APPENDIX B

AUGMENTED DICKEY-FULLER UNIT ROOT TEST

Many macroeconomic time series such as money supply, investment, interest rate and GDP may be characterized as having stochastic trends. In other words, these time series data contain unit-root, or integrated of order one, I (1), process. The non-stationary process, which contains stochastic trend, means that we have to first-difference time series data to correct it into a "difference-stationary series" or I(0). For convenience, we begin with a very simple stationary univariate model observed over the sequence of time t=1, 2,…,T as below:

$$y_t = \rho y_{t-1} + \varepsilon_t \qquad \text{where } |\rho| < 1$$

$$(1-\rho L)y_t = \varepsilon_t$$

where L is the lag operator i.e., $Ly_t = y_{t-1}$. The variable y_t is generated by its own past realizations together with a disturbance term ε_t. And ε_t represents the influence of all other variables excluded from the model, which is presumed to be random so that it has zero expected value $(E(\varepsilon_t) = 0)$, constant variance $(E(\varepsilon_t^2) = \sigma_\varepsilon^2)$, and is uncorrelated with its own past value $(E(\varepsilon_t \varepsilon_{t-1}) = 0)$.

Harris (1995) has provided the distinction between stationary and non-stationary variables as follows:

> "*A stationary series tend to return to its mean value and fluctuates around it within a more-or-less constant range (i.e., it has a finite variance), while a non-stationary series has a difference mean at different points in time and its variance increase with the sample size or the time, and goes to infinity as time approach to infinity*".

In AR(1) process, the difference of y_t will be stationary, or non-stationary, depending on the value of $|\rho|$:

If $|\rho| < 1$, the variable y_t is stationary as $t \to \infty$, If $|\rho| = 1$, the variable y_t is non-stationary, and exhibits a random walk process since the current value of y_t depends on its previous value and all disturbances accruing through time t; therefore the variance of y_t is time dependent,

If $|\rho| > 1$, the variable y_t is non-stationary and explosive (i.e., it will tend to either $\pm\infty$).

To test the presence of a unit root, the Dickey-Fuller (DF) approach. DF test tends to be more popular because of its simplicity and it is more general in nature.

The simplest form of the DF test is :

$$\Delta y_t = \rho^* y_{t-1} + \varepsilon_t \;\; ; \;\; \varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2) \qquad \text{(b.1)}$$

where $\rho^* = \rho - 1$

with the hypothesis that

$$\text{H}_0 : \rho^* = 0 \qquad \text{(non} - \text{stationary)}$$
$$\text{H}_1 : \rho^* < 0 \qquad \text{(stationary)}$$

The series is stationary when $\rho^* < 0 \; or \; \rho < 1.$ The standard approach to test such hypothesis is to construct a t-test; however, under the null hypothesis, the computed statistic does not possess the standard t-distribution. We need to use the Dickey-Fuller τ $(tau) - distribution$ which was computed by Monte Carlo simulations and report in Fuller (1976).

Testing for a unit root using (1) is based on the prior assumption that the underlying data generating process for y_t is a simple first-order autoregressive process with a zero and no trend component. However, if it is not known whether the overall mean of the series is zero, then it is better to allow a constant or a drift term $, \mu,$ to enter the regression model when testing for a unit root (equation b.2). Furthermore, when the underlying series is a process around a deterministic trend, then we must allow a time trend t to enter the regression model used to test for a unit root as well (equation b.3). Then, the forms of the DF test are as follows :

$$\Delta y_t = \mu + \rho^* y_{t-1} + \varepsilon_t \;\; ; \;\; \varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2) \qquad \text{(b.2)}$$

$$\Delta y_t = \mu + \gamma t + \rho^* y_{t-1} + \varepsilon_t \;\; ; \;\; \varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2) \qquad \text{(b.3)}$$

The appropriate critical values in this case are given by the DF distribution relating to τ_{μ} *and* τ_{τ}, respectively (in absolute value $\tau_{\tau}, \tau_{\mu}, \tau$).

When a simple AR(1) DF model is used while in fact y_t follows an AR(ρ) process, the error term will be autocorrelated because of the misspecification of the dynamic structure of y_t. Autocorrelated errors will invalidate the use of the DF distribution which are based on the assumption the ε_t is 'white-noise'. Therefore, Dickey and Fuller (1981) extend further to allow for moving-average (MA) parts in the ε_t. The 'Augmented Dickey-Fuller test' (ADF) can be written as follows:

$$\Delta y_t = p^* y_{t-1} + \sum_{i=1}^{k} \rho_i^* \Delta y_{t-i} + \mu + \gamma t + \varepsilon_t \qquad ; \varepsilon_t \sim i.i.d(0, \sigma_{\varepsilon}^2) \qquad \text{(b.4)}$$

To investigate for a unit root, we test the null hypothesis of $\rho^* = 0$ against the alternative hypothesis of $\rho^* < 0$. And the critical value for testing student- t ratio of ρ^* is provided in Dickey-Fuller's Table.

The problem of appropriate lag-length rises since too few lags may result in over-rejecting the null when it is true, while too many lags may reduce the power of the test. In other wards, the value of k should be relatively small so as to save degrees of freedom but large enough to capture the existence of autocorrelation in the error term.

An accepted solution to determine the appropriate lag length is the Akaike information criterion (AIC).

$$AIC = \left[\frac{ESS}{T} \right] e^{(2k/T)}$$

where ESS = residual sum of squares

T = number of observation

k = appropriate lag length (s).

The optimal lag length is chosen in order to minimize the AIC value. The basic idea of choosing k involves using a model selection procedure that test, whether an additional lag is significant in increasing the value of $\overline{R}^2$.