# 14

# INTRODUCTION TO NON-STATIONARY TIME SERIES

(outline draft)

## 14.1 Stationarity and examples of stationary processes

The purpose of this section is to give a warning about the problems encountered by those applying regression analysis to nonstationary time series data. First we need definitions of stationarity and nonstationarity. A time series $x_t$ is said to be stationary if its expected value and population variance are independent of time and if the population covariance between its values at time $t$ and $t+s$ depends on $s$ but not on time. An example of a stationary time series is an AR(1) process

$$x_t = \rho x_{t-1} + \varepsilon_t \tag{14.1}$$

provided that $-1 < \rho < 1$, where $\varepsilon_t$ is a random variable with 0 mean and constant variance and is not subject to autocorrelation. This can easily be demonstrated. If equation (14.1) is valid for time period $t$, it is also valid for time period $t$-1:

$$x_{t-1} = \rho x_{t-2} + \varepsilon_{t-1} \tag{14.2}$$

Substituting for $x_{t-1}$ in equation (14.1), one has

$$x_t = \rho^2 x_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \tag{14.3}$$

Continuing this process of lagging and substituting, one has

$$x_t = \rho^t x_0 + \rho^{t-1}\varepsilon_1 + ... + \rho\varepsilon_{t-1} + \varepsilon_t \tag{14.4}$$

For $t$ large enough, $\rho^t x_0$ tends to 0. The expected value of $x_t$ is then given by

$$E(x_t) = \rho^{t-1}E(\varepsilon_1) + ... + \rho E(\varepsilon_{t-1}) + E(\varepsilon_t) \tag{14.5}$$

Each of the expectations is 0, so the expected value of $x_t$ is 0 and thus independent of $t$.

The population variance of $x_t$ is given by the population variance of $(\rho^{t-1}\varepsilon_1 + ... + \rho\varepsilon_{t-1} + \varepsilon_t)$ since it is unaffected by the additive constant $\rho^t x_0$ (Variance Rule 4). Since $\varepsilon$ is not autocorrelated, the population covariance between $\varepsilon_t$ and $\varepsilon_s$ $(t \neq s)$ is 0 and the population variance of $x_t$ is given by

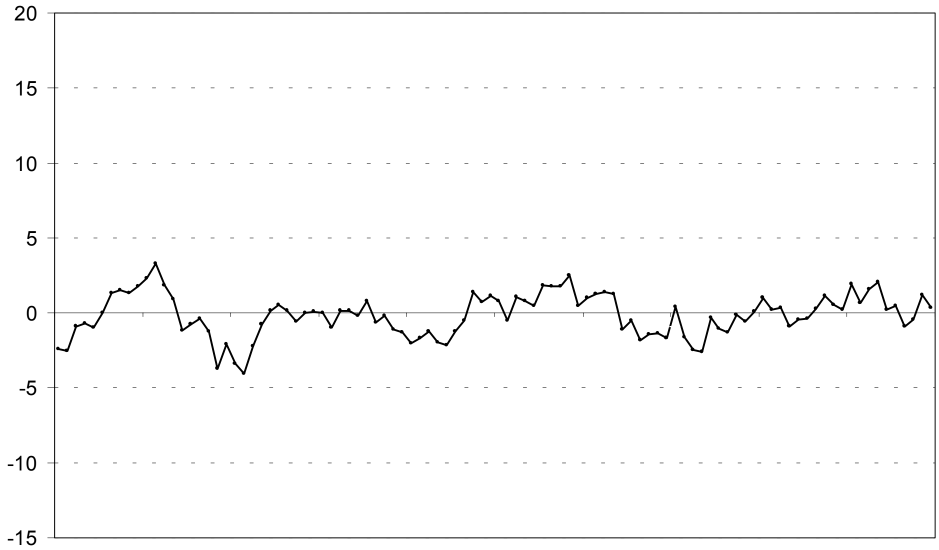


**Figure 14.1.** A Stationary Process

$$\begin{aligned}\sigma_{x_t}^2 &= \rho^{2t-2}\sigma_\varepsilon^2 + \ldots + \rho^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 \\ &= \frac{1-\rho^{2t}}{1-\rho^2}\sigma_\varepsilon^2\end{aligned} \quad (14.6)$$

The term $\rho^{2t}$ tends to 0 as $t$ becomes large and the variance is then independent of $t$. Similarly it can be shown that the population covariance of $x_t$ and $x_s$, $t > s$, is equal to $\rho^{t-s}\sigma_\varepsilon^2/(1 - \rho^2)$. This depends on the difference between $t$ and $s$ but is independent of $t$ itself. Figure 14.1 provides an example of this type of stationary process with $\rho = 0.7$.

A slightly more general version of the autoregressive process is

$$x_t = \rho x_{t-1} + \mu + \varepsilon_t \quad (14.7)$$

where $\mu$ is a constant. By lagging and substituting as before, one obtains

$$x_t = \rho^t x_0 + (\rho^{t-1} + \ldots + \rho^2 + \rho)\mu + \rho^{t-1}\varepsilon_1 + \ldots + \rho\varepsilon_{t-1} + \varepsilon_t \quad (14.8)$$

Provided that $\rho$ is less than 1, the series remains stationary, with expected value $\mu/(1 - \rho)$. The population variance of $x_t$ and the population covariance of $x_t$ and $x_s$ are unaffected by the inclusion of $\mu$.

## 14.2 Nonstationarity

In the previous examples, the condition $-1< \rho <1$ was crucial for stationarity. If $\rho$ is equal to 1, the original series becomes.

$$x_t = x_{t-1} + \varepsilon_t \tag{14.9}$$

This is an example of a nonstationary process known as a random walk. If it starts at $x_0$ at time 0, its value at time $t$ is given by

$$x_t = x_0 + \varepsilon_1 + \ldots + \varepsilon_t \tag{14.10}$$

In this case the expected value and population variance of $x$ do not have unconditional meanings. If the expectations are taken at time 0, the expected value at any future time $t$ is independent of $t$ (it is always equal to $x_0$); but its variance is given by

$$\begin{aligned} \sigma_\varepsilon^2 &= \sigma_\varepsilon^2 + \ldots + \sigma_\varepsilon^2 \\ &= t\sigma_\varepsilon^2 \end{aligned} \tag{14.11}$$

and so it is not independent of time. Figure 14.2 provides an example of a random walk.

In the more general version of the autoregressive process with the constant $\mu$, the process becomes what is known as a random walk with drift if $\rho$ equals 1:

$$x_t = x_{t-1} + \mu + \varepsilon_t \tag{14.12}$$

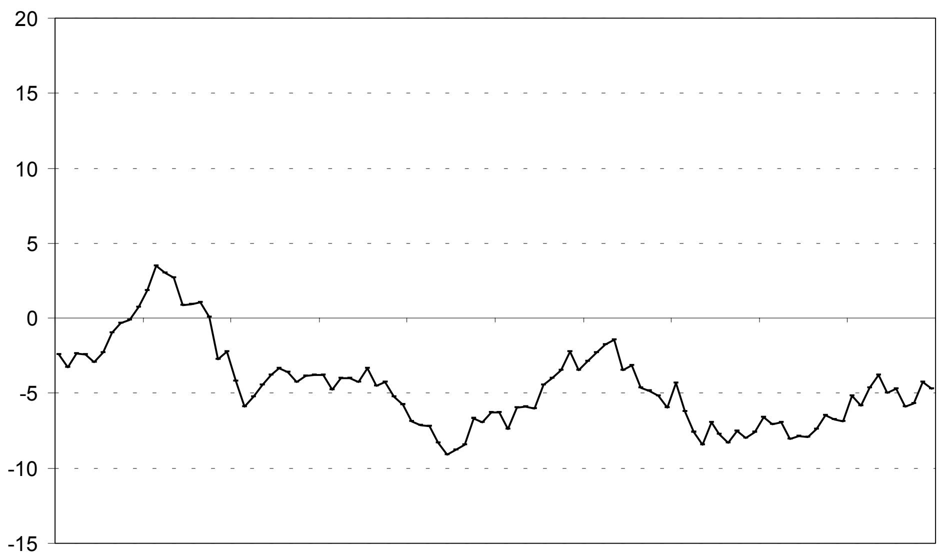


**Figure 14.2.** Random Walk

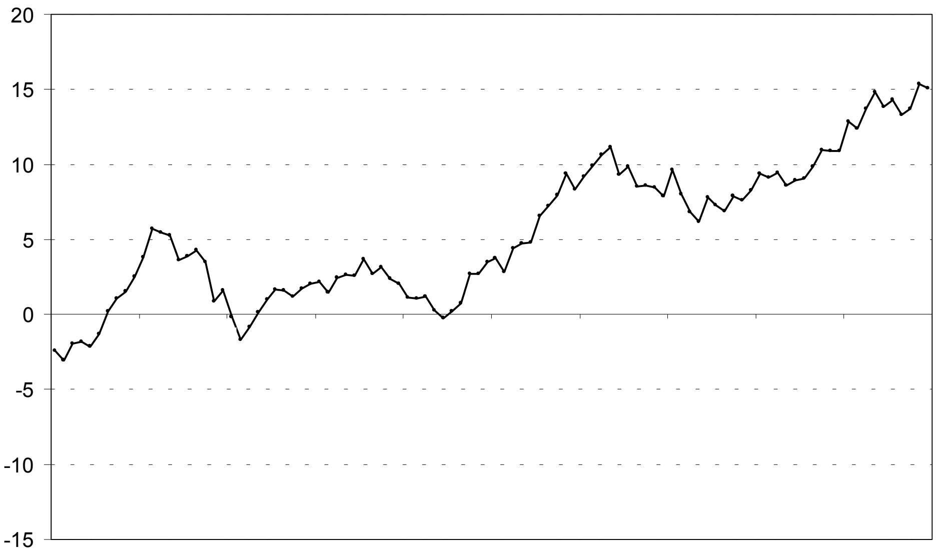


**Figure 14.2.** Random Walk with Drift

If the series starts at $x_0$ at time 0, $x_t$ is given by

$$x_t = x_0 + \mu t + \varepsilon_1 + \ldots + \varepsilon_t \tag{14.13}$$

Now the expectation of $x_t$ at time 0, $(x_0 + \mu t)$, is also a function of $t$. Figure 14.3 provides an example of a random walk with drift.

Random walks are not the only type of nonstationary process. Another common example of a nonstationary time series is one possessing a time trend:

$$x_t = \alpha + \beta t + \varepsilon_t \tag{14.14}$$

Its expected value at time $t$, $(\alpha + \beta t)$, is not independent of $t$ and its population variance is not defined.

### *Difference-Stationarity and Trend-Stationarity*

In the discussion that follows, a distinction will be made between difference-stationary and trend-stationary time series. If a nonstationary process can be transformed into a stationary one by differencing, it is said to be difference-stationary. A random walk, with or without drift, is an example. If $x_t$ is a random walk with drift, as in equation (14.12),

$$\Delta x_t = (x_t - x_{t-1}) = \mu + \varepsilon_t \tag{14.15}$$

This is a stationary process with population mean $\mu$ and variance $\sigma_\varepsilon^2$, both independent of time. If a time series can be transformed into a stationary process by differencing once, as in this case, it is described as integrated of order 1, I(1). If a time series can be made stationary by differencing twice, it is known as I(2), and so on. To complete the picture, a stationary process, which by definition needs no differencing, is described as I(0).

A nonstationary time series is described as being trend-stationary if it can be transformed into a stationary process by extracting a time trend. For example, the very simple model given by equation (14.14) can be detrended by fitting the equation

$$\hat{x}_t = a + bt \tag{14.16}$$

and defining a new variable

$$\tilde{x}_t = x_t - \hat{x}_t = x_t - a - bt \tag{14.17}$$

The new, detrended, variable is of course just the residuals from the regression of $x$ on $t$.

### *Spurious Regressions*

In a celebrated article, Granger and Newbold (1974) report the results of a Monte Carlo experiment in which two variables were generated as independent random walks and then one was regressed on the other. Obviously, a regression of one random walk on another ought not to yield significant results except as a matter of Type I error. Granger and Newbold ran the experiment 100 times and so, using a 5% significance test, one would anticipate that the slope coefficient would not be significantly different from 0 about 95 times, the other 5 times where it appeared to be significant being cases of Type I error. However they found that the slope coefficient had a significant $t$ statistic on 77 occasions. Using a more cautious 1% test made very little difference. The null hypothesis of no relationship was rejected on 70 occasions.

The reason for this is that random walks are nonstationary, and OLS, and variations like AR(1), yield invalid standard errors and test statistics when the regression model includes nonstationary variables. It is also likely that the estimates of the coefficients will be inconsistent.

## 14.3 Detection of Nonstationarity

Unfortunately for econometricians working with time series data, many economic time series appear to be of the I(1) type. It is therefore important to test time series for nonstationarity before attempting to use them in a regression model. Testing for nonstationarity is often described as testing for unit roots, for reasons that need not concern us here. The standard test, pioneered by Dickey and Fuller, is based on the model

$$x_t = \rho x_{t-1} + \beta t + \mu + \varepsilon_t \tag{14.18}$$

This process is difference-stationary if $\rho = 1$ and $\beta = 0$, trend-stationary if $\rho < 1$ and $\beta \neq 0$, and it also allows the possibility that the series is stationary after all ($\rho < 1$ and $\beta = 0$). One can allow for more complex dynamics by letting $x$ depend on further lagged values of itself, for example as in

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \beta t + \mu + \varepsilon_t \quad (14.19)$$

In this case the process will be stationary if $\rho_1 + \rho_2 < 1$ and $\beta = 0$. It will be difference-stationary if $\rho_1 + \rho_2 = 1$ and $\beta = 0$. It is convenient to rewrite the equation as

$$\Delta x_t = (\rho_1 + \rho_2 - 1)x_{t-1} - \rho_2 \Delta x_{t-1} + \beta t + \mu + \varepsilon_t \quad (14.20)$$

and test the null hypothesis that the coefficients of $x_{t-1}$ and $t$ are equal to 0. The test, or rather family of tests because there are various alternatives, are known as augmented Dickey-Fuller tests. The usual procedure is to perform a one-tailed $t$ test on the coefficient of $x_{t-1}$ under the null hypothesis that it is equal to 0 (that is, that there is a unit root and the series is nonstationary). The reason for performing a one-tailed test is that if the alternative hypothesis, stationarity, is correct, the coefficient will be negative. Under the null hypothesis, $t$ does not have its usual distribution and the critical value, for any given significance level, is higher than that shown in the standard tables. Critical values are provided in the appendix to this note.

## 14.4 Cointegration

If one or more time series in a model exhibit nonstationarity, conventional regression techniques are inappropriate. It is beyond the scope of this introduction to describe what should be done under these circumstances. The literature on the subject, although growing very rapidly, is still in its infancy. However, suppose that the model

$$y_t = \alpha + \beta x_t + u_t \quad (14.21)$$

is a correct specification. The error term $u_t$ can be thought of as measuring the deviation between the components of the model:

$$u_t = y_t - \alpha - \beta x_t \quad (14.22)$$

In the short run the divergence between the components will fluctuate, but if the model is genuinely correct there should be a limit to the divergence. Hence even if $y_t$ and $x_t$ are nonstationary, $u_t$ should be stationary. If this is found to be the case, $y_t$ and $x_t$ are said to be cointegrated and the relationship is interpreted as a long-run one. To test for cointegration, the first step is to check that the time series involved are of the same order of integration, for otherwise it is not possible for them to be cointegrated. The second is to run an OLS regression and to test the residuals for nonstationarity using a unit root test. If the null hypothesis is rejected in favor of stationarity, a cointegrating relationship may have been found. Note that, because the OLS regression has used nonstationary time series, the standard errors and other diagnostic statistics should be disregarded.

Exercises

14.1 (A repeat of Granger and Newbold's experiment). Construct two 100-observation random walks and regress one on the other. Does the $t$ statistic on the slope coefficient appear to be significant using a 5% test? Repeat the experiment several times (at least 5 times; 20 would be better) and note the frequency of Type I errors.

14.2 Test the logarithms of disposable personal income, expenditure on your commodity, and the relative price series for your commodity for difference-stationarity. Calculate the first differences and test these for difference-stationarity.

14.3 Run logarithmic regressions of expenditure on your commodity on disposable personal income and relative price and test for cointegration.

## *Reference*

Granger, C.W.J., and P. Newbold (1974) Spurious regressions in econometrics, *Journal of Econometrics* 2 (2), 111-120