1 Introduction

This chapter starts with a discussion of the motivation and scope of this book. Then, it introduces properties of unit root processes, relations between social science and unit roots, and some basic technical tools related to inferences on unit roots. It also provides an overview of subsequent chapters. Discussions on preliminary concepts and basic tools are brief because of the nature of this book, and the reader is referred to more specialized books such as Brockwell and Davis (1991), Davidson (1994), Hamilton (1994), Fuller (1976), and Serfling (1980).

1.1 Motivation and Scope of This Book

The last two decades or so have seen significant developments in the literature on unit roots. By the early 1980s, only a handful of papers had been written about unit roots, mostly by Professor Wayne Fuller and his coauthors. In those days, researchers in social science seldom used unit root tests for their empirical studies, and it was hard to find a graduate course on time series analysis offered by departments related to social science. Todays, the situation is radically different: there are many theoretical papers about unit roots, as the reference section of this book attests, and various procedures designed for testing for a unit root are often used in social science, particularly in economics. Naturally, commercial software for econometrics and statistics has incorporated many of the methods developed in the literature on unit roots.

Because so many unit root tests had been developed by the 1990s, some even thought that efforts dedicated to unit roots were excessive and unwarranted, as Maddala and Kim (1998, p. 488) succinctly quipped, “What we do not need is more unit root tests (each of which uses the Nelson–Plosser data as a guinea pig).” Nonetheless, because no one can predict with confidence the future direction of the world of knowledge, research on unit roots has continued to expand.

https://doi.org/10.1017/CBO9781316157824.003 Published online by Cambridge University Press

The main vehicle for the massive theoretical developments in unit root regressions and testing has been the functional central limit theory (FCLT) that Phillips (1986, 1987a) first introduced to the literature on unit roots. Although White (1958) and Dickey and Fuller (1979) developed asymptotic theory for the AR(1) model with a unit root, their methods require a normality assumption and are difficult to use for other types of regressions involving nonstationary regressors. By contrast, the FCLT allows us to employ general assumptions on the DGP of the model in use and can be applied to various linear and nonlinear regression models. In this sense, it is fair to say that the FCLT of Phillips (1986, 1987a) has played a pivotal role in the developments of limit theory for regressions with nonstationary time series.

There are several reasons why unit roots are important in economics, other disciplines of social science, and statistics. First, the validity of many economic propositions hinges on the presence or absence of unit roots. For example, real exchange rates should not have a unit root for the relative purchasing power parity to hold as discussed in Subsection 1.3.3. Some other examples are also given in Section 1.3. This is one of the reasons why unit root tests are so often used in economics.

Second, regressions and VARs require knowing about the univariate properties of the variables in use. If the variables are stationary, conventional theories on regressions and VARs can be used. But if they have a unit root, regressions can be spurious (cf. Granger and Newbold, 1974) unless those variables are cointegrated. In a VAR system, the presence of unit roots invites a host of nontrivial issues for such standard VAR procedures as the causality test, impulse response analysis, and forecast error variance decomposition, as analyzed in Toda and Phillips (1993) and Phillips (1998). Using differenced data is not necessarily the best option for conducting VAR analysis. Thus, testing for a unit root has almost always preceded regressions and VARs in economics, political science, and sociology, and those test results have routinely been reported. Because unit roots appear to be present at many key time series in social science (e.g., GDP, nominal interest rates, exchange rates, consumer sentiments, presidential approval rates, etc.), such preliminary specification testing has been performed faithfully in empirical time series analysis. In addition, a way of testing for cointegration or spuriousness of regressions is to check the presence of a unit root in the OLS residuals as suggested by Engle and Granger (1987). Unit root tests are again used for this purpose, although asymptotic theory for such tests is not dealt with in this book.

Third, as subsequent chapters show, diverse econometric and statistical theories have been applied to the AR model with unit roots. For many theoretical researchers, unit roots have been an important means with which they could test their econometric and statistical theories. Moreover, the level of generality, strength, and usefulness of those theories can be assessed when they are applied to the AR model with unit roots, which makes the model important to theorists in econometrics and statistics. The reader can learn about those theories from this book, which is an added benefit that this book can provide beyond knowledge on unit roots.

In light of the well-accepted importance of unit roots, it is no wonder that many researchers in social science and statistics want to learn about them. Indeed, knowledge on unit roots has become so essential for modern time series analysis that performing empirical time series analysis and understanding empirical literatures in social science are virtually impossible without it. But for those who want to study the literature on unit roots, it is difficult to know where to start because the literature is now so immense. Finding specific information on unit roots for each researcher's purposes is also hard for the same reason. These difficulties motivate an extensive, compact, nontechnical, and up-to-date book on unit roots. This book rests on this motivation and will be useful to those who want to study the literature on unit roots. From this book, the reader will be able to obtain the most comprehensive and up-to-date information on unit roots that he or she can then use to conduct empirical and theoretical research on unit roots.

This book covers research papers on unit roots from 1958 to the present time. The oldest paper this book discusses is White (1958) (see Subsection 2.2.1), and the most recent one is Gao and Robinson (2013) (see Subsection 4.2.2). More space is given to important papers such as those by Dickey and Fuller (1979) and Phillips (1987a), but lesser known papers are also discussed in detail if they are deemed to be based on novel and useful ideas. Because there are so many papers in the area of univariate unit roots alone, those on cointegration and multivariate unit roots are not included in this book and relegated to future works. This book may look incomplete because of this feature, but this choice was necessary to keep its length within reasonable bounds. This book tries to cover as many papers as possible to provide comprehensive information on unit roots to the reader and to record the developments of the literature on unit roots. Undoubtedly, however, some papers must have been neglected. This makes it necessary to put the word "almost" in the title of this book. But let me emphasize that I made a genuine effort to make this book as comprehensive as possible.

1.2 Properties of Unit Root Processes

The characteristic equation of the AR(1) model,

$$y_t = \alpha y_{t-1} + u_t, (t = 2, \ldots, T),$$

https://doi.org/10.1017/CBO9781316157824.003 Published online by Cambridge University Press

where $\{u_t\}$ is a white noise process with variance σ_u^2, is written as

$$1 - \alpha z = 0.$$

When the root of this equation is 1 or -1, the process is said to have a unit root. That is, if $\alpha = \pm 1$, $\{y_t\}$ has a unit root. In most applications in economics, the main concern is whether the coefficient α is equal to one. Hence, discussions in this section revolve around the case of $\alpha = 1$.

When $\alpha = 1$, $y_t = \sum_{i=1}^{t} u_i + y_0$, although we may write $y_t = \sum_{i=0}^{\infty} \alpha^i u_{t-i}$ when $|\alpha| < 1$. These representations can be used to show that the stochastic properties of $\{y_t\}$ with $\alpha = 1$ are remarkably different from those of $\{y_t\}$ with $|\alpha| < 1$. Engle and Granger (1987) summarize them as follows.

(i) When $\alpha = 1$, $\mathrm{Var}(y_t) \to \infty$ as $t \to \infty$ once y_0 is assumed to be a constant. When $|\alpha| < 1$, however, $\mathrm{Var}(y_t) = \frac{\sigma_u^2}{1-\alpha^2}$ for all t. These imply that the data become more variable as we collect more of them when $\alpha = 1$. But the data will move within a fixed range when $|\alpha| < 1$.

(ii) When $\alpha = 1$, an innovation (i.e., u_{t-i}, $i \geq 0$) has a permanent effect on the value of y_t that does not die out as the stochastic process progresses toward the future. When $|\alpha| < 1$, an innovation will lose its effect on the value of y_t eventually as we move forward into the future.

(iii) When $\alpha = 1$, $f_{yy}(0) = \infty$ where $f_{yy}(\cdot)$ denotes the spectral density of $\{y_t\}$. This means that $\{y_t\}$ has a strong long-run component. When $|\alpha| < 1$, the spectral density is finite at all frequencies.

(iv) When $\alpha = 1$, the expected time between crossings of $y = 0$ is infinite. Thus, $\{y_t\}$ has no tendency to return to its theoretical mean. When $|\alpha| < 1$, the expected time between crossings of $y = 0$ is finite, which implies that the process moves around its mean and has a tendency of mean reversion.

(v) When $\alpha = 1$, the theoretical autocorrelation at lag k converges to 1 for all k as $t \to \infty$. This means that the autocorrelation does not allow conventional interpretations when $\alpha = 1$. When $|\alpha| < 1$, the autocorrelation decreases steadily in magnitude as k increases.

In addition to these properties, the coefficient α also affects the variance of the forecasting error. Suppose that we forecast $y_{T+1}, y_{T+2}, \ldots,$ with a known value of the coefficient α. Then the optimal forecasts are $\hat{y}_{T+1} = \alpha y_T$, $\hat{y}_{T+2} = \alpha \hat{y}_{T+1}, \ldots,$ and the forecast error is defined by

$$\begin{aligned} y_{T+1} - \hat{y}_{T+1} &= u_{T+1} \\ y_{T+2} - \hat{y}_{T+2} &= u_{T+2} + \alpha u_{T+1} \\ &\vdots \end{aligned}$$

Thus, denoting the forecasting horizon as h, the variance of the forecasting error is $\sigma_u^2(1 + \alpha^2 + \cdots + \alpha^{2(h-1)})$ when $|\alpha| < 1$. But when $\alpha = 1$, it is $h\sigma_u^2$, which grows linearly with h, and is larger than that for the case $|\alpha| < 1$. This indicates that it becomes difficult to predict the future observations precisely when $\alpha = 1$.

We have assumed so far that $\{u_t\}$ is a white noise process. But essentially the same results hold true when $\{u_t\}$ is a stationary and invertible ARMA process, indicating that an ARMA model with a unit root has properties quite different from that without it.

1.3 Economics and Unit Roots

Discussions in the previous section indicate that the unit root case has quite distinctive characteristics. It is no wonder that researchers have exerted so much effort to study the AR process with a unit root. However, these characteristics alone do not explain the huge interest in the unit root AR model in the economics literature. This section delves into why the unit root case has attracted so much attention from economists.

1.3.1 Nelson and Plosser (1982)

It was Nelson and Plosser (1982) who brought the issue of nonstationarity to the forefront of economic research.[1] They investigated whether macroeconomic time series are characterized as stationary fluctuations around a deterministic trend or as unit root processes with drift. Using historical time series for the United States, they could not reject the hypothesis of a unit root with drift for most of them. Using these findings and an unobserved components model for output, they conclude that "macroeconomic models that focus on monetary disturbances as a source of purely transitory fluctuations may never be successful in explaining a large fraction of output variation and that stochastic variation due to real factors is an essential element of any model of macroeconomic fluctuations" (Nelson and Plosser, 1982, p. 139). In other words, they interpret presence of a unit root or a high level of persistence in real GNP as supporting evidence for real-business-cycle theory. However, this interpretation does not seem to be universally accepted. Romer (2001, p. 210) writes,

> Keynesian models do not require that persistence be low. To begin with, although they attribute the bulk of short-run fluctuations to aggregate demand disturbances, they do not assume that the processes that drive long-run growth

[1] Before Nelson and Plosser (1982) published their research, Altonji and Ashenfelter (1980) also applied the Dickey-Fuller test to the annual real wage data of the United States and the United Kingdom and could not reject the null hypothesis of a unit root.

> follow a deterministic trend; thus they allow at least one part of output movements to be highly persistent. More importantly, the part of fluctuations that is due to aggregate demand movements may also be persistent.

In other words, according to Romer, the presence of a unit root in real GDP should not be construed as evidence against Keynesian business-cycle models.

Some ascribed Nelson and Plosser's (1982) results to the low power of Dickey and Fuller's (1979) test they used, which prompted further studies seeking to improve the power of unit root tests (see Section 2.4). Nelson and Plosser's dataset was used extensively as an experimental object for some time whenever someone invented a new unit root test, and its extended version is used in Schotman and van Dijk (1991b). Subsequent similar studies have generally confirmed Nelson and Plosser's empirical results. However, as is seen in Sections 3.2 and 4.6, unit root tests accommodating structural changes and a Bayesian approach can yield somewhat different results.

1.3.2 Cointegration

Engle and Granger (1987) define that an $I(1)$ multiple time series $\{X_t\}$ is cointegrated if there exists a vector γ such that $\{\gamma' X_t\}$ becomes $I(0)$. The vector γ denotes a statistical equilibrium relationship among the elements of $\{X_t\}$ because $\{\gamma' X_t\}$ tends to return to its mean while each element of $\{X_t\}$ does not possess such a property. The concept of cointegration and related econometric tools have often been used in economics to model statistical equilibrium relationships among economic variables and to verify those relationships. In cointegration analysis, the first step is to test whether the variables of interest have a unit root. Thus, without exceptions, unit root tests are used in applications of cointegration, serving as specification tests, the results of which are used for subsequent analysis.

1.3.3 Purchasing Power Parity Hypothesis

The absolute law of one price postulates that the same good should have the same price across countries and is expressed by the relation

$$P_{it} = S_t P_{it}^*, \tag{1.1}$$

where P_{it} is the price of good i in terms of the domestic currency at time t, S_t is the domestic price of a unit of foreign currency at time t, and P_{it}^* is the price of good i in terms of the foreign currency at time t. Taking natural logarithms of relation (1.1), we obtain

$$p_{it} = s_t + p_{it}^*,$$

where lowercase letters denote the logarithms of the corresponding capital letters. Summing all the traded goods in each country with weights γ_i yields a relation

$$p_t = s_t + p_t^*, \tag{1.2}$$

where $p_t = \sum_{i=1}^{N} \gamma_i p_{it}$, $p_t^* = \sum_{i=1}^{N} \gamma_i p_{it}^*$, and $\sum_{i=1}^{N} \gamma_i = 1$. Because p_t and p_t^* can be considered as national price levels,[2] equation (1.2) indicates that the exchange rate is determined by the price levels of both countries and is called the absolute purchasing power parity (PPP) relation. The relative PPP hypothesis postulates that

$$\Delta p_t = \Delta s_t + \Delta p_t^*. \tag{1.3}$$

That is, changes in the nominal exchange rate should match those of the national price levels. The relative PPP holds if $q_t = s_t - p_t + p_t^*$, called the real exchange rate, is a constant. In reality, it is hard to expect that this relation holds in every t. But if relation (1.3) provides a reasonably good approximation to the real world, $\{q_t\}$ should be a stationary process with possibly a nonzero mean. In other words, there should not be a unit root in the real exchange rate $\{q_t\}$ for the relative PPP to hold. Empirical studies employing unit root tests have generally been unable to reject the null hypothesis of a unit root for real exchange rates (see section 3 of Sarno and Taylor, 2002).

1.3.4 Asset Prices

Samuelson (1965) shows that asset prices in an informationally efficient market follow the martingale process, which means that returns are unpredictable and that asset prices have a unit root. Although some evidence has emerged for the predictability of stock returns at a long horizon when variables such as term spread, dividend yield, and earnings/price ratio are used (e.g., Lettau and Ludvigson, 2001), it is now empirically well accepted that asset prices have a unit root. In a consumption-based asset pricing model without dividends, asset prices also follow the martingale process if investors are risk-neutral and if the discount factor is equal to one (see Cochrane, 2005, pp. 24–25).

1.3.5 Relative Mean Reversion in International Stock Markets

Mean reversion of asset prices refers to their tendency to return to a trend path. Fama and French (1988) and Poterba and Summers (1988) are the first works

[2] In practice, countries use different baskets of goods to formulate price indices. Moreover, it is more common to use arithmetic than geometric price indices. These aspects are disregarded in this relation. See Sarno and Taylor (2002) for further discussion.

that study mean reversion. Fama and French and Poterba and Summers use regression and the variance ratio, respectively. More recently, Balvers, Wu, and Gilliland (2000) study mean reversion using unit root tests. Their methods can be summarized as follows. Let $p_{i,t}$ denote the log of the total return index of the stock market in country i at the end of period t and assume that the evolution of $p_{i,t}$ is described by a mean-reverting process,

$$p_{i,t+1} - p_{i,t} = a_i + \lambda \left(p^*_{i,t+1} - p_{i,t}\right) + \varepsilon_{i,t+1}, \tag{1.4}$$

where $p^*_{i,t+1}$ is an unobserved fundamental value of the index, a_i is a positive constant, and $\varepsilon_{i,t+1}$ is a stationary disturbance with an unconditional mean of zero. Parameter λ is the speed of mean reversion and is assumed to be the same across countries. If $0 < \lambda < 1$, deviations of $p_{i,t}$ from its fundamental or trend value $p^*_{i,t+1}$ will be reversed over time. But if $\lambda = 0$, the log price follows a unit root process, and there is no mean reversion. Balvers, Wu, and Gilliland assume

$$p^*_{i,t} = p^*_{r,t} + z_i + \eta_{i,t}, \tag{1.5}$$

where r denotes a reference country, z_i is a constant, and $\eta_{i,t}$ is a stationary process with mean zero. Combining equations (1.4) and (1.5) eliminates $p^*_{i,t+1}$ and yields

$$r_{i,t+1} - r_{r,t+1} = \alpha_i - \lambda\left(p_{i,t} - p_{r,t}\right) + \omega_{i,t+1},$$

where $r_{i,t+1} = p_{i,t+1} - p_{i,t}$ is the log return on market i, $\alpha_i = a_i - a_r + \lambda z_i$, and $\omega_{i,t} = \varepsilon_{i,t} - \varepsilon_{r,t} + \lambda\eta_{i,t}$. Note that α_i is a constant and that $\omega_{i,t}$ is stationary with an unconditional mean of zero. In this formulation, no mean reversion (i.e., $\lambda = 0$) corresponds to the presence of a unit root in $\{p_{i,t} - p_{r,t}\}$. Thus, mean reversion can be tested using unit root tests. Balvers, Wu, and Gilliland report evidence of mean reversion in relative stock-index prices using stock-index data from 18 nations during the period 1969–1996.

1.3.6 Growth and Convergence

Economists have taken an interest in empirically investigating whether per capita outputs of nations converge to the same level, starting from the works of Baumol (1986) and DeLong (1988). Although these studies employ cross-sectional regressions, Quah (1994) and Bernard and Durlauf (1995) use a time series approach. In the latter works, two nations' per capita output converge if their difference is a stationary process with zero mean because this means that the difference is only transitory and fluctuates around zero. Thus, if there is a unit root in the difference, the convergence hypothesis is rejected. These works have generally rejected the convergence hypothesis.

1.3.7 Convergence of Real Interest Rates

Researchers in the field of international macroeconomics have been interested in testing for capital-market integration. One way of examining this issue is to study whether two nations' real interest rate differential follows a zero-mean stationary process. If it does, the two nations have essentially the same real interest rates, and their differences dissipate over time. Thus, the presence of a unit root in real interest rate differentials implies that the capital markets of the two nations are not fully integrated. This approach is taken, for example, in Herwartz and Roestel (2011).

1.3.8 Inflation Convergence

The issue of inflation convergence within European nations adopting the common currency euro has attracted much attention. This issue is important because it is related to whether the single monetary policy of the European Central Bank has succeeded in stabilizing the inflation rates of its member nations. Kocenda and Papell (1997) test this issue using panel unit root tests. Suppose that the i-th country's inflation rate, π_{it}, follows an AR(1) process:

$$\pi_{it} = \mu + \alpha \pi_{i,t-1} + u_{it}, (i = 1, \ldots, N). \tag{1.6}$$

The cross-sectional average of the inflation rates has the dynamics represented by

$$\overline{\pi}_{.t} = \mu + \alpha \overline{\pi}_{.t-1} + \overline{u}_{.t}, \tag{1.7}$$

where $z_{.t} = \frac{1}{N} \sum_{i=1}^{N} z_{it}$. Subtracting equation (1.7) from (1.6) yields

$$\pi_{it} - \overline{\pi}_{.t} = \alpha \left(\pi_{i,t-1} - \overline{\pi}_{.t-1} \right) + u_{it} - \overline{u}_{.t}.$$

If $|\alpha| < 1$, the difference between the i-th country's inflation rate and the average inflation rate is transitory; thus, it can be said that inflation rates converge. In contrast, if $\alpha = 1$, one can say that there is no inflation convergence. One can also use pairwise differences of inflation rates to examine inflation convergence as in the literature on growth convergence. In this case, the presence of a unit root implies divergence of the two nations' inflation rates. This approach is taken in Busetti, Forni, Harvey, and Venditti (2007).

1.3.9 Unemployment Hysteresis

Blanchard and Summers (1986, 1987) propose the concept of unemployment hysteresis, in which cyclical business fluctuations have permanent effects on the level of unemployment. If the unemployment-hysteresis hypothesis is accurate, high unemployment rates in an economy will persist unless government intervenes to correct them. That is, active government interventions in the labor market are supported by this hypothesis.

We discuss here how we can test the unemployment-hysteresis hypothesis following the framework of Brunello (1990) and Song and Wu (1997). To test this hypothesis, consider the Phillips curve,

$$P_t = E_{t-1}P_t - \beta(u_t - u_t^*) + \xi_t, \tag{1.8}$$

where P_t is the current inflation rate, $E_{t-1}P_t$ is the expected inflation rate of time t given information at time $t-1$, β is a constant, u_t is the current unemployment rate, u_t^* is the natural unemployment rate, and ξ_t is an error term. Assume that the natural rate is a function of past unemployment rates, which can be expressed by

$$u_t^* = c + \alpha u_{t-1} + \zeta_t, \tag{1.9}$$

where c and α are constants and ζ_t is an error term. Substituting (1.9) into (1.8), we obtain

$$u_t = c + \alpha u_{t-1} + \varepsilon_t,$$

where $\varepsilon_t = (E_{t-1}P_t - P_t + \xi_t)/\beta + \zeta_t$. If $\alpha = 1$, the unemployment rate has no mean reversion and wanders around without being anchored to a particular point. Thus, the unemployment-hysteresis hypothesis can be tested by testing the null hypothesis of a unit root. Brunello reports some evidence supporting the unemployment-hysteresis hypothesis using Japanese data, whereas Song and Wu find evidence against it using panel data from the United States.

1.4 Other Branches of Social Science and Unit Roots

1.4.1 Political Science and Unit Roots

There is conspicuously less data analysis in political science than in economics, most likely due to the data limitations in the discipline. Still, researchers in political science have used unit root tests in their work. Some economic propositions can be probed by testing for a unit root as we have seen in the last section. In political science, however, because the presence or absence of a unit root seldom carries any structural implications (with the exception of macropartisanship, as discussed later), tests for a unit root have usually been used to decide whether to difference the time series for subsequent regressions and VARs. This subsection presents several works in political science that use unit root tests, without discussing their empirical results in full detail.

Chowdhury (1991) and Heo and Eger (2005), among others, study the relationship between economic growth and military spending. They find from applying the augmented Dickey-Fuller test (see Subsection 2.3.2) that some key variables such as defense spending and military share of GDP have a unit root, and they take the difference of these variables for subsequent regression and VAR analyses. Their empirical results suggest that the relationship between economic growth and military spending cannot be generalized across countries.

Clarke and Stewart (1994) and Price and Sanders (1993) study U.S. presidential approval ratings and UK government popularity, respectively, and relate these variables to some economic variables. In their studies, the null hypothesis of a unit root is not rejected for U.S. presidential approval ratings and UK government popularity. Subsequently, Clarke and Stewart use the error-correction model (see Engle and Granger, 1987), and Price and Sanders the AR regression with some exogenous variables. As expected, a strong economy increases U.S. presidential approval ratings and UK government popularity.

Blood and Phillips (1995) investigate the relationships among four variables: headlines referring to a recession from the *New York Times*, consumer sentiment, a composite measure of leading economic indicators, and presidential popularity. All these variables are found to have a unit root. The main conclusion from their cointegration analysis and Granger-causality test is that recession headlines significantly influence consumer sentiment.

Green, Palmquist, and Schickler (1998) and Box-Steffensmeier and Smith (1996) study macropartisanship—the aggregate distribution of party identification. They find some evidence for a unit root in some U.S. macropartisanship data, and Box-Steffensmeier and Smith argue that they are well modeled by nonstationary fractionally integrated processes. In addition, using macropartisanship data, Meffert, Norpoth, and Ruhil (2001) test for a unit root in the presence of possible structural changes, although it is not certain exactly what tests they used.

1.4.2 Sociology and Unit Roots

As in political science, sociologists usually use tests for a unit root to decide whether to difference the time series for subsequent analyses. This subsection introduces a few papers published in sociology journals that use unit root tests.

Jacobs and Helms (2001) study how the progressivity of the income tax is influenced by civil rights activities in the United States. They measure progressivity of the income tax by the logged marginal tax rates of different income groups. Explanatory variables are the number of civil rights actions, the number of riots, the percentage of nonwhites, median family income and dummy variables for a large number of crimes, and the presence of Republican presidents. They difference all these data except the dummy variables before running regressions, because unit root tests show that they have a unit root. They find that civil rights activities lead to redistributive tax codes, but that riots reduce tax progressivity.

Kristal (2010) studies the dynamics of labor's share of national income in 16 industrialized nations during the period 1960–1995. Labor's share of national income and other variables are differenced because there is evidence for a unit root in these variables. Using the error-correction model, Kristal reports that labor's share of national income is largely explained by indicators for working-class organizational power in economic and political spheres, working-class power in the global sphere, and working-class integration in the intraclass sphere.

Jacobs and Kent (2007) investigate what factors explain executions in the United States. The dependent variables in their regression study are the number of yearly executions from 1951–1998 and the percentage of respondents who support the death penalty in surveys. Both of them are found to have a unit root and are differenced in their regressions. According to Jacobs and Kent, economic inequality and Republican political strength in the states lead to both greater public support and increased executions, whereas civil rights protests reduce public support for capital punishment.

1.5 Technical Tools

This section briefly introduces some key technical tools that have been used in the literature on unit roots. No doubt, many other technical tools have been used, but discussing all of them is beyond the scope of this book. For more details, the reader is referred to such monographs as Davidson (1994) and Serfling (1980).

1.5.1 Brownian Motion

A continuous-time stochastic process, $\{W(r), 0 \leq r \leq 1\}$, is called Brownian motion or a Wiener process if it satisfies the following conditions.

(i) $W(0) = 0$, almost surely.
(ii) For $0 \leq t_0 \leq t_1 \leq \ldots \leq t_k$, $W(t_1) - W(t_0), \ldots, W(t_k) - W(t_{k-1})$ are independent.
(iii) $W(t) - W(s)$ $(t > s)$ follows $\mathbf{N}(0, t - s)$.

Brownian motion has been used to represent limiting distributions of estimators and test statistics in the literature on unit roots since Solo (1984), Phillips (1987a), and Chan and Wei (1988) started using it. White (1958) also used it to represent the limiting distribution of the OLS estimator of the AR coefficient for the AR(1) model with a unit root, but he did not provide a formal proof for the representation.

1.5.2 Functional Central Limit Theorem

Let $S_t = \sum_{i=1}^{t} u_i$, where $u_t \sim \text{iid}(0, \sigma_u^2)$. The functional central limit theorem (FCLT) or invariance principle of Donsker (1951) states that

$$X_T(r) = \frac{1}{\sqrt{T}\sigma_u} S_{[Tr]} \Rightarrow W(r)\,(0 \leq r \leq 1) \text{ as } T \to \infty,$$

where $[Tr]$ denotes the integer part of Tr and $\Rightarrow$ weak convergence. In the FCLT, $X_T(r)$ and its weak limits are functions of r unlike in the usual central limit theorem. When $r = 1$, this becomes the Lindeberg-Lévy central limit theorem.

There are many extensions of Donsker's FCLT. For example, Billingsley (1999) and Herrndorf (1984) extend Donsker's FCLT to the cases of stationary and ergodic innovations and of weakly dependent and heterogeneously distributed innovations, respectively. The latter type of innovation has often been used in the literature on unit roots since Phillips (1987a) introduced it to the econometrics literature.

1.5.3 Continuous-Mapping Theorem

Suppose that $X_T \Rightarrow X$ as $T \to \infty$, where X_T is a sequence of random vectors and X a random vector. The continuous mapping theorem states that $g(X_T) \Rightarrow g(X)$ as $T \to \infty$ where the function $g(\cdot)$ is continuous with probability one. This theorem has often been used to derive limiting distributions of estimators and test statistics in the literature on unit roots.

1.5.4 Stochastic Integrals

The stochastic, or Itô integral, of the form $\int_0^1 W(r)\,\mathrm{d}W(r)$ appears often throughout this book. This integral does not allow for the use of the usual formula for the Riemann–Stieltjes integral and is not equal to $\left[\frac{1}{2}W^2(r)\right]_0^1 = \frac{1}{2}W^2(1)$. The reason for this is the excessive variability of $W(r)$. In other words, $W(r)$ is not of bounded variation so that the Riemann–Stieltjes integral does not exist with probability one. Stochastic integrals that include $\int_0^1 W(r)\,\mathrm{d}W(r)$ as a special case are constructed using high-level probability theory as in, for example, Karatzas and Shreve (1991). To evaluate the stochastic integral $\int_0^1 W(r)\,\mathrm{d}W(r)$, we need to use Itô's rule,

$$g(W(t)) - g(0) = \int_0^t g'(W(r))\,\mathrm{d}W(r) + \frac{1}{2}\int_0^t g''(W(r))\,\mathrm{d}r,$$

where $g : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable. For $g(W(r)) = W^2(r)$ and $t = 1$, this formula gives

$$\int_0^1 W(r)\,\mathrm{d}W(r) = \frac{1}{2}(W^2(1) - 1) = \frac{1}{2}(\chi^2(1) - 1).$$

1.5.5 Other Integrals Involving $W(r)$

For a nonstochastic function $G : \mathbb{R} \to \mathbb{R}$ with the property $\int_0^t G^2(r)\,\mathrm{d}r < \infty$, the integral $\int_0^t G(r)\,\mathrm{d}W(r)$ is equal in distribution to $\mathbf{N}(0, \int_0^t G^2(r)\,\mathrm{d}r)$. The proof for this result can be found in Arnold (1974, p. 77). For $G(r) = r$ and $t = 1$, this gives $\int_0^1 r\,\mathrm{d}W(r) = \mathbf{N}(0, \frac{1}{3})$. Likewise, $\int_0^1 \mathrm{d}W(r) = \mathbf{N}(0, 1)$.

We also use integrals such as $\int_0^1 W(r)\,\mathrm{d}r$, $\int_0^1 W^2(r)\,\mathrm{d}r$, and $\int_0^1 r W(r)\,\mathrm{d}r$. These are the usual Riemann–Stieltjes integrals for a fixed elementary event. Therefore, separate constructions of these integrals are not needed. Banerjee, Dolado, Galbraith, and Hendry (1993, p. 91) show that $\int_0^1 W(r)\,\mathrm{d}r = \mathbf{N}(0, \frac{1}{3})$ and that $\int_0^1 r W(r)\,\mathrm{d}r = \mathbf{N}(0, \frac{2}{15})$. The former result can also be proven using transformed Brownian motion as in Davidson (1994, p. 488).

1.6 Outline of Subsequent Chapters

This section outlines topics discussed in the subsequent chapters. Chapter 2 introduces basic methods for the inference on unit roots. It starts from the AR(1) model with a unit root or a near unit root and then introduces the AR and ARMA regression results with fractionally integrated errors. Dickey and Fuller's (1979) test for a unit root and its extensions are discussed next. Because Dickey and Fuller's (1979) test and its extensions are perceived to have low power, various attempts have been made to improve the power of unit root tests. Research results deriving from these efforts are introduced. Asymptotic theory for the AR models with negative and complex unit roots is also reported in this chapter.

Chapter 3 introduces inferential procedures for a unit root under model specifications that are different from the standard AR model. The topics of this chapter are unit root tests under structural changes in the nonstochastic regressors and the innovation variance, unit root tests with conditional heteroskedasticity, unit root tests in the presence of additive and innovational outliers, unit root distributions and testing under fat-tailed distributions, and unit root tests against nonlinear alternatives.

Chapter 4 introduces unit root tests against the alternatives of fractional integration, regression methods for the AR model that are robust to outliers, model-free tests for a unit root, bootstrapping methods, Bayesian inferential methods for a unit root, tests that take stationarity with or without structural changes as the null hypothesis, and tests for changing persistence.

Chapter 5 introduces a smorgasbord of topics that are relevant to unit roots, but are inappropriate to be included in the previous chapters. These include model selection, interval and point estimation for the AR model possibly with a unit root, improved estimation methods for the AR(1) model, distribution theory for the AR(1) model with unit roots, sampling frequency and tests for a unit root, and the effects of seasonal adjustments on unit root testing.

Chapter 6 is mostly about testing for seasonal unit roots, but it also discusses the periodic AR model, which is regarded as a viable alternative to the seasonal ARIMA model. The topics discussed in Chapter 6 include tests for seasonal unit roots, seasonal stationarity tests, seasonal unit root and stationarity tests under structural changes, periodic integration, and empirical evidence on seasonal unit roots.

Chapter 7 is about panel unit roots. It discusses methods of testing for unit roots and for stationarity using panel data. Topics discussed are unit root tests for independent panels, panel tests for the null of stationarity, unit root and stationarity tests under structural changes, unit root tests for cross-sectionally correlated panels, stationarity tests for cross-sectionally correlated panels, tests for seasonal panel unit roots, simulation studies, and miscellaneous related studies. Research results reported in Chapter 7 are newer than those in previous chapters, and the methods discussed there are being used extensively today.

In this book, unless otherwise stated, all the limits are taken under $T \to \infty$, with T denoting the number of time series observations. In Chapter 7, $N \to \infty$ is also used, where N is the number of cross-sectional observations.