Problem Set 2

Thais Takeuchi

May 25, 2023

Problem 1

1.1

Given that Y_t is a MA(1) process with a deterministic time trend, i.e.,

$$Y_t = \alpha + \delta \cdot t + \varepsilon_t + \theta \cdot \varepsilon_{t-1} \tag{1}$$

The dynamic multipliers s-periods ahead for this stochastic process, where $s \in 0 \cup \mathbb{N}$, can be expressed as follows:

- If $s = 0$,

  $$\frac{\partial Y_t}{\partial \varepsilon_t} = 1$$

  Since:

  $$Y_{t+1} = \alpha + \delta \cdot (t) + \varepsilon_t + \theta \cdot \varepsilon_t \tag{2}$$

- If $s = 1$,

  $$\frac{\partial Y_{t+1}}{\partial \varepsilon_t} = \theta$$

  Since:

  $$Y_{t+1} = \alpha + \delta \cdot (t+1) + \varepsilon_{t+1} + \theta \cdot \varepsilon_t \tag{3}$$

- If $s \geq 2$,

  $$\frac{\partial Y_{t+s}}{\partial \varepsilon_t} = 0$$

  Since:

  $$Y_{t+s} = \alpha + \delta \cdot (t+s) + \varepsilon_{t+s} + \theta \cdot \varepsilon_{t+s-1} \tag{4}$$

1.2

For any $s \in \{0\} \cup \mathbb{N}$, the dynamic multipliers s-periods ahead for the following random walk with drift[1]

$$Y_t = \delta + Y_{t-1} + \varepsilon_t \tag{5}$$

can be obtained through equation (5), through iteration. When $s = 1$:

$$\begin{aligned} Y_{t+1} &= \delta + Y_t + \varepsilon_{t+1} && (6)\\ &= \delta + \delta + Y_{t-1} + \varepsilon_t + \varepsilon_{t+1} && (7)\\ &= 2 \cdot \delta + Y_{t-1} + \varepsilon_t + \varepsilon_{t+1} && (8) \end{aligned}$$

[1] $\{Y_t\}$ is an $I(1)$ process with a drift, where $\{\varepsilon_t\}$ is a white noise process.

When $s = 2$ we have:

$$
\begin{aligned}
Y_{t+2} &= \delta + Y_{t+1} + \varepsilon_{t+2} && (9)\\
&= \delta + 2 \cdot \delta + Y_{t-1} + \varepsilon_t + \varepsilon_{t+1} + \varepsilon_{t+2} && (10)\\
&= 3 \cdot \delta + Y_{t-1} + \varepsilon_t + \varepsilon_{t+1} + \varepsilon_{t+2} && (11)
\end{aligned}
$$

For an arbitrary $s \in 0 \cup \mathbb{N}$:

$$
\begin{aligned}
Y_{t+s} &= s \cdot \delta + Y_t + \sum_{i=1}^{s} \varepsilon_{t+i} && (12)\\
&= s \cdot \delta + \delta + Y_{t-1} + \varepsilon_t + \sum_{i=1}^{s} \varepsilon_{t+i} && (13)\\
&= (s+1) \cdot \delta + Y_{t-1} + \sum_{i=0}^{s} \varepsilon_{t+i} && (14)
\end{aligned}
$$

Therefore, the dynamic multipliers s-periods ahead for this stochastic process can be expressed as:

$$\frac{\partial Y_{t+s}}{\partial \varepsilon_t} = 1$$

1.3

When a shock in ε_t in the MA(1) process with a deterministic time trend occurs, it is temporary because the ε_t are not correlated. Now, when we think about a random process with a drift we observe a unit root, i.e., a unit root refers to a situation where the auto-regressive coefficient of the lagged variable is equal to 1. This means that the variable continues to be influenced by past shocks indefinitely. As a result, the shocks have a lasting impact on the variable's behavior, leading to **infinite memory, explained by persistent and non-stationary nature of the variable's behavior**, illustrated by:

$$\frac{\partial Y_{t+s}}{\partial \varepsilon_t} = 1$$

i.e., changes in ε_t have permanent effect over Y_{t+s}. It is worth noting that the infinite memory property of unit root processes does not imply that all past shocks have equal importance. The impact of shocks tends to diminish as time passes, but the cumulative effect still remains, leading to the perception of infinite memory.

Problem 2[2]

2.1 and 2.2

Let

$$Y_t = \alpha + \delta \cdot t + \varepsilon_t \tag{15}$$

a deterministic time trend model, where $\{\varepsilon_t\}$ is i.i.d. In order to understand the result below, for a deterministic time trend model:

$$\frac{\delta_T - \delta_0}{\left\{ s_T^2 \cdot \begin{bmatrix} 0 & 1 \end{bmatrix} \left(\sum_{t=1}^{T} X_t X_t' \right)^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}^{1/2}} \xrightarrow{d} N(0,1)$$

we will use a Monte Carlo simulation with sample size $T = 10,000, \alpha = 0$ and $\delta = 1$. The number of Monte Carlo repetitions will be $M = 10,000$. For each Monte Carlo repetition, we will test " $H_0 : \delta = 1$ " at the 10% significance level using the usual OLS t-test. We will use this simulation to compute the rejection rate of this test.

When ε_t follows a t Student distribution with 5 degrees of freedom (Table 1 below) the test's rejection rate is 10.32%: Furthermore, when ε_t follows a t Student distribution with 1 degree of freedom (Table 2 below) the test's rejection rate is 9.85%:

[2] Problem 2 (including its code) was written in collaboration with **Matheus Junqueira**.

Table 1: **Test's rejection rate**

Degrees of Freedom	Rejection Rate
5	10.32%

Table 2: **Test's rejection rate**

Degrees of Freedom	Rejection Rate
1	9.85%

2.3

The answer for different rejection rates in item 2.1 and 2.2 are grounded in Hamilton's Time Series Analysis[3] proposition below:

Proposition 16.1: Let y_t be generated according to the simple deterministic time trend (15) where ε_1 is $i.i$.d. with $E\left(\varepsilon_1^2\right)=\sigma^2$ **and** $E\left(\varepsilon_1^4\right)<\infty$. Then

$$\left[\begin{array}{c} \sqrt{T}\left(\hat{\alpha}_T-\alpha\right) \\ T^{3/2}\left(\hat{\delta}_T-\delta\right) \end{array}\right] \xrightarrow{L} N\left(\left[\begin{array}{l} 0 \\ 0 \end{array}\right], \sigma^2\left[\begin{array}{ll} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{array}\right]^{-1}\right). \tag{16}$$

Note that the resulting estimate of the coefficient on the time trend $\left(\hat{\delta}_T\right)$ is superconsistent-not only does $\hat{\delta}_T \xrightarrow{p} \delta$, but even when multiplied by T, we still have

$$T\left(\hat{\delta}_T-\delta\right) \xrightarrow{l} 0$$

Moreover, according to Hamilton, a sequence of random variables $\{X_T\}_{T=1}^{x}$ is $O_p\left(T^{-k}\right)$ if for every $\varepsilon>0$ there exists an $M>0$ such that

$$P\left\{\left|X_T\right|>M/\left(T^k\right)\right\}<\varepsilon$$

In this case, $\hat{\delta}_T$ in (15) is $O_p\left(T^{-3/2}\right)$ such that, there exists a band $\pm M$ around $T^{3/2}\left(\hat{\delta}_T-\delta\right)$ that contains as much of the probability distribution as desired.

After 10,000 simulations, the proportion of simulations in which we reject the null hypothesis (H_0) for 1 and 5 degrees of freedom is different because the errors follow different distributions. When ε_t follows a t-distribution with 1 degree of freedom, we can say that, equivalently, ε_t follows a Cauchy distribution.

The **Cauchy distribution** is considered more conservative in hypothesis tests because of its heavy tails and lack of finite moments (requirement for validity of **Proposition 16.1**). These characteristics affect the behavior of the distribution and its impact on hypothesis testing.

1. **Heavy Tails:** The Cauchy distribution has heavier tails compared to other distributions such as the normal distribution. This means that extreme values are more likely to occur in the Cauchy distribution. As a result, the distribution has more probability mass in the tails, leading to a larger spread of values. In hypothesis testing, this can make it more difficult to reject the null hypothesis, as extreme values that would provide evidence against the null hypothesis occur more frequently under the Cauchy distribution.

2. **Lack of Finite Moments:** The Cauchy distribution does not have finite moments, including the mean and variance. This violates assumptions made in many statistical tests, including those based on the normal distribution. The lack of a defined mean and variance makes the Cauchy distribution less well-behaved and introduces greater uncertainty in estimating its parameters. In hypothesis testing, this lack of finite moments can lead to a loss of power, meaning that the test is less likely to detect true effects or differences between groups.

[3]Chapter 16, page 160.

The combination of heavy tails and lack of finite moments in the Cauchy distribution makes it more conservative in hypothesis testing. It means that the Cauchy distribution is less likely to reject the null hypothesis even when it is false, resulting in a higher probability of Type II errors (failing to reject a false null hypothesis). In other words, the Cauchy distribution requires stronger evidence to reject the null hypothesis compared to distributions with lighter tails and finite moments, making it a more **conservative** distribution in hypothesis testing, i.e., it **under rejects** H_0 regarding hypothesis testing. Therefore, in this case, we see a **9.85%** rejection rate.

On the other hand, when ε_t follows a t-distribution with **5 degrees of freedom**, this distribution converges to a **Normal distribution** as the number of observations $T \to \infty$. Consequently, the rejection rate approaches to 10% when we mention **large numbers of degrees of freedom** and $T \to \infty$. Therefore, we see a **10.32%** rejection rate, a similar rejection rate related to the **test's significance level**.

Despite the aforementioned issues, the estimator $\hat{\delta}_T$ **in (15) is consistent to** δ_0 **with a rate of convergence of** $O_p\left(T^{-3/2}\right)$**. In other words,** $\hat{\delta}_T$ **consistently estimates** δ_0 **regardless of the distribution of** ε_t.

Problem 3[4]

1a

To prove that:

$$T \cdot (\widehat{\rho_T} - 1) \xrightarrow{d} \frac{\frac{1}{2} \cdot ([W(1)]^2 - 1)}{\int_0^1 [W(r)]^2 dr} \tag{17}$$

we will estimate the AR(1) model given by:

$$Y_t = \rho Y_{t-1} + \varepsilon_t \tag{18}$$

through an OLS estimation of ρ. Hence, the OLS estimator will be:

$$\widehat{\rho_T} = \frac{\sum_{t=1}^T Y_t \cdot Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \tag{19}$$

From (17) and (19) we have:

$$\begin{aligned}
\widehat{\rho_T} &= \frac{\sum_{t=1}^T Y_t \cdot Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \\
&= \frac{\sum_{t=1}^T (\rho Y_{t-1} + \varepsilon_t) \cdot Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \\
&= \frac{\sum_{t=1}^T \rho Y_{t-1}^2 + \varepsilon_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \\
&= \frac{\sum_{t=1}^T \rho Y_{t-1}^2}{\sum_{t=1}^T Y_{t-1}^2} + \frac{\sum_{t=1}^T \varepsilon_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \\
&= \rho + \frac{\sum_{t=1}^T \varepsilon_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2}
\end{aligned} \tag{20}$$

Given that $\rho = 1$:

$$\widehat{\rho_T} = \rho + \frac{\sum_{t=1}^T \varepsilon_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \Rightarrow \widehat{\rho_T} - 1 = \frac{\sum_{t=1}^T \varepsilon_t Y_{t-1}}{\sum_{t=1}^T Y_{t-1}^2} \tag{21}$$

If we multiply the right-hand side and the left-hand side to T we observe:

$$T \cdot (\widehat{\rho_T} - 1) = \frac{T^{-1} \cdot \sum_{t=1}^T \varepsilon_t Y_{t-1}}{T^{-2} \cdot \sum_{t=1}^T Y_{t-1}^2} \tag{22}$$

[4] Problem 3 was written in collaboration with **Matheus Junqueira**.

Next step: we will look at the behaviour of the numerator and the denominator of equation (22) above:

- Numerator:

$$T^{-1} \cdot \sum_{t=1}^{T} \varepsilon_t Y_{t-1} \tag{23}$$

If we apply the CMT, according to Hamilton[5] we have the following:

$$\begin{aligned} Y_t^2 = (Y_{t-1} + \varepsilon_t)^2 = Y_{t-1}^2 + 2\varepsilon_t \cdot Y_{t-1} + \varepsilon_t^2 \\ \Rightarrow 2\varepsilon_t \cdot Y_{t-1} = Y_t^2 - Y_{t-1}^2 - \varepsilon_t^2 \\ \Rightarrow \varepsilon_t \cdot Y_{t-1} = \frac{Y_t^2 - Y_{t-1}^2}{2} - \frac{\varepsilon_t^2}{2} \end{aligned} \tag{24}$$

Then, summing over $t = 1, \cdots, N$ and given that $Y_0 = 0$ we have:

$$\sum_{t=1}^{T} \varepsilon_t \cdot Y_{t-1} = \frac{Y_t^2 - Y_0^2}{2} - \frac{\sum_{t=1}^{T} \varepsilon_t^2}{2} = \frac{Y_t^2}{2} - \frac{\sum_{t=1}^{T} \varepsilon_t^2}{2} \tag{25}$$

The Law of Large Numbers gives us:

$$\frac{T^{-1} \cdot \sum_{t=1}^{T} \varepsilon_t^2}{2} \xrightarrow{p} \frac{\sigma^2}{2} \tag{26}$$

and the Continuous Mapping Theorem[6] shows that:

$$\frac{T^{-1} \cdot Y_t^2}{2} \xrightarrow{L} \frac{\sigma^2 [W(1)]^2}{2} \tag{27}$$

Hence,

$$T^{-1} \sum_{t=1}^{T} \varepsilon_t Y_{t-1} \xrightarrow{L} \frac{\sigma^2([W(1)]^2 - 1)]}{2} \tag{28}$$

- Denominator

$$T^{-2} \cdot \sum_{t=1}^{T} Y_{t-1}^2 \tag{29}$$

Based on Continuous Mapping Theorem[7] we have:

$$T^{-2} \sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 [W(r)]^2 dr \tag{30}$$

Therefore, Equation (22) is continuous for both Equation (28) and Equation (30). Hence, the desired result is demonstrated:

$$T(\widehat{\rho_T} - 1) \xrightarrow{L} \frac{(1/2)([W(1)]^2 - 1)}{\int_0^1 [W(r)]^2 dr} \tag{31}$$

■

[5] Chapter 17, p.483 - 486

[6] We've used a functional on Brownian motion instead of the χ^2.

[7] Chapter 17, p.486 - Proposition 17.1

1b

To prove that

$$t_T = \frac{(\widehat{\rho_T} - 1)}{\widehat{\sigma}_{\widehat{\rho_T}}} \xrightarrow{L} \frac{(1/2)([W(1)]^2 - 1)}{(\int_0^1 [W(r)]^2 dr)^{\frac{1}{2}}} \tag{32}$$

We shall recall that

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}} = \frac{(\hat{\rho}_T - 1)}{\left\{ \frac{s_T^2}{\sum_{t=1}^T Y_{t-1}^2} \right\}^{1/2}} \tag{33}$$

is a statistic for testing $H_0 : \rho = 1$ based on the OLS t test, such that, $\widehat{\sigma}_{\widehat{\rho_T}}$ is the OLS standard error for the estimated coefficient and s_T^2 denotes the OLS estimate of the residual variance, given by the following two equations:

$$\hat{\sigma}_{\rho_T} = \left\{ \frac{s_T^2}{\sum_{t=1} Y_{t-1}^2} \right\}^{\frac{1}{2}} \tag{34}$$

$$s_T^2 = \frac{\sum_{t=1}^T (Y_t - \hat{\rho}_T Y_{t-1})^2}{T - 1} \tag{35}$$

We can rewrite (33) as:

$$\begin{aligned} t_T &= T\left(\hat{\rho}_T - 1\right) \left\{ \frac{T^{-2} \sum_{i=1}^r y_{i-1}^2}{s_T^2} \right\}^{\frac{1}{2}} \\ t_T &= \left[\frac{T^{-1} \cdot \sum_{t=1}^T Y_{t-1} \varepsilon_t}{T^{-2} \sum_{t=1}^T Y_{t-1}} \right] \cdot \left\{ \frac{T^{-2} \sum_{t=1}^r Y_{t-1}^2}{s_T^2} \right\}^{\frac{1}{2}} \\ t_T &= \frac{T^{-1} \sum_{t=1}^T Y_{t-1} \varepsilon_t}{[s_T^2]^{\frac{1}{2}} \left[T^{-2} \sum_{t=1}^T Y_{t-1}^2 \right]^{\frac{1}{2}}} \end{aligned} \tag{36}$$

Hence[8]

$$t_T \xrightarrow{L} \frac{(1/2)\sigma^2 \left\{ [W(1)]^2 - 1 \right\}}{\left\{ \sigma^2 \int_0^1 [W(r)]^2 dr \right\}^{1/2} \left\{ \sigma^2 \right\}^{1/2}} = \frac{(1/2) \left\{ [W(1)]^2 - 1 \right\}}{\left\{ \int_0^1 [W(r)]^2 dr \right\}^{1/2}} \tag{37}$$

■

2a

We will prove that:

$$T\left(\hat{\rho}_T - 1\right) \xrightarrow{L} \frac{\frac{1}{2} \left\{ [W(1)]^2 - 1 \right\} - W(1) \cdot \int_0^1 W(r) dr}{\int_0^1 [W(r)]^2 dr - \left[\int_0^1 W(r) dr \right]^2} \tag{38}$$

Firstly, the OLS estimation regarding the following equation

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t \tag{39}$$

asks for a matrix notation[9] given by

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T \end{bmatrix} = \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma Y_t \\ \Sigma Y_{t-1} Y_t \end{bmatrix} \tag{40}$$

[8]Since $\hat{\rho}_T$ is a consistent estimator we have that $s_T^2 \xrightarrow{p} \sigma^2$. Furthermore, as $T \to \infty$ we observe the result written above.

[9]To simplify notation let's denote Σ as $\sum_{t=1}^T$.

The OLS estimator in matrix form is:

$$\begin{bmatrix} \hat{\alpha}_T - 0 \\ \hat{\rho}_T - 1 \end{bmatrix} = \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma \varepsilon_t \\ \Sigma Y_{t-1} \varepsilon_t \end{bmatrix} \tag{41}$$

Worth mentioning Proposition 17.1[10] by Hamilton's Time Series Analysis:

Proposition 17.1: *Suppose that ξ_t follows a random walk without drift,*

$$\xi_t = \xi_{t-1} + u_t,$$

where $\xi_0 = 0$ and $\{u_t\}$ is an i.i.d. sequence with mean zero and variance σ^2. Then

(a) $T^{-1/2} \sum_{t=1}^{T} u_t \xrightarrow{L} \sigma \cdot W(1)$ [17.3.7];

(b) $T^{-1} \sum_{t=1}^{T} \xi_{t-1} u_t \xrightarrow{L} (1/2)\sigma^2\{[W(1)]^2 - 1\}$ [17.3.26];

(c) $T^{-3/2} \sum_{t=1}^{T} t u_t \xrightarrow{L} \sigma \cdot W(1) - \sigma \cdot \int_0^1 W(r)\, dr$ [17.3.19];

(d) $T^{-3/2} \sum_{t=1}^{T} \xi_{t-1} \xrightarrow{L} \sigma \cdot \int_0^1 W(r)\, dr$ [17.3.16];

(e) $T^{-2} \sum_{t=1}^{T} \xi_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 [W(r)]^2\, dr$ [17.3.22];

(f) $T^{-5/2} \sum_{t=1}^{T} t\xi_{t-1} \xrightarrow{L} \sigma \cdot \int_0^1 rW(r)\, dr$ [17.3.23];

(g) $T^{-3} \sum_{t=1}^{T} t\xi_{t-1}^2 \xrightarrow{L} \sigma^2 \cdot \int_0^1 r \cdot [W(r)]^2\, dr$ [17.3.24];

(h) $T^{-(\nu+1)} \sum_{t=1}^{T} t^{\nu} \rightarrow 1/(\nu + 1)$ *for* $\nu = 0, 1, \ldots$ [16.1.15].

Since Y_{t-1} has the same properties as the process in Proposition 17.1, we can apply the analysis carried out in the proposition and its respective properties to the process Y_{t-1}. Therefore, as the following terms have different rates of convergence we have the following results:

- $\Sigma Y_{t-1} = O_p\left(T^{3/2}\right)$
- $\Sigma Y_{t-1}\varepsilon_t = O_p(T)$
- $\Sigma Y_{t-1}^2 = O_p\left(T^2\right)$
- $\Sigma \varepsilon_t = O_p\left(T^{1/2}\right))$

Hence, based on the items above we can rewrite (41) as:

$$\begin{bmatrix} \hat{\alpha}_T - 0 \\ \hat{\rho}_T - 1 \end{bmatrix} = \begin{bmatrix} O_p(T) & O_p\left(T^{3/2}\right) \\ O_p\left(T^{3/2}\right) & O_p\left(T^2\right) \end{bmatrix}^{-1} \begin{bmatrix} O_p\left(T^{1/2}\right) \\ O_p(T) \end{bmatrix} \tag{42}$$

To find out $\hat{\alpha}_T$ and $\hat{\rho}_T$ limiting distributions, since these estimates have different rates of convergence, we will use a scaling matrix $\mathbf{Y}_T$ in the following matrix manipulations to make our life easier in order to prove the desired result:

$$\mathbf{Y}_T \equiv \begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix} \tag{43}$$

[10] Chapter 17, page 486

Hence,

$$\begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} \hat{\alpha}_T \\ \hat{\rho}_T - 1 \end{bmatrix} = \left\{ \begin{bmatrix} T^{-1/2} & 0 \\ 0 & T^{-1} \end{bmatrix} \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix} \begin{bmatrix} T_{-1/2} & 0 \\ 0 & T^{-1} \end{bmatrix} \right\}^{-1} \times \left\{ \begin{bmatrix} T_{-1/2} & 0 \\ 0 & T^{-1} \end{bmatrix} \begin{bmatrix} \Sigma\varepsilon_t \\ \Sigma Y_{t-1}\varepsilon_t \end{bmatrix} \right\} \tag{44}$$

$$\begin{bmatrix} T^{1/2}\hat{\alpha}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} = \begin{bmatrix} 1 & T^{-3/2}\Sigma Y_{t-1} \\ T^{-3/2}\Sigma Y_{t-1} & T^{-2}\Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-1/2}\Sigma\varepsilon_t \\ T^{-1}\Sigma Y_{t-1}\varepsilon_t \end{bmatrix} \tag{45}$$

Worth mentioning that:

- $$\begin{bmatrix} 1 & T^{-3/2}\Sigma Y_{t-1} \\ T^{-3/2}\Sigma Y_{t-1} & T^{-2}\Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \tag{46}$$

in equation 45 can be written[11] as, according to Hamilton's proposition 17.1:

$$\begin{aligned} \begin{bmatrix} 1 & T^{-3/2}\Sigma Y_{t-1} \\ T^{-3/2}\Sigma Y_{t-1} & T^{-2}\Sigma Y_{t-1}^2 \end{bmatrix} &\xrightarrow{L} \begin{bmatrix} 1 & \sigma \cdot \int W(r)\,dr \\ \sigma \cdot \int W(r)\,dr & \sigma^2 \cdot \int [W(r)]^2\,dr \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} 1 & \int W(r)\,dr \\ W(r)\,dr & \int [W(r)]^2\,dr \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}, \end{aligned} \tag{47}$$

Moreover, still in equation 45 we have:

- $$\begin{bmatrix} T^{-1/2}\Sigma\varepsilon_t \\ T^{-1}\Sigma Y_{t-1}\varepsilon_t \end{bmatrix} \tag{48}$$

which can be written as:

$$\begin{bmatrix} T^{-1/2}\Sigma\varepsilon_t \\ T^{-1}\Sigma Y_{t-1}\varepsilon_t \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \sigma \cdot W(1) \\ (1/2)\sigma^2 \left\{ [W(1)]^2 - 1 \right\} \end{bmatrix} = \sigma \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} W(1) \\ (1/2)\left\{ [W(1)]^2 - 1 \right\} \end{bmatrix} \tag{49}$$

due to its asymptotic result. Talking about asymptotic result, when (47) and (49) are joined in (45) we have:

$$\begin{aligned} \begin{bmatrix} T^{1/2}\boldsymbol{\hat{\alpha}}_T \\ T(\hat{\rho}_T - 1) \end{bmatrix} &\xrightarrow{L} \sigma \cdot \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int [W(r)]^2 dr \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} W(1) \\ (1/2)\left\{ [W(1)]^2 - 1 \right\} \end{bmatrix} \\ &= \begin{bmatrix} \sigma & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int [W(r)]^2 dr \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} W(1) \\ (1/2)\left\{ [W(1)]^2 - 1 \right\} \end{bmatrix} \end{aligned} \tag{50}$$

Furthermore:

$$\begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int [W(r)]^2 dr \end{bmatrix}^{-1} = \Delta^{-1} \begin{bmatrix} \int [W(r)]^2 dr & -\int W(r)dr \\ -\int W(r)dr & 1 \end{bmatrix}, \tag{51}$$

where

$$\Delta \equiv \int [W(r)]^2 dr - \left[\int W(r)dr \right]^2 \tag{52}$$

Therefore, $\widehat{\rho_T}$ in the vector expression in (50) states that:

$$T(\hat{\rho}_T - 1) \xrightarrow{L} \frac{\frac{1}{2}\left\{ [W(1)]^2 - 1 \right\} - W(1) \cdot \int_0^1 W(r)dr}{\int_0^1 [W(r)]^2 dr - \left[\int_0^1 W(r)dr \right]^2} \tag{53}$$

■

[11] $\int$ is equivalent to $\int_0^1$

2b

In order to prove this result:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\hat{\rho}_T}} = \frac{T(\hat{\rho}_T - 1)}{(T^2\hat{\sigma}^2_{\hat{\rho}_T})^{\frac{1}{2}}} \xrightarrow{L} \frac{\frac{1}{2}\left\{[W(1)]^2 - 1\right\} - W(1)\cdot\int_0^1 W(r)dr}{\left\{\int_0^1 [W(r)]^2 dr - \left[\int_0^1 W(r)dr\right]^2\right\}^{1/2}} \tag{54}$$

we will use the following OLS t-test based on $H_0 : \rho = 1$:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\hat{\rho}_T}} = \frac{T(\hat{\rho}_T - 1)}{(T^2\hat{\sigma}^2_{\hat{\rho}_T})^{\frac{1}{2}}} \tag{55}$$

such that:

$$\begin{aligned} \hat{\sigma}^2_{\rho_T} &= s_T^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ s_T^2 &= (T-2)^{-1} \sum_{t=1}^{T} (Y_t - \hat{\alpha}_T - \hat{\rho}_T Y_{t-1})^2 \end{aligned} \tag{56}$$

When we multiply both sides of $\hat{\sigma}^2_{\rho_T}$'s equation we have:

$$\begin{aligned} T^2 \cdot \hat{\sigma}^2_{\hat{\rho}_T} &= s_T^2 \begin{bmatrix} 0 & T \end{bmatrix} \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ T \end{bmatrix} \\ &= s_T^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \mathbf{Y}_T \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \mathbf{Y}_T \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \tag{57}$$

Recall that:

$$\begin{aligned} &\mathbf{Y}_T \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \mathbf{Y}_T \\ &= \left\{ \mathbf{Y}_T^{-1} \begin{bmatrix} T & \Sigma Y_{t-1} \\ \Sigma Y_{t-1} & \Sigma Y_{t-1}^2 \end{bmatrix} \mathbf{Y}_T^{-1} \right\}^{-1} \\ &= \begin{bmatrix} 1 & T^{-3/2}\Sigma Y_{t-1} \\ T^{-3/2}\Sigma Y_{t-1} & T^{-2}\Sigma Y_{t-1}^2 \end{bmatrix}^{-1} \\ &\xrightarrow{L} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}^{-1} \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int W(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix}^{-1}. \end{aligned} \tag{58}$$

Then, asymptotically:

$$T^2 \cdot \hat{\sigma}^2_{\hat{\rho}_T} \xrightarrow{L} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \int W(r)dr \\ \int W(r)dr & \int W(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\int [W(r)]^2 dr - \left[\int W(r)dr\right]^2} \tag{59}$$

Therefore, we have the desired result when the Continuous Mapping Theorem and the Functional Central Limit Theorem are applied to (59), which results in the asymptotic distribution of the *OLS t* test:

$$t_T = \frac{\hat{\rho}_T - 1}{\hat{\sigma}_{\hat{\rho}_T}} = \frac{T(\hat{\rho}_T - 1)}{(T^2\hat{\sigma}^2_{\hat{\rho}_T})^{\frac{1}{2}}} \xrightarrow{L} \frac{\frac{1}{2}\left\{[W(1)]^2 - 1\right\} - W(1)\cdot\int_0^1 W(r)dr}{\left\{\int_0^1 [W(r)]^2 dr - \left[\int_0^1 W(r)dr\right]^2\right\}^{1/2}} \tag{60}$$

■

3.3

We will prove that:

$$\begin{aligned} \begin{bmatrix} T^{-1/2}\Sigma u_t \\ T^{-3/2}\Sigma y_{t-1}u_t \end{bmatrix} &\xrightarrow{p} \begin{bmatrix} T^{-1/2}\Sigma u_t \\ T^{-3/2}\Sigma \alpha(t-1)u_t \end{bmatrix} \\ &\xrightarrow{L} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix} \right) \quad = \quad \mathrm{N}\left(0, \sigma^2 \mathbf{Q}\right) \end{aligned}$$

based on the random walk with drift true process:

$$Y_t = \alpha + Y_{t-1} + \varepsilon_t \tag{61}$$

such that, the true value of α is not zero. It has an effect on the asymptotic distribution of $\hat{\alpha}$ and $\hat{\rho}$. To see why, note that [17.4.38] implies that

$$Y_t = Y_0 + \alpha t + (\varepsilon_1 + \cdots + \varepsilon_t) = Y_0 + \alpha t + \xi_t$$

where

$$\xi_t \equiv \varepsilon_1 + \cdots + \varepsilon_t \quad \text{for } t = 1, 2, \ldots, T$$

with $\xi_0 \equiv 0$. Then,

$$\sum_{i=1}^{T} Y_{t-1} = \sum_{i=1}^{T} \left[Y_0 + \alpha(t-1) + \xi_{t-1}\right]. \tag{62}$$

Based on Hamilton's proposition 17.1 we can see that (62) can be written as:

$$\sum_{i=1}^{T} Y_{t-1} = \underbrace{\sum_{t=1}^{T} Y_0}_{O_p(T)} + \underbrace{\sum_{i=1}^{T} \alpha(t-1)}_{O_p(T^2)} + \underbrace{\sum_{i=1}^{T} \xi_{t-1}}_{O_p\left(T^{3/2}\right)} \tag{63}$$

Notice that the term:

$$\underbrace{\sum_{i=1}^{T} \alpha(t-1)}_{O_p(T^2)}$$

asymptotically dominates the terms

$$\underbrace{\sum_{t=1}^{T} Y_0}_{O_p(T)}$$

and

$$\underbrace{\sum_{i=1}^{T} \xi_{t-1}}_{O_p\left(T^{3/2}\right)}$$

leading to:

$$\begin{aligned} T^{-2} \sum_{t=1}^{T} Y_{t-1} &= T^{-1} Y_0 + T^{-2} \sum_{t=1}^{T} \alpha(t-1) + T^{-1/2} \left\{ T^{-3/2} \sum_{t=1}^{T} \xi_{t-1} \right\} \\ &\xrightarrow{p} 0 + \alpha/2 + 0. \end{aligned} \tag{64}$$

which is followed by:

$$\begin{aligned} \sum_{t=1}^{T} Y_{t-1}^2 &= \sum_{t=1}^{T} \left[Y_0 + \alpha(t-1) + \xi_{t-1}\right]^2 \\ &= \underbrace{\sum_{t=1}^{T} Y_0^2}_{O_p(T)} + \underbrace{\sum_{t=1}^{T} \alpha^2 (t-1)^2}_{O_p(T^3)} + \underbrace{\sum_{t=1}^{T} \xi_{t-1}^2}_{O_p(T^2)} \\ &+ \underbrace{\sum_{t=1}^{T} 2Y_0 \alpha(t-1)}_{O_p(T^2)} + \underbrace{\sum_{t=1}^{T} 2Y_0 \xi_{t-1}}_{O_p\left(T^{3/2}\right)} + \underbrace{\sum_{t=1}^{T} 2\alpha(t-1)\xi_{t-1}}_{O_p\left(T^{3/2}\right)}. \end{aligned} \tag{65}$$

Notice that $\lim_{T\to\infty} \frac{\alpha^2(t-1)^2}{T^3} \neq 0$ is the only term that survives asymptotically, then:

$$T^{-3}\sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{p} \alpha^2/3 \tag{66}$$

Furthermore,

$$\begin{aligned}\sum_{t=1}^{T} Y_{t-1}\varepsilon_t &= \sum_{t=1}^{T}\left[Y_0 + \alpha(t-1) + \xi_{t-1}\right]\varepsilon_t \\ &= \underbrace{Y_0\sum_{t=1}^{T}\varepsilon_t}_{o_p(T^2)} + \underbrace{\sum_{t=1}^{T}\alpha(t-1)\varepsilon_t}_{O_p\left(T^{3/2}\right)} + \underbrace{\sum_{t=1}^{T}\xi_{t-1}\varepsilon_t}_{O_p\left(T^{3/2}\right)}\end{aligned} \tag{67}$$

and

$$T^{-3/2}\sum_{t=1}^{T} Y_{t-1}\varepsilon_t \xrightarrow{p} T^{-3/2}\sum_{t=1}^{T}\alpha(t-1)\varepsilon_t \tag{68}$$

Resulting in

$$\begin{bmatrix} \hat{\alpha}_T - \alpha \\ \hat{\rho}_T - 1 \end{bmatrix} = \begin{bmatrix} O_p(T) & O_p\left(T^2\right) \\ O_p\left(T^2\right) & O_p\left(T^3\right) \end{bmatrix}^{-1} \begin{bmatrix} O_p\left(T^{1/2}\right) \\ O_p\left(T^{3/2}\right) \end{bmatrix}. \tag{69}$$

Also, $\mathbf{Y}_T$ is given by:

$$\mathbf{Y}_T \equiv \begin{bmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{bmatrix} \tag{70}$$

such that:

$$\begin{bmatrix} 1 & T^{-2}\Sigma Y_{t-1} \\ T^{-2}\Sigma Y_{t-1} & T^{-2}\Sigma Y_{1-1}^2 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix} \equiv \mathbf{Q}. \tag{71}$$

Therefore

$$\begin{aligned}\begin{bmatrix} T^{-1/2}\Sigma u_t \\ T^{-3/2}\Sigma y_{t-1}u_t \end{bmatrix} &\xrightarrow{p} \begin{bmatrix} T^{-1/2}\Sigma u_t \\ T^{-3/2}\Sigma\alpha(t-1)u_t \end{bmatrix} \\ &\xrightarrow{L} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2\begin{bmatrix} 1 & \alpha/2 \\ \alpha/2 & \alpha^2/3 \end{bmatrix}\right) \quad = \quad \mathrm{N}\left(0, \sigma^2\mathbf{Q}\right)\end{aligned} \tag{72}$$

■

Problem 4

According to Hamilton's Time Series Analysis[12] and the slides provided by the teacher[13]:
Let $\{Y_t\}$ be a AR(p) stochastic process possibly with a trend term, i.e.,

$$Y_t = \alpha + \delta\cdot t + \phi_1\cdot Y_{t-1} + \ldots \phi_p\cdot Y_{t-p} + \varepsilon_t$$

We want to test wether $\{Y_t\}$ has one unit root, i.e., if the polynomial function

$$1 - \phi_1\cdot z - \phi_2\cdot z^2 - \ldots - \phi_p\cdot z^p \tag{73}$$

has one and only one unit root. To test for the presence of one unit root, we implement the Augment Dickey-Fuller test. It uses the following estimating regressions:

1. No Drift and No Deterministic Time Trend
2. With Drift but No Deterministic Time Trend
3. With Drift and a Deterministic Time Trend

[12] Chapter 17
[13] Lecture 02C unit root models

We start by testing the most restrictive model, i.e., the model with drift and a deterministic time trend, illustrated by:

$$\Delta Y_t = \gamma \cdot Y_{t-1} + \delta \cdot t + \alpha + \sum_{i=2}^{p} \beta_i \cdot \Delta Y_{t-i+1} + \varepsilon_t \tag{74}$$

$$\rightarrow (\phi 2) \quad H_0 : \rho = 1 \quad \text{and} \quad \delta = 0 \quad \text{and} \quad \alpha = 0 \tag{75}$$

$$\rightarrow (\phi 3) \quad H_0 : \rho = 1 \quad \text{and} \quad \delta = 0 \tag{76}$$

$$\rightarrow (\tau 3) \quad H_0 : \rho = 1 \tag{77}$$

such that:
Test (ϕ_2): Presence of unit root, absence of drift and absence time trend under the null.
Test (ϕ_3): Presence of a unit root and absence of time trend under the null.
Test (τ_3): Presence of a unit root under the null.

The time series for corn production (tonnes) in the U.S, using the data between 1950 and 2021 can be illustrated as follows:

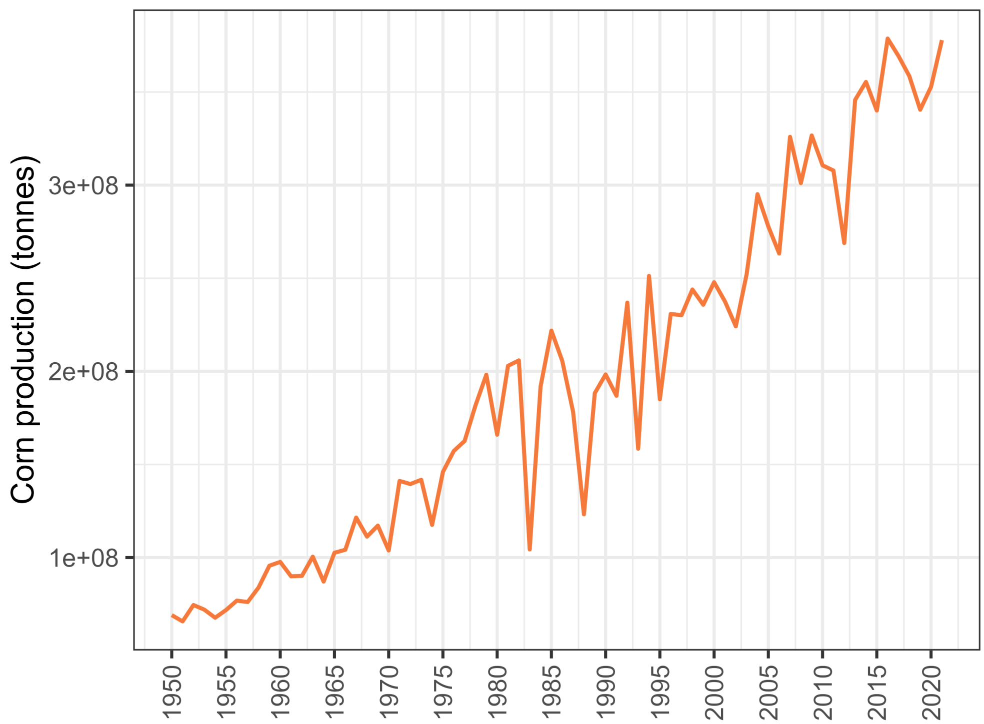

Figure 1: Corn production between 1950 to 2021

Upon observing the graph, it becomes apparent that the **corn production does not fluctuate around a mean value**. This characteristic does not align with a unit root process, as the unit root process implies that, the best predictor for tomorrow is today's realization. To formally assess this, we conduct the **Augmented Dickey-Fuller (ADF) test with drift and deterministic time trend** because, graphically, we can notice a deterministic trend in the series illustrated above.

First, we ran a test using Bayesin Information Criterion (BIC) to choose the number of lags for the dependent variable and we see from the Table 3 that we used information criterion that indicates lag of one.

Table 3: **Regression with drift and time trend**

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	26840735.72	8353042.59	3.2133	0.0020**
Lag 1	-0.6705	0.1543	-4.3468	0.0000***
Time Trend	2981665.75	683939.07	4.36	0.0000***
1^{st} Difference	-0.1808	0.1207	-1.4976	0.1390

Significance levels: 0 '***' — 0.001 '**' — 0.01 '*'

Table 4 provides information on the test statistic values and critical values for the test regarding 1%, 5% and 10% rejection rate, respectively,

Table 4: **ADF Test**

	Statistic	Critical Value (1%)	Critical Value (5%)	Critical Value (10%)	Rejection
τ_3	-4.347	-4.040	-3.450	-3.150	Yes
ϕ_2	7.864	6.500	4.880	4.160	Yes
ϕ_3	9.598	8.730	6.490	5.470	Yes

which leads us to the following conclusions about the Augmented Dickey-Fuller test, also, based on the R code below:

```
#-------------------------------------------------------------------------------
# Augmented Dickey-Fuller Test: Drift and Time Trend
# Run the test using BIC to choose the number of lags
df_drift <- ur.df(
  y = dt_corn$corn_prod, # Vector to be tested for unit root
  type = 'trend',
  selectlags = c("BIC") # Lag selection can be achieved according to "BIC"
)

# The first test statistic is tau3 in the slides, while the second one is phi2 and the third
# one is phi3. Their critical values are reported at the bottom.
print(summary(df_drift))
```

Figure 2: Augmented Dickey-Fuller Test: Drift and Time Trend

Thus, R indicates that we are conducting a **one-tailed test**. In this case, when the test statistic value exceeds the critical value reported in Table 4, we reject the null hypothesis., which is confirmed in the rejection column in Table 4 for all tests. Then, for the following tests:

- ϕ_2: Under H_0, the hypothesis indicating the presence of a unit root, absence of drift, and absence of a time trend is **rejected**.
- ϕ_3: Under H_0, the hypothesis indicating the presence of a unit root and absence of drift is **rejected**.
- τ_3: Under H_0 the presence of a unit root is **rejected**.

In this scenario, the **process does not exhibit a unit root as the non-stationarity of the process is attributed to the presence of the time trend component**.