} \tag{14}$$

$$\Delta r_t^l = -\alpha_l(e_{t-1}) + \sum_{i=1}^{m} a_{21,i}\Delta r_{t-i}^s + \sum_{i=1}^{n} a_{22,i}\Delta r_{t-i}^l + \eta_{lt} \tag{15}$$

Equations (14) and (15) can be estimated consistently by OLS and standard statistical inference on the coefficient estimates can be performed because all the variables in the error-correction model are stationary. The choice of the appropriate number of lags to use in each case can be determined by some lag length criteria or by dropping insignificant lags from the estimated equations.

To summarize, suppose we have two time series Y_t and X_t. First test to see if both series are I(1) processes by using, say, an ADF test. If both series are found to be I(1), there is the possibility that they share a common unit root component in which case they are cointegrated. Second, test for cointegration using the Engle-Granger procedure by estimating a regression of the form

$$Y_t = a_0 + \beta X_t + \varepsilon_t$$

and save the estimated residuals. Denote the OLS parameter estimates as $\hat{a}_0$ and $\hat{\beta}$ and the estimated residuals as e_t. Perform an ADF test (without a constant) on the estimated residuals. If the estimated residuals are found to be stationary, conclude that the two series are cointegrated. The cointegrating vector is $(1 \quad -\hat{\beta})$. The long-run equilibrium relationship is:

$$Y_t = \hat{a}_0 + \hat{\beta} X_t$$

and the deviation from long-run equilibrium is:

$$e_t = Y_t - \hat{a}_0 - \hat{\beta} X_t$$

Third, there exists an error-correction representation which shows how the variables adjust to last period's deviation from the long-run cointegrating relationship. Specifically, the error correction model is of the form:

$$\Delta Y_t = k_1 + \alpha_y(e_{t-1}) + \sum_{i=1}^{p} a_{11,i}\Delta X_t + \sum_{i=1}^{q} a_{12,i}\Delta Y_t + \eta_{yt} \tag{16}$$

$$\Delta X_t = k_2 + \alpha_x(e_{t-1}) + \sum_{i=1}^{m} a_{21,i}\Delta X_{t-1} + \sum_{i=1}^{n} a_{22,i}\Delta Y_{t-1} + \eta_{xt} \tag{17}$$

The two equations in the error-correction model can be estimated consistently by OLS. Standard statistical inference on the coefficient estimates can be performed because all the variables that appear in equations (16) and (17) are stationary. Finally, if the OLS estimate of α_y is zero, Y_t is said to be weakly exogenous as it does not adjust to deviations from the long-run equilibrium relationship. All of the adjustment to equilibrium occurs through changes in X_t since $\alpha_x \neq 0$. (If both α_y and α_x were estimated to be zero, the long-run equibrium relationship does not appear and the model is not one of error correction or cointegration). Similarly, if $\alpha_x = 0$ and $\alpha_y \neq 0$, X_t is weakly exogenous.

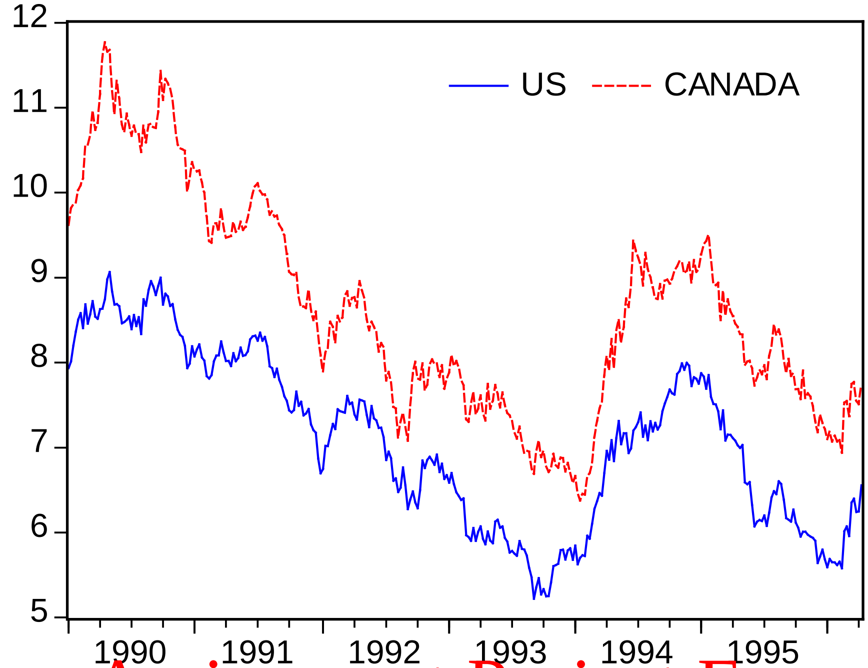
Note that if the data for Y_t and X_t are generated according to equations (7a) and (7b), respectively, Y_t and X_t are cointegrated and X_t would appear in the error correction model as weakly exogenous ($\hat{\alpha}_x = 0$) since Y_t moves with X_t but not conversely.

5. An Application

Figure 1 shows a graph of the yield to maturity on 10-year U.S. and Canadian government bonds. The data are weekly and cover the period from the 2nd of January 1990 to the 9th of April 1996. Both series appear to wander suggesting that both are I(1) processes. However, they seem to wander together suggesting that the series share a common stochastic trend so that they are cointegrated.

The first step is to perform a unit root test on each yield to see whether they are I(1) processes. For the U.S., the ADF test statistic was –1.8066 with a p-value of 0.3770. (In the ADF regression, an intercept but no time trend was included and the number of lagged changes of the interest rate on the right-hand side of the regression was four). Thus, we cannot reject the null of a unit root in the U.S. 10-year government bond yield. For Canada, the ADF test statistic was –1.6087 with a p-value of 0.4770. (The ADF regression also contained an intercept but no time trend and four lagged changes of the interest rate on the right-hand side). Thus, the null of a unit root in the Canadian interest rate cannot be rejected.

Figure 1: U.S. and Canadian 10-year Government Bond Yields



The second step is to test for cointegration between the two I(1) interest rates using the Engle-Granger procedure. We estimate the following regression by OLS and test whether the estimated residuals are stationary, that is, whether they are I(0). The results of estimating the regression

$$r_t^{can} = a + \beta r_t^{us} + \varepsilon_t \tag{18}$$

are shown in Table 2 where RCAN denotes r_t^{can} and RUS denotes r_t^{us}.

Table 2: OLS Estimates of Regression Equation

Dependent Variable: RCAN
Method: Least Squares
Sample: 1/02/1990 4/09/1996
Included observations: 328

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.09599	0.142373	-0.67424	0.5006
RUS	1.229138	0.01986	61.88914	0.0000

R-squared	0.921564	Mean dependent var	8.629399
Adjusted R-squared	0.921324	S.D. dependent var	1.280855
S.E. of regression	0.359271	Akaike info criterion	0.7966
Sum squared resid	42.0787	Schwarz criterion	0.819728
Log likelihood	-128.642	F-statistic	3830.265
Durbin-Watson stat	0.133664	Prob(F-statistic)	0.0000

It appears that the Canadian interest rate moves more than one-for-one with the U.S. rate as $\hat{\beta} = 1.23$.

In EViews, the residuals from this regression were saved as the series *resid01*. Table 3 reports the results of the ADF test on the residuals (*resid01*). In the ADF regression, an intercept was not included because the residuals have a mean of zero by construction (as they come from a regression equation). The ADF t-statistic is –3.329. According to Table 1, the 5% critical value is –3.368 for a sample of 200 observations, which is closest to the sample size of 327 which applies here. Thus the ADF t-statistic is, for practical purposes, statistically significant at the 5% level. (It is statistically significant at the 10% level as the critical value from Table 1 is –3.067). We conclude that the null hypothesis of no cointegration can be rejected at approximately the 5% level. Thus, there appears to be a cointegrating relationship between the two interest rates and that relationship is characterized by the estimated cointegrating vector $\beta' = (1 \quad -1.23)$.

Table 3: Adjusted Dickey-Fuller Test on Residuals

Null Hypothesis: RESID01 has a unit root				
Exogenous: None				
Lag Length: 0 (Automatic based on AIC, MAXLAG=16)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.32914	0.0009
Test critical values:	1% level		-2.57205	
	5% level		-1.9418	
	10% level		-1.61605	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(RESID01)				
Method: Least Squares				
Sample (adjusted): 1/09/1990 4/09/1996				
Included observations: 327 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID01(-1)	-0.06633	0.019925	-3.32914	0.001
R-squared	0.032861	Mean dependent var		-0.00057
Adjusted R-squared	0.032861	S.D. dependent var		0.131349
S.E. of regression	0.129172	Akaike info criterion		-1.25228
Sum squared resid	5.439477	Schwarz criterion		-1.24069
Log likelihood	205.7484	Durbin-Watson stat		1.905795

Having established that the evidence favours cointegration, an error-correction model is estimated for the change in each interest rate series. The error-correction model for each interest rate series exists by the Granger Representation theorem. In each error-correction equation, a lag length of four was initially chosen for both lagged interest rate changes. The OLS estimates of the error-correction equation for the U.S. are shown in Table 4. (Note that DRUS and DRCAN denote the first difference of the U.S. and Canadian interest rates, respectively). For the U.S., the lagged error-correction coefficient (the coefficient on the one-period lag of *resid01*) is statistically insignificant. This indicates that the US interest rate does not adjust to last period's deviation from the long-run equilibrium or cointegrating relationship. In this case, the U.S. rate is said to be weakly exogenous. This means that the adjustment to restore long-run equilibrium following a deviation from it last period must take place through adjustments in the Canadian interest rate. Note also that the change in the U.S. rate is influenced by the change in the U.S. rate the previous period and by the three-period lagged change in the Canadian rate, as both are statistically significant. These reflect how the U.S. rate adjusts in the short-run.

Table 4: Estimates of the Error-Correction Model for the U.S.

Dependent Variable: DRUS
Method: Least Squares
Sample (adjusted): 2/06/1990 4/09/1996
Included observations: 323 after adjustments
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00528	0.007692	-0.68687	0.4927
RESID01(-1)	0.006206	0.023264	0.266746	0.7898
DRUS(-1)	-0.19797	0.080688	-2.45347	0.0147
DRUS(-2)	-0.00601	0.07925	-0.07577	0.9397
DRUS(-3)	0.032939	0.087539	0.376275	0.707
DRUS(-4)	0.116009	0.08124	1.427981	0.1543
DRCAN(-1)	0.020198	0.067421	0.299577	0.7647
DRCAN(-2)	0.021274	0.0623	0.341479	0.733
DRCAN(-3)	0.168386	0.059074	2.850419	0.0047
DRCAN(-4)	-0.00029	0.061107	-0.00473	0.9962

R-squared	0.087686	Mean dependent var	-0.00604
Adjusted R-squared	0.061454	S.D. dependent var	0.142604
S.E. of regression	0.138153	Akaike info criterion	-1.09044
Sum squared resid	5.973979	Schwarz criterion	-0.97349
Log likelihood	186.1065	F-statistic	3.34264
Durbin-Watson stat	1.990908	Prob(F-statistic)	0.000637

The OLS estimates of the error-correction equation for Canada are shown in Table 5. By contrast, the lagged error-correction term (that is, the first period lag of *resid01*) is statistically significant. If it were not, given that it is statistically insignificant in the U.S. equation, then the model is not one of cointegration. For a two-variable setup, in at least one error-correction equation, the one-period lagged error-correction term must be statistically significant; otherwise there is not a cointegrating relationship between the two variables. Let us write out the first part of this equation in detail:

$$\Delta r_t^{can} = -0.064(r_{t-1}^{can} - 1.229 r_{t-1}^{us} + 0.096) + other\ terms \tag{19}$$

The expression in brackets is the deviation from long-run equilibrium last period. If this deviation is positive, the Canadian interest rate adjusts downwards this period because the coefficient on the error-correction term (or the coefficient on the deviation from equilibrium) is negative: it is –0.064. Suppose the deviation from long-run equilibrium is positive because last period the Canadian interest rate was high relative to the U.S. rate. The U.S. rate does not adjust to this disequilibrium because it is weakly exogenous. However, the Canadian rate adjusts this and subsequent periods, and according to equation (19), it decreases until equilibrium is restored, namely, until $r_t^{can} - 1.229 r_t^{us} + 0.096 = 0$ for some future *t*, all else unchanged.

Table 5: Estimates of the Error-Correction Model for Canada

Dependent Variable: DRCAN
Method: Least Squares
Sample (adjusted): 2/06/1990 4/09/1996
Included observations: 323 after adjustments
White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00681	0.009996	-0.6811	0.4963
RESID01(-1)	-0.06438	0.03227	-1.9951	0.0469
DRUS(-1)	-0.34573	0.103049	-3.35501	0.0009
DRUS(-2)	-0.20027	0.11548	-1.73426	0.0839
DRUS(-3)	0.059842	0.108565	0.551207	0.5819
DRUS(-4)	0.253656	0.109759	2.311021	0.0215
DRCAN(-1)	0.03981	0.087219	0.456435	0.6484
DRCAN(-2)	0.106725	0.081667	1.306822	0.1922
DRCAN(-3)	0.170655	0.080884	2.109884	0.0357
DRCAN(-4)	-0.00609	0.080609	-0.07557	0.9398

R-squared	0.117062	Mean dependent var	-0.00709
Adjusted R-squared	0.091674	S.D. dependent var	0.188659
S.E. of regression	0.179804	Akaike info criterion	-0.56343
Sum squared resid	10.1191	Schwarz criterion	-0.44648
Log likelihood	100.9941	F-statistic	4.610925
Durbin-Watson stat	2.017559	Prob(F-statistic)	0.00001

Suppose, on the other hand, the deviation from long-run equilibrium last period is positive because the U.S. rate is high relative to the Canadian rate. Again, the U.S. rate does not adjust to the disequilibrium last period because it is weakly exogenous. The Canadian rate increases, however, to restore equilibrium. To see this, suppose the U.S. rate last period increased by 1% but the Canadian rate was unchanged. According to equation (19), the Canadian rate this period will increase by –0.064(–1.229(1%)) = 0.0787%. The adjustment is somewhat slow because the speed of adjustment coefficient of –0.064 is small. Thus, the Canadian rate will increase over time to restore the cointegrating relationship between the U.S. and Canadian rates.

Note that, in the short-run, changes in the Canadian interest rate are influenced by the first and fourth lagged change in the U.S. rate and also by the third lagged change in the Canadian rate as each is statistically significant at the 5% level.

Finally, to obtain a more parsimonious error-correction equation for the change in the U.S. and Canadian rates, respectively, one can delete the insignificant variables on the right-hand side of each equation and then re-estimate the equation.