STAT443 Time Series Analysis Project

Ziqiu Zhu
20519999

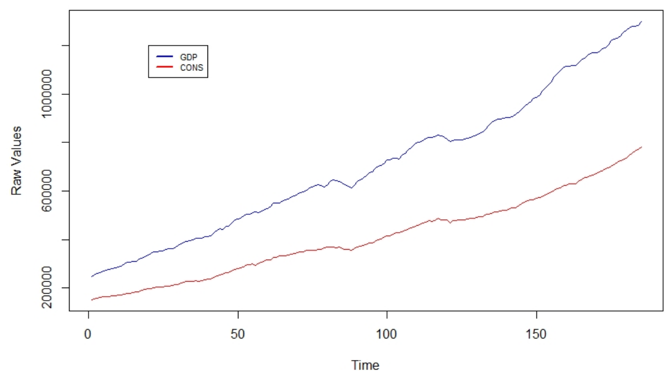
July 24 2018

1 Raw Data Series

For the first 5 sections, this paper aims to analyze using R program, the raw time series: (1) W_{1t}, 185 quarterly observations of the seasonally adjusted Canadian GDP series from Q1 1961 to Q1 of 2007, and (2) W_{2t}, 185 quarterly observations of the seasonally adjusted Canadian Person Expenditure on Consumer Goods and Services (CONS) data from the same time period.

The following figure is a graphical representation of W_{1t} and W_{2t}:

Figure 1: Plot of GDP and CONS time series from 1961 Q1 to 2007 Q1



2 Loglinear Law and Decomposition Law

Our objective is to model W_{1t} using simple laws (i.e. linear law), however the raw data series by itself it does not obey linear laws. In order to account for this, we use the *Loglinear Law*, or rather we take natural logarithmic transformation of W_{1t}, define $X_t = \ln(W_{1t})$, so that it we can easily describe it using linear laws. Using *Decomposition Law*, $X_t = \ln(W_{1t})$ can be decomposed as

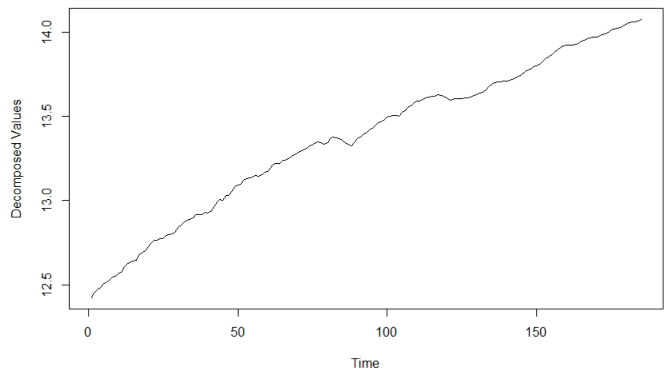
$$X_t = T_t + Y_t + S_t$$

where T_t is the trend, Y_t is the cycle, and S_t is the seasonal component. By *Stationary Law*, Y_t is stationary, but X_t,T_t, and S_t are not.

Figure 2 below is a graphical representation of the $X_t = \ln(W_{1t})$:

From figure 2, the post-decomposed data series shows no significant seasonal trends, as well as a positive linear trend over time, indicating the growth of Canadian GDP since 1961. Two significant drops can be visibly seen from the plot: the first at 82-89 indices (1981 Q2 to 1982 Q4 time stamps), and the second at the 120-122 indices (1990 Q3 to 1991 Q1 time stamps), which corresponds to the recessions Canada faced during those times.

Figure 2: Plot of Decomposed GDP Times Series X_t



2.1 Trend Stationary Model

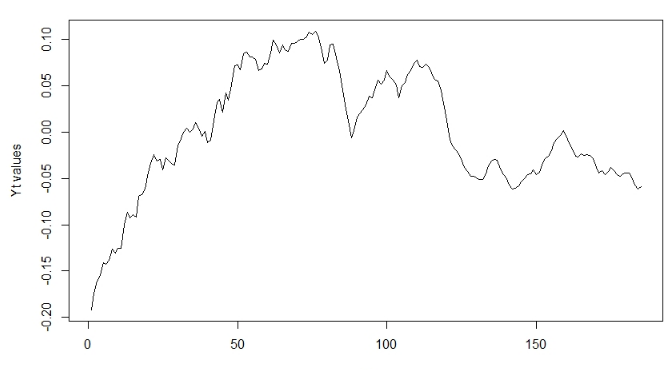
One way to model the decomposed data series X_t, is by imagining that W_t grows deterministically (i.e. $W_t = W_0 e^{ut}$). Ignoring the seasonal effects as observed on figure 2, we obtain $X_t = \ln(Wt) = \alpha + \mu t$. We obtain this model using R by regressing the decomposed data series on a constant and time t. The model we obtain from the decomposed data series using the trend stationary approach (where Y_t is the error term) is:

$$\widehat{X_t} = 12.60655 + 0.00826t + Y_t$$

(t) (0.00971) (0.00009)
n = [185], F-ratio = [8328], *RSS*=[0.7917], R^2 = [0.9785]

Below is a plot of the error terms Y_t:

Figure 3: Plot of Residuals Y_t from the Trend Stationary Model



2.2 Difference Stationary Model

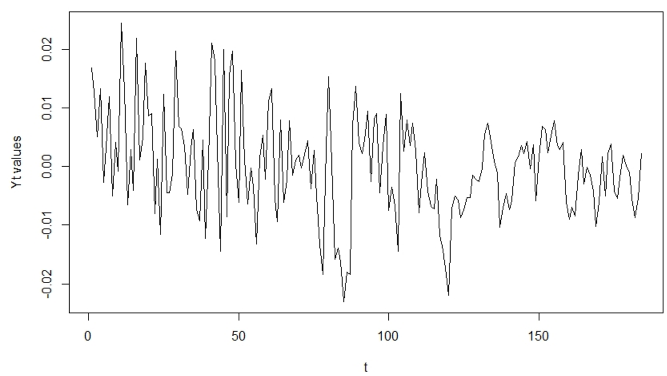
Another way to model the decomposed data series is by using *Difference Stationary Law* — applying Δ to a deterministic trend $T_t = \alpha + \mu t$ and the seasonal trend S_t makes it stationary. Then $\Delta X_t = \Delta Y_t + \Delta T_t + \Delta S_t$ is stationary. We obtain this model by differencing the series $\Delta X_t = X_t - X_{t-1} = \mu + Y_t$ and then regressing ΔX_t on a Constant. The model we obtain from the decomposed data series using the difference stationary approach (where Y_t is the error term) is:

$$\widehat{\Delta X_t} = 0.00899 + Y_t$$

(t) (0.00065)
n = [184], *RSS*=[0.0142]

Below is a plot of the error terms Y_t:

Figure 4: Plot of Residuals Y_t from the Difference Stationary Model



3 Fitting an Autoregressive AR(p) and Forecasting

In this section we answer problems 3 and 4 of the project. We first define a p^{th} order autoregressive model *AR(p)* by:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \alpha t$$

In order to determine the optimal number of lags to model Y_t, the cycle, we use the Bayesian Information Criterion (BIC), which maximizes the trade-off between fit and parsimony given by:

BIC $\propto$ maximized log-likelihood minus the number of parameters

The BIC is similar to the Akaike Information Criterion, except the BIC estimator emphasizes a higher weight on its parameters, resulting in using $\ln(N)$ to replace 2:

$$BIC(k) = \ln(\hat{\sigma}_k^2) + \frac{(\ln(N) * k)}{N}$$

3.1 Trend Stationary Model

We begin by approximating an appropriate p lags for the trend stationary model we obtain in Section 2.2 using the Bayesian Information Criterion, which results in p = 2. Table 1 below shows the results of the BIC(k) estimates using up to 8 lags. the minimized value at p = 2 is bolded.

Autoregressive model AR(2)

To get the AR(2) desired, instead of regressing, we used R's built-in function ARIMA with a 2 level degree of differencing in order approximate an AR(2) model. This allows us to keep the number of observations to 185 instead of 184 from a normal linear regression. Then this allows our AR(p) model to be as follows:

$$\widehat{Y_t} = 1.333Y_{t-1} + -0.338Y_{t-2} + \alpha_t$$

(t) (0.06998) (0.07044)
n = [185], *RSS*=[0.01278]

Given that $\hat{\phi_1}$ = 1.333, $\hat{\phi_2}$ = -0.338, we can use R's built in function of ARMAacf to extract the ρ_1 and ρ_2 autocorrelation values of the model through the formula

$$\rho\hat{(1)} = \frac{\hat{\phi_1}}{1-\hat{\phi_2}} = 0.99673, \quad \rho\hat{(2)} = \frac{\hat{\phi_1}^2}{1-\hat{\phi_2}} + \hat{\phi_2} = 0.99128$$

We can then calculate the other autocorrelation values using the formula:

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2), \ \ where \ \ \rho(0) = 1 \ \ and \ \ \rho(-k) = \rho(k)$$

which is tabulated into table 1

It follows then that:

$$\gamma(0)^{0.5} = \frac{\hat{\sigma^2}}{1-\phi_1\rho(1)-\phi_2\rho(2)} = 0.00831$$

where

$$\hat{\sigma^2} = \frac{RSS}{N} = 6.90826 * 10^{-5}$$

As for the ψ_k values, the moving average weights can be found iteratively using the formula:

$$\psi_k = \phi_1\psi_{k-1} + \phi_2\psi_{k-2}, \ where \ \psi_0 = 1, \ \psi_k = 0 \ for \ k < 0$$

The ψ_l values are then tabulated in table 1.

To find $E_t[\Delta X_{t+k}]$ for a TS model where t is the last observation in the sample, we use the result from Sampson's Time Series Analysis book;

$$E_t[\Delta X_{t+k}] = E_t[Y_{t+k}] - E_t[Y_{t+k-1}] + \mu \qquad and \qquad Var_t[\Delta X_{t+k}] = \sigma^2(1 + \sum_{j=0}^{k-1}(\psi_j - \psi_{j-1})^2)$$

Where $E_t[Y_{t+k}] = \phi_1 E_t[Y_{t+k-1}] + \phi_2 E_t[Y_{t+k-2}]$ and a 95% confidence interval for ΔX_{t+k} is

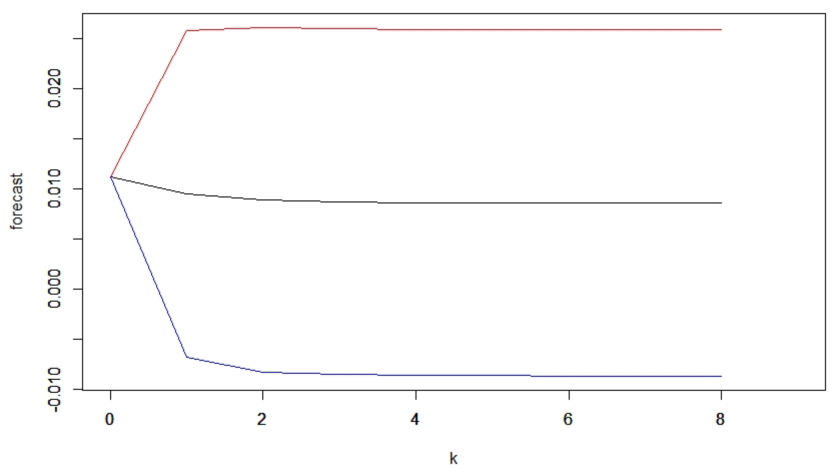
$$E_t[\Delta X_{t+k}] \pm 1.96\sqrt{Var_t[\Delta X_{t+k}]}$$

These results are also tabulated into table 1 below, and a graphical representation of the TS forecasts can be found in figure 5.

Table 1: Trend Stationary BIC(K), $\psi, \phi, \rho, E[Y_{t+k}], E_t[\Delta X_{t+k}], Var[\Delta X_{t+k}]$

Lag k	*BIC(k)*	ψ_k	$\rho(k)$	$E[Y_{t+k}]$	$E[\Delta X_{t+k}]$	$Var[\Delta X_{t+k}]$
0	-5.454	1.000	1.000	-0.059	0.01119	0.0000x10^{-5}
1	-9.435	1.333	0.997	-0.057	0.00951	6.908x10^{-5}
2	**-9.524**	1.440	0.991	-0.057	0.00893	7.675x10^{-5}
3	-9.500	1.470	0.985	-0.056	0.00874	7.7541x10^{-5}
4	-9.488	1.473	0.979	-0.056	0.00867	7.7602x10^{-5}
5	-9.460	1.468	0.972	-0.055	0.00864	7.7603x10^{-5}
6	-9.433	1.460	0.966	- 0.055	0.00863	7.7605x10^{-5}
7	-9.410	1.451	0.959	-0.055	0.00863	7.7609x10^{-5}
8	-9.388	1.441	0.953	-0.054	0.00863	7.7615x10^{-5}

Figure 5: Plot of Trend Stationary forecasts of ΔX_{t+k} and 95% Confidence intervals



3.2 Difference Stationary Model

In order to approximate the forecasts for the DS model, we first need an ARIMA model to forecast Y_t. Using the BIC, we found the optimal p that fits an AR(p) process to be p = 1 (highlighted in table 2 below). However, as per project instructions, we will be using p = 2.

Similar to the Trend stationary model, we fit the AR(2) model using R's built-in ARIMA function instead of fitting a linear regression and preserve the number of observations of 184. Following the fitted function, we approximate :

$$\hat{Y}_t = 0.311Y_{t-1} + 0.0582Y_{t-2} + \alpha_t$$

(t) (0.07385) (0.07421)

n = [184], *RSS*=[0.01263], $\hat{\sigma}^2 = \frac{RSS}{n} = 6.6852x10^{-5}$

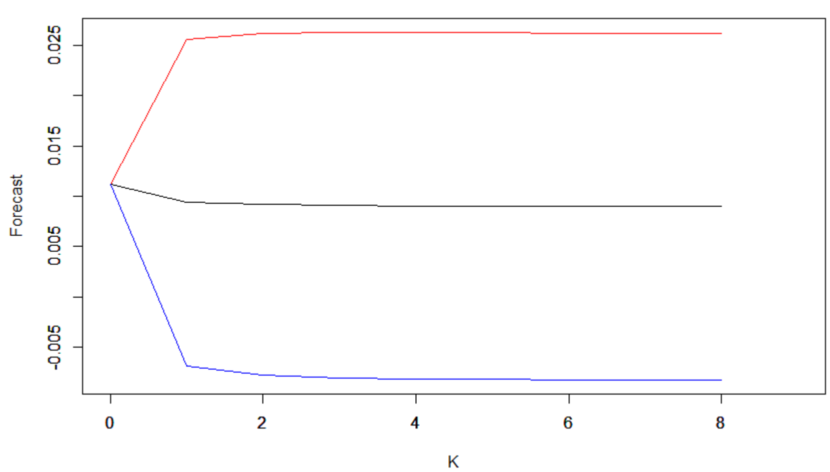
We the same formulas as in the trend stationary models to calculate ψ_k and $E_t[Y_{t+k}]$ as for the different stationary models (tabulated into table 2 below), which will then be used to forecast the growth rate. For the difference stationary model, the growth rate is defined as $\Delta X_{t+k} = \mu + Y_{t+k}$ Then $E_t[\Delta X_{t+k}] = \mu + E_t[Y_{t+k}]$ and $Var_t[\Delta X_{t+k}] = \sigma^2 \sum_{j=0}^{k-1} \psi_j^2$ per the results in Sampson's book.

Table 2: Difference Stationary BIC(K), $\hat{\psi}$, $E[Y_{t+k}], E_t[\Delta X_{t+k}], Var[\Delta X_{t+k}]$

Lag k	*BIC(k)*	$\hat{\psi}_k$	$E[Y_{t+k}]$	$E[\Delta X_{t+k}]$	$Var[\Delta X_{t+k}]$
0	-9.469	1.000	2.2066x10^{-3}	0.01119	0
1	**-9.555**	0.3110	3.7086x10^{-4}	0.00936	6.8652x10^{-5}
2	-9.530	0.1549	0.991x10^{-4}	0.00923	7.5292x10^{-5}
3	-9.517	0.0663	9.7354x10^{-5}	0.00908	7.6939x10^{-5}
4	-9.489	0.0296	4.4450x10^{-5}	0.00903	7.7241x10^{-5}
5	-9.462	0.0131	1.9486x10^{-5}	0.00901	7.7301x10^{-5}
6	-9.438	0.0058	8.6455x10^{-6}	0.00899	7.7312x10^{-5}
7	-9.415	0.0026	3.8221x10^{-6}	0.00899	7.7315x10^{-5}
8	-9.396	0.0011	1.6915x10^{-6}	0.00899	7.7315x10^{-5}

These forecasted values are relatively similar to the forecasted values for the TS model.

Figure 6: Plot of Difference Stationary forecasts of ΔX_{t+k} and 95% Confidence intervals



4 Augmented Dickey-Fuller test

Since we have been working with the two models: TS and DS, we want to know which is the more optimal model. In order to assess the the comparative appropriateness of the models, we perform an Augmented Dickey-Fuller Test on X_t

We define π = $-\phi(1)$ = $-(1 - \phi_1 - \phi_2 - ... - \phi_p)$. We want to test if X_t has a unit root, or in other words, if π = 0. Then our test becomes:

$$H_0\text{: } \pi = 0 \qquad\qquad H_A\text{: } -2 < \pi < 0$$

In other words, our null hypothesis tests if X_t is difference stationary, the alternative hypothesis is that X_t is trend stationary. The test statistic: $ADF_{stat} = \frac{\hat{\pi}}{se(\hat{\pi})}$ follows a Dickey Fuller distribution under H_0. Per the instruction of the project, we perform the test using 5 lags. Using the "tseries" package in R and performing the adf.test on X_t with a lag k=5, we obtain ADF_{stat} = -2.6929, and the p-value is 0.2865. As such, we do not reject the null hypothesis and conclude that a X_t is difference stationary.

5 Box-Jenkins Identification and Diagnostics Tests

The Box-Jenkins Identification analyzes the behavior of $\rho(k)$ and ϕ_{kk} to determine the most accurate model. They key to interpreting if the residuals Y_t follows an AR(p) or an MA(q) is through table 3 below.

Table 3: Box Jenkins Identification Guidelines

Model	$\rho(k)$	ϕ_{kk}
AR(p)	Damped Exponential	Cut-off at k=p
MA(q)	Cut-off at k=q	Damped Exponential
ARMA(p,q)	Damped Exponential	Damped Exponential

As per R's built-in partial autocorrelation functions (ϕ_{kk}) and autocorrelation ($\rho(k)$)functions "pacf" and "acf", we calculate $\rho(k)$ and ϕ_{kk} using the formulas below:

$$\rho(k) = Corr[Y_t, Y_{t-k}] = \frac{Cov[Y_t, Y_{t-k}]}{Var[Y_t]^{0.5} Var[Y_{t-k}]^{0.5}} = \frac{\gamma(k)}{\gamma(0)}$$

The partial autocorrelation functions are calculated using the Yule-Walker equation:

$$\phi_k = R_k^{-1} r_k$$

where ϕ_k is a vector with the k elements $\phi_{k1}, \phi_{k2}, ..., \phi_{kk}$,
r_k is a vector with k elements $\gamma(1), \gamma(2), ...\gamma(k)$ and

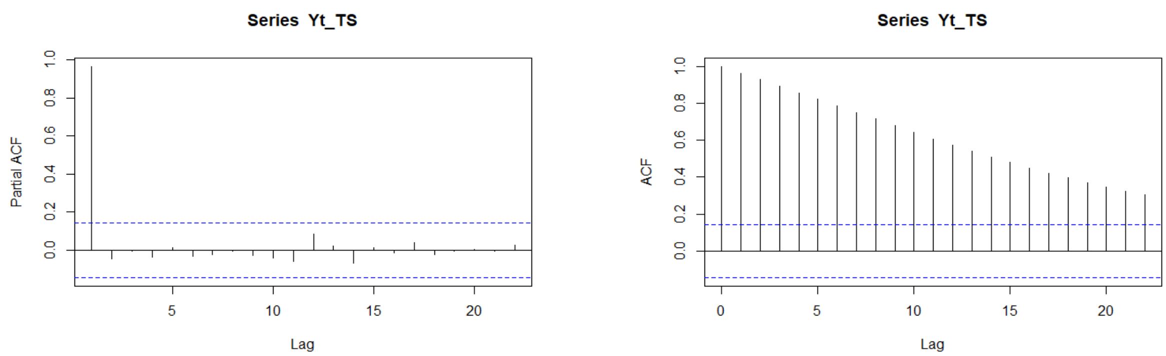
R_k is the kxk matrix

$$\begin{pmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{pmatrix}$$

which gives us the optimal forecast weights given all available information. Note that ϕ_{kk} or $\rho(k)$ is significant if their respective absolute values are greater than $\frac{2}{\sqrt{T}} = \frac{2}{185} = 0.1470$

5.1 Trend Stationary Model

Figure 7: TS Box Jenkins identification using PACF and ACF plots



(a) Plot of Partial Autocorrelation Functions

(b) Plot of Autocorrelation Functions

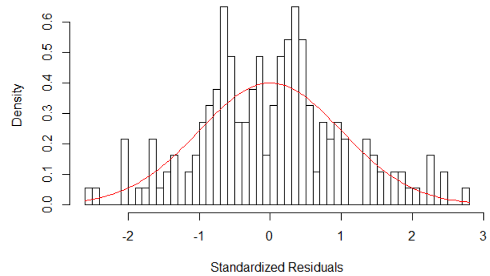
From the graphs, we can easily see that the residuals Y_t follow an AR(1) according to the Box Jenkins identification table, table 3e. Then using R's built-in "arima" function, we get:

$$\hat{Y}_t = 0.9974 Y_{t-1}$$

(t) (0.035)
n = [185], *RSS*=[0.0143] σ^2 = [7.769e-05]

Then plotting a histogram for the standardized residuals $z_t = \frac{\epsilon_t - \hat{\epsilon_t}}{\sigma}$, we have the graphical representation:

Figure 8: Plot of Standardized Residuals



Testing for Normality: For a correct model, the error terms should follow a Normal Distribution. Informally, we can observe the standardized residual plot, which is plotted against a standard Normal curve in red. From inspection of the graph, the standardized residuals seems to follow the Normal fairly well.

Formally, we can use the *Jaques-Bera* test which calculates the skewness k_k3 an kurtosis k_4 which can be measured as:

$$k_3 = E[z^3] = \frac{1}{T}\sum_{t=1}^{T}\hat{z^3} \quad k_4 = \frac{1}{T}\sum_{t=1}^{T}\hat{z^4}$$

The data is normal if k_3 = 0 and k_4 = 3 so we perform our Jaques Bera test with the hypothesis

H_0: k_3 = 0 and k_4 = 3 vs. H_A: $k_3 \neq 0$ and $k_4 \neq 3$

Our test statistic is JB = $\hat{t_3^2} + \hat{t_4^2} = N(\frac{\hat{k_3^2}}{6} + \frac{(k_4-3)^2}{24})$ ~χ_2^2
where $\hat{t_3^2}$ ~N[0,1] and $\hat{t_4^2}$ ~N[0,1]

The calculations lead to k_3 = 0.1772, k_4 = 3.1673, JB = 1.1837 which is less than the critical value of 5.9915 for a χ_2^2 at a 5% level of significance. So we do not reject the null hypothesis.

Box-Pierce Portamanteau Test: This is a residual-based test to assess the residual a_t's independence by testing for the correlation between the error terms. Although zero correlation does not mean independence, independence does imply that there is zero correlation. So with M = $\sqrt{N}$ = 14, we can perform the Box-Pierce test:

H_0: $\rho_a(k)$ = $\frac{E[a_t a_{t+k}]}{\sigma^2}$ = 0 for k = 1,2,...,M vs. H_A: $\rho_a(k) \neq 0$ where $\rho_a\hat{(}k)$ = $\frac{\sum_{t=1}^{N-|k|} a_t a_{t+k}}{\sum_{t=1}^{N} a_t^2}$

The test statistic is $Q = N(\hat{\rho_a}(1)^2 + \hat{\rho_a}(2)^2 + ... + \hat{\rho_a}(M)^2)$ ~χ_M^2 Using R's built-in "Box.test" with a lag of 14, the p-value is calculated to be 0.0001092 so we reject the null hypothesis of independence between the error terms for the TS AR(1) model.

Testing for Overfitting (r = 4): Given that Box Jenkins has resulted in a selected AR(p) = AR(1):
Y_t = $\phi_1 Y_{t-1}$, a test for overfitting tests if Y_t is also an AR(p+r). This is done using the log likelihood

ratio test. In the case of this paper, given r = 4, we test for an AR(5) $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \phi_5 Y_{t-5}$ under the null hypothesis that H_0: $\phi_2 = \phi_3 = \phi_4 = \phi_5$ = 0 under the test statistic

$$\Lambda = N * \ln \frac{\hat{\sigma}_p^2}{\hat{\sigma}_{p+r}^2} \sim\chi_r^2$$

The fitted AR(5) model using R is $Y_t = 1.3041Y_{t-1} - 0.2833Y_{t-2} + 0.1062Y_{t-3} - 0.1345Y_{t-4} + 0.0016Y_{t-5}$ with coefficient standard errors of (0.0738), (0.1211), (0.1224), (0.1209), and (0.0747) respectively. $\hat{\sigma^2_{p+r}}$ = 6.764e-05. Calculating the test statistic, we have Λ = 25.6283 which is greater than the critical value for X_4^2 = 9.4877, so we reject the null hypothesis and say that there is evidence of overfitting. This may explain the fact that the selected AR(p) model using BIC was p=2 which is greater than the p=1 model selected from the Box Jenkins identification.

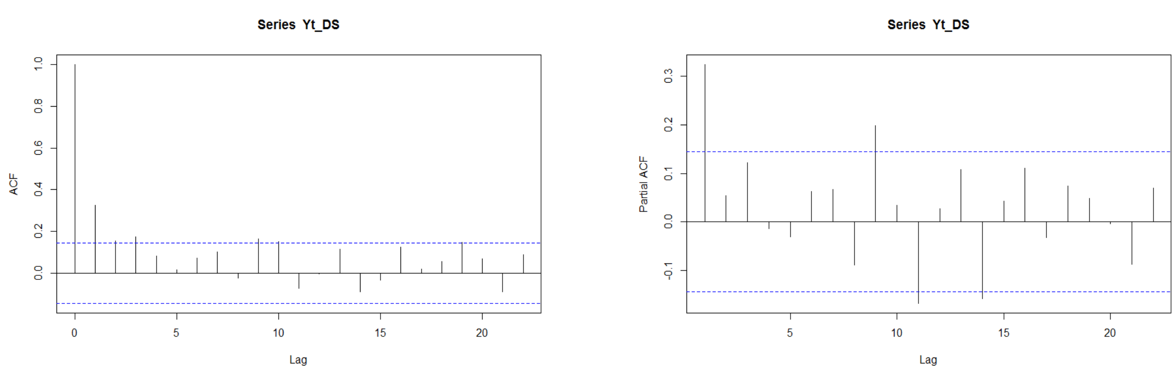
Testing for Non-Linear Dependence: Although we figured out that the error terms are not independent of each other from the Box Pierce test, we can solidify our findings using ARCH(q), autogressive conditional heteroskedasticity model defined as:

$$a_t = z_t * (\sigma^2 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + ... + \alpha_q a_{t-q}^2)^{1/2}$$

$where\ z_t \sim i.i.N[0, 1]$
Under Engle(1982)'s test on H_0, the test statistic is N x R^2 $\sim\chi_q^2$ Since our objective is to test for an ARCHC(6), we test the null hypothesis $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_6 = 0$ in $a_t = z_t * (\sigma^2 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + ... + \alpha_6 a_{t-6}^2)^{1/2}$ where $z_t \sim i.i.N[0, 1]$. We regress $\hat{a}_t^2$ on a constant and $\hat{a}_{t-1}^2, \hat{a}_{t-2}^2, \hat{a}_{t-3}^2, \hat{a}_{t-4}^2, \hat{a}_{t-5}^2, \hat{a}_{t-6}^2$ as
$\hat{a}_t^2 = \beta_0 \hat{a}_{t-1}^2 + \beta_1 \hat{a}_{t-2}^2 + \beta_3 \hat{a}_{t-3}^2 + \beta_4 \hat{a}_{t-4}^2 + \beta_5 \hat{a}_{t-5}^2 + \beta_6 \hat{a}_{t-6}^2 + error$ with R^2 = 0.3540 Since N = 185, our test statistic is then $N * R^2$ = 65.4832 which exceeds the critical value $\chi_6^2 = 12.5916$, meaning we reject the null hypothesis. Then our diagnostic test suggests that this model has failed.

5.2 Difference Stationary Models

Figure 9: DS Box Jenkins identification using PACF and ACF plots



(a) Plot of Partial Autocorrelation Functions (b) Plot of Autocorrelation Functions

Box Jenkins Identification: We can see from the plots that there are no obvious damped exponential autocorrelation graphs, meaning it's difficult to apply the Box Jenkins Identification.

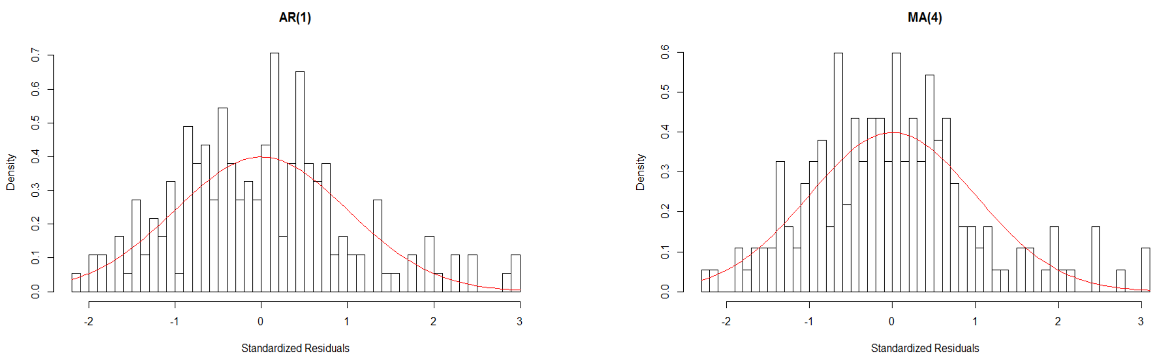
Instead, we test for two models: MA(4) and AR(1) since these are the cutoffs shown in the graphs above at the significance factor of $\frac{2}{\sqrt{184}} = 0.1474$

5.2.1 AR(1) and MA(4) models

Using R's arima function, we fit the AR(1) model to be $\hat{Y}_t = 0.3295Y_{t-1}$ with a coefficient standard error of 0.0701, n = 185 and *RSS* = 0.0127, $\hat{\sigma}^2_R$ = 6.888437*10^{-05}

Similarly, we fit the MA(4) model to be $\hat{Y}_t = a_t + 0.3054a_{t-1} + 0.1223a_{t-2} + 0.1586a_{t-3} + 0.0822a_{t-4}$, coefficient standard errors of 0.0735, 0.0774, 0.0776, 0.0869, n = 184, *RSS* = 0.01244387, $\hat{\sigma}^2_R$ = 6.762972*10^{-05}

Figure 10: Plot of Standardized Residuals for AR(1) and MA(4)



(a) Plot of AR(1) standardized residuals

(b) Plot of MA(4) standardized residuals

Normality Test: Similar to the TS model, we can informally predict that the standardized residuals do not follow a standard normal distribution. Formally, for the AR(1) model, we look at $\hat{k}_3 = 0.5223$, $\hat{k}_4 = 3.3510$, *JB* = 9.3099 which is greater than the critical value of 5.9915, so we reject the null hypothesis that the standardized residuals of the DS AR(1) model follows a standard normal.

Similarly for the MA(4) model, we calculate $\hat{k}_3 = 0.5507$, $\hat{k}_4 = 3.4952$, *JB* = 11.1800 which is again higher than the critical value of 5.9915, so we also reject the null hypothesis that the standardized residuals of the DS MA(4) model follows a standard normal.

Box-Pierce Portamanteau Test: We calculate for the AR(1) that the test statistic Q is 31.255 and p-value is 0.005105 so we reject the null hypothesis that the error terms are not correlated. The autocorrelation functions are non-zero for k $\neq$ 0.

However, for the MA(4) model, it reaches a different result. The test statistic, Q, is calculated to be 20.343 with a degree of freedom of 14. The respective p-value is then 0.1197 so we do not reject the null hypothesis and we can say that there is evidence that the error terms are not correlated.

Testing for Overfitting (r = 4): For the DS AR(1) model, we overfit by approximating the DS AR(5) model using arima, which gives us $Y_t = 0.3044Y_{t-1} + 0.0255Y_{t-2} + 0.1277Y_{t-3} +$

$0.0019Y_{t-4} - 0.0339Y_{t-5}$ with coefficient standard errors of {0.0739, 0.0773, 0.0767, 0.0776, 0.0743}, n = 184, *RSS* = 0.0124197, and $\hat{\sigma}_U^2$ = 6.7498*10^{-05}. Performing the likelihood ratio test, we have $\Lambda = 3.7605$ which is less than the critical value of 9.4877 (following a chi-squared distribution with a degree of freedom of 4. Then we do not reject the null hypothesis, which is evidence that this model is incorrect.

For the DS MA(4) model, we overfit by approximating the DS MA(8) model using arima, which gives us $\hat{Y}_t = a_t + 0.3317a_{t-1} + 0.1467a_{t-2} + 0.1272a_{t-3} + 0.1072a_{t-4} + 0.0180a_{t-5} - 0.0315a_{t-6} + 0.0855a_{t-7} - 0.0776a_{t-8}$ with coefficient standard errors of {0.0728, 0.0866, 0.0868, 0.0678, 0.0896, 0.0997, 0.0884} respectively, n = 184, *RSS* = 0.01213877, $\hat{\sigma}_U^2$ = 6.597158*10^{-05}. Performing the likelihood ratio test, we have $\Lambda = 4.59233$ which is less than the critical value of 9.4877 following a chi-squared distribution with a degree of freedom of 4. Then we do not reject the null hypothesis, which is evidence that this model is also incorrect.

Testing for Non-Linear Dependence: Just like the TS approach, we test using Engle(1982)'s test of H_0 using an ARCH(6). So we then regress the DS AR(1) $\hat{a}_t^2$ on a constant and $\hat{a}_{t-1}^2, \hat{a}_{t-2}^2, \hat{a}_{t-3}^2, \hat{a}_{t-4}^2, \hat{a}_{t-5}^2, \hat{a}_{t-6}^2$ as $\hat{a}_t^2 = \beta_0\hat{a}_{t-1}^2 + \beta_1\hat{a}_{t-2}^2 + \beta_3\hat{a}_{t-3}^2 + \beta_4\hat{a}_{t-4}^2 + \beta_5\hat{a}_{t-5}^2 + \beta_6\hat{a}_{t-6}^2 + error$ with R^2 = 0.3679 Since N = 185, our test statistic is then $N * R^2$ = 67.6849 which exceeds the critical value $\chi_6^2 = 12.5916$, meaning we reject the null hypothesis. Then our diagnostic test suggests that this model has failed.

We regress the MA(4) model in the same way, and we get R^2 = 0.3605. The test statistic is 66.33104 which is greater than the same critical value as above, so we reject the null hypothesis. Our diagnostic test then also suggests this model has failed.

6 S&P Financial Time Series

In this section, we aim to analyze the monthly data for the S&P500 stock prices, P_t from January 1939 to December 1992, totaling 648 observations. In order to determine if the series follows a random walk, we check if

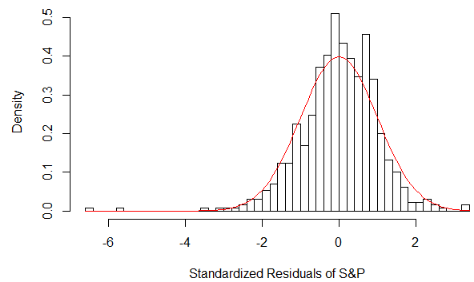
$$\ln(P_t) = \delta + \ln(P_{t-1}) + a_t \text{ with } a_t \sim i.i.N[0, \sigma^2]$$

Using linear regression, we find the model to be $\ln(P_t) = 0.005455 + 1.000014\ln(P_{t-1})$. The standard errors of the coefficients are {0.000014, 0.001638} with t-statics of {0.0784, 610.662} respectively. Additionally, n = 647, F-ratio = 3.729*10^5, R^2 = 0.9983, and *RSS*=1.2276.

Testing Normality: Going straight to the JB test, we calculate the standardized residuals $z_t = \frac{a_t - \hat{a}_t}{\hat{\sigma}^2}$, $\hat{k}_3$, and $\hat{k}_4$ and find the test statistic *JB* = $\hat{t}_3^2 + \hat{t}_4^2 = N(\frac{\hat{k_3^2}}{6} + \frac{(k_4-3)^2}{24})$ = 546.6699 which is higher than the critical value of 5.99 so we reject the null hypothesis of normalty. However, it is worth noticing that in figure 11, there are several severe outliers in the data that may cause the standardized residuals to not be normal. Although we can remove those points from our data, it creates a problem by challenging the assumption of stationarity.

We examine the autocorrelation functions of a_t in table 4 where we suspect something is wrong with the model if $|\hat{\rho}_a(k)| > \frac{2}{\sqrt{N}} = \frac{2}{\sqrt{647}} = 0.07863$ where $\hat{\rho}_a(k) = \frac{\sum_{t=1}^{N-|k|} a_t a_{t+k}}{\sum_{t=1}^{N} a_t^2}$. Notice at k = 5,

Figure 11: Plot of S&P Standardized Residuals



we see that $|\hat{\rho}_a(5)| > \frac{2}{\sqrt{N}} = 0.07863$ so we reject the null that a_t is white noise. Then we can say that the series does not follow a random walk.

Table 4: Autocorrelation of $\hat{a}_t$ for S&P Data

k =	1	2	3	4	5	6	7
$\hat{\rho}_a(k)$ =	-0.0116	-0.0135	0.0042	0.0208	**0.0825**	-0.0560	-0.0097

As for the autocorrelation functions of a_t^2, we estimate it using a similar formula as for a_t: $\frac{\sum_{t=1}^{N-|k|} a_t^2 a_{t+k}^2}{\sum_{t=1}^{N} a_t^2}$. The values are tabulated into table 5 below.

Table 5: Autocorrelation of $\hat{a}_t^2$ for S&P Data

k =	1	2	3	4	5	6	7
$\hat{\rho}_{a^2}(k)$ =	0.075	0.023	0.014	0.015	-0.009	0.046	-0.046

GARCH(1,1) Model Many financial series are subject to changes in volatility, and therefore there are many models that try to capture this behavior. Bollerslev (1986)'s generalized ARCH model (GARCH) added more flexible lag structure to Engel (1982)'s ARCH by expressing σ_t^2 as a linear function of its lags and past squared innovations. Then GARCH(p,q) is given by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

where σ_t^2 is the variance of a_t conditional on the information at time t - p (F_t), $\frac{\sqrt{\sigma_t}}{a_t}|F_t$ ~N[0,1], where $\sigma_t^2 = E[a_t^2|a_{t-1}, ...a_{t-p}]$, the $\omega > 0$, $\alpha_i > 0$ and $\beta_i > 0$. Using R's "garchFit" function from the "fGarch" package and following project instructions, we fit a_t on GARCH(1,1) which returned:

$$\hat{\sigma}_t^2 = 9.487 * 10^{-5} + 0.0607 a_t^2 + 0.8884 \sigma_{t-1}^2$$

(t) $(3.926 * 10^{-5})(2.026 * 10^{-2})\ (2.867 * 10^{-2})$

The p-values on a t-test of the coefficients are 0.0156, 0.0027, and 0.000 so all coefficients shows evidence of significance.

7 eGarch

In this section we attempt to model $\ln(P_t)$ using a model that displays non-linearity or conditional heteroskedascity that hasn't been previously mentioned. The modeling will be done in R using the package "rugarch".

The rugarch is a general R package for univariate financial time series analysis. Each model of the package consists of the mean equation (ARFIMA models with ARCH-M characteristics and exogenous variables) and volatility equation (CHICAGO, 2015).

The selected model is Nelson (1991)'s exponential GARCH model written as in "fugarch" as:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^{m} \zeta_i v_{it} + \sum_{k=1}^{q} g(\epsilon_{t-k}) + \sum_{j=1}^{p} \beta_j \ln(\sigma_{t-j}^2)$$

$$\text{with } \; g(\epsilon_{t-k}) = \alpha_k \epsilon_{t-k} + \gamma_k(|\epsilon_{t-k} - E|\epsilon_{t-k}|), \; \sigma_t \epsilon_t = a_t$$

where the logarithm of the conditional variance is a linear function of a constant, its past values, and a function of the lagged ϵ_t's, av_{it} denotes the exogeneous variables, and α_i identifies the leverage-effect.

The idea of the exponential GARCH is based on one limitation of GARCH model; the GARCH model accounts for the magnitude of the past innovations, but not its relative direction (positive or negative). The basis behind the eGarch model then is to account for the sign and size effect, and the related asymmetric response of the conditional variance. The reason that this model was selected to model $\ln(P_t)$ is because we want to account for the leverage effects on the S&P, which like any stock, is assumed to react to positive and negative news.

The model: For the sake of simplicity, the mean model ARIMA order will be (0,0), so the mean-equation is a constant. Using several GARCH candidate models from (2,2), (1,2), (2,1), and (1,1) for the eGARCH variance model, modeled onto $\ln(P_t)$, we lead to the following results:

eGarch(1,1):

$$\ln(\sigma_t^2) = -0.3071 - 0.0070\epsilon_{t-1} + 1.3573(|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)) + 0.9405\ln(\sigma_{t-1}^2)$$

(t) (0.1037) (0.0244) (1.3573) (0.0193)
n = [648], *RSS* = 825.0856, Pr(>|t|) = {0.0000, 0.8775, 0.0000, 0.0000}

eGarch(1,2):

$$\ln(\sigma_t^2) = -0.3492 - 0.0039\epsilon_{t-1} + 1.5647(|\epsilon_{t-1}| - E(|\epsilon_{t-1}|)) + 0.8221\ln(\sigma_{t-1}^2) + 0.1153\ln(\sigma_{t-2}^2)$$

(t) (0.4582) (0.0523) (1.8156) (2.1107) (2.1919)
n = [648] *RSS* = 825.086, Pr(>|t|) = {0.4461, 0.9404, 0.3888, 0.6969, 0.9580}

eGarch(2,1):

$$\ln(\sigma_t^2) = -0.2150 - 0.3243\epsilon_{t-1} + 1.5417(|\epsilon_{t-1}| - E(|\epsilon_{t-1}|))$$
$$+ 0.32543\epsilon_{t-2} - 0.5466(|\epsilon_{t-12}| - E(|\epsilon_{t-2}|)) + 0.96187\ln(\sigma_{t-1}^2)$$

(t) (0.1325) (0.1853) (0.7802) (0.1816) (0.2417)
(0.0178)

n = [648] *RSS* = 833.774, Pr(>|t|) = {0.1045 0.0801, 0.0481, 0.0731, 0.0237, 0.0000}

Similarly, we test for eGarch(2,2), but out of constraint for space, we summarize it that *RSS* = 850.94 for the model. Since eGarch(1,1) had the lowest mean residual squared, it will be the model that we will use for our data.

Diagnostics Test: Just because eGarch(1,1) seemed to be the best model out of the candidate pool does not mean that it is a good model. Using the same diagnostics tests we've established in section 5, we perform the following tests:

Jaques-Bera: We calculate k_3 = -0.2006, k_4 = 2.1189, and find the test statistic JB = 25.3044 which is greater than the critical value of 5.99 so we reject the null hypothesis of normality of the error terms of the model.

Box-Pierce: We calculate the box pierce test statistic using R's "Box.test" on a lag of $\sqrt(648) = 26$ which resulted in a test statistic of 14477 and a p-value of 2.2*10^{-16} so we reject the null hypothesis that the error terms are not correlated.

ARCH(6) test: We calculate the R^2 values on a error terms of the eGarch(1,1) model with a lag of 6, which leads to a value of 0.3540. The test statistic is then 648 * 0.3540 = 229.3683 which is higher than the critical value of 12, so we reject the null hypothesis which suggests that this model has failed.

Although the diagnostics test we used for the eGarch(1,1) has all failed, and we believe this to be an inappropriate model for $\ln(P_t)$, it could also be that the diagnostics tests used for the eGarch model is inappropriate. The R output suggests performing different diagnostic tests such as: Weighted Ljung-Box Test on Standardized Residuals, Weighted ARCH LM Tests, Nyblom stability test, Sign Bias Test, and Adjusted Pearson Goodness-of-Fit Test that was not covered in this paper. However, from the standpoint of what is covered in this paper, it seems that the eGarch(1,1) is not an appropriate model.