Econ 722: Time Series Assignment 3

Name(s):Benjamin O. Harrison

1. We will learn how to do Monte Carlo simulations for a simple AR model. You can use any software of your preference. Just post the output (I do not want to see your code but you are welcome to post that as well if you want).

   (a) Generate a sample of T = 100 observations from an AR(1) with $\rho = 0.5$. Using 5,000 Monte Carlo simulations, compute the distribution of the OLS estimate for ρ.

   Answer
   Figure 1 represents the distribution of the $\hat{\rho}$ for 5000 MC simulations. One will observe that the distribution is normally distributed but the OLS estimate is a bit biased.

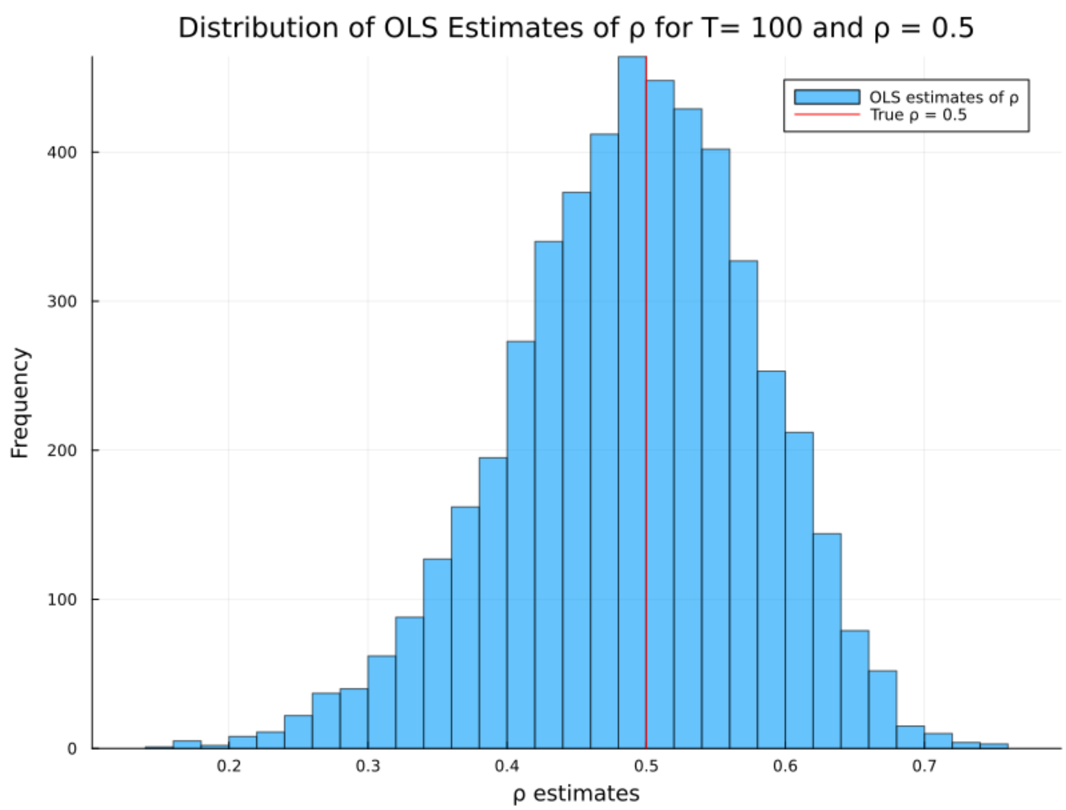

Figure 1: The distribution of $\hat{\rho}$ estimates using 5000 MC simulations.

   (b) Do the same as in a. but this time compute the empirical distribution for the $t - statistics$ for the test $\rho = 0.5$.

Answer
In this part, we first compute the HAC standard errors of $\hat{\rho}$, which is done using JULIA. Given this HAC standard error, we compute the t statistics which is given as

$$T = \frac{\hat{\rho} - \rho}{\sqrt{V_{\hat{\rho}}}}$$

. This is done for 5000 MC simulations and one can observe from Figure 2 that the T statistics follow a normal distribution with mean 0 and variance 1.

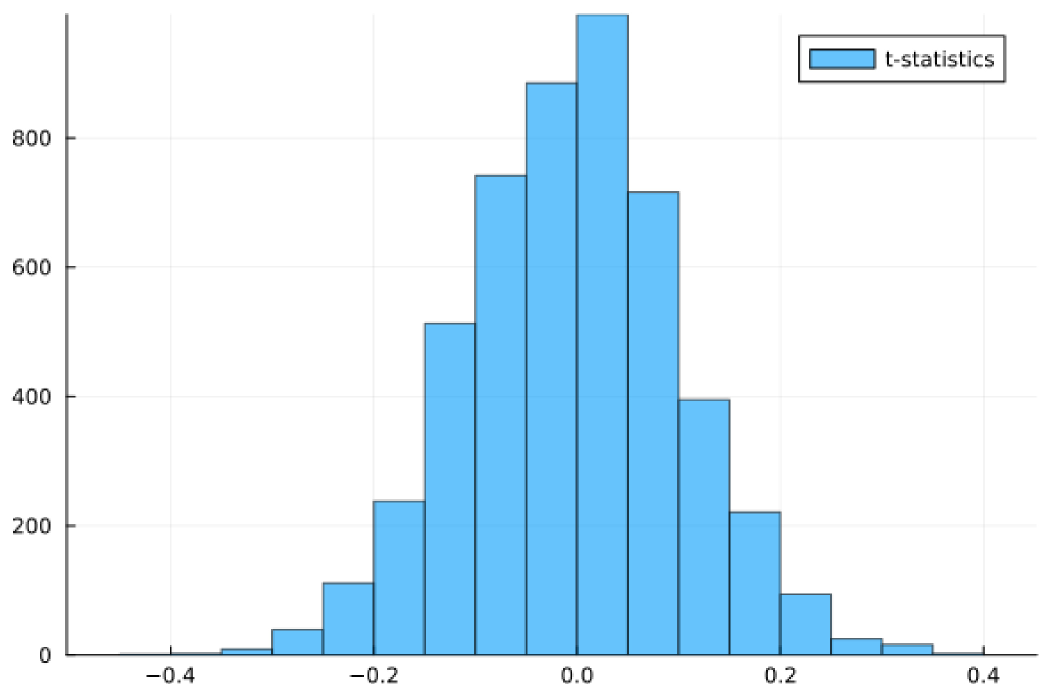

Figure 2: The distribution of $T - stat$ estimates using 5000 MC simulations.

(c) Do the same for $\rho = 1$ and the t-statistics for the test $\rho = 1$.
Answer
Figure 3 computes the distribution of $\hat{\rho}$ and the T statistics when the true parameter $\rho = 1$. One can observe that distribution for $\hat{\rho}$ is skwed to the left and truncated at 1. This is because the process diverges for any $\rho > 1$ and the OLS estimate does not exist. So the mass of the estimates is around 0.9 to 1.

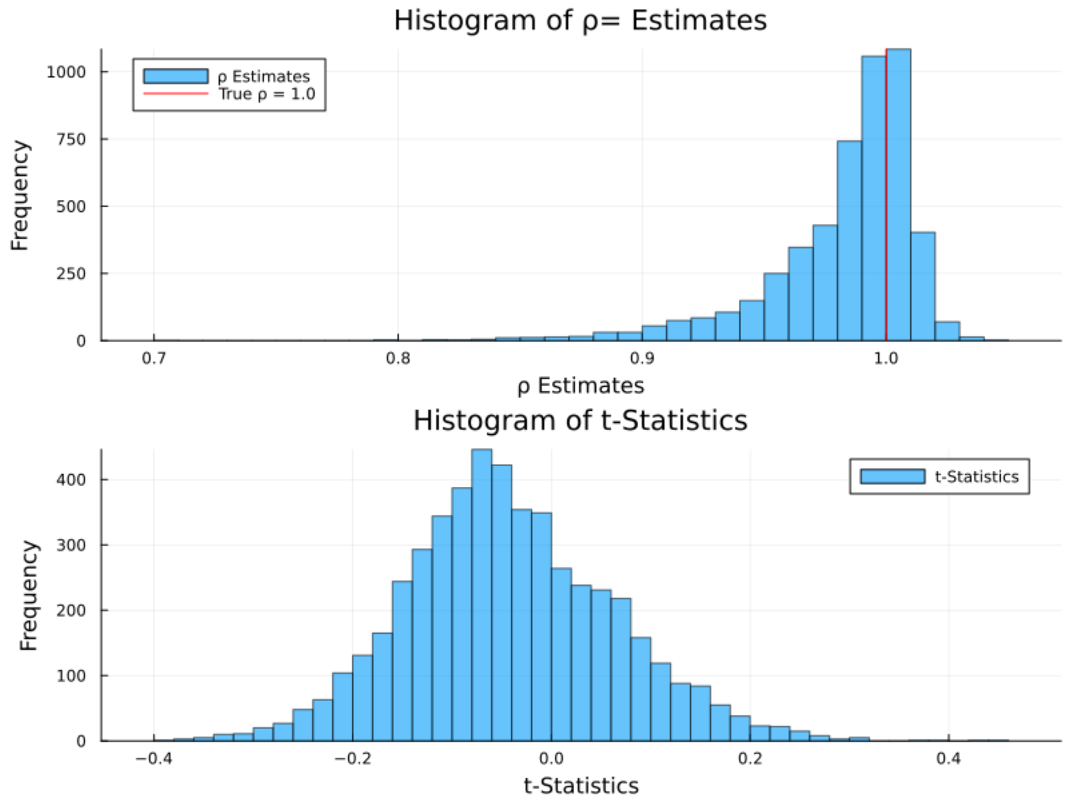

Figure 3: The distribution of T_stat and $\hat{\rho}$ estimates using 5000 MC simulations.

(d) Repeat a,b, and c for T = 1000. How do things change?

Answer
One main difference in the $\hat{\rho}$ distribution between T=100 and T=1000 is that the bias in the distribution reduces. This demonstrates that the OLS estimate is consistent that as we increase the sample size, the estimator converges to the true estimator. As can be seen in Figure 4.

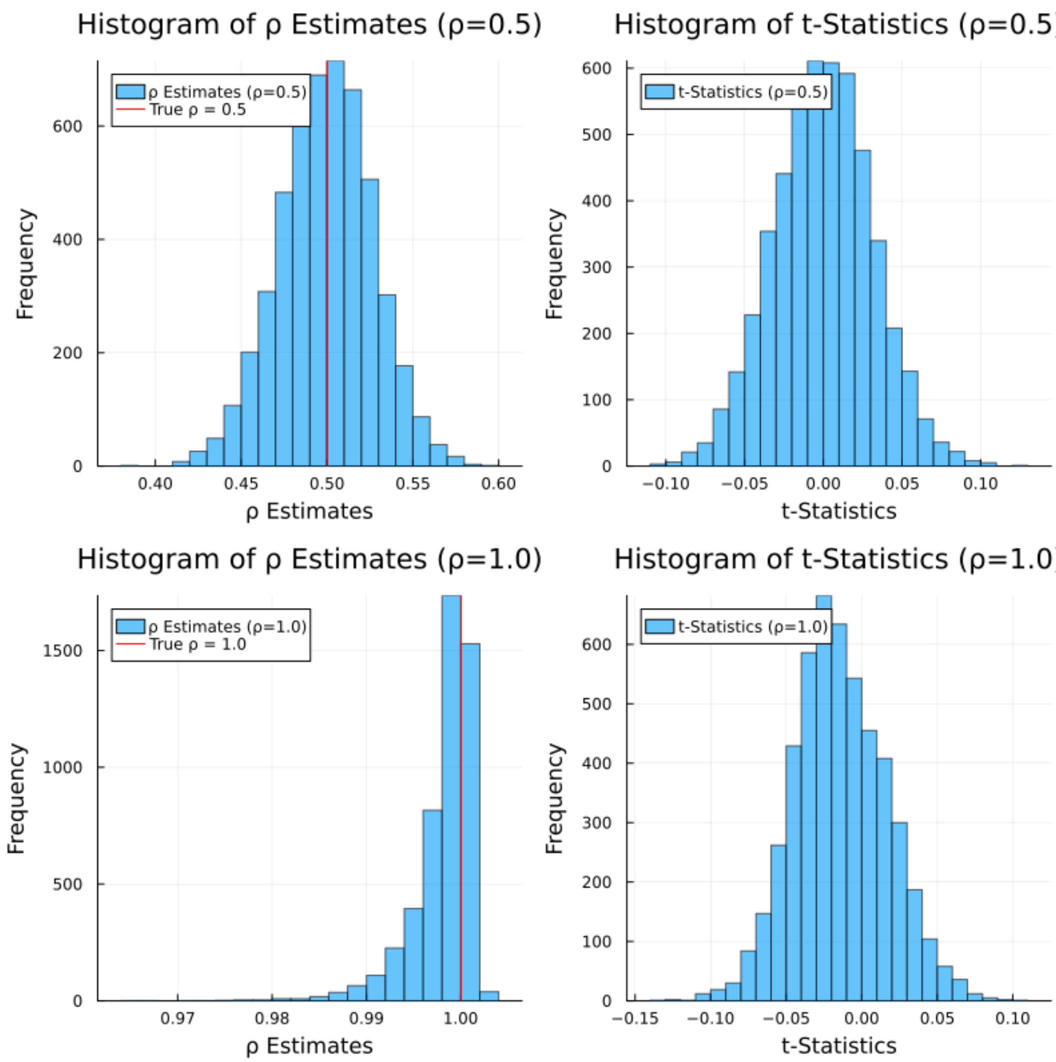

Figure 4: The distribution of $T_s tat$ and $\hat{\rho}$ estimates using 5000 MC simulations with T = 1000.

(e) Repeat a and b and c for T = 100 and T = 1000 and $\rho = 0.90$ and $\rho = 0.99$. Compare your results and discuss.

Answer

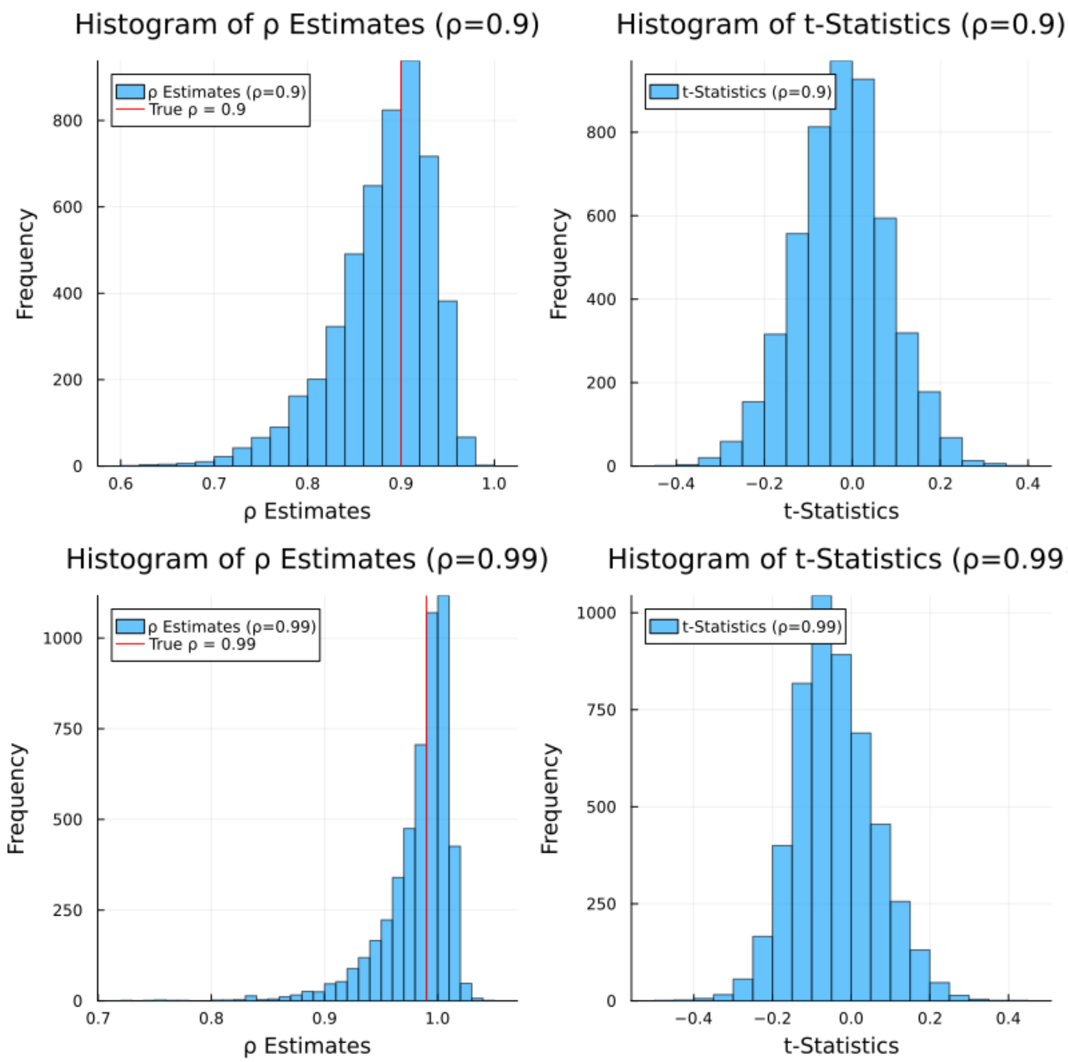

Figure 5: The distribution of T_stat and $\hat{\rho}$ estimates using 5000 MC simulations with T = 1000.

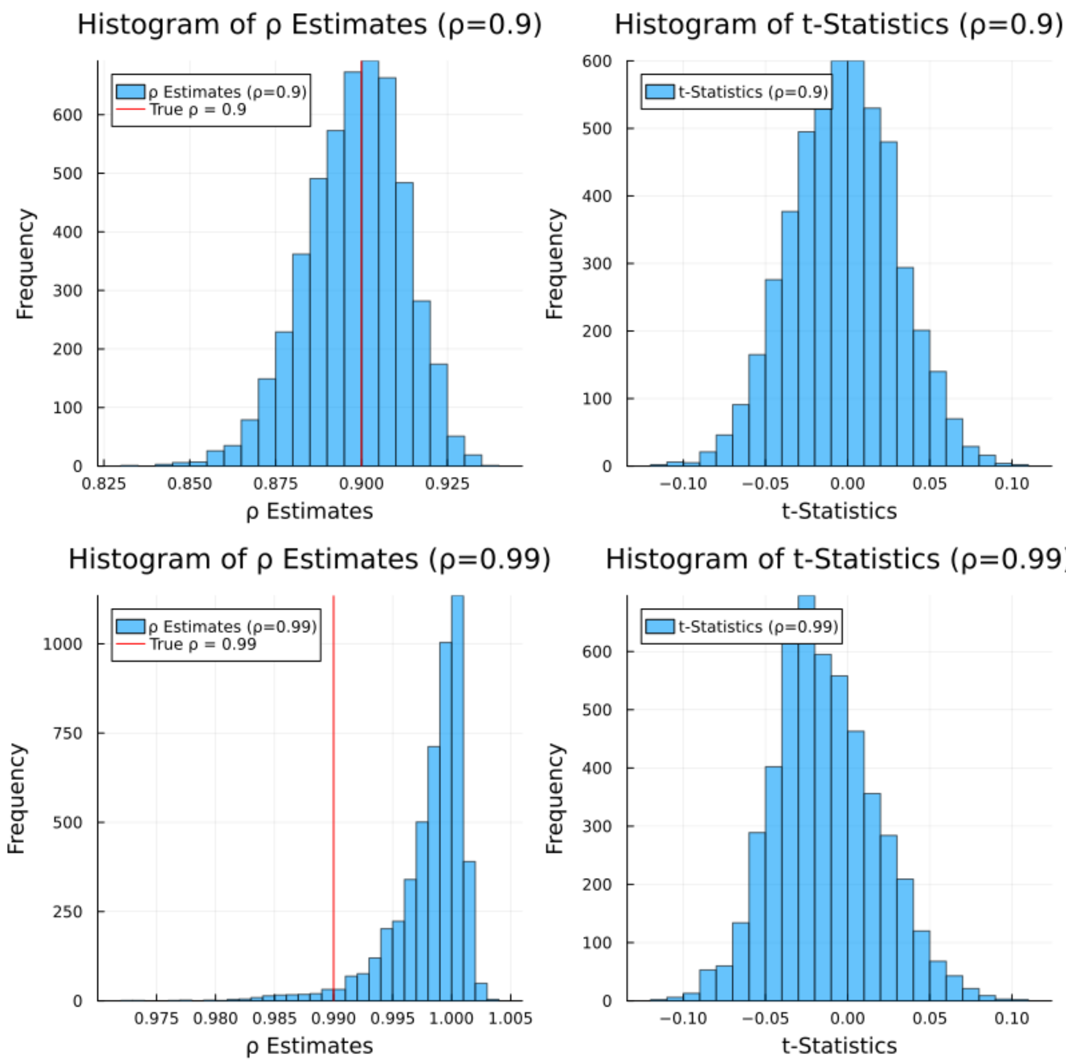

Figure 6: The distribution of $T_s tat$ and $\hat{\rho}$ estimates using 5000 MC simulations with T = 1000.

2. Go to the FRED database (the data base of the Federal Reserve). It contains all sorts of macro data for the US. We want to look at a measure of GDP for the US. Choose one and download the data.

   (a) Using your program of choice, estimate the autocorrelation and partial autocorrelation function for USGDP. What can you say just looking at the correlogram?
   Answer:
   Figure 7 represents the ACF and PACF plots of USGDP. It is interesting to see that the ACF for as many lags as 20 have higher predictive power but the PACF actually shows that the first two lags are the most relevant lags with predictive power. This is because the USGDP has trend behavior and so each previous lag is highly correlated but the first two lags seems to contain all the information needed.

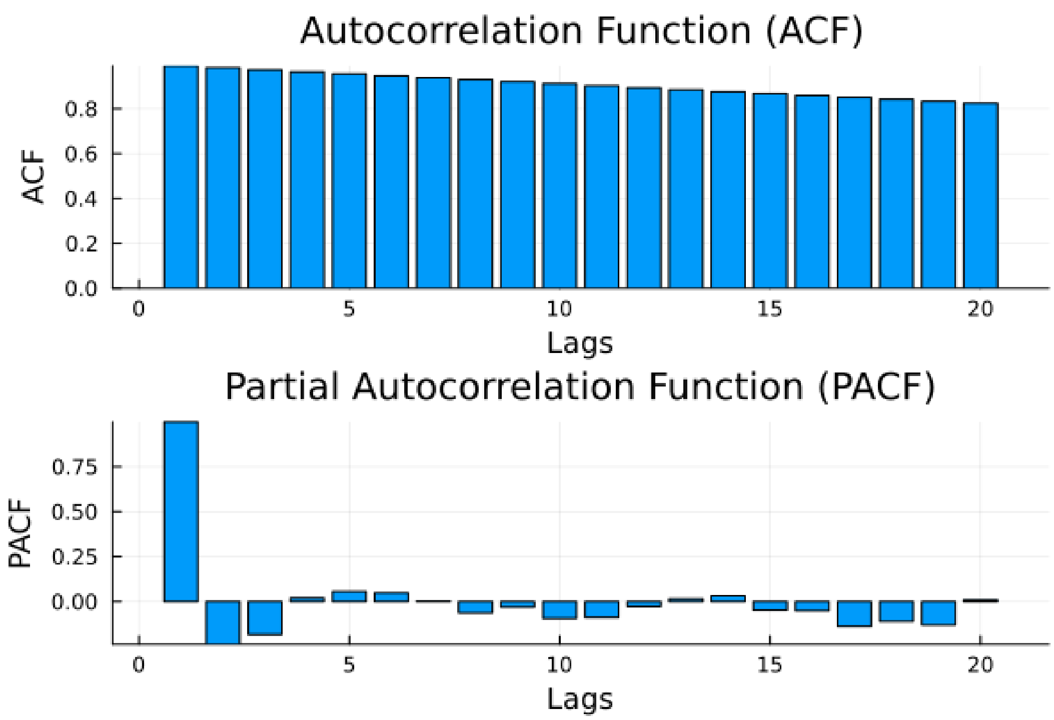

Figure 7: The ACF and PACF of the log GDP of US.

   (b) Estimate the model that is suggested by the correlogram and compute the one and two-step-ahead forecasts. (If you have more than one model in mind, compute the forecasts with both models).

   Answer:
   By observing the ACF, one may believe in using an AR(p) model with p = 20 lags. But the PACF tells us that only the first two lags are relevant. One way to go about

this is to let the data decide the lags that are relevant by employing the Bayesian criteria information which allows us to run the AR model for say 10 lags and choose the optimal lag. By employing this method, the BIC criteria provide the optimal number of lags to use is 2 which is shown from the PACF plot. Figure **??** shows the estimated model with the optimal lag.
We then compute the one-step and two-step given as:

$$\hat{y_{t+1}} = \hat{\alpha_0} + \hat{\alpha_1} y_t + \hat{\alpha_2} y_{t-1} = 10.2554$$
$$\hat{y_{t+2}} = \hat{\alpha_0} + \hat{\alpha_1} \hat{y_{t+1}} + \hat{\alpha_2} y_t = 10.2659$$

Table 1: Model Coefficients

Variable	Coef.	Std. Error	t	Pr(>|t|)	Lower 95%	Upper 95%
x1	0.0219297	0.00425687	5.15	<1e-06	0.0135532	0.0303063
x2	-0.237099	0.0554976	-4.27	<1e-04	-0.346305	-0.127892
x3	1.23584	0.0555836	22.23	<1e-65	1.12647	1.34522

(c) Estimate a AR(1) model for USGDP. What is the value of the root?

Table 2: AR(1) Model Coefficients

Variable	Coef.	Std. Error	t	Pr(>|t|)	Lower 95%	Upper 95%
x1	0.0282028	0.0040521	6.96	<1e-10	0.0202294	0.0361762
x2	0.99842	0.000493421	2023.46	<1e-99	0.997449	0.999391

$$\implies root = \frac{1}{\alpha} = 1.0015$$

(d) Test for the presence of a unit root in USGDP using any test we talked about in class. Do you reject or not? Make sure you think about which is the relevant model we should test (deterministic, no deterministic etc.).

Answer:
We compute the Augmented Dickey fuller test as a test for unit root where we test the null that $\alpha = 1$. From the summary of the test below, we obtain a pvalue of 0.995 which is very high, showing that we fail to reject the null. Hence the GDP has a unit root.

- **Population details:**
  - Parameter of interest: Coefficient on lagged non-differenced variable

- Value under H_0: 0
- Point estimate: 0.000485644

- **Test summary:**
  - Outcome with 95% confidence: Fail to reject H_0
  - p-value: 0.9953
- **Details:**
  - Sample size in regression: 308
  - Number of lags: 1
  - ADF statistic: 0.131724
  - Critical values at 1%, 5%, and 10%: $[-3.98827, -3.42475, -3.13543]$

3. You have two possible models for your data. Model "a" is a trend stationary model of the form $y_t = \alpha + \beta t + \varepsilon_t + v\varepsilon_{t-1}$. Model "b" is the unit root model $(1-L)y_t = \delta + \varepsilon_t + v\varepsilon_{t-1}$

   (a) Compute the s-period ahead forecast error and MSE for both models. What is the difference between the forecast errors and the MSE for the two models, as the forecast horizon gets large?

   Answer:
   Consider the model given by :

   $$y_t = \alpha + \beta t + \epsilon_t + \vartheta\epsilon_{t-1}$$

   where ϵ_t is white noise with mean zero and variance σ^2. The forecast for y_{t+s} based on information at time t is:

   $$\hat{y}_{t+s} = \alpha + \beta(t+s)$$

   Since ϵ_t and ϵ_{t-1} are not predictable beyond time t, the forecast error is:

   $$e_{t+s} = y_{t+s} - \hat{y}_{t+s} = \epsilon_{t+s} + \vartheta\epsilon_{t+s-1}$$

   We compute the MSE as the expected value of the squared forecast error:

   $$\text{MSE}_\text{a}(s) = \mathbb{E}[e_{t+s}^2] = \mathbb{E}[(\epsilon_{t+s} + \vartheta\epsilon_{t+s-1})^2]$$

   Since ϵ_t is white noise, we have:

   $$\text{MSE}_\text{a}(s) = \mathbb{E}[\epsilon_{t+s}^2] + \vartheta^2\mathbb{E}[\epsilon_{t+s-1}^2] = \sigma^2(1+\vartheta^2)$$

   Thus, the forecast error for model a remains constant at $\sigma^2(1+\vartheta^2)$ for all forecast horizons.

Now consider the unit root model:

$$(1 - L)y_t = \delta + \epsilon_t + \vartheta\epsilon_{t-1}$$

which can be written as:

$$y_t = y_{t-1} + \delta + \epsilon_t + \vartheta\epsilon_{t-1}$$

The forecast for y_{t+s} based on information at time t is:

$$\hat{y}_{t+s} = y_t + s\delta$$

The forecast error is then:

$$e_{t+s} = y_{t+s} - \hat{y}_{t+s} = \sum_{i=1}^{s} \epsilon_{t+i} + \vartheta\epsilon_{t+s-1}$$

Similarly, the MSE is given as:

$$\text{MSE}_\text{b}(s) = \mathbb{E}\left[\left(\sum_{i=1}^{s} \epsilon_{t+i} + \vartheta\epsilon_{t+s-1}\right)^2\right]$$

Expanding this expression:

$$\text{MSE}_\text{b}(s) = \sum_{i=1}^{s} \mathbb{E}[\epsilon_{t+i}^2] + \vartheta^2\mathbb{E}[\epsilon_{t+s-1}^2]$$

Since ϵ_t is white noise:

$$\text{MSE}_\text{b}(s) = s\sigma^2 + \vartheta^2\sigma^2 = \sigma^2(s + \vartheta^2)$$

Thus, the forecast error increases with the forecast horizon s.

In conclusion, we can observe that in the trend stationary model a, the MSE is constant at $\sigma^2(1 + \vartheta^2)$, meaning the forecast accuracy does not worsen with time. However, with the unit root model b, the MSE grows linearly with the forecast horizon s, $\sigma^2(s + \vartheta^2)$, indicating that the forecasts become less reliable as s increases.