4
Unit Root Tests

4.1 Introduction

Many economic and financial time series exhibit trending behavior or non-stationarity in the mean. Leading examples are asset prices, exchange rates and the levels of macroeconomic aggregates like real GDP. An important econometric task is determining the most appropriate form of the trend in the data. For example, in ARMA modeling the data must be transformed to stationary form prior to analysis. If the data are trending, then some form of trend removal is required.

Two common trend removal or de-trending procedures are first differencing and time-trend regression. First differencing is appropriate for $I(1)$ time series and time-trend regression is appropriate for trend stationary $I(0)$ time series. Unit root tests can be used to determine if trending data should be first differenced or regressed on deterministic functions of time to render the data stationary. Moreover, economic and finance theory often suggests the existence of long-run equilibrium relationships among nonstationary time series variables. If these variables are $I(1)$, then cointegration techniques can be used to model these long-run relations. Hence, pre-testing for unit roots is often a first step in the cointegration modeling discussed in Chapter 12. Finally, a common trading strategy in finance involves exploiting mean-reverting behavior among the prices of pairs of assets. Unit root tests can be used to determine which pairs of assets appear to exhibit mean-reverting behavior.

This chapter is organized as follows. Section 4.2 reviews $I(1)$ and trend stationary $I(0)$ time series and motivates the unit root and stationary tests described in the chapter. Section 4.3 describes the class of autoregressive unit root tests made popular by David Dickey, Wayne Fuller, Pierre Perron and Peter Phillips. Section 4.4 describes the stationarity tests of Kwiatkowski, Phillips, Schmidt and Shinn (1992). Section 4.5 discusses some problems associated with traditional unit root and stationarity tests, and Section 4.6 presents some recently developed so-called "efficient unit root tests" that overcome some of the deficiencies of traditional unit root tests.

In this chapter, the technical details of unit root and stationarity tests are kept to a minimum. Excellent technical treatments of nonstationary time series may be found in Hamilton (1994), Hatanaka (1995), Fuller (1996) and the many papers by Peter Phillips. Useful surveys on issues associated with unit root testing are given in Stock (1994), Maddala and Kim (1998) and Phillips and Xiao (1998).

4.2 Testing for Nonstationarity and Stationarity

To understand the econometric issues associated with unit root and stationarity tests, consider the stylized trend-cycle decomposition of a time series y_t:

$$\begin{aligned} y_t &= TD_t + z_t \\ TD_t &= \kappa + \delta t \\ z_t &= \phi z_{t-1} + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma^2) \end{aligned}$$

where TD_t is a deterministic linear trend and z_t is an AR(1) process. If $|\phi| < 1$ then y_t is $I(0)$ about the deterministic trend TD_t. If $\phi = 1$, then $z_t = z_{t-1} + \varepsilon_t = z_0 + \sum_{j=1}^{t} \varepsilon_j$, a stochastic trend and y_t is $I(1)$ with drift. Simulated $I(1)$ and $I(0)$ data with $\kappa = 5$ and $\delta = 0.1$ are illustrated in Figure 4.1. The $I(0)$ data with trend follows the trend $TD_t = 5 + 0.1t$ very closely and exhibits trend reversion. In contrast, the $I(1)$ data follows an upward drift but does not necessarily revert to TD_t.

Autoregressive *unit root tests* are based on testing the null hypothesis that $\phi = 1$ (difference stationary) against the alternative hypothesis that $\phi < 1$ (trend stationary). They are called unit root tests because under the null hypothesis the autoregressive polynomial of z_t, $\phi(z) = (1 - \phi z) = 0$, has a root equal to unity.

Stationarity tests take the null hypothesis that y_t is trend stationary. If y_t is then first differenced it becomes

$$\begin{aligned} \Delta y_t &= \delta + \Delta z_t \\ \Delta z_t &= \phi \Delta z_{t-1} + \varepsilon_t - \varepsilon_{t-1} \end{aligned}$$

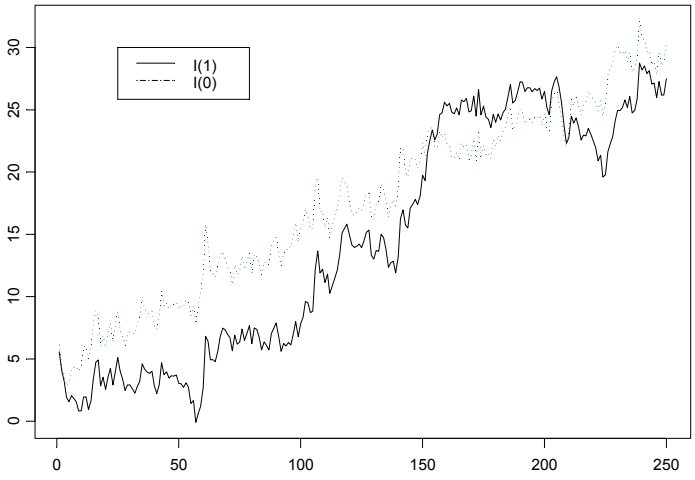

FIGURE 4.1. Simulated trend stationary ($I(0)$) and difference stationary ($I(1)$) processes.

Notice that first differencing y_t, when it is trend stationary, produces a unit moving average root in the ARMA representation of Δz_t. That is, the ARMA representation for Δz_t is the non-invertible ARMA(1,1) model

$$\Delta z_t = \phi \Delta z_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$$

with $\theta = -1$. This result is known as *overdifferencing*. Formally, stationarity tests are based on testing for a unit moving average root in Δz_t.

Unit root and stationarity test statistics have nonstandard and nonnormal asymptotic distributions under their respective null hypotheses. To complicate matters further, the limiting distributions of the test statistics are affected by the inclusion of deterministic terms in the test regressions. These distributions are functions of standard Brownian motion (Wiener process), and critical values must be tabulated by simulation techniques. MacKinnon (1996) provides response surface algorithms for determining these critical values, and various `S+FinMetrics` functions use these algorithms for computing critical values and p-values.

4.3 Autoregressive Unit Root Tests

To illustrate the important statistical issues associated with autoregressive unit root tests, consider the simple AR(1) model

$$y_t = \phi y_{t-1} + \varepsilon_t, \text{ where } \varepsilon_t \sim WN(0, \sigma^2)$$

The hypotheses of interest are

$$\begin{aligned} H_0 &: \phi = 1 \text{ (unit root in } \phi(z) = 0) \Rightarrow y_t \sim I(1) \\ H_1 &: |\phi| < 1 \Rightarrow y_t \sim I(0) \end{aligned}$$

The test statistic is

$$t_{\phi=1} = \frac{\hat{\phi} - 1}{SE(\hat{\phi})}$$

where $\hat{\phi}$ is the least squares estimate and $SE(\hat{\phi})$ is the usual standard error estimate[1]. The test is a one-sided left tail test. If $\{y_t\}$ is stationary (i.e., $|\phi| < 1$) then it can be shown (c.f. Hamilton (1994) pg. 216)

$$\sqrt{T}(\hat{\phi} - \phi) \xrightarrow{d} N(0, (1 - \phi^2))$$

or

$$\hat{\phi} \overset{A}{\sim} N\left(\phi, \frac{1}{T}(1 - \phi^2)\right)$$

and it follows that $t_{\phi=1} \overset{A}{\sim} N(0, 1)$. However, under the null hypothesis of nonstationarity the above result gives

$$\hat{\phi} \overset{A}{\sim} N(1, 0)$$

which clearly does not make any sense. The problem is that under the unit root null, $\{y_t\}$ is not stationary and ergodic, and the usual sample moments do not converge to fixed constants. Instead, Phillips (1987) showed that the sample moments of $\{y_t\}$ converge to random functions of Brownian

[1]The AR(1) model may be re-written as $\Delta y_t = \pi y_{t-1} + u_t$ where $\pi = \phi - 1$. Testing $\phi = 1$ is then equivalent to testing $\pi = 0$. Unit root tests are often computed using this alternative regression and the `S+FinMetrics` function `unitroot` follows this convention.

motion[2]:

$$T^{-3/2}\sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \sigma \int_0^1 W(r)dr$$
$$T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 W(r)^2 dr$$
$$T^{-1}\sum_{t=1}^{T} y_{t-1}\varepsilon_t \xrightarrow{d} \sigma^2 \int_0^1 W(r)dW(r)$$

where $W(r)$ denotes a standard Brownian motion (Wiener process) defined on the unit interval. Using the above results Phillips showed that under the unit root null $H_0 : \phi = 1$

$$T(\hat{\phi} - 1) \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr} \tag{4.1}$$

$$t_{\phi=1} \xrightarrow{d} \frac{\int_0^1 W(r)dW(r)}{\left(\int_0^1 W(r)^2 dr\right)^{1/2}} \tag{4.2}$$

The above yield some surprising results:

- $\hat{\phi}$ is *super-consistent*; that is, $\hat{\phi} \xrightarrow{p} \phi$ at rate T instead of the usual rate $T^{1/2}$.
- $\hat{\phi}$ is not asymptotically normally distributed and $t_{\phi=1}$ is not asymptotically standard normal.
- The limiting distribution of $t_{\phi=1}$ is called the Dickey-Fuller (DF) distribution and does not have a closed form representation. Consequently, quantiles of the distribution must be computed by numerical approximation or by simulation[3].
- Since the *normalized bias* $T(\hat{\phi} - 1)$ has a well defined limiting distribution that does not depend on nuisance parameters it can also be used as a test statistic for the null hypothesis $H_0 : \phi = 1$.

[2] A Wiener process $W(\cdot)$ is a continuous-time stochastic process, associating each date $r \in [0, 1]$ a scalar random variable $W(r)$ that satisfies: (1) $W(0) = 0$; (2) for any dates $0 \leq t_1 \leq \cdots \leq t_k \leq 1$ the changes $W(t_2) - W(t_1), W(t_3) - W(t_2), \ldots, W(t_k) - W(t_{k-1})$ are independent normal with $W(s) - W(t) \sim N(0, (s - t))$; (3) $W(s)$ is continuous in s.

[3] Dickey and Fuller (1979) first considered the unit root tests and derived the asymptotic distribution of $t_{\phi=1}$. However, their representation did not utilize functions of Wiener processes.

4.3.1 Simulating the DF and Normalized Bias Distributions

As mentioned above, the DF and normalized bias distributions must be obtained by simulation methods. To illustrate, the following S-PLUS function `wiener` produces one random draw from the functions of Brownian motion that appear in the limiting distributions of $t_{\phi=1}$ and $T(\hat{\phi}-1)$:

```
wiener = function(nobs) {
   e = rnorm(nobs)
   y = cumsum(e)
   ym1 = y[1:(nobs-1)]
   intW2 = nobs^(-2) * sum(ym1^2)
   intWdW = nobs^(-1) * sum(ym1*e[2:nobs])
   ans = list(intW2=intW2,
              intWdW=intWdW)
   ans
}
```

A simple loop then produces the simulated distributions:

```
> nobs = 1000
> nsim = 1000
> NB = rep(0,nsim)
> DF = rep(0,nsim)
> for (i in 1:nsim) {
+    BN.moments = wiener(nobs)
+    NB[i] = BN.moments$intWdW/BN.moments$intW2
+    DF[i] = BN.moments$intWdW/sqrt(BN.moments$intW2)
}
```

Figure 4.2 shows the histograms and density estimates of the simulated distributions. The DF density is slightly left-skewed relative to the standard normal, and the normalized bias density is highly left skewed and non-normal. Since the alternative is one-sided, the test statistics reject if they are sufficiently negative. For the DF and normalized bias densities the empirical 1%, 5% and 10% quantiles are

```
> quantile(DF,probs=c(0.01,0.05,0.1))
     1%     5%    10%
 -2.451 -1.992 -1.603
> quantile(NB,probs=c(0.01,0.05,0.1))
     1%    5%    10%
 -11.94 -8.56 -5.641
```

For comparison purposes, note that the 5% quantile from the standard normal distribution is -1.645.

The simulation of critical values and p-values from (4.1) and (4.2) is straightforward but time consuming. The `punitroot` and `qunitroot` func-

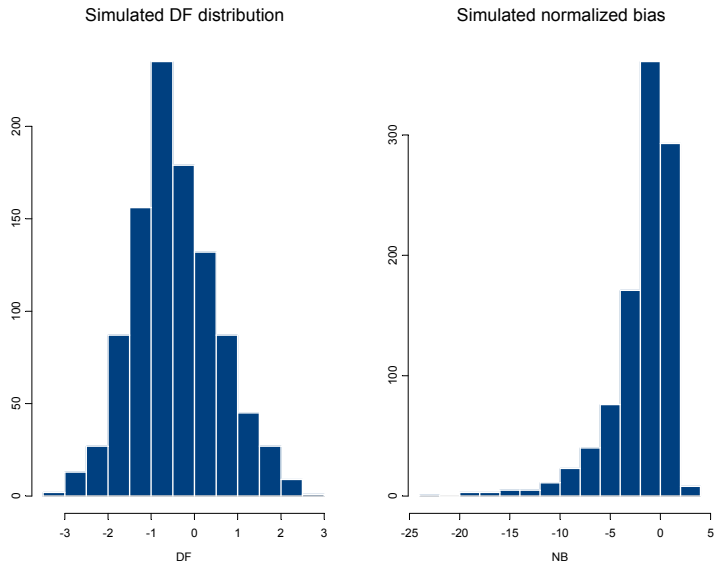

FIGURE 4.2. Histograms of simulated DF and normalized bias distributions.

tions in `S+FinMetrics` produce p-values and quantiles of the DF and normalized bias distributions based on MacKinnon's (1996) response surface methodology. The advantage of the response surface methodology is that accurate p-values and quantiles appropriate for a given sample size can be produced. For example, the 1%, 5% and 10% quantiles for (4.2) and (4.1) based on a sample size of 100 are

```
> qunitroot(c(0.01,0.05,0.10), trend="nc", statistic="t",
+ n.sample=100)
[1] -2.588 -1.944 -1.615
> qunitroot(c(0.01,0.05,0.10), trend="nc", statistic="n",
+ n.sample=100)
[1] -13.086  -7.787  -5.565
```

The argument `trend="nc"` specifies that no constant is included in the test regression. Other valid options are `trend="c"` for constant only and `trend="ct"` for constant and trend. These trend cases are explained below. To specify the normalized bias distribution, set `statistic="n"`. For asymptotic quantiles set `n.sample=0`.

Similarly, the p-value of -1.645 based on the DF distribution for a sample size of 100 is computed as

```
> punitroot(-1.645, trend="nc", statistic="t")
[1] 0.0945
```

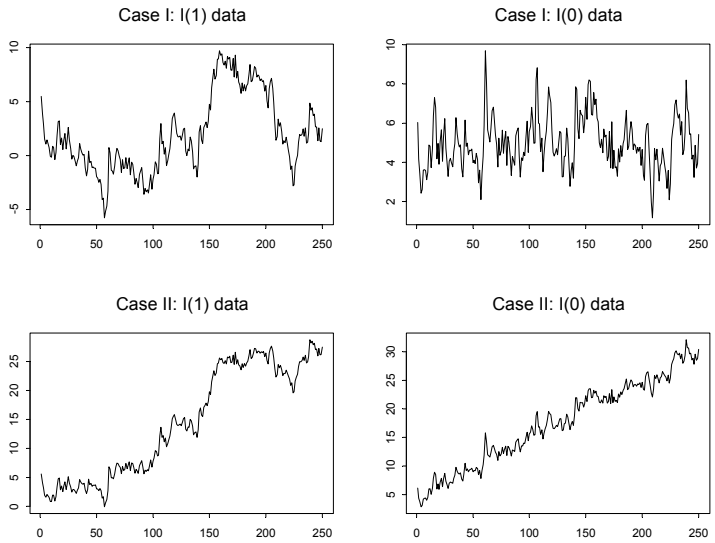

FIGURE 4.3. Simulated $I(1)$ and $I(0)$ data under trend cases I and II.

4.3.2 Trend Cases

When testing for unit roots, it is crucial to specify the null and alternative hypotheses appropriately to characterize the trend properties of the data at hand. For example, if the observed data does not exhibit an increasing or decreasing trend, then the appropriate null and alternative hypotheses should reflect this. The trend properties of the data *under the alternative hypothesis* will determine the form of the test regression used. Furthermore, the type of deterministic terms in the test regression will influence the asymptotic distributions of the unit root test statistics. The two most common trend cases are summarized below and illustrated in Figure 4.3.

Case I: Constant Only

The test regression is

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

and includes a constant to capture the nonzero mean under the alternative. The hypotheses to be tested are

$$\begin{aligned} H_0 &: \phi = 1 \Rightarrow y_t \sim I(1) \text{ without drift} \\ H_1 &: |\phi| < 1 \Rightarrow y_t \sim I(0) \text{ with nonzero mean} \end{aligned}$$

This formulation is appropriate for non-trending financial series like interest rates, exchange rates, and spreads. The test statistics $t_{\phi=1}$ and $T(\hat{\phi} - 1)$

are computed from the above regression. Under $H_0 : \phi = 1$ the asymptotic distributions of these test statistics are different from (4.2) and (4.1) and are influenced by the presence but not the coefficient value of the constant in the test regression. Quantiles and p-values for these distributions can be computed using the `S+FinMetrics` functions `punitroot` and `qunitroot` with the `trend="c"` option:

```
> qunitroot(c(0.01,0.05,0.10), trend="c", statistic="t",
+ n.sample=100)
[1] -3.497 -2.891 -2.582
> qunitroot(c(0.01,0.05,0.10), trend="c", statistic="n",
+ n.sample=100)
[1] -19.49 -13.53 -10.88
> punitroot(-1.645, trend="c", statistic="t", n.sample=100)
[1] 0.456
> punitroot(-1.645, trend="c", statistic="n", n.sample=100)
[1] 0.8172
```

For a sample size of 100, the 5% left tail critical values for $t_{\phi=1}$ and $T(\hat{\phi} - 1)$ are -2.891 and -13.53, respectively, and are quite a bit smaller than the 5% critical values computed when `trend="nc"`. Hence, inclusion of a constant pushes the distributions of $t_{\phi=1}$ and $T(\hat{\phi} - 1)$ to the left.

Case II: Constant and Time Trend

The test regression is

$$y_t = c + \delta t + \phi y_{t-1} + \varepsilon_t$$

and includes a constant and deterministic time trend to capture the deterministic trend under the alternative. The hypotheses to be tested are

$$\begin{aligned} H_0 &: \phi = 1 \Rightarrow y_t \sim I(1) \text{ with drift} \\ H_1 &: |\phi| < 1 \Rightarrow y_t \sim I(0) \text{ with deterministic time trend} \end{aligned}$$

This formulation is appropriate for trending time series like asset prices or the levels of macroeconomic aggregates like real GDP. The test statistics $t_{\phi=1}$ and $T(\hat{\phi} - 1)$ are computed from the above regression. Under $H_0 : \phi = 1$ the asymptotic distributions of these test statistics are different from (4.2) and (4.1) and are influenced by the presence but not the coefficient values of the constant and time trend in the test regression. Quantiles and p-values for these distributions can be computed using the `S+FinMetrics` functions `punitroot` and `qunitroot` with the `trend="ct"` option:

```
> qunitroot(c(0.01,0.05,0.10), trend="ct", statistic="t",
+ n.sample=100)
[1] -4.052 -3.455 -3.153
```

```
> qunitroot(c(0.01,0.05,0.10), trend="ct", statistic="n",
+ n.sample=100)
[1] -27.17 -20.47 -17.35
> punitroot(-1.645, trend="ct", statistic="t", n.sample=100)
[1] 0.7679
> punitroot(-1.645, trend="ct", statistic="n", n.sample=100)
[1] 0.9769
```

Notice that the inclusion of a constant and trend in the test regression further shifts the distributions of $t_{\phi=1}$ and $T(\hat{\phi}-1)$ to the left. For a sample size of 100, the 5% left tail critical values for $t_{\phi=1}$ and $T(\hat{\phi}-1)$ are now -3.455 and -20.471.

4.3.3 Dickey-Fuller Unit Root Tests

The unit root tests described above are valid if the time series y_t is well characterized by an AR(1) with white noise errors. Many financial time series, however, have a more complicated dynamic structure than is captured by a simple AR(1) model. Said and Dickey (1984) augment the basic autoregressive unit root test to accommodate general ARMA(p, q) models with unknown orders and their test is referred to as the *augmented Dickey-Fuller* (ADF) test. The ADF test tests the null hypothesis that a time series y_t is $I(1)$ against the alternative that it is $I(0)$, assuming that the dynamics in the data have an ARMA structure. The ADF test is based on estimating the test regression

$$y_t = \boldsymbol{\beta}'\mathbf{D}_t + \phi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t \tag{4.3}$$

where $\mathbf{D}_t$ is a vector of deterministic terms (constant, trend etc.). The p lagged difference terms, Δy_{t-j}, are used to approximate the ARMA structure of the errors, and the value of p is set so that the error ε_t is serially uncorrelated. The error term is also assumed to be homoskedastic. The specification of the deterministic terms depends on the assumed behavior of y_t under the alternative hypothesis of trend stationarity as described in the previous section. Under the null hypothesis, y_t is $I(1)$ which implies that $\phi = 1$. The ADF t-statistic and normalized bias statistic are based on the least squares estimates of (4.3) and are given by

$$\begin{aligned} ADF_t &= t_{\phi=1} = \frac{\hat{\phi}-1}{SE(\phi)} \\ ADF_n &= \frac{T(\hat{\phi}-1)}{1-\hat{\psi}_1-\cdots-\hat{\psi}_p} \end{aligned}$$

An alternative formulation of the ADF test regression is

$$\Delta y_t = \boldsymbol{\beta}'\mathbf{D}_t + \pi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \varepsilon_t \tag{4.4}$$

where $\pi = \phi - 1$. Under the null hypothesis, Δy_t is $I(0)$ which implies that $\pi = 0$. The ADF t-statistic is then the usual t-statistic for testing $\pi = 0$ and the ADF normalized bias statistic is $T\hat{\pi}/(1 - \hat{\psi}_1 - \cdots - \hat{\psi}_p)$. The test regression (4.4) is often used in practice because the ADF t-statistic is the usual t-statistic reported for testing the significance of the coefficient y_{t-1}. The `S+FinMetrics` function `unitroot` follows this convention.

Choosing the Lag Length for the ADF Test

An important practical issue for the implementation of the ADF test is the specification of the lag length p. If p is too small then the remaining serial correlation in the errors will bias the test. If p is too large then the power of the test will suffer. Ng and Perron (1995) suggest the following data dependent lag length selection procedure that results in stable size of the test and minimal power loss. First, set an upper bound $p_{\max}$ for p. Next, estimate the ADF test regression with $p = p_{\max}$. If the absolute value of the t-statistic for testing the significance of the last lagged difference is greater than 1.6 then set $p = p_{\max}$ and perform the unit root test. Otherwise, reduce the lag length by one and repeat the process.

A useful rule of thumb for determining $p_{\max}$, suggested by Schwert (1989), is

$$p_{\max} = \left[12 \cdot \left(\frac{T}{100} \right)^{1/4} \right] \tag{4.5}$$

where $[x]$ denotes the integer part of x. This choice allows $p_{\max}$ to grow with the sample so that the ADF test regressions (4.3) and (4.4) are valid if the errors follow an ARMA process with unknown order.

Example 19 *Testing for a unit root in exchange rate data using ADF tests*

To illustrate the ADF test procedure, consider testing for a unit root in the logarithm of the US/CA monthly spot exchange rate, denoted s_t, over the 30 year period 1976 - 1996. Figure 4.4 shows $s_t, \Delta s_t$ as well as the sample autocorrelations for these series. The data and plots are created with the `S-PLUS` commands

```
> uscn.spot = lexrates.dat[,"USCNS"]
> uscn.spot@title = "Log US/CN spot exchange rate"
> par(mfrow=c(2,2))
> plot.timeSeries(uscn.spot, reference.grid=F,
+ main="Log of US/CN spot exchange rate")
```

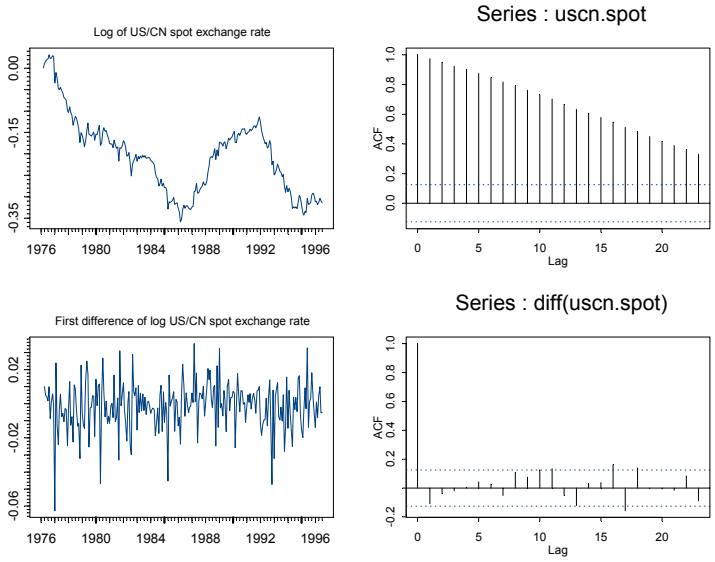

FIGURE 4.4. US/CN spot rate, first difference and SACF.

```
> xx = acf(uscn.spot)
> plot.timeSeries(diff(uscn.spot), reference.grid=F,
+ main="First difference of log US/CN spot exchange rate")
> xx = acf(diff(uscn.spot))
```

Clearly, s_t exhibits random walk like behavior with no apparent positive or negative drift. However, Δs_t behaves very much like a white noise process. The appropriate trend specification is to include a constant in the test regression. Regarding the maximum lag length for the Ng-Perron procedure, given the lack of serial correlation in Δs_t a conservative choice is $p_{\max} = 6$. The ADF t-statistic computed from the test regression with a constant and $p = 6$ lags can be computed using the `S+FinMetrics` function `unitroot` as follows

```
> adft.out = unitroot(uscn.spot, trend="c", statistic="t",
+ method="adf", lags=6)
> class(adft.out)
[1] "unitroot"
```

The output of `unitroot` is an object of class "`unitroot`" for which there are `print` and `summary` methods. Typing the name of the object invokes the `print` method and displays the basic test result

```
> adft.out
Test for Unit Root: Augmented DF Test
```

```
Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -2.6
        P-value: 0.09427

Coefficients:
    lag1    lag2    lag3    lag4    lag5    lag6 constant
 -0.0280 -0.1188 -0.0584 -0.0327 -0.0019  0.0430 -0.0075

Degrees of freedom: 239 total; 232 residual
Time period: from Aug 1976 to Jun 1996
Residual standard error: 0.01386
```

With $p = 6$ the ADF t-statistic is -2.6 and has a p-value (computed using `punitroot`) of 0.094. Hence we do not reject the unit root null at the 9.4% level. The small p-value here may be due to the inclusion of superfluous lags. To see the significance of the lags in the test regression, use the `summary` method

```
> summary(adft.out)
Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t test
 Test Statistic: -2.6
        P-value: 0.09427

Coefficients:
           Value Std. Error t value Pr(>|t|)
    lag1 -0.0280  0.0108    -2.6004  0.0099
    lag2 -0.1188  0.0646    -1.8407  0.0669
    lag3 -0.0584  0.0650    -0.8983  0.3700
    lag4 -0.0327  0.0651    -0.5018  0.6163
    lag5 -0.0019  0.0651    -0.0293  0.9766
    lag6  0.0430  0.0645     0.6662  0.5060
constant -0.0075  0.0024    -3.0982  0.0022

Regression Diagnostics:

         R-Squared 0.0462
Adjusted R-Squared 0.0215
Durbin-Watson Stat 2.0033

Residual standard error: 0.01386 on 235 degrees of freedom
F-statistic: 1.874 on 6 and 232 degrees of freedom, the
```

```
p-value is 0.08619
Time period: from Aug 1976 to Jun 1996
```

The results indicate that too many lags have been included in the test regression. Following the Ng-Perron backward selection procedure $p = 2$ lags are selected. The results are

```
> adft.out = unitroot(uscn.spot, trend="c", lags=2)
> summary(adft.out)
Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t test
 Test Statistic: -2.115
        P-value: 0.2392

Coefficients:
           Value Std. Error t value Pr(>|t|)
    lag1 -0.0214  0.0101    -2.1146  0.0355
    lag2 -0.1047  0.0635    -1.6476  0.1007
constant -0.0058  0.0022    -2.6001  0.0099

Regression Diagnostics:

         R-Squared 0.0299
Adjusted R-Squared 0.0218
Durbin-Watson Stat 2.0145

Residual standard error: 0.01378 on 239 degrees of freedom
F-statistic: 3.694 on 2 and 240 degrees of freedom, the
p-value is 0.02629
Time period: from Apr 1976 to Jun 1996
```

With 2 lags the ADF t-statistic is -2.115, the p-value 0.239 and we have greater evidence for a unit root in s_t. A similar result is found with the ADF normalized bias statistic

```
> adfn.out = unitroot(uscn.spot, trend="c", lags=2,
+ statistic="n")
> adfn.out
Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: normalized test
 Test Statistic: -5.193
        P-value: 0.4129
```

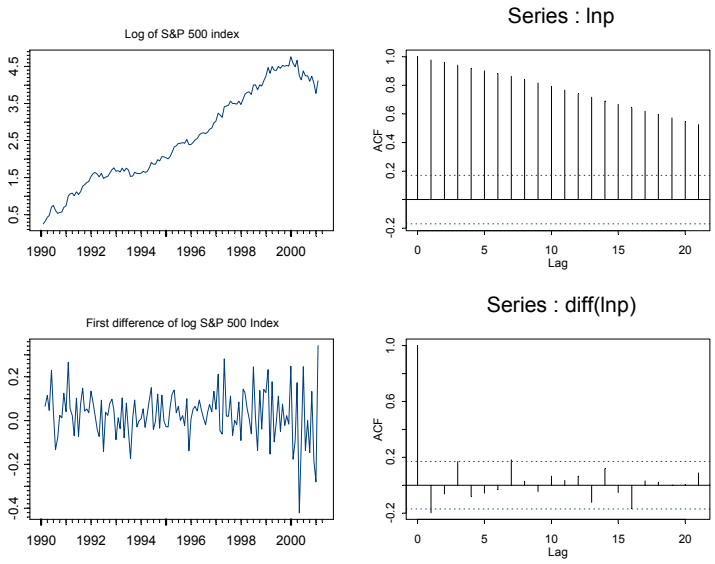

FIGURE 4.5. Log prices on the S&P 500 index, first difference and SACF.

```
Coefficients:
    lag1    lag2 constant
 -0.0214 -0.1047 -0.0058

Degrees of freedom: 243 total; 240 residual
Time period: from Apr 1976 to Jun 1996
Residual standard error: 0.01378
```

Example 20 *Testing for a unit root in log stock prices*

The log levels of asset prices are usually treated as $I(1)$ with drift. Indeed, the random walk model of stock prices is a special case of an $I(1)$ process. Consider testing for a unit root in the log of the monthly S&P 500 index, p_t, over the period January 1990 through January 2001. The data is taken from the `S+FinMetrics` "`timeSeries`" `singleIndex.dat`. The data and various plots are created with the `S-PLUS` commands

```
> lnp = log(singleIndex.dat[,1])
> lnp@title = "Log of S&P 500 Index"
> par(mfrow=c(2,2))
> plot.timeSeries(lnp, reference.grid=F,
+ main="Log of S&P 500 index")
> acf.plot(acf(lnp,plot=F))
> plot.timeSeries(diff(lnp), reference.grid=F,
```

```
+ main="First difference of log S&P 500 Index")
> acf.plot(acf(diff(lnp),plot=F))
```

and are illustrated in Figure 4.5. Clearly, the p_t is nonstationary due to the positive trend. Also, there appears to be some negative autocorrelation at lag one in Δp_t. The null hypothesis to be tested is that p_t is $I(1)$ with drift, and the alternative is that the p_t is $I(0)$ about a deterministic time trend. The ADF t-statistic to test these hypotheses is computed with a constant and time trend in the test regression and four lags of Δp_t (selecting using the Ng-Perron backward selection method)

```
> adft.out = unitroot(lnp, trend="ct", lags=4)
> summary(adft.out)

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
   Type of Test: t test
 Test Statistic: -1.315
        P-value: 0.8798

Coefficients:
           Value Std. Error t value Pr(>|t|)
    lag1 -0.0540  0.0410    -1.3150  0.1910
    lag2 -0.1869  0.0978    -1.9111  0.0583
    lag3 -0.0460  0.0995    -0.4627  0.6444
    lag4  0.1939  0.0971     1.9964  0.0481
constant  0.1678  0.1040     1.6128  0.1094
    time  0.0015  0.0014     1.0743  0.2848

Regression Diagnostics:

         R-Squared 0.1016
Adjusted R-Squared 0.0651
Durbin-Watson Stat 1.9544

Residual standard error: 0.1087 on 125 degrees of freedom
F-statistic: 2.783 on 5 and 123 degrees of freedom, the
p-value is 0.0204
Time period: from May 1990 to Jan 2001
```

$ADF_t = -1.315$ and has a p-value of 0.8798, so one clearly does not reject the null that p_t is $I(1)$ with drift.

4.3.4 Phillips-Perron Unit Root Tests

Phillips and Perron (1988) developed a number of unit root tests that have become popular in the analysis of financial time series. The Phillips-Perron (PP) unit root tests differ from the ADF tests mainly in how they deal with serial correlation and heteroskedasticity in the errors. In particular, where the ADF tests use a parametric autoregression to approximate the ARMA structure of the errors in the test regression, the PP tests ignore any serial correlation in the test regression. The test regression for the PP tests is

$$\Delta y_t = \boldsymbol{\beta}'\mathbf{D}_t + \pi y_{t-1} + u_t$$

where u_t is $I(0)$ and may be heteroskedastic. The PP tests correct for any serial correlation and heteroskedasticity in the errors u_t of the test regression by directly modifying the test statistics $t_{\pi=0}$ and $T\hat{\pi}$. These modified statistics, denoted Z_t and Z_π, are given by

$$\begin{aligned} Z_t &= \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2}\right)^{1/2} \cdot t_{\pi=0} - \frac{1}{2}\left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right) \cdot \left(\frac{T \cdot SE(\hat{\pi})}{\hat{\sigma}^2}\right) \\ Z_\pi &= T\hat{\pi} - \frac{1}{2}\frac{T^2 \cdot SE(\hat{\pi})}{\hat{\sigma}^2}(\hat{\lambda}^2 - \hat{\sigma}^2) \end{aligned}$$

The terms $\hat{\sigma}^2$ and $\hat{\lambda}^2$ are consistent estimates of the variance parameters

$$\begin{aligned} \sigma^2 &= \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E[u_t^2] \\ \lambda^2 &= \lim_{T\to\infty} \sum_{t=1}^{T} E\left[T^{-1}S_T^2\right] \end{aligned}$$

where $S_T = \sum_{t=1}^{T} u_t$. The sample variance of the least squares residual $\hat{u}_t$ is a consistent estimate of σ^2, and the Newey-West long-run variance estimate of u_t using $\hat{u}_t$ is a consistent estimate of λ^2.

Under the null hypothesis that $\pi = 0$, the PP Z_t and Z_π statistics have the same asymptotic distributions as the ADF t-statistic and normalized bias statistics. One advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroskedasticity in the error term u_t. Another advantage is that the user does not have to specify a lag length for the test regression.

Example 21 *Testing for a unit root in exchange rates using the PP tests*

Recall the arguments for the `S+FinMetrics` `unitroot` function are

```
> args(unitroot)
function(x, trend = "c", method = "adf",
```

```
statistic = "t",lags = 1, bandwidth = NULL,
window = "bartlett", asymptotic = F, na.rm = F)
```

The PP statistics may be computed by specifying the optional argument `method="pp"`. When `method="pp"` is chosen, the argument `window` specifies the weight function and the argument `bandwidth` determines the lag truncation parameter used in computing the long-run variance parameter λ^2. The default bandwidth is the integer part of $(4 \cdot (T/100))^{2/9}$ where T is the sample size.

Now, consider testing for a unit root in the log of the US/CN spot exchange rate using the PP Z_t and Z_π statistics:

```
> unitroot(uscn.spot, trend="c", method="pp")
Test for Unit Root: Phillips-Perron Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -1.97
        P-value: 0.2999

Coefficients:
    lag1 constant
 -0.0202 -0.0054

Degrees of freedom: 244 total; 242 residual
Time period: from Mar 1976 to Jun 1996
Residual standard error: 0.01383

> unitroot(uscn.spot, trend="c", method="pp", statistic="n")
Test for Unit Root: Phillips-Perron Test

Null Hypothesis: there is a unit root
   Type of Test: normalized test
 Test Statistic: -4.245
        P-value: 0.5087

Coefficients:
    lag1 constant
 -0.0202 -0.0054

Degrees of freedom: 244 total; 242 residual
Time period: from Mar 1976 to Jun 1996
Residual standard error: 0.01383
```

As with the ADF tests, the PP tests do not reject the null that the log of the US/CN spot rate is $I(1)$ at any reasonable significance level.

4.4 Stationarity Tests

The ADF and PP unit root tests are for the null hypothesis that a time series y_t is $I(1)$. Stationarity tests, on the other hand, are for the null that y_t is $I(0)$. The most commonly used stationarity test, the KPSS test, is due to Kwiatkowski, Phillips, Schmidt and Shin (1992). They derive their test by starting with the model

$$\begin{aligned} y_t &= \boldsymbol{\beta}'\mathbf{D}_t + \mu_t + u_t \\ \mu_t &= \mu_{t-1} + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \end{aligned} \tag{4.6}$$

where $\mathbf{D}_t$ contains deterministic components (constant or constant plus time trend), u_t is $I(0)$ and may be heteroskedastic. Notice that μ_t is a pure random walk with innovation variance σ_ε^2. The null hypothesis that y_t is $I(0)$ is formulated as $H_0 : \sigma_\varepsilon^2 = 0$, which implies that μ_t is a constant. Although not directly apparent, this null hypothesis also implies a unit moving average root in the ARMA representation of Δy_t. The KPSS test statistic is the Lagrange multiplier (LM) or score statistic for testing $\sigma_\varepsilon^2 = 0$ against the alternative that $\sigma_\varepsilon^2 > 0$ and is given by

$$KPSS = \left(T^{-2} \sum_{t=1}^{T} \hat{S}_t^2 \right) / \hat{\lambda}^2 \tag{4.7}$$

where $\hat{S}_t = \sum_{j=1}^{t} \hat{u}_j$, $\hat{u}_t$ is the residual of a regression of y_t on $\mathbf{D}_t$ and $\hat{\lambda}^2$ is a consistent estimate of the long-run variance of u_t using $\hat{u}_t$. Under the null that y_t is $I(0)$, Kwiatkowski, Phillips, Schmidt and Shin show that KPSS converges to a function of standard Brownian motion that depends on the form of the deterministic terms $\mathbf{D}_t$ but not their coefficient values $\boldsymbol{\beta}$. In particular, if $\mathbf{D}_t = 1$ then

$$KPSS \xrightarrow{d} \int_0^1 V_1(r)dr \tag{4.8}$$

where $V_1(r) = W(r) - rW(1)$ and $W(r)$ is a standard Brownian motion for $r \in [0, 1]$. If $\mathbf{D}_t = (1, t)'$ then

$$KPSS \xrightarrow{d} \int_0^1 V_2(r)dr \tag{4.9}$$

where $V_2(r) = W(r) + r(2 - 3r)W(1) + 6r(r^2 - 1)\int_0^1 W(s)ds$. Critical values from the asymptotic distributions (4.8) and (4.9) must be obtained by simulation methods, and these are summarized in Table 4.1.

The stationary test is a one-sided right-tailed test so that one rejects the null of stationarity at the $100 \cdot \alpha\%$ level if the KPSS test statistic (4.7) is greater than the $100 \cdot (1 - \alpha)\%$ quantile from the appropriate asymptotic distribution (4.8) or (4.9).

	Right tail quantiles				
Distribution	0.90	0.925	0.950	0.975	0.99
$\int_0^1 V_1(r)dr$	0.349	0.396	0.446	0.592	0.762
$\int_0^1 V_2(r)dr$	0.120	0.133	0.149	0.184	0.229

TABLE 4.1. Quantiles of the distribution of the KPSS statistic

4.4.1 Simulating the KPSS Distributions

The distributions in (4.8) and (4.9) may be simulated using methods similar to those used to simulate the DF distribution. The following S-PLUS code is used to create the quantiles in Table 4.1:

```
wiener2 = function(nobs) {
   e = rnorm(nobs)
# create detrended errors
   e1 = e - mean(e)
   e2 = residuals(OLS(e~seq(1,nobs)))
# compute simulated Brownian Bridges
   y1 = cumsum(e1)
   y2 = cumsum(e2)
   intW2.1 = nobs^(-2) * sum(y1^2)
   intW2.2 = nobs^(-2) * sum(y2^2)
   ans = list(intW2.1=intW2.1,
              intW2.2=intW2.2)
   ans
}
#
# simulate KPSS distributions
#
> nobs = 1000
> nsim = 10000
> KPSS1 = rep(0,nsim)
> KPSS2 = rep(0,nsim)
> for (i in 1:nsim) {
   BN.moments = wiener2(nobs)
   KPSS1[i] = BN.moments$intW2.1
   KPSS2[i] = BN.moments$intW2.2
}
#
# compute quantiles of distribution
#
> quantile(KPSS1, probs=c(0.90,0.925,0.95,0.975,0.99))
   90.0%   92.5%   95.0%   97.5%   99.0%
 0.34914 0.39634 0.46643 0.59155 0.76174
> quantile(KPSS2, probs=c(0.90,0.925,0.95,0.975,0.99))
```

```
   90.0%  92.5%   95.0%  97.5%   99.0%
0.12003 0.1325 0.14907 0.1841 0.22923
```

Currently, only asymptotic critical values are available for the KPSS test.

4.4.2 Testing for Stationarity Using the S+FinMetrics Function stationaryTest

The S+FinMetrics function stationaryTest may be used to test the null hypothesis that a time series y_t is $I(0)$ based on the KPSS statistic (4.7). The function stationaryTest has arguments

```
> args(stationaryTest)
function(x, trend = "c", bandwidth = NULL, na.rm = F)
```

where x represents a univariate vector or "timeSeries". The argument trend specifies the deterministic trend component in (4.6) and valid choices are "c" for a constant and "ct" for a constant and time trend. The argument bandwidth determines the lag truncation parameter used in computing the long-run variance parameter λ^2. The default bandwidth is the integer part of $(4 \cdot (T/100))^{2/9}$ where T is the sample size. The output of stationaryTest is an object of class "stationaryTest" for which there is only a print method. The use of stationaryTest is illustrated with the following example.

Example 22 *Testing for stationarity in exchange rates*

Consider the US/CN spot exchange data used in the previous examples. To test the null that s_t is $I(0)$, the KPSS statistic is computed using a constant in (4.6):

```
> kpss.out = stationaryTest(uscn.spot, trend="c")
> class(kpss.out)
[1] "stationaryTest"
> kpss.out

Test for Stationarity: KPSS Test

Null Hypothesis: stationary around a constant

Test Statistics:
    USCNS
 1.6411**

 * : significant at 5% level
** : significant at 1% level
```

```
Total Observ.: 245
   Bandwidth : 5
```

The KPSS statistic is 1.641 and is greater than the 99% quantile, 0.762, from Table.4.1. Therefore, the null that s_t is $I(0)$ is rejected at the 1% level.

4.5 Some Problems with Unit Root Tests

The ADF and PP tests are asymptotically equivalent but may differ substantially in finite samples due to the different ways in which they correct for serial correlation in the test regression. In particular, Schwert (1989) finds that if Δy_t has an ARMA representation with a large and negative MA component, then the ADF and PP tests are severely size distorted (reject $I(1)$ null much too often when it is true) and that the PP tests are more size distorted than the ADF tests. Recently, Perron and Ng (1996) have suggested useful modifications to the PP tests to mitigate this size distortion. Caner and Killian (2001) have found similar problems with the KPSS test.

In general, the ADF and PP tests have very low power against $I(0)$ alternatives that are close to being $I(1)$. That is, unit root tests cannot distinguish highly persistent stationary processes from nonstationary processes very well. Also, the power of unit root tests diminish as deterministic terms are added to the test regressions. That is, tests that include a constant and trend in the test regression have less power than tests that only include a constant in the test regression. For maximum power against very persistent alternatives the recent tests proposed by Elliot, Rothenberg and Stock (1996) and Ng and Perron (2001) should be used. These tests are described in the next section.

4.6 Efficient Unit Root Tests

Assume that T observations are generated by

$$y_t = \boldsymbol{\beta}'\mathbf{D}_t + u_t,\ u_t = \phi u_{t-1} + v_t$$

where $\mathbf{D}_t$ represents a vector of deterministic terms, $E[u_0] < \infty$, and v_t is a 1-summable linear process with long-run variance λ^2. Typically $D_t = 1$ or $\mathbf{D}_t = [1, t]$. Consider testing the null hypothesis $\phi = 1$ versus $|\phi| < 1$. If the distribution of the data were known then the Neyman-Pearson Lemma gives the test with best power against any point alternative $\bar{\phi}$. The power of this optimal test plotted as a function of $\bar{\phi}$ gives an upper bound (envelope) for the power of any test based on the same distribution of the

data. An undesirable feature of this power envelope is that it depends on the specific value of $\bar{\phi}$, so that there is no uniformly most power full test that can be used against all alternatives $|\phi| < 1$. However, using asymptotic approximations based on the local-to-unity alternative $\phi = 1 + c/T$, for $c < 0$, Elliot, Rothenberg and Stock (2001) (hereafter ERS) derive a class of test statistics that come very close to the power envelope for a wide range of alternatives. These tests are referred to as *efficient unit root tests*, and they can have substantially higher power than the ADF or PP unit root tests especially when ϕ is close to unity.

4.6.1 Point Optimal Tests

The starting point for the class of efficient tests is the feasible test statistic that is optimal (asymptotically) for the point alternative $\bar{\phi} = 1 - \bar{c}/T$, $\bar{c} < 0$. This test is constructed as follows. Define the $T-$dimensional column vector $\mathbf{y}_\phi$ and $T \times q$ dimensional matrix $\mathbf{D}_\phi$ by

$$\begin{aligned} \mathbf{y}_\phi &= (y_1, y_2 - \phi y_1, \ldots, y_T - \phi y_{T-1})' \\ \mathbf{D}_\phi &= (\mathbf{D}_1', \mathbf{D}_2' - \phi \mathbf{D}_1', \ldots, \mathbf{D}_T' - \phi \mathbf{D}_{T-1}')' \end{aligned}$$

All of the elements of $\mathbf{y}_\phi$ and $\mathbf{D}_\phi$, except the first, are quasi-differenced using the operator $1 - \phi L$. Next, for any value of ϕ, define $S(\phi)$ as the sum of squared residuals from a least squares regression of $\mathbf{y}_\phi$ on $\mathbf{D}_\phi$. That is,

$$S(\phi) = \tilde{\mathbf{y}}_\phi' \tilde{\mathbf{y}}_\phi$$

where $\tilde{\mathbf{y}}_\phi = \mathbf{y}_\phi - \mathbf{D}_\phi \hat{\boldsymbol{\beta}}_\phi$ and $\hat{\boldsymbol{\beta}}_\phi = (\mathbf{D}_\phi' \mathbf{D}_\phi)^{-1} \mathbf{D}_\phi' \mathbf{y}_\phi$. ERS show that the feasible *point optimal unit root test* against the alternative $\bar{\phi} = 1 - \bar{c}/T$ has the form

$$P_T = \left[S(\bar{\phi}) - \bar{\phi} S(1) \right] / \hat{\lambda}^2 \tag{4.10}$$

where $\hat{\lambda}^2$ is a consistent estimate of λ^2. ERS derive the asymptotic distribution of P_T for $D_t = 1$ and $D_t = (1, t)$ and provide asymptotic and finite sample critical values for tests of size 1%, 2.5%, 5% and 10%[4].

Through a series of simulation experiments, ERS discover that if $\bar{\phi} = 1 + \bar{c}/T$ is chosen such that the power of P_T is tangent to the power envelope at 50% power then the overall power of P_T, for a wide range of $\bar{\phi}$ values less than unity, is close to the power envelope. For a given sample size T, the value of $\bar{\phi}$ that results in P_T having 50% power depends on $\bar{c}$ and the form of the deterministic terms in $\mathbf{D}_t$. ERS show that if $D_t = 1$ then $\bar{c} = -7$, and if $\mathbf{D}_t = (1, t)$ then $\bar{c} = -13.5$.

The ERS P_T statistic may be computed using the function `unitroot` with `method="ers"`.

[4] These critical values are given in ERS Table I panels A and B.

4.6.2 DF-GLS Tests

In the construction of the ERS feasible point optimal test (4.10), the unknown parameters $\boldsymbol{\beta}$ of the trend function are efficiently estimated under the alternative model with $\bar{\phi} = 1 + \bar{c}/T$. That is, $\hat{\boldsymbol{\beta}}_{\bar{\phi}} = (\mathbf{D}_{\bar{\phi}}'\mathbf{D}_{\bar{\phi}})^{-1}\mathbf{D}_{\bar{\phi}}'\mathbf{y}_{\bar{\phi}}$. ERS use this insight to derive an efficient version of the ADF t-statistic, which they call the *DF-GLS test*. They construct this t-statistic as follows. First, using the trend parameters $\hat{\boldsymbol{\beta}}_{\bar{\phi}}$ estimated under the alternative, define the detrended data

$$y_t^d = y_t - \hat{\boldsymbol{\beta}}_{\bar{\phi}}'\mathbf{D}_t$$

ERS call this detrending procedure *GLS detrending*[5]. Next, using the GLS detrended data, estimate by least squares the ADF test regression which omits the deterministic terms

$$\Delta y_t^d = \pi y_{t-1}^d + \sum_{j=1}^{p} \psi_j \Delta y_{t-j}^d + \varepsilon_t \tag{4.11}$$

and compute the t-statistic for testing $\pi = 0$. When $D_t = 1$, ERS show that the asymptotic distribution of the DF-GLS test is the same as the ADF t-test, but has higher asymptotic power (against local alternatives) than the DF t-test. Furthermore, ERS show that the DF-GLS test has essentially the same asymptotic power as the ERS point optimal test when $\bar{c} = -7$. When $\mathbf{D}_t = (1, t)$ the asymptotic distribution of the DF-GLS test, however, is different from the ADF t-test. ERS and Ng and Perron (2001) provide critical values for the DF-GLS test in this case. ERS show that the DF-GLS test has the same asymptotic power as the ERS point optimal test with $c = -13.5$, and has higher power than the DF t-test against local alternatives.

The DF-GLS t-test may be computed using the function `unitroot` with `method="dfgls"`.

4.6.3 Modified Efficient PP Tests

Ng and Perron (2001) use the GLS detrending procedure of ERS to create efficient versions of the modified PP tests of Perron and Ng (1996). These *efficient modified PP tests* do not exhibit the severe size distortions of the PP tests for errors with large negative MA or AR roots, and they can have substantially higher power than the PP tests especially when ϕ is close to unity.

[5]For deterministicly trending trend data with ergodic-stationary deviations from trend, Grenander's Theorem gives the result that least squares estimates of the trend parameters ignoring serial correlation are asymptotically equivalent to the generalized least squares estimates that incorporate serial correlation.

Using the GLS detrended data y_t^d, the efficient modified PP tests are defined as

$$\begin{aligned}
\overline{MZ}_\alpha &= (T^{-1}y_T^d - \hat{\lambda}^2)\left(2T^{-2}\sum_{t=1}^{T} y_{t-1}^d\right)^{-1} \\
\overline{MSB} &= \left(T^{-2}\sum_{t=1}^{T} y_{t-1}^d/\hat{\lambda}^2\right)^{1/2} \\
\overline{MZ}_t &= \overline{MZ}_\alpha \times \overline{MSB}
\end{aligned}$$

The statistics $\overline{MZ}_\alpha$ and $\overline{MZ}_t$ are efficient versions of the PP Z_α and Z_t tests that have much smaller size distortions in the presence of negative moving average errors. Ng and Perron derive the asymptotic distributions of these statistics under the local alternative $\phi = 1 - c/T$ for $D_t = 1$ and $\mathbf{D}_t = (1, t)$. In particular, they show that the asymptotic distribution of $\overline{MZ}_t$ is the same as the DF-GLS t-test.

The statistic $\overline{MZ}_t$ may be computed using the function `unitroot` with `method="mpp"`.

4.6.4 Estimating λ^2

Ng and Perron (2001) emphasize that the estimation of the long-run variance λ^2 has important implications for the finite sample behavior of the ERS point optimal test and the efficient modified PP tests. They stress that an autoregressive estimate of λ^2 should be used to achieve stable finite sample size. They recommend estimating λ^2 from the ADF test regression (4.11) based on the GLS detrended data:

$$\hat{\lambda}_{AR} = \frac{\hat{\sigma}_p^2}{\left(1 - \hat{\psi}(1)\right)^2}$$

where $\hat{\psi}(1) = \sum_{j=1}^{p} \hat{\psi}_j$ and $\hat{\sigma}_p^2 = (T-p)^{-1}\sum_{t=p+1}^{T} \hat{\varepsilon}_t^2$ are obtained from (4.11) by least squares estimation.

4.6.5 Choosing Lag Lengths to Achieve Good Size and Power

Ng and Perron also stress that good size and power properties of the all the efficient unit root tests rely on the proper choice of the lag length p used for specifying the test regression (4.11). They argue, however, that traditional model selection criteria such as AIC and BIC are not well suited for determining p with integrated data. Ng and Perron suggest the *modified information criteria* (MIC) that selects p as $p_{mic} = \arg\min_{p \le p_{\max}} MIC(p)$

where

$$\begin{aligned} MIC(p) &= \ln(\hat{\sigma}_p^2) + \frac{C_T(\tau_T(p)+p)}{T-p_{\max}} \\ \tau_T(p) &= \frac{\hat{\pi}^2 \sum_{t=p_{\max}+1}^{T} y_{t-1}^d}{\hat{\sigma}_p^2} \\ \hat{\sigma}_p^2 &= \frac{1}{T-p_{\max}} \sum_{t=p_{\max}+1}^{T} \hat{\varepsilon}_t^2 \end{aligned}$$

with $C_T > 0$ and $C_T/T \to 0$ as $T \to \infty$. The maximum lag, $p_{\max}$, may be set using (4.5). The modified AIC (MAIC) results when $C_T = 2$, and the modified BIC (MBIC) results when $C_T = \ln(T - p_{\max})$. Through a series of simulation experiments, Ng and Perron recommend selecting the lag length p by minimizing the MAIC.

Example 23 *Efficient unit root tests*

To illustrate the efficient unit root tests, consider testing for a unit root in the 30-day interest rate differential formed from the difference between monthly US and UK spot exchange rates:

```
> fd = lexrates.dat[,"USUKS"] - lexrates.dat[,"USUKF"]
> colIds(fd) = "USUKFD"
> fd@title = "US/UK 30-day interest rate differential"
```

The interest rate differential, its SACF, and the SACF of its first difference are depicted in Figure 4.6. The graphs clearly show that the interest rate differential has a high degree of persistence, and that there is little persistence in the first difference.

The ERS P_T test, DF-GLS t-test and Ng-Perron $\overline{MZ}_t$ test all with $D_t = 1$ may be computed using the function `unitroot` as follows:

```
> ers = unitroot(fd,trend="c",method="ers",max.lags=12)
> dfgls = unitroot(fd,trend="c",method="dfgls",max.lags=12)
> mpp = unitroot(fd,trend="c",method="mpp",max.lags=12)
```

Since the optional argument `lags` is omitted, the lag length for the test regression (4.11) is determined by minimizing the MAIC with $p_{\max} = 12$ set by the optional argument `max.lags=12`. The results of the efficient unit root tests are:

```
> ers.test

Test for Unit Root: Elliott-Rothenberg-Stock Test

Null Hypothesis: there is a unit root
 Test Statistic: 1.772**
```

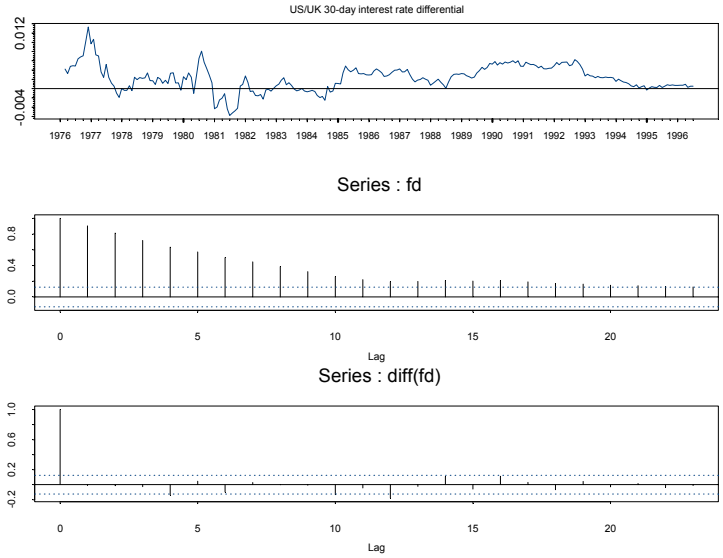

FIGURE 4.6. 30-day US/UK interest rate differential.

```
 * : significant at 5% level
** : significant at 1% level

Coefficients:
  lag1
 -0.07

Degrees of freedom: 244 total; 243 residual
Time period: from Mar 1976 to Jun 1996
Residual standard error: 0.00116

> dfgls.test

Test for Unit Root: DF Test with GLS detrending

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -2.9205**
 * : significant at 5% level
** : significant at 1% level

Coefficients:
  lag1
```

```
 -0.07

Degrees of freedom: 244 total; 243 residual
Time period: from Mar 1976 to Jun 1996
Residual standard error: 0.00116

> mpp.test

Test for Unit Root: Modified Phillips-Perron Test

Null Hypothesis: there is a unit root
   Type of Test: t-test
 Test Statistic: -2.8226**
 * : significant at 5% level
** : significant at 1% level

Coefficients:
  lag1
 -0.07

Degrees of freedom: 244 total; 243 residual
Time period: from Mar 1976 to Jun 1996
Residual standard error: 0.00116
```

Minimizing the MAIC gives $p = 0$, and with this lag length all tests reject the null hypothesis of a unit root at the 1% level.

4.7 References

[1] Caner, M. and L. Kilian (2001). "Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate," *Journal of International Money and Finance*, 20, 639-657.

[2] Dickey, D. and W. Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association,* 74, 427-431.

[3] Dickey, D. and W. Fuller (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057-1072.

[4] Elliot, G., T.J. Rothenberg, and J.H. Stock (1996). "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813-836.

[5] Fuller, W. (1996). *Introduction to Statistical Time Series, Second Edition.* John Wiley, New York.

[6] Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, New Jersey.

[7] Hatanaka, T. (1995). *Time-Series-Based Econometrics: Unit Roots and Co-Integration*. Oxford University Press, Oxford.

[8] Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992). "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics,* 54, 159-178.

[9] MacKinnon, J. (1996). "Numerical Distribution Functions for Unit Root and Cointegration Tests," *Journal of Applied Econometrics*, 11, 601-618.

[10] Maddala, G.S. and I.-M. Kim (1998). *Unit Roots, Cointegration and Structural Change*. Oxford University Press, Oxford.

[11] Ng, S., and P. Perron (1995). "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association*, 90, 268-281.

[12] Ng, S., and P. Perron (2001). "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69, 1519-1554.

[13] Perron, P. and S. Ng. (1996). "Useful Modifications to Some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties," *Review of Economic Studies*, 63, 435-463.

[14] Phillips, P.C.B. (1987). "Time Series Regression with a Unit Root," *Econometrica*, 55, 227-301.

[15] Phillips, P.C.B. and P. Perron (1988). "Testing for Unit Roots in Time Series Regression," *Biometrika*, 75, 335-346.

[16] Phillips, P.C.B. and Z. Xiao (1998). "A Primer on Unit Root Testing," *Journal of Economic Surveys*, 12, 423-470.

[17] Schwert, W. (1989). "Test for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics*, 7, 147-159.

[18] Said, S.E. and D. Dickey (1984). "Testing for Unit Roots in Autoregressive Moving-Average Models with Unknown Order," *Biometrika*, 71, 599-607.

[19] Stock, J.H. (1994). "Units Roots, Structural Breaks and Trends," in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics, Volume IV*. North Holland, New York.