Chapter 13

Introduction to nonstationary time series

Overview

This chapter begins by defining the concepts of stationarity and nonstationarity as applied to univariate time series and, in the case of nonstationary series, the concepts of difference-stationarity and trend-stationarity. It next describes the consequences of nonstationarity for models fitted using nonstationary time-series data and gives an account of the Granger–Newbold Monte Carlo experiment with random walks. Next the two main methods of detecting nonstationarity in time series are described, the graphical approach using correlograms and the more formal approach using Augmented Dickey–Fuller unit root tests. This leads to the topic of cointegration. The chapter concludes with a discussion of methods for fitting models using nonstationary time series: detrending, differencing, and error-correction models.

Further material

Generalization of the Augmented Dickey Fuller test for unit roots

In Section 13.3 it was shown how a Dickey–Fuller test could be used to detect a unit root in the process

$$X_t = \beta_1 + \beta_2 X_{t-1} + \gamma t + \varepsilon_t$$

and an Augmented Dickey–Fuller test could be used for the same purpose when the process included an additional lagged value of X:

$$X_t = \beta_1 + \beta_2 X_{t-1} + \beta_3 X_{t-2} + \gamma t + \varepsilon_t$$

In principle the process may have further lags, the general form being

$$X_t = \beta_1 + \sum_{s=1}^{p} \beta_{s+1} X_{t-s} + \gamma t + \varepsilon_t$$

One condition for stationarity is that the sum of the coefficients of the lagged X variables should be less than one. Writing our test statistic θ as

$$\theta = \sum_{s=1}^{p} \beta_{s+1} - 1$$

the null hypothesis of nonstationarity is H_0: $\theta = 0$ and the alternative hypothesis of stationarity is H_1: $\theta < 0$. Two issues now arise. One is how to reparameterize the specification so that we can obtain a direct estimate of θ. The other is how to determine the appropriate value of p.

From the definition of θ, we have

$$\beta_2 = \theta + 1 - \sum_{s=2}^{p} \beta_{s+1}$$

Substituting this into the original specification, we have

$$X_t = \beta_1 + \left(\theta + 1 - \sum_{s=2}^{p} \beta_{s+1}\right) X_{t-1} + \sum_{s=2}^{p} \beta_{s+1} X_{t-s} + \gamma t + \varepsilon_t$$

Hence

$$
\begin{aligned}
X_t - X_{t-1} &= \beta_1 + \theta X_{t-1} - X_{t-1}\sum_{s=2}^{p}\beta_{s+1} + \sum_{s=2}^{p}\beta_{s+1}X_{t-s} + \gamma t + \varepsilon_t \\
&= \beta_1 + \theta X_{t-1} - X_{t-1}\sum_{s=2}^{p}\beta_{s+1} + X_{t-2}\sum_{s=2}^{p}\beta_{s+1} - X_{t-2}\sum_{s=2}^{p}\beta_{s+1} + \sum_{s=2}^{p}\beta_{s+1}X_{t-s} + \gamma t + \varepsilon_t \\
&= \beta_1 + \theta X_{t-1} - \sum_{s=2}^{p}\beta_{s+1}\left(X_{t-1} - X_{t-2}\right) - X_{t-2}\sum_{s=3}^{p}\beta_{s+1} + \sum_{s=3}^{p}\beta_{s+1}X_{t-s} + \gamma t + \varepsilon_t \\
&= \beta_1 + \theta X_{t-1} - \sum_{s=2}^{p}\beta_{s+1}\left(X_{t-1} - X_{t-2}\right) - \sum_{s=3}^{p}\beta_{s+1}\left(X_{t-2} - X_{t-3}\right) - \ldots - \beta_p\left(X_{t-p+1} - X_{t-p}\right) + \gamma t + \varepsilon_t
\end{aligned}
$$

Thus the reparameterized regression model may be written

$$\Delta X_t = \beta_1 + \theta X_{t-1} - \delta_1 \Delta X_{t-1} - \delta_2 \Delta X_{t-2} \ldots - \delta_{p-1}\Delta X_{t-p+1} + \gamma t + \varepsilon_t$$

where

$$\delta_q = \sum_{s=q+1}^{p}\beta_{s+1}$$

and $\Delta X_{t-q} = X_{t-q} - X_{t-q-1}$. The parameter of interest is, of course, the coefficient of X_{t-1}.

There now arises the question of how to determine the appropriate number of lagged values of X in the original specification or, equivalently, of ΔX in the reparameterized specification. Looking directly at the goodness of fit, as measured by R^2 or *RSS*, does not provide an answer. We have seen that R^2 will increase and *RSS* will decrease when additional variables, even irrelevant ones, are included in the regression specification. $\bar{R}^2$, 'adjusted' R^2, discussed in Section 3.5, is one measure of goodness of fit that attempts to allow for this effect, but it is unsatisfactory. Newer measures are the Bayes Information Criterion (BIC) and the Akaike Information Criterion (AIC). The BIC (also known as the Schwarz Information Criterion) and the AIC have become popular for helping to determine the appropriate number of lags in time series analysis in general and unit root tests in particular. Indeed the latest version of EViews includes the BIC/Schwarz as the default option when testing for unit roots.

The BIC and AIC are defined by

$$\text{BIC} = \log\frac{RSS}{T} + \frac{k}{T}\log T$$

and

$$\text{AIC} = \log\frac{RSS}{T} + \frac{2k}{T}$$

where k is the number of parameters in the regression specification. For both information criteria, the optimal regression specification is the one that minimizes the statistic. For both, the first term will decrease when additional terms are included in the regression specification, but the second term will increase. Since $\log T > 2$ for $T > 7$, increasing the number of parameters is penalized more heavily in the BIC than the AIC, with the consequence that in time series analysis the BIC tends to produce specifications with fewer lags. It can be shown that the BIC provides consistent estimates of the lag length, while the AIC does not, but for finite samples neither has an obvious advantage and both are used in practice.

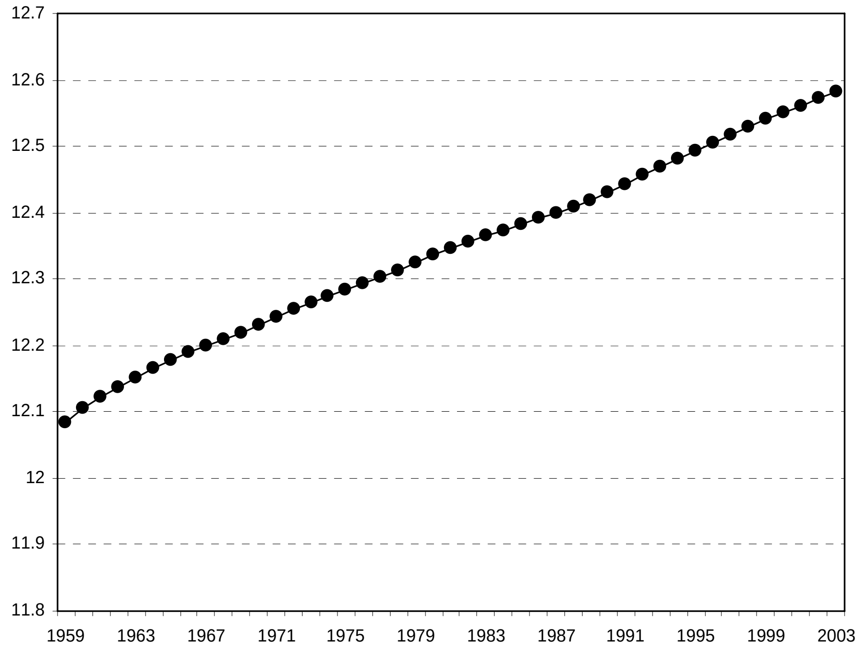
Learning outcomes

After working through the corresponding chapter in the text, studying the corresponding slideshows, and doing the starred exercises in the text and the additional exercises in this guide, you should be able to:

- explain what is meant by stationarity and nonstationarity.
- explain what is meant by a random walk and a random walk with drift
- derive the condition for the stationarity of an AR(1) process
- explain what is meant by an integrated process and its order of integration
- explain why Granger and Newbold obtained the results that they did
- explain what is depicted by a correlogram.
- perform an Augmented Dickey–Fuller unit root test to test a time series for nonstationarity
- test whether a set of time series are cointegrated.
- construct an error-correction model and describe its advantages over detrending and differencing.

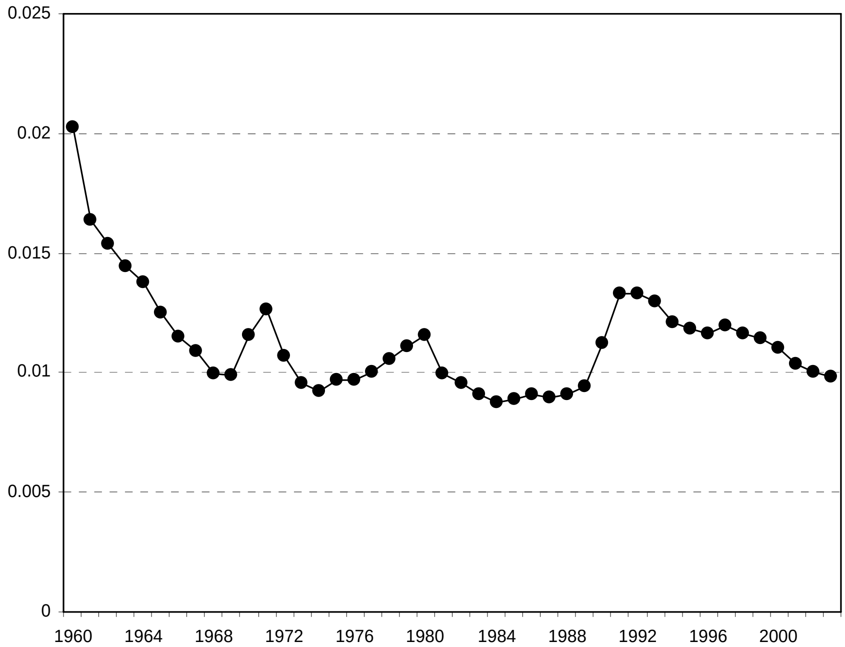
Additional exercises

A13.1 The figure plots the logarithm of the US population for the period 1959–2003. It is obviously nonstationary. Discuss whether it is more likely to be difference-stationary or trend-stationary.



Logarithm of the US population

A13.2 The figure plots the first difference of the logarithm of the US population for the period 1959–2003. Explain why the vertical axis measures the proportional growth rate. Comment on whether the series appears to be stationary or nonstationary.



Logarithm of the US population, first difference

A13.3 The regression output shows the results of ADF unit root tests on the logarithm of the US population, and its difference, for the period 1959–2003. Comment on the results and state whether they confirm or contradict your conclusions in Exercises A13.2.

```
      Augmented Dickey-Fuller Unit Root Test on LGPOP
============================================================
Null Hypothesis: LGPOP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Fixed)
============================================================
                                       t-Statistic  Prob.*
============================================================
Augmented Dickey-Fuller test statistic   -2.030967   0.5682
Test critical values1% level             -4.186481
                    5% level             -3.518090
                   10% level             -3.189732
============================================================
*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LGPOP)
Method: Least Squares
Sample(adjusted): 1961 2003
Included observations: 43 after adjusting endpoints
============================================================
     Variable      Coefficient Std. Error t-Statistic  Prob.
============================================================
     LGPOP(-1)      -0.047182   0.023231  -2.030967   0.0491
   D(LGPOP(-1))      0.687772   0.058979   11.66139   0.0000
         C           0.574028   0.281358   2.040209   0.0481
   @TREND(1959)      0.000507   0.000246   2.060295   0.0461
============================================================
R-squared            0.839263    Mean dependent var 0.011080
Adjusted R-squared   0.826898    S.D. dependent var 0.001804
S.E. of regression   0.000750    Akaike info criter-11.46327
Sum squared resid    2.20E-05    Schwarz criterion -11.29944
Log likelihood       250.4603    F-statistic        67.87724
Durbin-Watson stat   1.164933    Prob(F-statistic)  0.000000
============================================================
```

```
      Augmented Dickey-Fuller Unit Root Test on DLGPOP
============================================================
Null Hypothesis: DLGPOP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Fixed)
============================================================
                                        t-Statistic  Prob.*
============================================================
Augmented Dickey-Fuller test statistic   -2.513668   0.3203
Test critical values1% level             -4.192337
                    5% level             -3.520787
                   10% level             -3.191277
============================================================
*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLGPOP)
Method: Least Squares
Sample(adjusted): 1962 2003
Included observations: 42 after adjusting endpoints
============================================================
     Variable      Coefficient Std. Error t-Statistic  Prob.
============================================================
    DLGPOP(-1)      -0.161563   0.064274  -2.513668   0.0163
   D(DLGPOP(-1))     0.294717   0.117766   2.502573   0.0167
         C           0.001714   0.000796   2.152327   0.0378
   @TREND(1959)     -1.32E-07   9.72E-06  -0.013543   0.9893
============================================================
R-squared            0.320511    Mean dependent var-0.000156
Adjusted R-squared   0.266867    S.D. dependent var 0.000827
S.E. of regression   0.000708    Akaike info criter-11.57806
Sum squared resid    1.90E-05    Schwarz criterion -11.41257
Log likelihood       247.1393    F-statistic        5.974780
Durbin-Watson stat   1.574084    Prob(F-statistic)  0.001932
============================================================
```

A13.4 A researcher believes that a time series is generated by the process

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{9.1}$$

where ε_t is a white noise series generated randomly from a normal distribution with mean zero, constant variance, and no autocorrelation. Explain why the null hypothesis for a test of nonstationarity is that the series is nonstationary, rather than stationary.

A13.5 A researcher correctly believes that a time series is generated by the process

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{9.1}$$

where ε_t is a white noise series generated randomly from a normal distribution with mean zero, constant variance, and no autocorrelation. Unknown to the researcher, the true value of ρ is 0.7. The researcher uses a unit root test to test the series for nonstationarity. The output is shown. Discuss the result of the test.

```
        Augmented Dickey-Fuller Unit Root Test on X
============================================================
ADF Test Statistic -2.528841     1%   Critical Value*-3.6289
                                 5%   Critical Value -2.9472
                                 10% Critical Value  -2.6118
============================================================
*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(X)
Method: Least Squares
Sample(adjusted): 2 36
Included observations: 35 after adjusting endpoints
============================================================
     Variable      Coefficient Std. Error t-Statistic  Prob.
============================================================
       X(-1)        -0.379661   0.150132  -2.528841   0.0164
         C           0.222066   0.203435   1.091580   0.2829
============================================================
R-squared            0.162331    Mean dependent var-0.052372
Adjusted R-squared   0.136947    S.D. dependent var 1.095782
S.E. of regression   1.017988    Akaike info criteri2.928979
Sum squared resid    34.19792    Schwarz criterion  3.017856
Log likelihood      -49.25714    F-statistic        6.395035
Durbin-Watson stat   1.965388    Prob(F-statistic)  0.016406
============================================================
```

A13.6 Test of cointegration. Perform a logarithmic regression of expenditure on your commodity on income, relative price, and population. Save the residuals and test them for stationarity. (*Note*: the critical values in the regression output do not apply to tests of cointegration. For the correct critical values, see the text.)

A13.7 A variable Y_t is generated by the autoregressive process

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \varepsilon_t$$

where $\beta_2 = 1$ and ε_t satisfies the regression model assumptions. A second variable Z_t is generated as the lagged value of Y_t:

$$Z_t = Y_{t-1}$$

Show that Y and Z are nonstationary processes. Show that nevertheless they are cointegrated.

A13.8 X_t and Z_t are independent I(1) (integrated of order 1) time series. W_t is a stationary time series. Y_t is generated as the sum of X_t, Z_t, and W_t. Not knowing this, a researcher regresses Y_t on X_t and Z_t. Explain whether he would he find a cointegrating relationship.

A13.9 Two random walks RA_t and RB_t, and two stationary processes SA_t and SB_t are generated by the following processes

$$RA_t = RA_{t-1} + \varepsilon_{1t}$$

$$RB_t = RB_{t-1} + \varepsilon_{2t}$$

$$SA_t = \rho_A SA_{t-1} + \varepsilon_{3t} \qquad 0 < \rho_A < 1$$

$$SB_t = \rho_B SB_{t-1} + \varepsilon_{4t} \qquad 0 < \rho_B < 1$$

where ε_{1t}, ε_{2t}, ε_{3t}, and ε_{4t}, are iid N(0,1) (independently and identically distributed from a normal distribution with mean 0 and variance 1).

(a) Two series XA_t and XB_t are generated as

$$XA_t = RA_t + SA_t$$

$$XB_t = RB_t + SB_t$$

Explain whether it is possible for XA_t and XB_t to be stationary.
Explain whether it is possible for them to be cointegrated.

(b) Two series YA_t and YB_t are generated as

$$YA_t = RA_t + SA_t$$

$$YB_t = RA_t + SB_t$$

Explain whether it is possible for YA_t and YB_t to be cointegrated.

(c) Two series ZA_t and ZB_t are generated as

$$ZA_t = RA_t + RB_t + SA_t$$

$$ZB_t = RA_t - RB_t + SB_t$$

Explain whether it is possible for ZA_t and ZB_t to be stationary.
Explain whether it is possible for them to be cointegrated.

Answers to the starred exercises

13.1 Demonstrate that the MA(1) process

$$X_t = \varepsilon_t + \alpha_2\varepsilon_{t-1}$$

is stationary. Does the result generalize to higher-order MA processes?

Answer: The expected value of X_t is zero and therefore independent of time:

$$E(X_t) = E(\varepsilon_t + \alpha_2\varepsilon_{t-1}) = E(\varepsilon_t) + \alpha_2 E(\varepsilon_{t-1}) = 0 + 0 = 0$$

Since ε_t and ε_{t-1} are uncorrelated,

$$\sigma_{X_t}^2 = \sigma_{\varepsilon_t}^2 + \alpha_2^2\sigma_{\varepsilon_{t-1}}^2$$

and this is independent of time. Finally, because

$$X_{t-1} = \varepsilon_{t-1} + \alpha_2\varepsilon_{t-2},$$

the population covariance of X_t and X_{t-1} is given by

$$\sigma_{X_t X_{t-1}} = \alpha_2\sigma_\varepsilon^2$$

This is fixed and independent of time. The population covariance between X_t and X_{t-s} is zero for all $s > 1$ since then X_t and X_{t-1} have no elements in common. Thus the third condition for stationarity is also satisfied.

All MA processes are stationary, the general proof being a simple extension of that for the MA(1) case.

13.2 Demonstrate that the AR(1) process (13.2), with $\beta_2 < 1$, is stationary for finite samples if X_0 is generated as a random variable with appropriate mean and variance.

Answer: The process will be stationary if X_0 is generated as a random variable with mean 0 and variance $\frac{1}{1-\beta_2^2}\sigma_\varepsilon^2$, for then both the expected value and the variance of X_t will be independent of t, even for finite samples. The AR(1) process (13.2) is

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

Lagging and substituting, it was shown that

$$X_t = \beta_2^t X_0 + \beta_2^{t-1}\varepsilon_1 + ... + \beta_2 \varepsilon_{t-1} + \varepsilon_t .$$

It follows that

$$E(X_t) = \beta_2^t E(X_0) + \beta_2^{t-1} E(\varepsilon_1) + ... + \beta_2 E(\varepsilon_{t-1}) + E(\varepsilon_t) = \beta_2^t E(X_0)$$

Hence if X_0 is generated as a random variable with mean 0, $E(X_t) = 0$ and is independent of t. Of course, setting $X_0 = 0$ would have the same effect. The reason that we make X_0 a random variable is that we wish to choose its variance so as to make that of X_t independent of time as well.

The variance of X_t is given by

$$\begin{aligned}\sigma_{X_t}^2 &= \beta_2^{2t}\sigma_{X_0}^2 + \beta_2^{2t-2}\sigma_\varepsilon^2 + ... + \beta_2^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= \beta_2^{2t}\sigma_{X_0}^2 + \frac{1-\beta_2^{2t}}{1-\beta_2^2}\sigma_\varepsilon^2\end{aligned}$$

This is equal to $\frac{1}{1-\beta_2^2}\sigma_\varepsilon^2$, and hence is independent of t, if $\sigma_{X_0}^2 = \frac{1}{1-\beta_2^2}\sigma_\varepsilon^2$.

13.7 Suppose that a series is generated as

$$X_t = \beta_2 X_{t-1} + \varepsilon_t$$

with β_2 is equal to $1 - \delta$, where δ is small. Demonstrate that, if δ is small enough that terms involving δ^2 may be neglected, the variance may be approximated as

$$\begin{aligned}\sigma_{X_t}^2 &= (1-[2t-2]\delta) + ... + (1-2\delta) + 1)\sigma_\varepsilon^2 \\ &= (1-[t-1]\delta)t\sigma_\varepsilon^2\end{aligned}$$

and draw your conclusions concerning the properties of the time series.

Answer:

$$X_t = \beta_2^t X_0 + \beta_2^{t-1}\varepsilon_1 + ... + \varepsilon_t$$

Hence

$$\begin{aligned}\sigma_{X_t}^2 &= (\beta_2^{2t-2} + ... + \beta_2^2 + 1)\sigma_\varepsilon^2 \\ &= ([1-\delta]^{2t-2} + ... + [1-\delta]^2 + 1)\sigma_\varepsilon^2 \\ &= (1-[2t-2]\delta) + ... + [1-2\delta] + 1)\sigma_\varepsilon^2\end{aligned}$$

assuming that δ is so small that terms involving δ^2 may be neglected. (Note that the expansion of $(1 + x)^n$ is $(1 + nx + \frac{n(n-1)}{2}x^2 + ...)$ and if x is so small that terms involving x^2 and higher powers of x may be neglected, the expansion reduces to $1 + nx$.). Thus

$$\sigma_{X_t}^2 = (t - 2\delta[t-1+...+1])\sigma_\varepsilon^2$$
$$= (t - \delta t[t-1])\sigma_\varepsilon^2$$
$$= (1-[t-1]\delta)t\sigma_\varepsilon^2$$

It follows that, for finite t¸ the variance is a function of t and hence that the series exhibits nonstationary behavior for finite t, even though it is stationary.

13.8 Demonstrate that if the disturbance term in (13.30) is u_t, where u_t is generated by an AR(1) process, the appropriate specification for the Augmented Dickey–Fuller test is given by equation (13.32).

Answer: Let the process (13.29) be rewritten

$$X_t = \lambda_1 + \lambda_2 X_{t-1} + \theta t + u_t,$$

with u_t subject to the AR(1) process

$$u_t = \rho u_{t-1} + \varepsilon_t.$$

Lagging (13.29) one period and multiplying through by ρ, we have

$$\rho X_{t-1} = \lambda_1\rho + \lambda_2\rho X_{t-2} + \rho\theta(t-1) + \rho u_{t-1}.$$

Subtracting this from the equation for X_t, and rearranging, we obtain

$$X_t = \lambda_1(1-\rho) + \rho\theta + (\lambda_2 + \rho)X_{t-1} - \lambda_2\rho X_{t-2} + \theta(1-\rho)t + \varepsilon_t.$$

Thus we obtain the model

$$X_t = \beta_1 + \beta_2 X_{t-1} + \beta_3 X_{t-2} + \gamma t + \varepsilon_t$$

with redefinitions of the parameters. The condition for stationarity is $\beta_2 + \beta_3 < 1$. The process will be non-explosively nonstationary if $\beta_2 + \beta_3 = 1$. Subtracting X_{t-1} from both sides, and adding and subtracting $\beta_3 X_{t-1}$ on the right side, we have

$$X_t - X_{t-1} = \beta_1 + \beta_2 X_{t-1} - X_{t-1} + \beta_3 X_{t-1} - \beta_3 X_{t-1} + \beta_3 X_{t-2} + \gamma t + \varepsilon_t$$

Hence we obtain

$$\Delta X_t = \beta_1 + (\beta_2 + \beta_3 - 1)X_{t-1} - \beta_3 \Delta X_{t-1} + \gamma t + \varepsilon_t,$$

and the test is on the coefficient of X_{t-1}, with H_0: $\beta_2 + \beta_3 - 1 = 0$ being the null hypothesis of nonstationarity and $\beta_2 + \beta_3 - 1 < 0$ being the alternative hypothesis of stationarity.

Answers to the additional exercises

A13.1 The population series exhibits steady growth and is therefore obviously nonstationary. The growth is partly due to an excess of births over deaths and partly due to immigration. The question is whether variations in these factors are likely to be offsetting in the sense that a relatively large birth/death excess one year is somehow automatically counterbalanced by a relatively small one in a subsequent year, or that a relatively large rate of immigration one year stimulates a reaction that leads to a relatively small one later. Such compensating mechanisms do not seem to exist, so trend-stationarity may be ruled out. Population is a very good example of an integrated series with the effects of shocks being permanently incorporated in its level.

A13.2 It is difficult to come to any firm conclusion regarding this series. At first sight it looks like a random walk. On closer inspection, you will notice that after an initial decline in the first few years, the series appears to be stationary, with a high degree of correlation. The series is too short to allow one to discriminate between the two possibilities.

A13.3 As expected, given that the series is evidently nonstationary, the coefficient of *LGPOP*(–1), –0.05, is close to zero and not significant.

When we difference the series, the coefficient of *DLGPOP*(–1) is –0.16 and not significant, even at the 5 percent level. One possibility, which does not seem plausible, is that the population series is I(2). It is more likely that it is I(1), the first difference being stationary but highly autocorrelated.

A13.4 If the process is nonstationary, $\rho = 1$. If it is stationary, it could lie anywhere in the range $-1 < \rho < 1$. We must have a specific value for the null hypothesis. Hence we are forced to use nonstationarity as the null hypothesis, despite the inconvenience of having to compute alternative critical values of t.

A13.5 The model has been rewritten

$$X_t - X_{t-1} = (\rho - 1)X_{t-1} + \varepsilon_t$$

so that the coefficient of X_{t-1} is zero under the null hypothesis of nonstationarity. We see that the null hypothesis is not rejected at any significance level, despite the fact that we know that the series is stationary. However, the estimate of the coefficient of X_{t-1}, –0.38, is not particularly close to zero. It implies an estimate of 0.67 for ρ, close to the actual value. This is a common outcome. Unit root tests generally have low power, making it generally difficult or impossible to discriminate between nonstationary processes and highly autocorrelated stationary processes.

A13.6 Where the hypothetical cointegrating relationship has a constant but no trend, as in the present case, the critical values of t are –3.34 and –3.90 at the 5 and 1 percent levels, respectively (Davidson and MacKinnon, 1993). Hence the test indicates that we have a cointegrating relationship only for *DENT* and then only at the 5 percent level. However, one knows in advance that the residuals are likely to be highly autocorrelated. Many of the coefficients are greater than 0.2 in absolute terms and perfectly compatible with a hypothesis of highly autocorrelated stationarity.

Test of cointegration

	b_2	*s.e.*	t		b_2	*s.e.*	t
ADM	–0.09	0.06	–1.69	*GASO*	–0.08	0.05	–1.62
BOOK	–0.17	0.08	–2.24	*HOUS*	–0.31	0.12	–2.52
BUSI	–0.23	0.09	–2.40	*LEGL*	–0.26	0.10	–2.59
CLOT	–0.41	0.13	–3.17	*MAGS*	–0.39	0.13	–3.03
DENT	–0.51	0.15	–3.51	*MASS*	–0.07	0.05	–1.48
DOC	–0.35	0.12	–2.99	*OPHT*	–0.14	0.08	–1.86
FLOW	–0.22	0.10	–2.14	*RELG*	–0.17	0.07	–2.35
FOOD	–0.29	0.11	–2.61	*TELE*	–0.22	0.09	–2.35
FURN	–0.32	0.10	–3.29	*TOB*	–0.16	0.10	–1.66
GAS	–0.24	0.09	–2.79	*TOYS*	–0.17	0.09	–1.96

A13.7 The expected value of Y_t is $\beta_1 t + Y_0$, and thus it is not independent of t, one of the conditions for stationarity. Similarly for Z_t. However

$$Y_t - \beta_1 - \beta_2 Z_{t-1} = \varepsilon_t$$

and is therefore I(0).

A13.8 $$Y_t - X_t - Z_t = W_t$$

Since W_t is stationary, the left side of the equation is a cointegrating relationship.

A13.9 (a) A combination of a nonstationary process and a stationary one is nonstationary. Hence both X_A and X_B are nonstationary.

Since the nonstationary components of X_A and X_B are unrelated, there is no linear combination that is stationary, and so the series are not cointegrated.

(b) $$YA_t - YB_t = SA_t - SB_t$$

This is a cointegrating relationship for YA_t and YB_t since $SA_t - SB_t$ is stationary.

(c) No linear combination of RA_t and RB_t can be stationary since they are independent random walks, and so ZA_t and ZB_t are both nonstationary.

No linear combination of ZA_t and ZB_t can eliminate both RA_t and RB_t, so there is no cointegrating relationship.