Chapter 6: Cointegration

Financial Econometric Modeling

Stan Hurn, Vance Martin, Peter C.B. Phillips and Jun Yu

Oxford University Press, 2020

Introduction

- An important implication of the discussion of stochastic trends and unit roots is that a nonstationary time series can be rendered stationary by differencing.
- An alternative method to achieve stationarity is to form linear combinations of nonstationary series to generate a stationary series.
- The ability to generate stationary time series as a linear combination of nonstationary time series is known as cointegration and the weights applied to each of the series in the combination are known as the cointegrating parameters.

Importance

The existence of cointegration between nonstationary time series has important theoretical, statistical and dynamic implications.

(i) A number of theoretical models in finance can be couched within a cointegrating framework.

(ii) Estimates of the parameters in the cointegrating equations converge to their population values at a rate faster than is the case for stationary variables, a property known as super-consistency.

(iii) Modelling a system of cointegrated variables allows for the joint specification of long-run and short-run dynamics of financial variables in terms of the VECM.

One important model is the present value model.

Present Value Model

US Equity prices

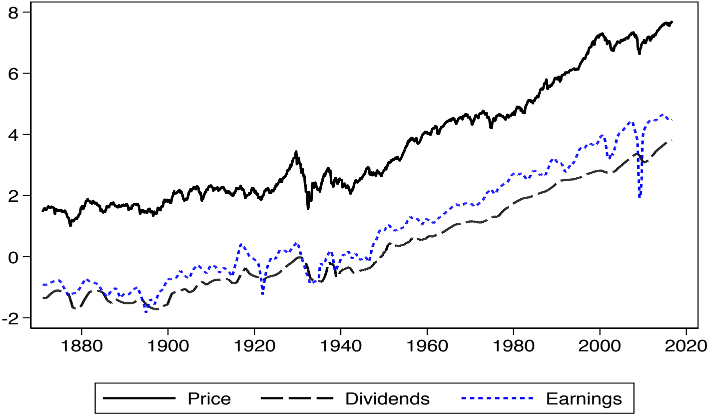


Figure: Time series plots of the logarithms of monthly United States real equity prices, real dividends and real earnings per share for the period February 1871 to September 2016.

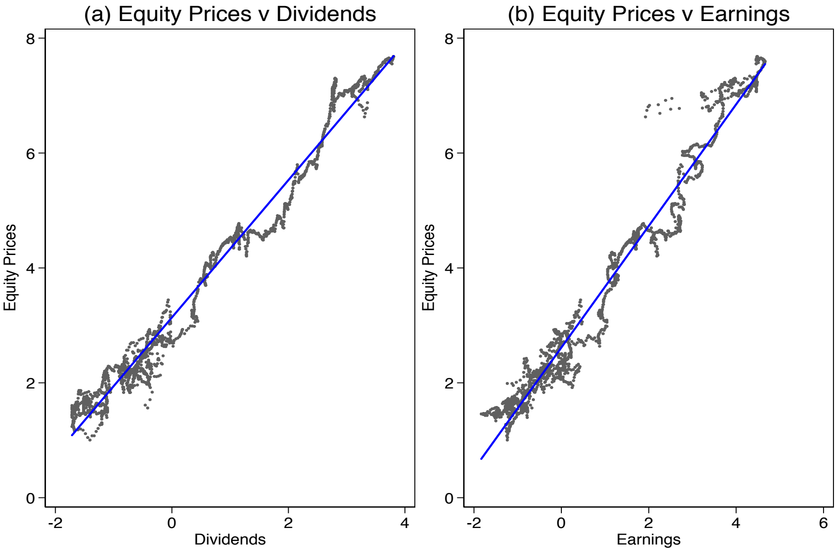
Present value model

- Even though equity prices and dividends are potentially $I(1)$, the present value model of equity prices suggests a theoretical link between the behaviour of equity prices and dividends.
- This relationship is expressed as

$$p_t = \beta_0 + \beta_d d_t + u_t \,,$$

  where p_t is the log equity price, d_t is the log dividend, u_t is a disturbance term and β_0 and β_d are unknown parameters.
- Both p_t and d_t are well-modelled as $I(1)$ processes, but the present value model indicates that the differences $p_t - \beta_0 - \beta_d d_t = u_t$ are simply transient shocks that do not disturb the nature of the relationship over time. The linkage between p_t and d_t is therefore regarded as a long run (or permanent) relationship between trending $I(1)$ series and the residuals u_t are viewed as transient shocks or $I(0)$.
- The linear combination of the $I(1)$ variables p_t and d_t which results in the $I(0)$ variable given by u_t is known as a cointegrating relationship or simply as cointegration. The reduction of the trending $I(1)$ character of p_t and d_t to the transient $I(0)$ character of u_t is the essential condition of cointegration. When these conditions hold, then the present value equation is known as a cointegrating equation and the parameters β_0 and β_d are the cointegrating coefficients.

An alternative view – scatter plot



Interpretation

- Superimposed on the scatter plot is an estimate of the present value model

$$p_t = 3.1375 + 1.1957\, d_t + \widehat{u}_t,$$

  obtained by regressing p_t on a constant and d_t.
- Even though p_t and d_t are both nonstationary, prices and dividends never deviate too far away from the line given by $3.1375 + 1.1957\, d_t$, suggesting that this is a (dynamic) equilibrium line.
- If there was no cointegration the scatter diagram would have points evenly scattered over 2-dimensions. In other words, cointegration has the effect of compressing equity prices and dividends from 2 dimensions to 1.

Equilibrium Adjustment and Error Correction Dynamics

Equilibrium

- This interpretation of the scatter diagram also suggests that actual movements in p_t can be decomposed into a long-run component representing the (dynamic) equilibrium price as determined by dividends and a short-run component, representing the temporary deviations of p_t from its long-run.
- This decomposition is expressed as

$$\underbrace{p_t}_{\text{Actual}} = \underbrace{\beta_0 + \beta_d d_t}_{\text{Long-run}} + \underbrace{u_t}_{\text{Short-run}}$$

Equilibrium adjustment

- To understand the equilibrating mechanisms of the present value model consider the effects of a shock at time t , as given by u_t, assuming that dividends d_t are not affected initially.
  1. For a positive shock, $u_t > 0$, equities are overvalued relative to the long-run level with the stock price p_t operating above its long-run equilibrium price
  2. Fo a negative shock, $u_t < 0$, equities are undervalued with p_t operating below the long-run price.
- An equilibrium relationship implies that any shock to the system will result in an adjustment taking place in such a way that equilibrium is eventually restored

Equilibrium adjustment

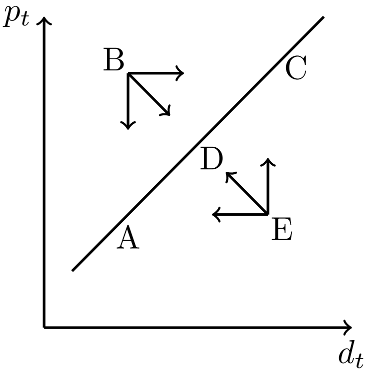


Figure: Phase diagram to demonstrate the equilibrium adjustment if two variables are cointegrated.

Potentical scenarios for a positive shock to u_t

(i) **Equity prices adjust**

$$p_{t+1} - p_t = \delta_1 + \alpha_1(p_t - \beta_0 - \beta_d d_t) + v_{1t+1}.$$

Given that equity prices need to adjust downwards in this scenario to restore equilibrium, the adjustment parameter α_1 satisfies the restriction $\alpha_1 < 0$.

(ii) **Dividends adjust**

$$\Delta d_{t+1} = \delta_2 + \alpha_2(p_t - \beta_0 - \beta_d d_t) + v_{2t+1}.$$

Given that dividends need to increase to restore equilibrium in this scenario, the adjustment parameter α_2 satisfies the restriction $\alpha_2 > 0$.

(iii) **Equity prices and dividends adjust**
In this scenario p_t decreases and d_t increases. The relative strength of the movements in equity prices and dividends is determined by the relative magnitudes of the adjustment parameters α_1 and α_2.

Present value model in the US

- To gauge the relative strength of the adjustment parameters, α_1 and α_2, in restoring equilibrium in the present value model for the United States, the equations for prices and dividend adjustment are estimated using US data.
- Letting $\widehat{u}_{t-1}$ represent the ordinary least squares residuals from the cointegrating regression, the parameter estimates of the error correction model are

$$\Delta p_t = 0.0035 - 0.0011\, \widehat{u}_{t-1} + \widehat{v}_{1t}$$
$$\Delta d_t = 0.0029 + 0.0078\, \widehat{u}_{t-1} + \widehat{v}_{2t}\,.$$

- The signs of the adjustment parameter estimates are consistent with the dynamics. Moreover, since

$$|\widehat{\alpha}_2| = 0.0078 \; > |\widehat{\alpha}_1| = 0.0011,$$

dividends are the main driving force in restoring equilibrium in the system after a shock to the equity price.

Error correction dynamics

- Consider the system

$$\Delta p_t = \delta_1 + \alpha_1(p_{t-1} - \beta_0 - \beta_d d_{t-1}) + v_{1t}$$
$$\Delta d_t = \delta_2 + \alpha_2(p_{t-1} - \beta_0 - \beta_d d_{t-1}) + v_{2t}.$$

- The long run relationship

$$u_{t-1} = p_{t-1} - \beta_0 - \beta_d d_{t-1}$$

  is common to both equations.

- We can therefore write the system as

$$\Delta p_t = \delta_1 + \alpha_1(u_{t-1}) + v_{1t}$$
$$\Delta d_t = \delta_2 + \alpha_2(u_{t-1}) + v_{2t}.$$

- This form is known as the error correction model because the system behaves in such a way as to correct the equilibrium errors, u_{t-1}.

Vector error correction models (VECM)

1. In general, the simplest bivariate vector error correction model or VECM is given by

$$\begin{aligned}\Delta y_{1t} &= \delta_1 + \alpha_1(y_{1t-1} - \beta_0 - \beta_2 y_{2t-1}) + v_{1t} \\ \Delta y_{2t} &= \delta_2 + \alpha_2(y_{1t-1} - \beta_0 - \beta_2 y_{2t-1}) + v_{2t},\end{aligned} \tag{1}$$

2. To allow the bivariate VECM to have additional short-run dynamics the model may be respecified as

$$\begin{aligned}\Delta y_{1t} = \delta_1 + \alpha_1(y_{1t-1} - \beta_0 - \beta_2 y_{2t-1}) \\ + \sum_{i=1}^{k-1} \gamma_{11i}\Delta y_{1t-i} + \sum_{i=1}^{k-1} \gamma_{12i}\Delta y_{2t-i} + v_{1t} \\ \Delta y_{2t} = \delta_2 + \alpha_2(y_{1t-1} - \beta_0 - \beta_2 y_{2t-1}) \\ + \sum_{i=1}^{k-1} \gamma_{21i}\Delta y_{1t-i} + \sum_{i=1}^{k-1} \gamma_{22i}\Delta y_{2t-i} + v_{2t},\end{aligned}$$

where k controls the length of the lag structure in the transient dynamics.

VECM in matrix notation

The extensions discussed above are conveniently combined by expressing the VECM in matrix notation. Define the following $(N \times 1)$ vectors

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Nt} \end{bmatrix} \; \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} \; \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \; \beta = \begin{bmatrix} 1 \\ -\beta_2 \\ \vdots \\ -\beta_N \end{bmatrix} \; v_t = \begin{bmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{Nt} \end{bmatrix},$$

so that the VECM can be written as

$$\Delta y_t = \delta + \alpha(\beta' y_{t-1} - \beta_0) + v_t,$$

where β_0 is a scalar. To capture the proposed extensions to the VECM, a general specification is given by

$$\Delta y_t = \delta + \alpha(\beta' y_{t-1} - \beta_0 - \phi t) + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \Psi x_t + v_t,$$

where ϕ is a scalar, Γ_i is a $(N \times N)$ matrix of parameters at lag i, and Ψ is a $(N \times K)$ matrix of parameters associated with the K stationary variables contained in x_t.

Relationship with VARs

Consider the bivariate VECM

$$\begin{aligned} y_{1t} - y_{1t-1} &= \alpha_1(y_{1t-1} - \beta_2 y_{2t-1}) + v_{1t} \\ y_{2t} - y_{2t-1} &= \alpha_2(y_{1t-1} - \beta_2 y_{2t-1}) + v_{2t}, \end{aligned}$$

Now re-express each equation in terms of the levels of the variables as

$$\begin{aligned} y_{1t} &= (1 + \alpha_1)y_{1t-1} - \alpha_1\beta_2 y_{2t-1} + v_{1t} \\ y_{2t} &= \alpha_2 y_{1t-1} + (1 - \alpha_2\beta_2)y_{2t-1} + v_{2t}, \end{aligned}$$

or

$$\begin{aligned} y_{1t} &= \phi_{11} y_{1t-1} + \phi_{12} y_{2t-1} + v_{1t} \\ y_{2t} &= \phi_{21} y_{1t-1} + \phi_{22} y_{2t-1} + v_{2t}. \end{aligned}$$

The parameters of the VECM are related to those in the VAR by the restrictions

$$\phi_{11} = 1 + \alpha_1, \quad \phi_{12} = -\alpha_1\beta_2, \quad \phi_{21} = \alpha_2, \quad \phi_{22} = 1 - \alpha_2\beta_2.$$

Three important details concerning VARs and VECMs

(i) The VECM is a restricted VAR containing 3 parameters whereas the unrestricted VAR contains 4 parameters. This difference in the number of unknown parameters is encapsulated by the set of restrictions which arise from the y_{it} variables jointly having the same long-run equation.

(ii) The VAR in levels is an appropriate general specification because it allows for both the VECM system by way of specialisation of the system as well as cases where there is no cointegration. But when there is no cointegration between y_{1t} and y_{2t}, the adjustment parameters $\alpha_1 = \alpha_2 = 0$ and there are no equilibrating forces present to return the system towards a long-run equilibrium. Imposing the restrictions $\alpha_1 = \alpha_2 = 0$ reduces the model to a simple VAR in first differences.

(iii) For a VARs with k lagged dependent variables the corresponding VECM contains only $k - 1$ additional lagged dependent variables, all specified in terms of differences. This means that the information criteria used to determine the optimal lag structure of a VAR are also valid for determining the optimal lag length of a VECM.

The possible relationships between VECM and VAR models may be summarised using a bivariate VAR(1) model expressed as

$$\Delta y_t = \Phi y_{t-1} + v_t .$$

Three versions of this model are possible, each depending on the rank of the Φ.

Full rank case ($r = 2$)

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} .$$

Reduced rank case ($r = 1$)

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} [1, -\beta_2] \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} .$$

In this case there is a cointegrating linear combination $y_{1t-1} - \beta_2 y_{2t-1} = \beta' y_{t-1}$ between the two variables in levels.

Zero rank case ($r = 0$)

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} .$$

In this case the model is a VAR in first differences, each variable is $I(1)$.

Estimating the Parameters of VECMs

Two Estimation Methods

- Two methods for unknown parameters of the VECM are discussed, namely,
  1. the fully modified (FM-OLS) regression-based estimator proposed by Phillips and Hansen (1990); and
  2. the Johansen reduced rank regression estimator (Johansen, 1988)
- These estimators both possess two important properties.
  1. They are consistent for the true population parameters at the rate T, which is considerably faster than the $\sqrt{T}$ rate at which coefficients are estimated in stationary time series models. This faster rate of convergence is sometimes referred to as super-consistency. It arises because the variables $\{y_{1t}, y_{2t}, \cdots, y_{Nt}\}$ are nonstationary $I(1)$ series, whose stochastically trending nature facilitates estimation of the cointegrating relationships.
  2. The asymptotic distributions of these two estimators of the cointegrating parameters can be used to conduct hypothesis tests in much the same way as for models based on stationary variables. This important feature is not true for all cointegrating regression estimators.

FM-OLS

- Consider the present value model

$$\begin{aligned} p_t &= \beta_0 + \beta_d d_t + u_t \\ \Delta d_t &= v_t. \end{aligned} \tag{2}$$

  The first equation is the cointegrating equation while the second equation makes explicit the unit root stochastic trend property of dividends. The disturbances u_t and v_t are $I(0)$: the former because of cointegration and the latter because dividends are assumed to be difference stationary.
- Estimating cointegrating equation by ordinary least squares delivers super-consistent parameter estimates, inferences based on these estimates are generally invalid for several reasons:
  (i) d_t is an endogenous regressor;
  (ii) u_t and v_t are correlated;
  (iii) u_t and v_t are autocorrelated.
- The FM estimator modifies least squares by taking into account these general features of the generating mechanism of the data. The endogeneity is addressed by making a correction for endogeneity and the autocorrelation is addressed by means of a serial correlation correction.

Implementing FM-OLS

1. Estimate the cointegrating equation by ordinary least squares to obtain $\widehat{\beta}_0$ and $\widehat{\beta}_d$ and the residuals $\widehat{u}_t$, as well as $\widehat{v}_t = v_t = \Delta d_t$.
2. Compute the long-run covariance matrices

$$\begin{aligned} \Omega &= \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} = \sum_{j=-\infty}^{\infty} \Gamma_j \qquad \text{[2-sided]} \\ \Lambda &= \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \sum_{j=0}^{\infty} \Gamma_j, \qquad \text{[1-sided]} \end{aligned} \tag{3}$$

   where $\Gamma_j = E(w_t w'_{t-j})$ with $w_t = (u_t, v_t)'$ and with the lag length m selected according to some fixed value or data-determined according to some optimal selection criterion.
3. The FM estimator has the following explicit form

$$\widehat{\beta}_{FM} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_d \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^{T} d_t \\ \sum_{t=1}^{T} d_t & \sum_{t=1}^{T} d_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} p_t^+ \\ \sum_{t=1}^{T} (d_t p_t^+ - \widehat{c}) \end{bmatrix}, \tag{4}$$

   where $p_t^+ = p_t - \widehat{\rho}\Delta d_t$, $\widehat{\rho} = \widehat{\omega}_{12}\widehat{\omega}_{22}^{-1}$ and $\widehat{c} = \widehat{\lambda}_{12} - \widehat{\rho}\widehat{\lambda}_{22}$

Some intuition

There are two correction terms that modify the usual ordinary least squares formulae and correct for endogeneity and serial dependence .

1. The estimator $\widehat{\rho} = \widehat{\omega}_{12}\widehat{\omega}_{22}^{-1}$ is an estimator of the long-run regression coefficient $\omega_{12}\omega_{22}^{-1}$ of u_t on v_t. The correction term $p_t^+ = p_t - \widehat{\rho}\Delta d_t$ then adjusts the dependent variable p_t for its long-run joint dependence on d_t. This is the endogeneity correction.
2. The quantity $\widehat{c} = \widehat{\lambda}_{12} - \widehat{\rho}\widehat{\lambda}_{22}$ provides an adjustment that compensates for the contemporaneous and serial dependence of the price equation errors u_t with the errors $\Delta d_t = v_t$ in the dividend equation. This adjustment leads to the modification of sample moment $\sum_1^T (d_t p_t^+ - \widehat{c})$ that appears in the final member of the right side of the FM-OLS estimator.

Combining the adjustments (i) and (ii) leads to the fully modified version of least squares regression.

U.S. data on equity prices and dividends

- OLS estimates of the cointegrating regression

$$p_t = 3.1375 + 1.1957\, d_t + \widehat{u}_t$$

- FM-OLS estimates of the cointegrating regression

$$p_t = \underset{(0.038)}{3.127} + \underset{(0.023)}{1.195}\, d_t + \widehat{u}_t\,.$$

  The long-run covariance matrices are computed using a lag length of $m = 8$.
- The FM-OLS parameter estimates are qualitatively very similar to the OLS parameter estimates which reflects the fact that both estimators generate super-consistent parameter estimates. However, the standard errors associated with the FM-OLS estimates are robust to endogeneity and serial dependence by virtue of their construction.

Johansen's approach

- The Johansen approach yields estimates for all of the parameters in an explicit parametric form of the VECM, including both the long-run cointegrating parameters and the short-run error-correction parameters.
- This is in contrast to the semi-parametric regression-based FM estimator which yields estimates of only the long-run parameters because the transient dynamics are treated nonparametrically by FM-OLS.
- Consider VECM of equity prices and dividends with the short-run dynamics captured by an additional lagged difference in the two variables

$$\begin{aligned} p_t &= \beta_0 + \beta_d d_t + u_t \\ \Delta p_t &= \delta_1 + \alpha_1 (p_{t-1} - \beta_0 - \beta_d d_{t-1}) + \gamma_{11} \Delta p_{t-1} + \gamma_{12} \Delta d_{t-1} + v_{1t} \\ \Delta d_t &= \delta_2 + \alpha_2 (p_{t-1} - \beta_0 - \beta_d d_{t-1}) + \gamma_{21} \Delta p_{t-1} + \gamma_{22} \Delta d_{t-1} + v_{2t} \,, \end{aligned}$$

  where $v_t = (v_{1t}, v_{2t})'$ is a (2×1) vector of the VECM disturbances.
- Johansen develops a computationally efficient algorithm to estimate all the parameters simultaneously.

The Johansen estimator

1. Construct the residuals from 4 auxiliary regressions by regressing $\{\Delta p_t, \Delta d_t, p_{t-1}, d_{t-1}\}$ on $\{1, \Delta p_{t-1}, \Delta d_{t-1}\}$.
2. Use the residuals in the previous step to perform an eigen decomposition and compute $\widehat{\beta}_0$ and $\widehat{\beta}_d$ using the eigenvector corresponding to the largest eigenvalue.
3. Construct the cointegrating residuals $\widehat{u}_t = p_t - \widehat{\beta}_0 - \widehat{\beta}_d d_t$ and estimate the VECM by regressing $\{\Delta p_t, \Delta d_t\}$ in turn, on the regressors $\left\{1, \widehat{u}_{t-1}, \Delta p_{t-1}, \Delta d_{t-1}\right\}$.

The Johansen estimator is based on the maximum likelihood principle (see Chapter 10).

U.S. price and dividend data

The estimated cointegrating equation is

$$p_t = 3.4034 + \underset{(0.0303)}{1.1773}\, d_t + \widehat{u}_t.$$

Variable	Δp_t	Δd_t
Constant	0.0004	0.0010
	(0.0013)	(0.0002)
$\widehat{u}_{t-1}$	- 0.0068	0.0024
	(0.0032)	(0.0004)
Δp_{t-1}	0.2901	0.0006
	(0.0231)	(0.0030)
Δd_{t-1}	0.1366	0.8792
	(0.0841)	(0.0110)

Testing in Cointegrated Systems

Residual based tests of cointegration

- Consider the present value model

$$p_t = \beta_0 + \beta_d d_t + u_t,$$

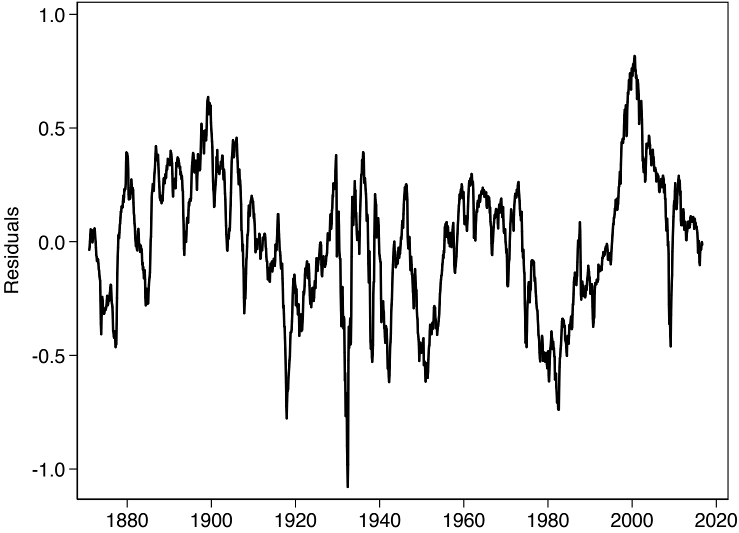
where the log equity price p_t, and log dividend d_t are $I(1)$ and u_t is the disturbance term which will be $I(0)$ when p_t and d_t are cointegrated.

- Anatural way to test for cointegration is a two-step procedure consisting of estimating the cointegrating equationby ordinary least squares in the first step and testing the residuals for stationarity in the second step.
- As the unit root test treats the null hypothesis as nonstationary, the null hypothesis corresponds to no cointegration whereas the alternative hypothesis of stationarity corresponds to cointegration:

$$\begin{array}{lll} H_0 : & \text{No Cointegration} & [u_t \text{ is nonstationary}] \\ H_1 : & \text{Cointegration.} & [u_t \text{ is stationary}] \end{array}$$

- The distribution of the residual-based cointegration test is non-standard, as might be expected because the procedure tests for a unit root null hypothesis.

Residuals from equity prices and dividends cointegrating regression



Tests for cointegration

Residual-based tests of cointegration between United States equity prices and dividends using equation (31). The test regression has no constant term with lags running from 0 to 4.

Lags	Augmented Dickey-Fuller	Phillips-Perron
0	−3.028	−3.028
1	−4.287	−3.493
2	−4.029	−3.680
3	−4.037	−3.770
4	−4.317	−3.849

(i) Critical values are, respectively: 1% −2.580; 5% −1.950; and 10% −1.620.

(ii) In the ADF test, 'Lags' represents the number of lagged differences of the dependent variable included in the regression. In the Phillips-Perron test, 'Lags' represents the number of Newey-West lags employed in the estimates of the nonparametric components of the test statistic.

Johansen tests of cointegration

Consider again the VECM

$$\Delta p_t = \delta_1 + \alpha_1(p_{t-1} - \beta_0 + \beta_d d_{t-1}) + \gamma_{11}\Delta p_{t-1} + \gamma_{12}\Delta d_{t-1} + v_{1t}$$
$$\Delta d_t = \delta_2 + \alpha_2(p_{t-1} - \beta_0 + \beta_d d_{t-1}) + \gamma_{21}\Delta p_{t-1} + \gamma_{22}\Delta d_{t-1} + v_{2t}.$$

The Johansen method is based on testing a sequence of different hypotheses.

1. All variables are nonstationary, that is p_t and d_t are $I(1)$ and the error correction term $u_t = p_t - \beta_0 + \beta_d d_t$ is also $I(1)$. The null and alternative hypotheses are:

   Stage 1 $\quad H_0$: No cointegration, all variables are $I(1)$
   $\quad H_1$: 1 or more cointegrating equations

2. Tests the restrictions arising from $r = 1$ so that p_t and d_t are $I(1)$ but and the error correction term $u_t = p_t - \beta_0 + \beta_d d_t$ is $I(0)$.

   Stage 2 $\quad H_0$: 1 cointegrating equation
   $\quad H_1$: All variables are stationary $I(0)$

3. Tests the restrictions arising from $r = 2$ so that all variables are stationary.

Trace test

- The test statistic is computed from the estimated eigenvalues of the Johansen estimator of the VECM, $\widehat{\lambda}_i, i = 1, 2, \cdots, N$, ordered from highest to lowest.
- For an N-dimensional system the form of the statistic is

$$TRACE = -(T - k) \sum_{i=r+1}^{N} \log(1 - \widehat{\lambda}_i).$$

- The test statistic is called the trace test because the trace of a matrix is determined by the number of non-zero eigenvalues. Large values of the trace statistic relative to the critical value result in rejection of the null hypothesis.
- The subscript r represents the rank of the system. Under the null hypothesis in Stage 1 the rank of the system is $r = 0$ with the statistic representing a joint test that all N eigenvalues are zero. The null hypothesis in Stage 2 is that $r = 1$. The statistic in this case is a test that the smallest eigenvalue is zero.
- For higher dimensional VECMs with $N > 2$, there may be more than a single cointegrating long-run equation connecting the nonstationary variables up to a possible maximum of $N - 1$ cointegrating equations.

Trace tests and U.S. prices and dividends

Johansen test of cointegration between United States equity prices and dividends.

Null Hypothesis	Eigenvalue	Statistic	5% CV	p value
$r = 0$: No cointegration	0.0207	37.9330	15.41	0.000
$r = 1$: 1 cointegrating equation	0.0008	1.3910	3.76	0.238

This sequence of tests provides strong support for the present value model because cointegration between equity prices and dividends is confirmed, thereby complementing the results obtained using the residual-based tests for cointegration.

Parameter testing

- When there is cointegration in the system, the FM-OLS and reduced-rank regression estimators deliver super consistent parameter estimates of the cointegrating parameters associated with the $I(1)$ variables together with standard errors that can be used to construct asymptotically valid t tests concerning these parameters.
- In testing the parameters of a VECM three broad sets of hypotheses can be explored:
  1. hypotheses relating to the cointegrating parameters;
  2. hypotheses relating to the equilibrium adjustment parameters; and
  3. hypotheses relating to the transient dynamic parameters.
- The distributions of these t tests are, in fact, asymptotically normally distributed as $N(0, 1)$ despite the component variables being nonstationary. This property arises because when it is correctly specified the VECM system is expressed in stationary variable form involving either first differences of $I(1)$ variables or long-run equilibrium errors such as u_{t-1}.