Unit Root Testing Overview Until July 2009

Basil Ibrahim

July 3, 2009

Introduction

With the emerging importance of anlayzing time series, it has become essential to find/derive mathematical models that well represent these data sets. A useful starting point for obtaining the optimal model is to perform diagnostic tests on the time series, giving us a better "feel" on the data set, which could help point us to suitable candidate models. For instance, if we have a time series with a cyclical pattern, which could be seen in the form of ups and downs if plotted, one can consider using SARMA or PARMA models to represent the data. Another example can be a time series with an increasing volatility trend over time, implying heteroskedasticity; and so, one would be inclined to using ARCH and GARCH models. Therefore, diagnostic testing of time series helps analysts narrow down their modelling choosing criteria.

This report provides an overview on what can be considered as a diagnostic test: testing for stationarity. There are a lot of desirable properties in having a stationary time series; here when we say stationary, we mean weakly stationary, that is a constant mean and time independent covariance. And so, a main objective is in testing whether the time series needs to be made stationary (by differencing).

This report is organized as follows: I begin by explaining what a unit roots is, as testing for them (there could be more than one) tells us whether or not the time series is stationary. I then give some examples of unit roots in a list of time series models. Next, I explain the relationship between having a unit root or more in a time series and whether or not it is stationary. Then I proceed to the different tests throughout the history of the literarture until now (July 2009, the date this report is published). Finally I conclude with possible expansions for research.

1 Definition of a Unit Root

In algebra, a root of a polynomial is a value such that when plugged into that polynomial, the result is zero. Specifically, a unit root of a polynomial is defined as a value satisfying this and also equal to one. However, it also important to note that roots lying on the unit circle the real-complex plane (e.g. $(1, 0)$, $(0, -i)$, $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}i)$) also bear significant consequences on determining whether or not a time series is stationary.

In discrete time series, Dhrymes in 1981 showed that the algebra of polynomial functions on lag operators, $f(B)$ for instance, is isomorphic to the algebra of ordinary polynomial functions, $f(x)$ *(Franses, p.32)*. And so, one will find that discrete time series models not only can be represented in terms of operator polynomial functions, but also manipulated and treated as ordinary algebraic ones. Examples on this are shown in the next section.

2 Examples of Unit Circle Roots in Time Series

1) AR(1): $y_t = y_{t-1} + \epsilon_t \Leftrightarrow (1-B)y_t = \epsilon_t$.
   Solve: $(1-x) = 0 \Rightarrow x = 1 \Rightarrow$We have a unit root.

2) MA(1): $y_t = \epsilon_t + \epsilon_{t-1} \Leftrightarrow y_t = (1+B)\epsilon_t$.
   Solve: $(1+x) = 0 \Rightarrow x = -1 \Rightarrow$We have a root on the unit circle.

3) AR(2): $y_t = 0.3y_{t-1} + 0.7y_{t-2} + \epsilon_t \Leftrightarrow (1 - 0.3B - 0.7B^2)y_t = \epsilon_t$.
   Solve: $(1 - 0.3x - 0.7x^2) = 0 \Rightarrow x_1 = 1, x_2 = -10/7 \Rightarrow$We have a unit root.

4) AR(2): $y_t = -y_{t-2} + \epsilon_t \Leftrightarrow (1 + 7B^2)y_t = \epsilon_t$.
   Solve: $(1 + x^2) = 0 \Rightarrow x_1 = i, x_2 = -i \Rightarrow$We have two roots on the unit circle.

5) AR(3): $y_t = y_{t-3} - y_{t-2} + y_{t-1} + \epsilon_t \Leftrightarrow (1 - B^3 + B^2 - B)y_t = \epsilon_t$.
   Solve: $(1 - x^3 + x^2 - x) = 0 \Rightarrow x_1 = 1, x_2 = i, x_3 = -i$
   $\Rightarrow$We have one unit root and two other roots on the unit circle.

Now that one has an idea about the algebra of the lag operators and how they are used to determine the roots of the model, the next section relates roots on the unit circle - most importantly the unit root - to stationarity.

3 Unit Root and Stationarity

Here, one has to distinguish between different types of discrete time series models, because stationarity can be always exist in some of them, regardless of having a unit root. The two principle models are AR and MA models:

For AR based models (e.g AR(p), ARI(p,d), ARMA(p,q), VAR(p),etc.), having at least one root on the unit circle (not necessarily a unit root) in the AR component implies non-stationarity, because the variance of y_t becomes a function of time, $t\sigma^2$ (weak stationarity assumes a constant variance σ^2).

For MA based models, stationarity is not an issue (provided the other components do not exhibit nonstationarity), since the MA component is always stationary. However, having one root or more on the unit circle in the MA component can be used as an indicator of overdifferencing *(Brockwell and Davis, p.196)*. Therefore, the main focus of the literature is towards unit root tests in the AR component.

The main reason I believe why unit root tests are much more popular than tests for roots on the unit circle in general is because having a unit root simply implies that differencing will remedy the stationarity issue (unless there are other issues such as nonlinearity and heteroskedasticity, which can also be tested for and remedied in some cases), while having a root on the unit circle other than one involves cyclical patterns (Franses denotes them as seasonal unit roots, *(Franses, p.106)*). These roots can be tested for using seasonal dummies in standard unit root tests *(Franses, p.115)* or in some cases handled through parameter reduction. The following example demonstrates the latter issue (derived from a similar problem in Franses, p.105):

$\Delta_1\Delta_4 y_t = 2\epsilon_t - 3\epsilon_{t-1} + 3\epsilon_{t-2} - 3\epsilon_{t-3} + \epsilon_{t-4} = (2 - 3B + 3B^2 - 3B^3 + B^4)\epsilon_t,$

Where $\Delta_1\Delta_4 = (1 - B)(1 - B^4) = (1 - B)(1 - B)(1 + B)(1 - iB)(1 + iB)$.

The solutions to the polynomial $(2 - 3x + 3x^2 - 3x^3 + x^4) = 0$ are

$x_1 = 1, x_2 = 2, x_3 = i, x_4 = -i$. Therefore, both sides of the equation can be written as:

$(1 - B)(1 - B)(1 + B)(1 - iB)(1 + iB)y_t = (1 - B)(2 - B)(1 - iB)(1 + iB)\epsilon_t.$

Similarly to polynomial functions, we can cancel out terms, ending up with:

$(1-B)(1+B)y_t = (2-B)\epsilon_t \Rightarrow y_t = y_{t-2} + 2\epsilon_t - \epsilon_{t-1}$, which is a more parsimonious model.

In summary, testing for unit roots in AR based models is a fundamental method in examining the stationarity state of the underlying model. There are stationarity tests for models incorporating seasonality, which involve other roots in the unit circle. These can be used as further diagnostic tests on the time series, alongside other tests such as for heteroskedasticity and nonlinearity. We now proceed to the main overview of the literature of unit root testing until this date, July 2009.

4 Unit Root Testing

As we are focusing on the AR component of the model, $\phi_1 y_{t-1} + ... + \phi_p y_{t-p}$, we are testing if $\sum_{i=0}^{p} \phi_i = 1$ (because we are examining the values of the ϕ_i's that make the polynomial $(1 - \phi_1 x - ... - \phi_p x^p) = 0$ when $x = 1$, which is true if and only if the sum of the ϕ_i's is 1).

Before proceeding to the different tests, it is worth emphasizing the importance of unit root tests as a diagnostic tools. As mentioned before, unit root testing helps us decide whether or not we should difference data. According to Diebold and Kilian, "forecasters face three choices: always difference the data, never difference, or use a unit-root pretest". In their paper, they conclude that pretesting for a unit root (i.e. testing before modelling) has better forecasting results than simply adopting the philosophy of differencing the data from the start. Thus, unit root testing is a very useful diagnostic tool for choosing models with better forecasting power.

Now we proceed to the different tests:

4.1 Dickey-Fuller Unit Root Tests (1984)

4.1.1 Simple Case: Dickey Fuller (DF) Test

Consider an AR(1) model:

$y_t - \mu = \phi_1(y_{t-1} - \mu) + \epsilon_t$, where $\epsilon_t \sim WN(0, \sigma^2)$,

with $H_0 : \phi_1 = 1$
$H_a : |\phi_1| < 1$

(Note: from my knowledge, explosive time series, $|\phi_1| > 1$, do not exist in practice).

The above equation can be rewritten in this form:

$\Delta y_t = y_t - y_{t-1} = \mu(1 - \phi_1) + (\phi_1 - 1)y_{t-1} + \epsilon_t = a_0 + a_1 y_{t-1} + \epsilon_t.$

And so the null hypothesis becomes $a_1 = 0$.

After estimating a_0 and a_1, using standard OLS regression, one calculates the t-ratio: $\hat{\tau}_\mu = \hat{a}_1/\widehat{SE}(\hat{a}_1)$, where the denominator is the estimated standard error of $\hat{a}_1$. Through simulation, Dickey and Fuller were able to calculate critical values to which this t-ratio can be compared (for instance, with $\mu = 0$ and $n > 500$ the critical value is -0.286 at a 0.05 significance level).

We now proceed to a more general case: AR(p).

4.1.2 General Case: Augmented Dickey Fuller (ADF) Test

Consider an AR(p) model:

$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + ... + \phi_p(y_{t-p} - \mu) + \epsilon_t$, where $\epsilon_t \sim WN(0, \sigma^2)$,

with $H_0 : \sum_{i=1}^{p} \phi_i = 1$
$H_a : |\sum_{i=1}^{p} \phi_i| < 1$

With some algebra, we can get:

$\Delta y_t = a_0 + a_1 y_{t-1} + a_2 \Delta y_{t-2} + ... + a_p \Delta y_{t-(p-1)} + \epsilon_t$, with

$a_0 = \mu(1 - \sum_{i=1}^{p} \phi_i)$, $a_1 = \{\sum_{i=1}^{p} \phi_i\} - 1$, $a_j = -\sum_{i=j}^{p} \phi_i$ for $j = 2, ..., p$.
Similarly to the simple case, the null hypothesis becomes $a_1 = 0$.

As done in the previous case, one calculates the t-ratio using OLS regression and compares it to the same critical values used in the DF test.

Another generalization of the DF test is to incorporate a linear trend.

4.1.3 ADF with a Linear Trend

Consider an AR(p) model with a linear trend:

$y_t - \mu - \delta t = \phi_1(y_{t-1} - \mu - \delta(t-1)) + \phi_2(y_{t-2} - \mu - \delta(t-2)) + ... + \phi_p(y_{t-p} - \mu - \delta(t-p)) + \epsilon_t$,
where $\epsilon_t \sim WN(0, \sigma^2)$,

with $H_0 : \sum_{i=1}^{p} \phi_i = 1$
$H_a : |\sum_{i=1}^{p} \phi_i| < 1$

One can also examine the joint hypothesis:

$H_0 : \sum_{i=1}^{p} \phi_i = 1$ and $\delta(1 - \sum_{i=1}^{p} \phi_i) = 0$
$H_a : H_0$ is not true.

However, the main interest lies in the first null hypothesis.

With some algebra, we get:

$\Delta y_t = a_0 + \delta^* t + a_1 y_{t-1} + a_2 \Delta y_{t-2} + ... + a_p \Delta y_{t-(p-1)} + \epsilon_t$, with

$a_0 = \mu(1 - \sum_{i=1}^{p} \phi_i) + \delta \sum_{i=1}^{p} i\phi_i)$, $\delta^* = \delta(1 - \sum_{i=1}^{p} \phi_i))$, $a_1 = \{\sum_{i=1}^{p} \phi_i\} - 1$, $a_j = -\sum_{i=j}^{p} \phi_i$ for $j = 2, ..., p$.

The first H_0 becomes $a_1 = 0$, and the second one becomes $a_1 = 0$ and $\delta^* = 0$.

To test for the first H_0, one uses OLS regression and calculates the t-ratio, as shown previously. The table derived by Dickey and Fuller includes critical values for AR(p) models with a linear trend. There are four possibilities: no trend and no drift, trend and no drift, no trend and drift, trend and drift.

Regarding the second H_0, Dickey and Fuller in 1981 propose a joint F-test *(Franses, p.83)*, however, as mentioned previously, the main interest is on the former H_0.

Remarks on ADF:

1) It is essential to specify a suitable number of lags, p, because if p is too small the model suffers a higher bias ("remaining serial correlation in the errors will bias the test" *(Zivot and Wang, p.121)*). On the other hand, if p is too large, then the power of the test suffers, leading to an increase in the chances of committing a Type II error. Thus, one must find an appropriate number of lags to avoid these issues.

Schwert (1989) proposed the following criterion: Set a ceiling $p_{max} = floor(12\sqrt[4]{T/100})$, and find the $p \leq p_{max}$ that gives the smallest t-ratio.

2) One can extend the ADF test to ARMA(p,q) models, however, there lies a risk that AR and MA components have common roots, leading to an elimination, as shown in an example above.

The following test has more flexibility on the AR component, especially regarding the specification of the lags.

4.2 Phillips-Perron (PP) Unit Root Test (1988)

Recall that:

ADF: $\Delta y_t = a_0 + \delta^* t + a_1 y_{t-1} + a_2 \Delta y_{t-2} + ... + a_p \Delta y_{t-(p-1)} + \epsilon_t$, where $\epsilon_t \sim WN(0, \sigma^2)$.

Here, we have:

PP: $\Delta y_t = a_0 + \delta^* t + a_1 y_{t-1} + u_t$, where u_t can be heteroskedastic.

These are the main Steps:

1) Apply OLS regression on the model and estimate the parameters.

2) Instead of calculating the t-ratio $\widehat{\tau}_\mu$, calculate the statistic Z_t, which incorporates the possible heteroskedastic nature of the error term.

Z_t is equal to $\left(\frac{\hat{\sigma}}{\hat{\lambda}}\right)\hat{\tau}_\mu - \left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{2\hat{\lambda}^2}\right)\left(\frac{n\widehat{SE}(\hat{a}_1)}{\hat{\sigma}^2}\right)$, where $\hat{\lambda}^2$ and $\hat{\sigma}^2$ are variance parameters *(Zivot and Wang, p.127)*.

3) Compare Z_t to the ADF critical values (same limiting distribution).

This test has two advantages over the ADF test:

1) PP is robust to general forms of heteroskedasticity of u_t, which is a restricted assumption in the ADF test.

2) Specifying the number of lags, p, is critical in the ADF test; however, here, one need not worry about this issue.

It is important to note that the PP test is not a dominant test over the ADF test; there are circumstances where the ADF test yields better results (such as the situation where the AR or MA component of the model has a root close to -1).

The following section takes a different testing approach altogether.

4.3 Stationarity Tests

In the ADF and PP tests, the null hypothesis is that we have a unit root ($a_1 = 0$) and the alternate hypothesis is that we do not. Stationarity tests switch these hypotheses:

H_0 : we do not have a unit root, H_a : we do. The most common stationarity test is the KPSS test (Kwiatowski, Phillips, Schmidt and Shin in 1992). The test is as follows:

We have the model: $y_t - \mu - \delta t = u_t + v_t$;

where $u_t = u_{t-1} + \epsilon_t$, with $\epsilon_t \sim WN(0, \sigma^2)$.

Note that u_t is a pure random walk (since we have a unit root), with variance $t\sigma^2$. We have:
$H_0 : \sigma^2 = 0 \Leftrightarrow u_t$ is a constant.
$H_a : \sigma^2 > 0$.

Steps *(Zivot and Wang, p.129)*:

1) Calculate KPSS statistic: $KPSS = \left(n^{-2} \sum_{t=1}^{n} \hat{S}_t^2 \right) / \hat{\lambda}^2$, where $\hat{S}_t = \sum_{i=1}^{t} \hat{u}_i$.

2) Compare KPSS to critical values (obtained by simulation).

Even though these tests provide a variety of alternative methods in detecting unit roots, and hence requiring the model to be differenced at least once, there are some limitations, leading experts to use/create more powerful unit root tests. Before proceeding to these tests, I highlight the main constraints the ADF, PP and KPSS tests pose *(Zivot and Wang, p.132)*:

1) In ARMA(p,q) models, if there are large negative components in the MA part, these tests tend to reject H_0 too often (i.e. tendency of committing a Type I error is too large). This is more severe for the PP test, which is one limitation of the PP test over the ADF test (the PP test also suffers this problem, when the AR component is close to -1).

2) The ADF and PP tests have very low power if the model is highly persistent (i.e.

$\sum_{i=1}^{p} \phi_i \approx 1$). This means that they have difficulty in distinguishing between stationary and long memory processes.

3) Adding a drift term and/or a trend to the model shifts the critical values in the ADF table to the left, which reduces the power of the ADF and PP tests (this is because the chances of committing a Type II error increase). Thus, more efficient unit root tests have been created, in order to tackle such issues. The following section lists some of these tests.

4.4 Efficient Unit Root Tests

By knowing the the distribution of the data, one can find a test amongst a class of tests with the best power (using the Neyman-Pearson Lemma). This is done by comparing the test to an upper bound, known as a power envelope.

In their paper "Efficient Tests for an Autoregressive Unit Root", Elliott, Rothenberg and Stock derive a class of statistics that come very close to the power envelope. This class was called efficient unit root tests, with $H_a : \sum_{i=1}^{p} = 1 + c/t$, with $c < 0$.

It has been shown that these tests in dominate the ADF, PP and KPSS tests, especially when the time series is highly persistent (i.e. $\sum_{i=1}^{p} \phi_i \approx 1$).

Here are the most famous types of efficient unit root tests:

4.4.1 Point Optimal (Invariant) Tests

A constraint with the ADF and PP tests is that when the time series has an unknown mean and/or linear trend, this class of tests is substantially more powerful than the standard ADF and PP tests. As the word "invariant" suggests, these tests are not significantly distorted (unlike the standard tests) by adding a drift and/or a linear trend to the model *(Elliot et. al, p.813)*.

4.4.2 DF-GLS Test

This is a modified, more efficient version of the ADF t-statistic. One begins by detrending the data, which is done by subtracting the trend terms from the original model. Then one applies GLS on the detrended model and calculates the t-ratio *(Zivot and Wang, p.134)*. Finally, one compares this ratio to the ADF critical values. A similar analogue is also available for the PP test.

4.4.3 Modified Efficient PP Tests (Efficient M-Tests)

Similar to the DF-GLS test, "Ng and Perron (2001) use the GLS detrending procedure [...] to create efficient versions of Perron and Ng (1996)" (Note that the original PP test was created in 1988, then in 1996 a modified version, known as M-tests was created) *(Zivot and Wang, p.134)*.

Although the M-tests are more sensitive in distinguishing between persistent time series and time series with a unit root *(Perron and Ng, p.435)*, the issue of rejecting the null hypothesis too often due to large negative AR or MA components (close to -1) has not been significantly reduced. However, when using the GLS detrending procedure, one finds a significant improvement: they do not "exhibit the severe size distortions of the PP tests for errors with large negative MA or AR roots" *(Zivot and Wang,p.134)*.

These are the main tests derived and used up until about the end of 2003. The next section gives an overview to more modern and power tests. Before proceeding into more up-to-date unit root tests, I give some remarks about these efficient unit root tests:

1) These tests work best for large sample sizes, as the residual autocorrelation can be captured and the asymptotes are more accurate *(Elliot et al., p.830)*.

2) Although efficient unit root tests are rendered "more powerful," they may actually worsen forecast accuracy for certain ranges of $\sum\limits_{i=1}^{p} \phi_i$ *(Diebold and Kilian, p.10)*. Rather, the literature suggests that there can exist a trade-off between finding the true model and accurately forecasting n-steps ahead.

3) Similar to the ADF case, proper choice of number of lags is critical for giving the test a good power. In this case, using criteria such as the AIC and BIC are not appropriate for determining the optimal lag length, p. Instead, one can use the MIC (modified information criteria), suggested by Ng and Perron.

We now proceed to more recent (up until July 2009) and effective tests.

4.5 More Recent Tests

1) Mixture tests that incorporate heteroskedasticity, such as a weighted average between an ADF statistic and a standard normal statistic. The higher the heteroskedasticity, the higher the weight is on the standard normal statistic *(Gospodinov and Tao, p.1)*.
Unit root tests with GARCH errors *(Gospodinov and Tao, p.1)*: Dickey-Fuller distribution is still valid, but conservative.

The main challenge is that these tests require nonlinear estimation and the "asymptotic distribution depends on nuisance parameters which involves additional computational for obtaining critical values," which suffer from severe size distortions. Note that in a GARCH setup, the unit root test is focused on the AR(p) component of the model, not the error terms (since the error terms can have autoregressive lags in a GARCH setup).

2) **Asymptotic Refinements Based on Bootstrap Theory:** In the recent literature (2003-2009), there have emerged many articles on tests using bootstrapping techniques.

In their paper, Gospodinov and Tao (2009) use a bootstrap method to obtain more effective unit root tests models with GARCH(1,1) errors. Cavaliere and Taylor use wild bootstrapping on GLS-ADF and Efficient M-Tests, showing a better performance than the original tests, even under small sample sizes. There are many other papers on different types of AR based models that use bootstrapping techniques to derive unit root tests more powerful than the standard ones.

After talking extensively of unit root tests for univariate AR based models, the next section briefly mentions how this is done for multivariate AR based models.

4.6 Unit Root Tests on Multivariate AR Models

Consider a VARMA(p,q) model:

$\Phi_p(B)Y_t = \mu + \Theta_q(B)e_t,$

where Y_t, μ and e_t are n-by-1; $\Phi_p(B)$ is p-by-n;$\Theta_q(B)$ is q-by-n.

For stationarity to hold, solutions to $|\Phi_p(x)| = 0$ must lie outside the unit circle *(Franses, p.196)*. Otherwise, the model contains unit roots.

A special case of a multivariate model with unit roots is a model whose linear combination of the Y_t variables is stationary. This is known as cointegration. A couple of famous tests for cointegration are:

1) **The Engel-Granger Test (1987):** It is based on the ADF test. The main restriction of this test is that the Y_t variables are assumed to be of the same order of integration (i.e. they all need to be differenced d times to become stationary).

2) **The Johansen Test (1988):** This test is more flexible than the previous one. The Y_t variables can have different orders of integration.

Testing for cointegration is useful as linear combinations are quite simple to deal with. And so, as a diagnostic tool on multivariate data, examining cointegration in a model is a very useful starting point that can help eliminate non-stationarity if found present in the data set.

A more complex series of tests is panel unit root tests. According to Pesaran et al., there (until this date, July 2009) two generations of panel unit root tests. For the first generation (2002-2003), the error terms were assumed to be cross-sectionally uncorrelated. The second generation (2004-present) allows for cross-sectional correlation of the error terms with a residual structure *(Pesaran et al., p.2)*.

The null hypothesis of these tests is that all cross-sections have a unit root, but are not cointegrated. As an extension to the ADF test, the Pesaran test statistic uses ADF average t-ratios over cross-section units to compare to critical values obtained by simulation.

Testing for cointegration and panel unit root tests are the main trend in multivariate data. As time goes by, more robust types of tests (e.g. using bootstrapping techniques, weighted tests, etc.) will emerge, tackling more complex types of multivariate models (e.g. models with heteroskedastic error terms). The final section highlights some possible topics for future research one can contribute to the literature.

5 Topics for Further Research

1) Using more complex models such as GARCH(p,q), FARSIMA and ANN (Artificial Neural Networks), which all contain AR components, but whose parameters and asymptotic distributions are very challenging to estimate or derive.

2) Using more powerful bootstrapping techniques, which can improve the power of the testing for a unit root. It will be especially useful for small sample sizes and nonlinear trends.

3) As explained earlier, unit roots basically indicate that the data need to be differenced. A possible topic is to examine time series that become stationary if we take the ratio of each two consecutive data points. However, here the isomorphic property of the lag operator with polynomial functions may not be of much use. Hence, one can redefine a different set of time series models (other than AR and MA), which can exploit some sort of isomorphism between lag points and algebraic functions.

To conclude, unit root testing is an extremely useful tool to help experts understand the time series, and, if necessary, make it mathematically tractable for accurate modelling and prediction.

References

[1] BROCKWELL, PETER J. AND RICHARD A. DAVIS (1998), *Introduction to Time Series Forecasting.* Springer Texts in Statistics, New York, Second Edition.

[2] CAVALIERE, GIUSEPPE AND A.M. ROBERT TAYLOR (2009), *Bootstrap M Unit Root Tests.* Econometric Reviews, 28:5, 293-421.

[3] DIEBOLD, FRANCIS X. AND LUTZ KILIAN (1991), *Unit Root Tests are Useful for Selecting Forecasting Models.* University of Pennsylvania, Philadephia.

[4] ELLIOTT, GRAHAM AND THOMAS J. ROTHENBERG and JAMES H. STOCK (July 1999), *Efficient Tests for Autoregressive Unit Root.* The Econometric Society, Vol. 64, No. 4, 813-836. <http://www.jstor.org/stable/2171846>.

[5] FRANSES, PHILIP HANS (1998), *Time Series Models for Business and Economic Forecasting.* Cambridge University Press, United Kingdom.

[6] GOSPOSDINOV, NIKOLAY AND YE TAO (Mar. 2009), *Bootstrap Unit Root Tests in Models with GARCH(1,1) Errors.* University of Concordia, Montreal.

[7] PERRON, PIERRE AND SERENA NG (1996), *Useful Modifications to Some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties.* The Review of Economic Studies Limited.

[8] PESARAN, M. HASHEM AND L. VANESSA SMITH AND TAKASHI YAGAMATA (Dec. 2007), *Panel Unit Root Tests in the Presence of a Multifactor Error Structure.* Economic and Social Research Council.

[9] ZIVOT, ERIC AND JIAHUI WANG (2003), *Unit Root Tests.* Business and Economics, 111-139.