ECON144-Lab2

Andrew Grove and Thomas Santosa

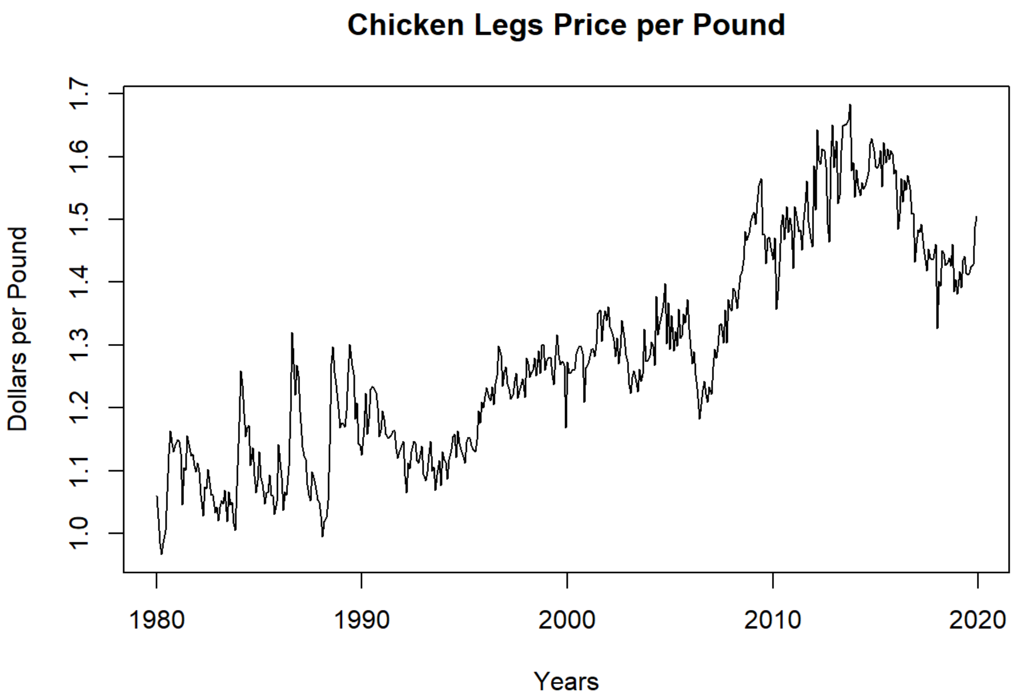
304785991 and 504797813

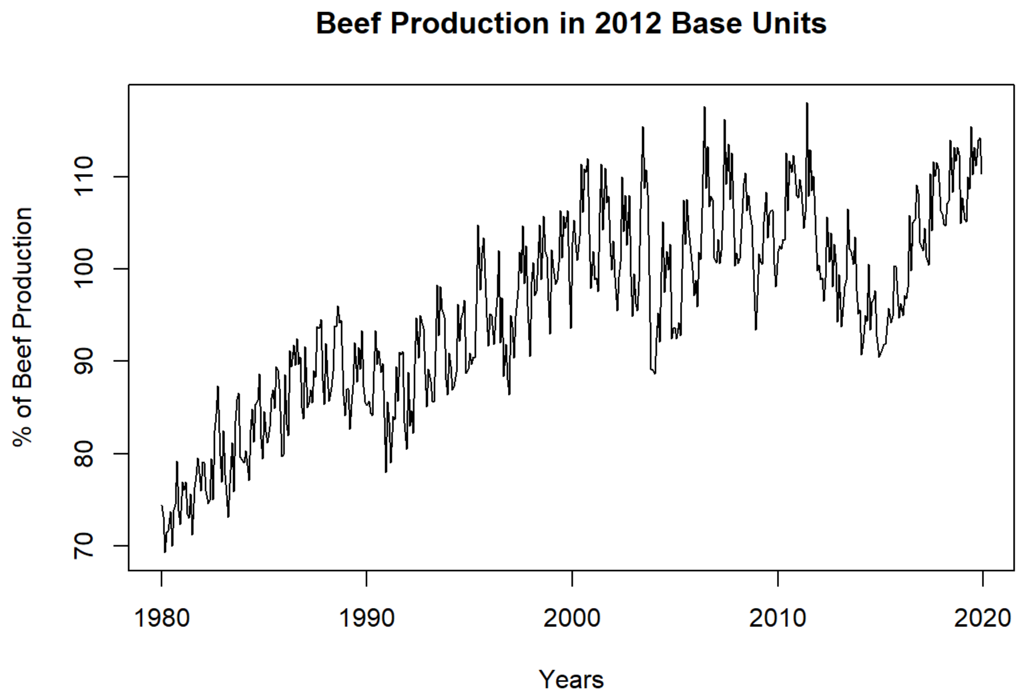
Part I

We chose two datasets from Economagic that we felt had a trend, seasonality, and cyclc component. The first is a monthly US average of the price of chicken legs per pound, measured in USD. The second is a monthly measurement of the amount of Beef Production in the United States, measured in percentages with the base percentage being the average month in 2012. We can see that in general, both have been increasing over time with occassional dips. One thing that pulled our attention is the chicken price is inversely associated with beef production, causing us to think that there might be a possible relation that VaR could find. This relation would make sense, since if beef production drops, demand for chicken should increase, which could increase the price.

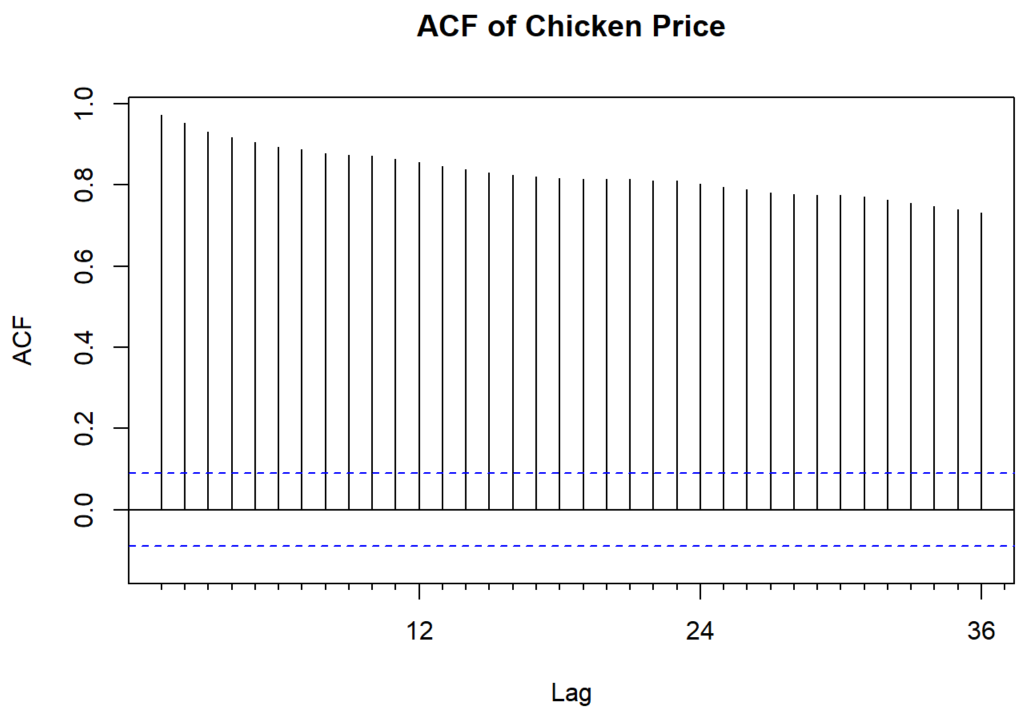
Part II

a.

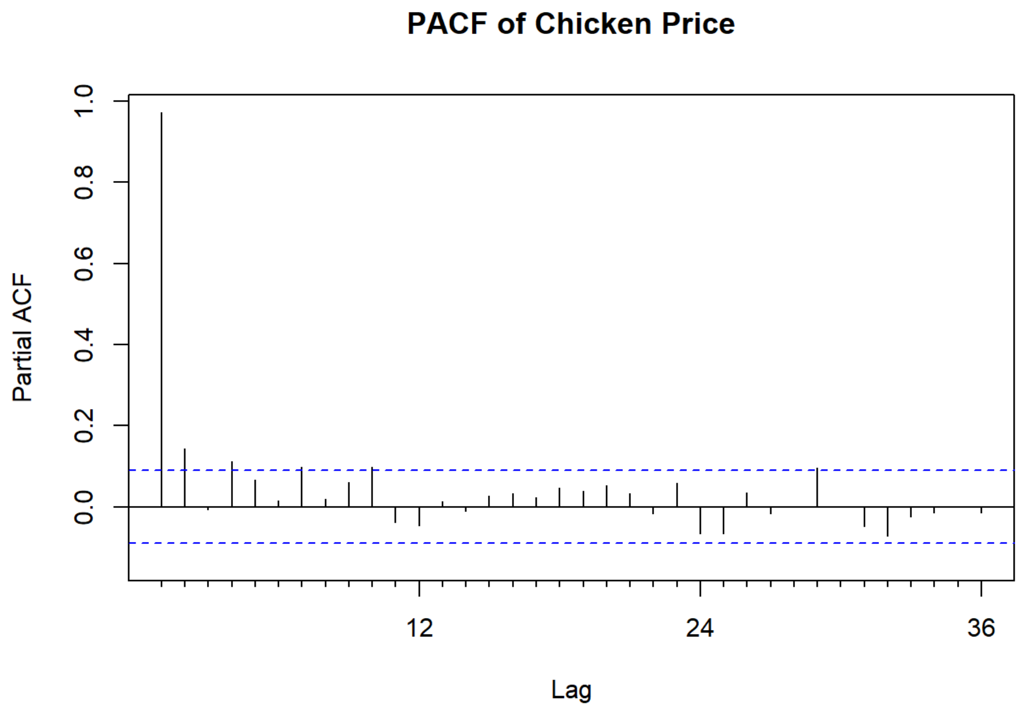




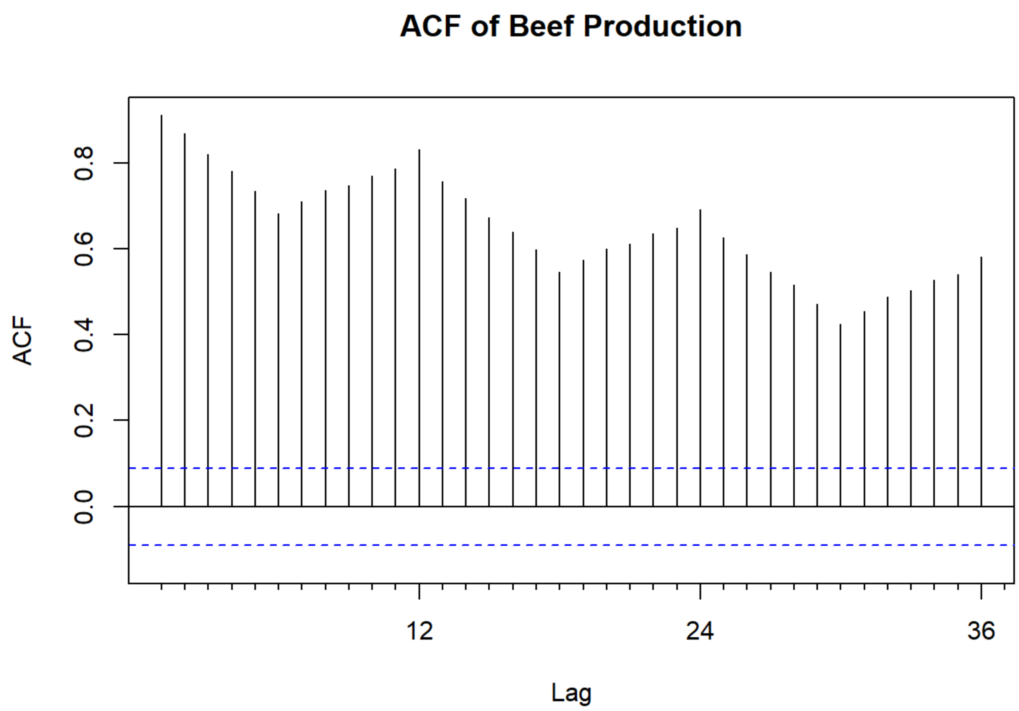
Beef Production in 2012 Base Units
% of Beef Production
70
80
90
100
110
1980
1990
2000
2010
2020
Years



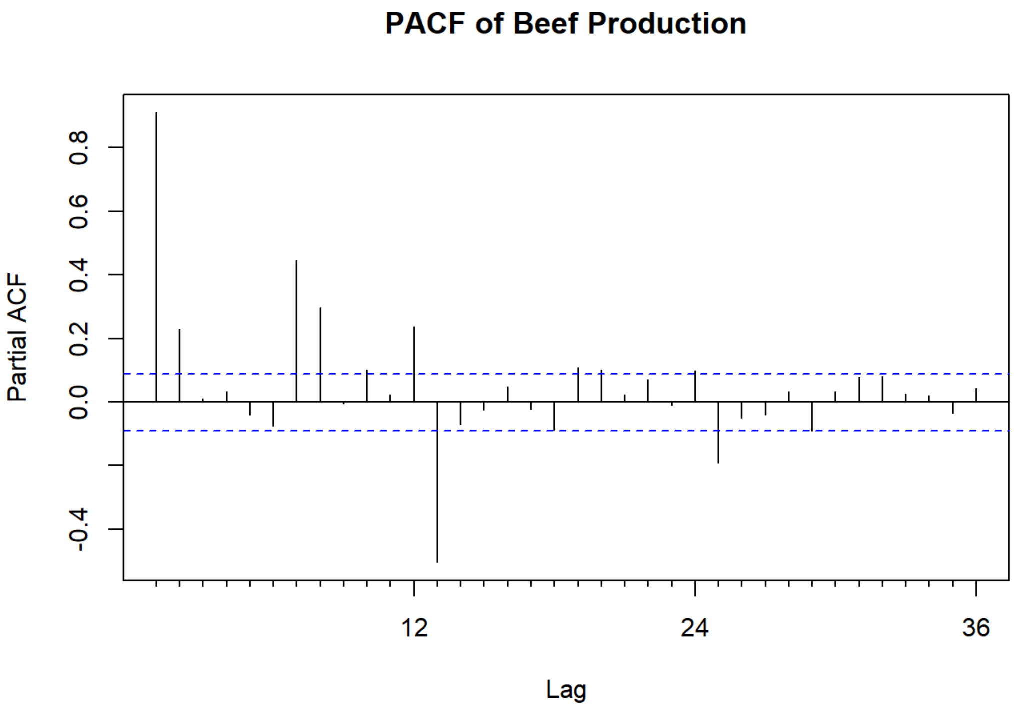
ACF of Chicken Price
ACF
0.0
0.2
0.4
0.6
0.8
1.0
12
24
36
Lag



PACF of Chicken Price
Partial ACF
1.0
0.8
0.6
0.4
0.2
0.0
12
24
36
Lag



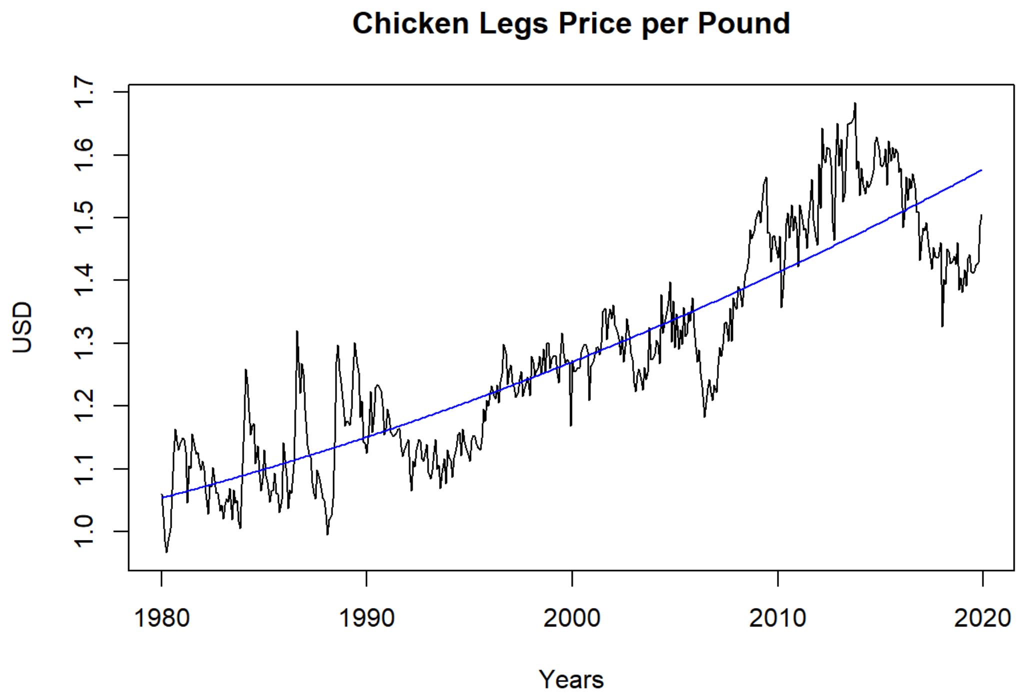
ACF of Beef Production
ACF
0.8
0.6
0.4
0.2
0.0
12
24
36
Lag

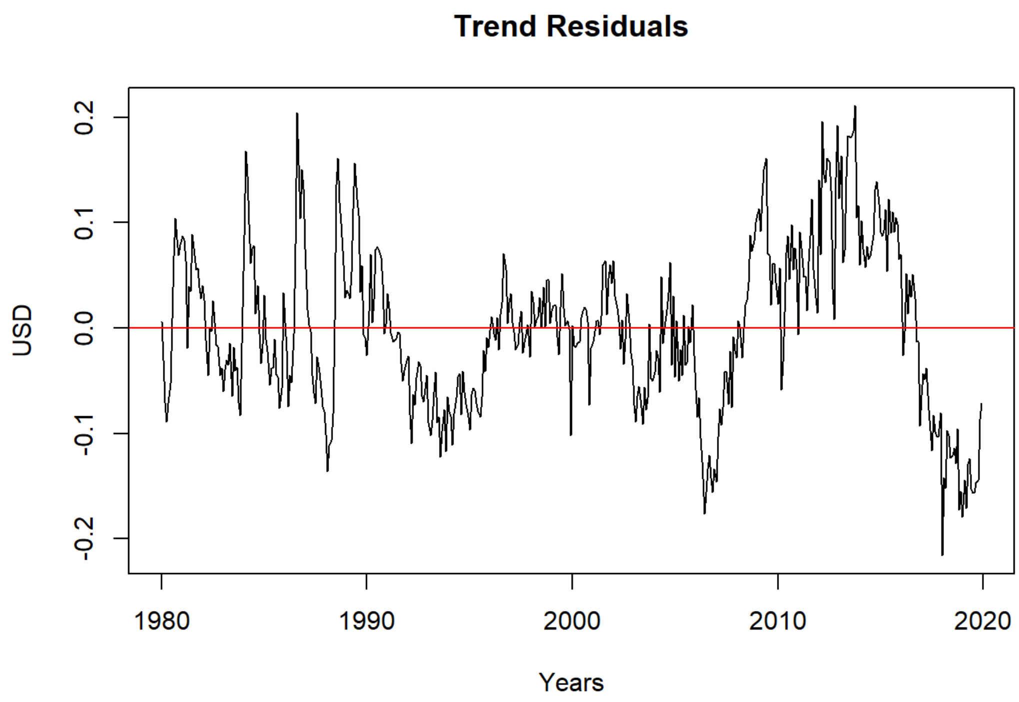


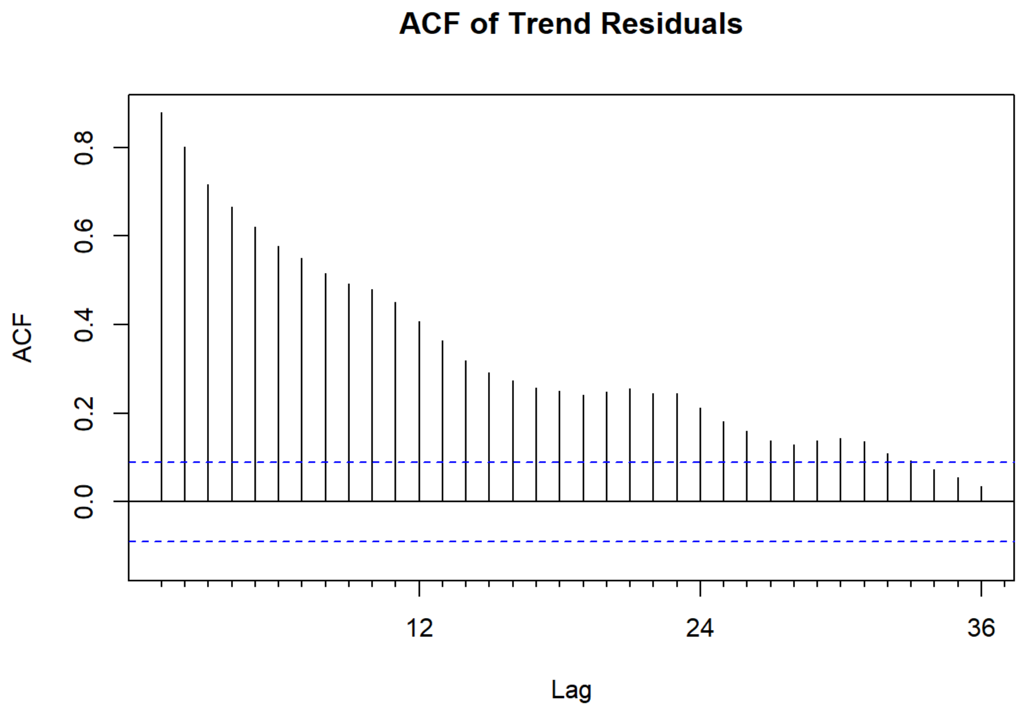
b.

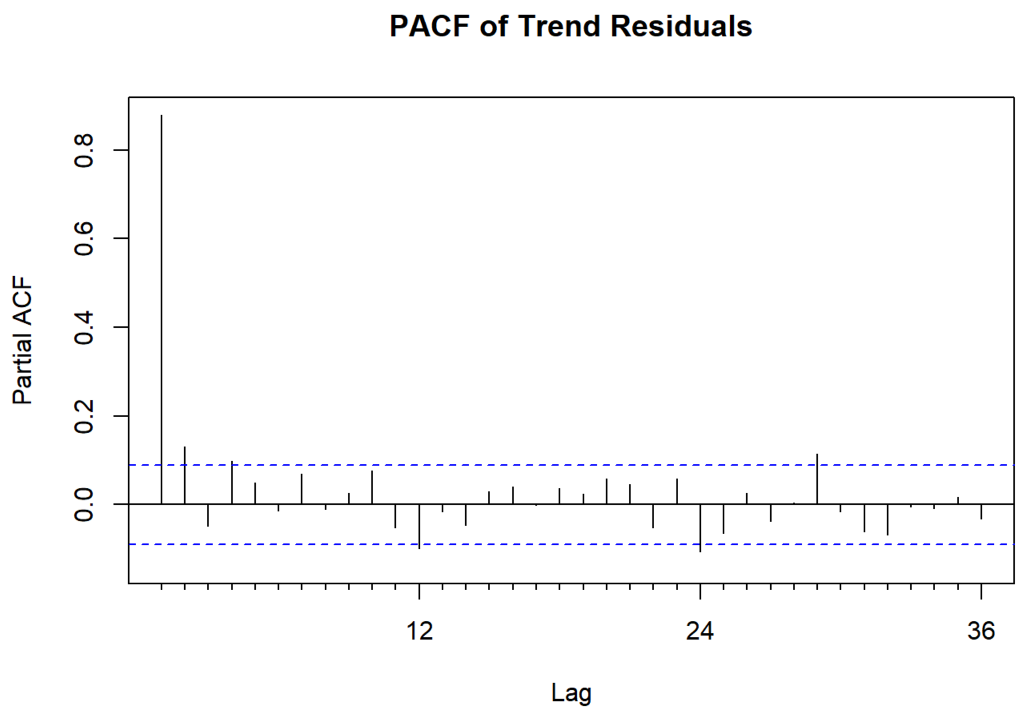
Model for Chicken

```
##
## Call:
## tslm(formula = chicken ~ t + t2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.215934 -0.055663 -0.003002  0.055437  0.210892
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.250e+02  1.185e+02   3.585 0.000371 ***
## t           -4.368e-01  1.185e-01  -3.685 0.000255 ***
## t2           1.125e-04  2.964e-05   3.795 0.000166 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07744 on 477 degrees of freedom
## Multiple R-squared:  0.7943, Adjusted R-squared:  0.7934
## F-statistic: 920.8 on 2 and 477 DF,  p-value: < 2.2e-16
```


Chicken Legs Price per Pound
USD
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1980
1990
2000
2010
2020
Years


Trend Residuals
USD
-0.2
-0.1
0.0
0.1
0.2
1980
1990
2000
2010
2020
Years

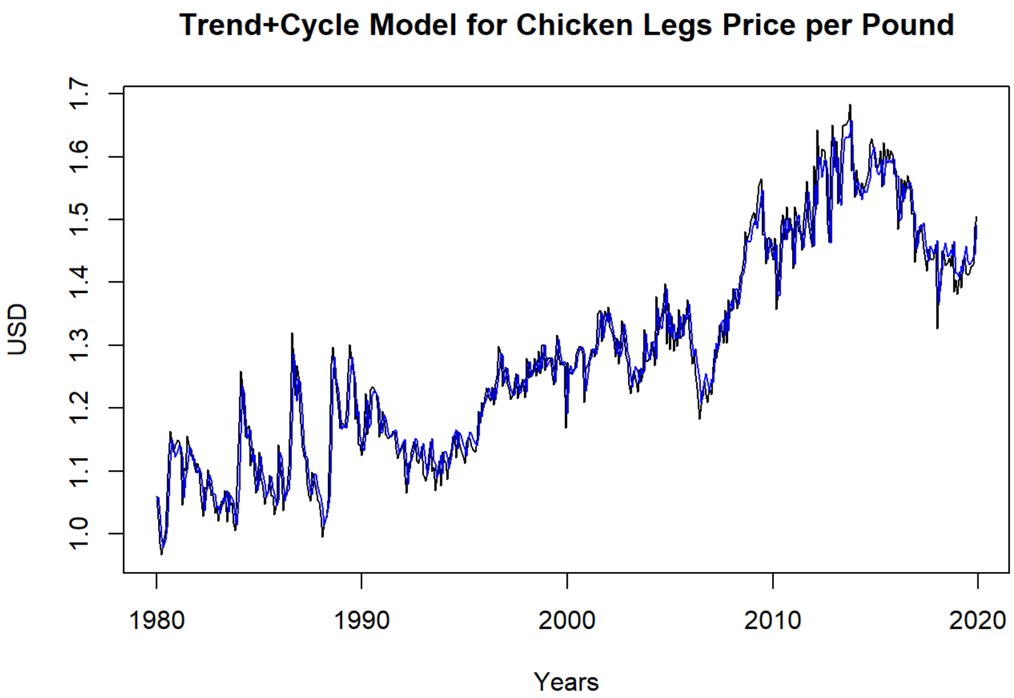


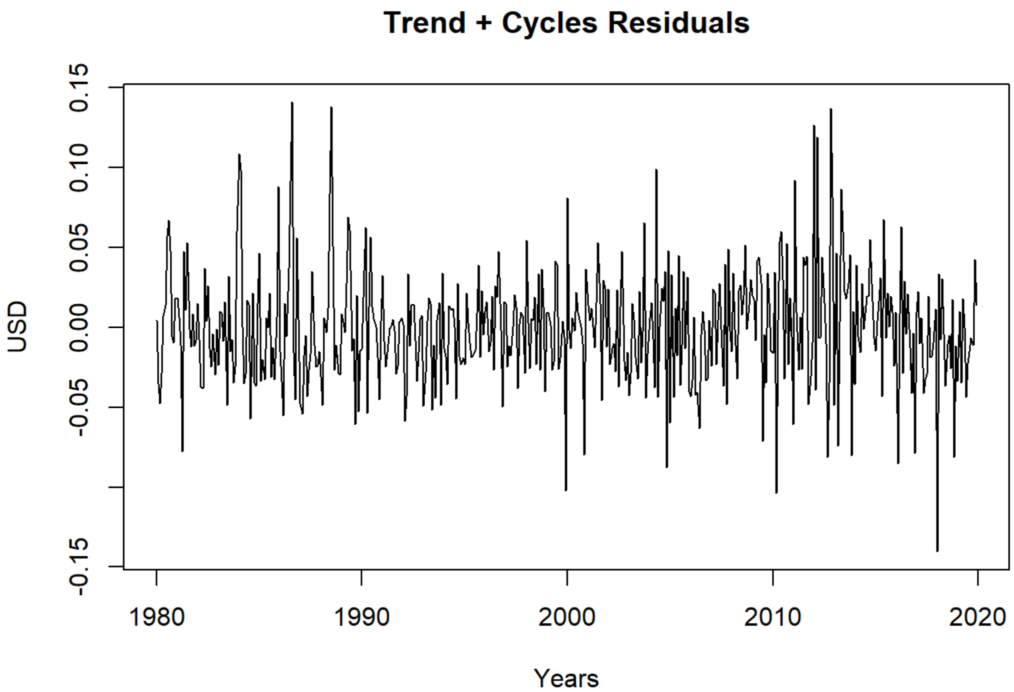


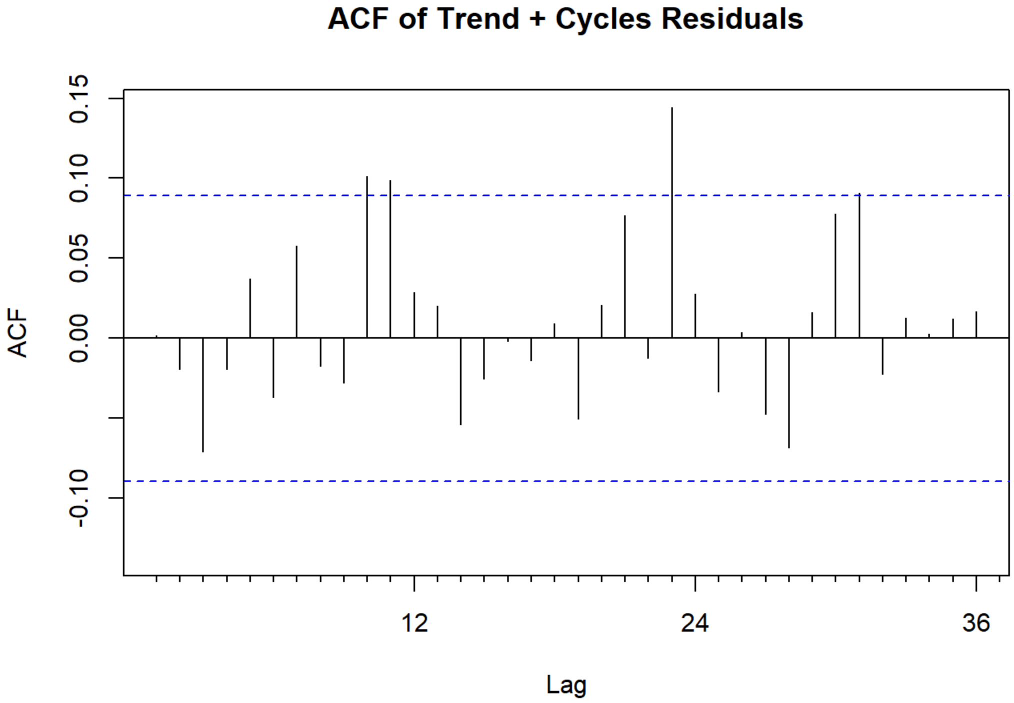
For the chicken data, we decided to use a quadratic function since that is what the graph's general shape most looks like. We can see in the graph of the residuals, that the trend does a somewhat good job of centering the data and removing the overall trend. Note that cubic also works very well with this data, however, we ran into issues when trying to use the cubic trend along with ARIMA (due to computationally singular errors).

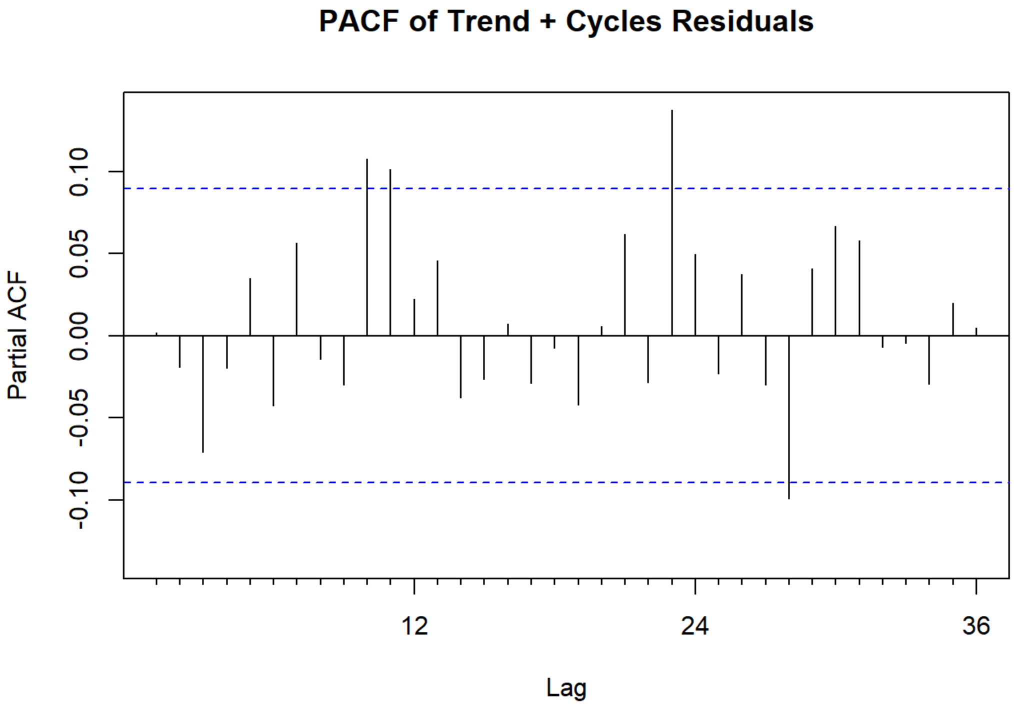
```
##           [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] -1802.987 -1801.364 -1807.745 -1805.765 -1803.937
## [2,] -1803.558 -1801.567 -1805.775 -1804.528 -1802.604
## [3,] -1801.563 -1802.804 -1803.964 -1802.075 -1801.929
## [4,] -1805.348 -1803.408 -1809.367 -1807.397 -1806.569
## [5,] -1803.715 -1802.119 -1807.402 -1806.601 -1804.964
```

```
##            [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] -1777.944 -1772.147 -1774.355 -1768.201 -1762.199
## [2,] -1774.342 -1768.176 -1768.211 -1762.790 -1756.692
## [3,] -1768.172 -1765.240 -1762.226 -1756.163 -1751.844
## [4,] -1767.784 -1761.670 -1763.455 -1757.311 -1752.310
## [5,] -1761.977 -1756.208 -1757.316 -1752.342 -1746.531
```





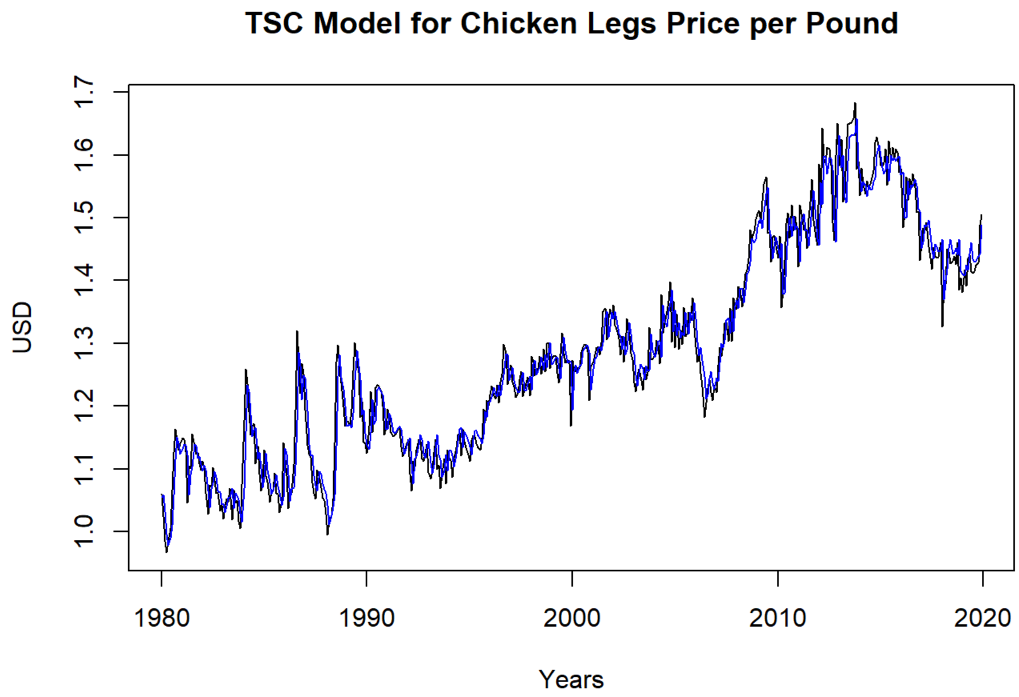


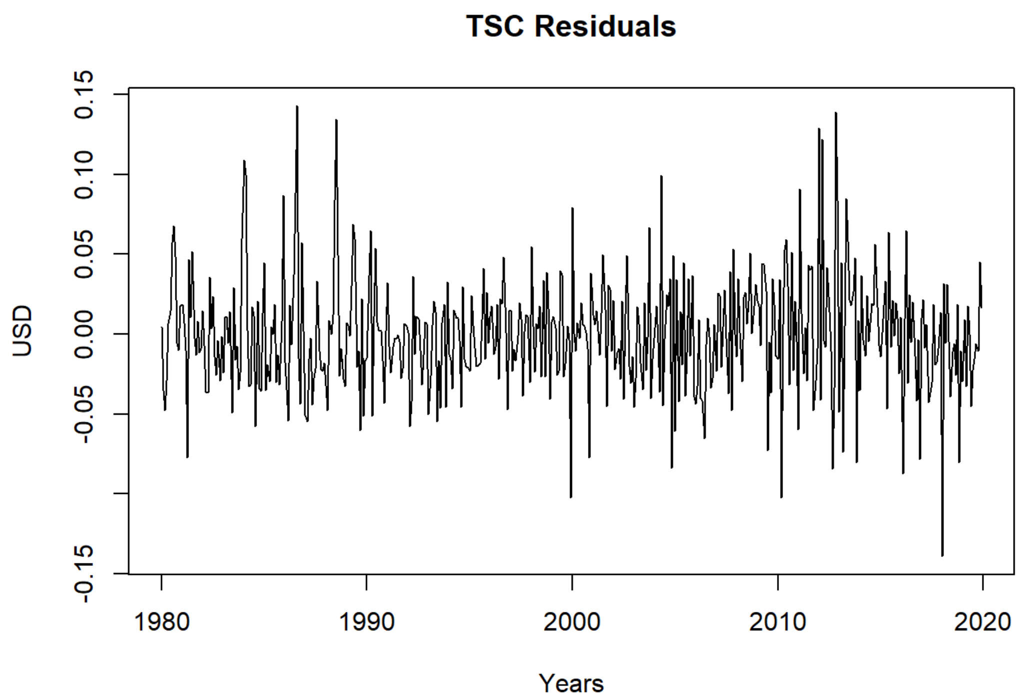


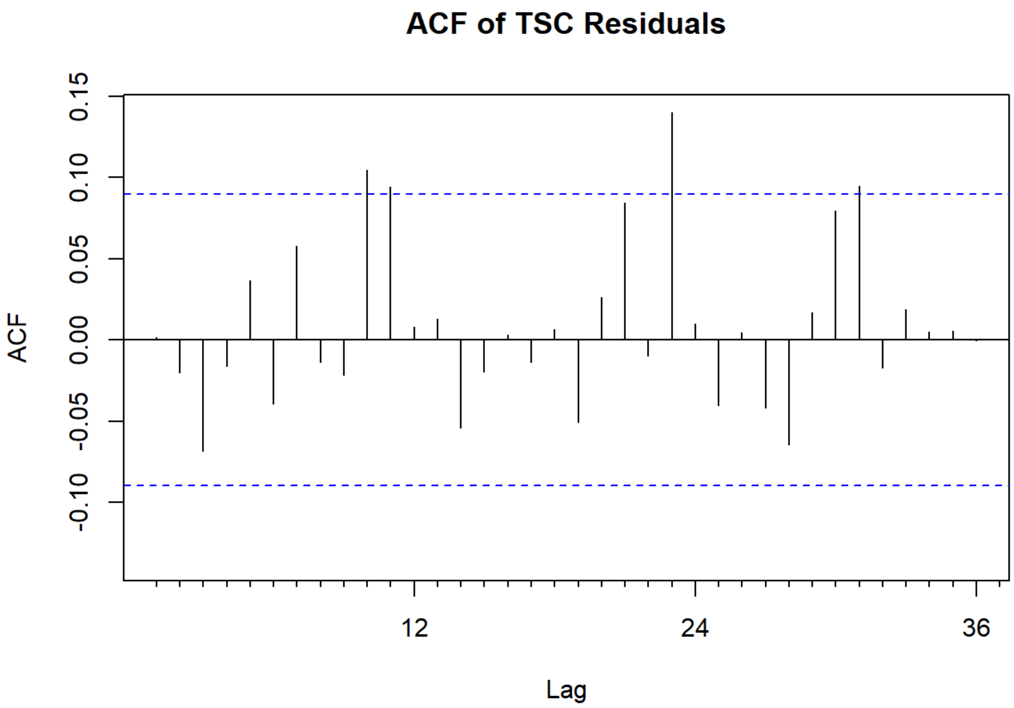
We then attempted to introduce a cyclical component. In order to do this, we optimized using AIC and BIC different order of ARMA, whose values are shown above. Note that the row corresponds with the order of AR and the column corresponds with the order of MA. The AIC and BIC are not congruent unfortunately. In that case, we will choose to use an ARMA(3,1) going forward, based on the lowest AIC value (and a medium-ly low BIC). We will use this to account for cyclicality.

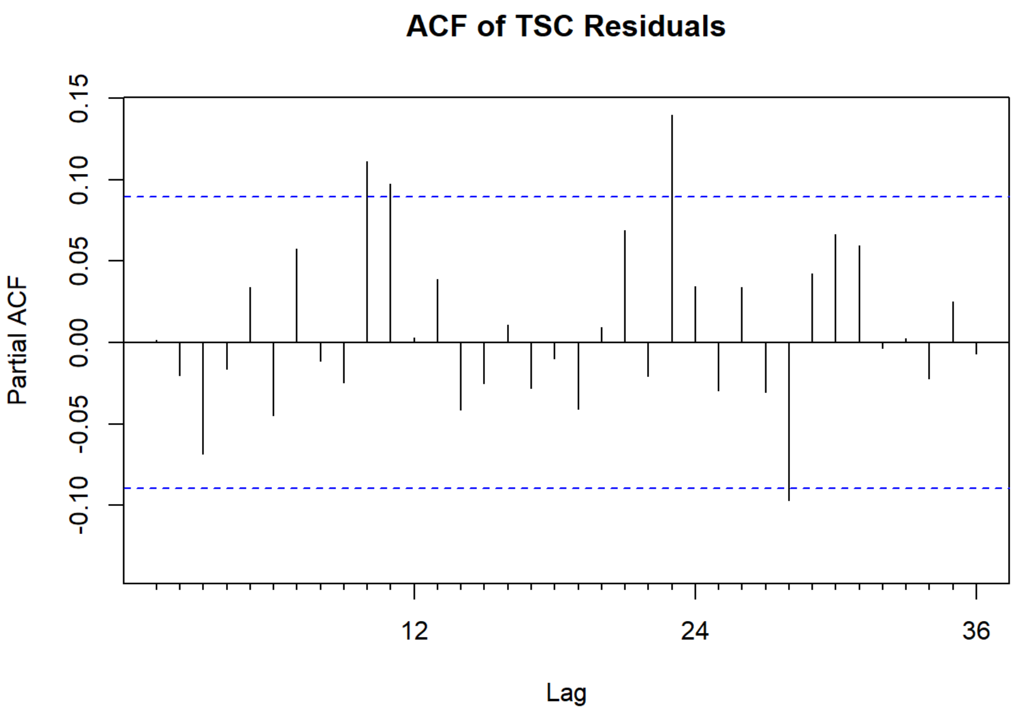
```
##           [,1]      [,2]
## [1,] -1803.558 -1801.942
## [2,] -1801.965 -1800.574
```

```
##           [,1]      [,2]
## [1,] -1774.342 -1768.552
## [2,] -1768.574 -1763.010
```


TSC Model for Chicken Legs Price per Pound
USD
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1980
1990
2000
2010
2020
Years


TSC Residuals
USD
-0.15
-0.05
0.00
0.05
0.10
0.15
1980
1990
2000
2010
2020
Years

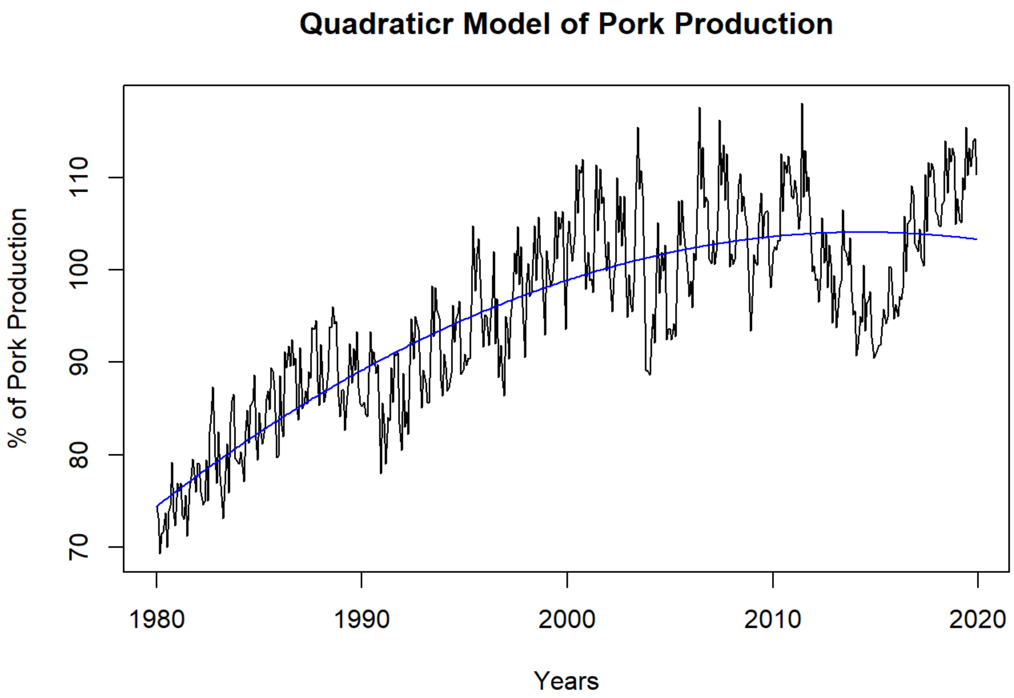


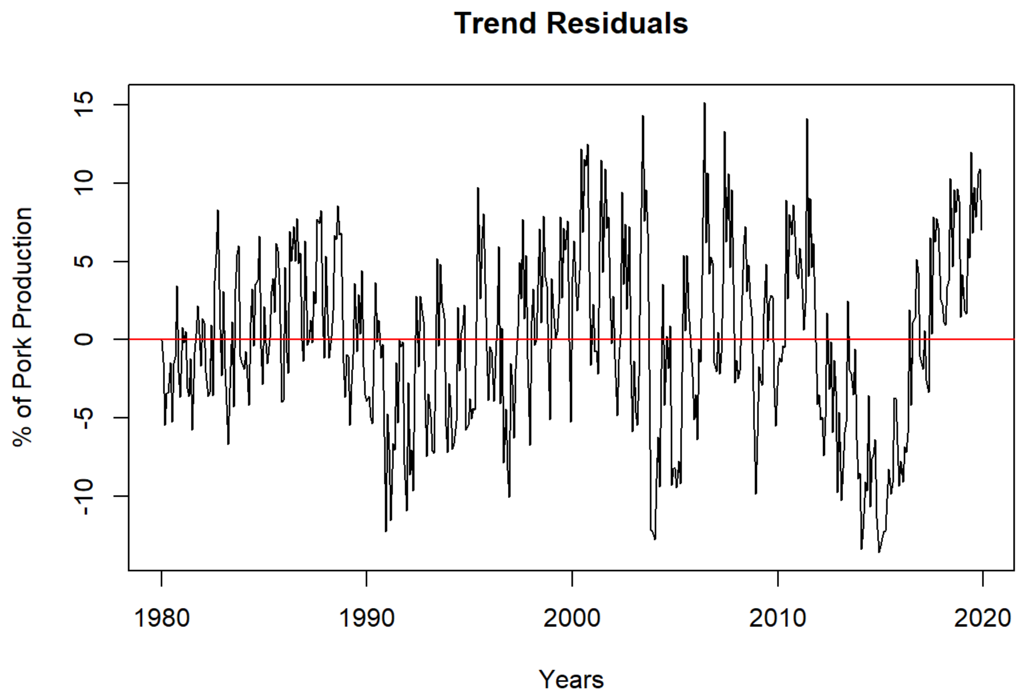


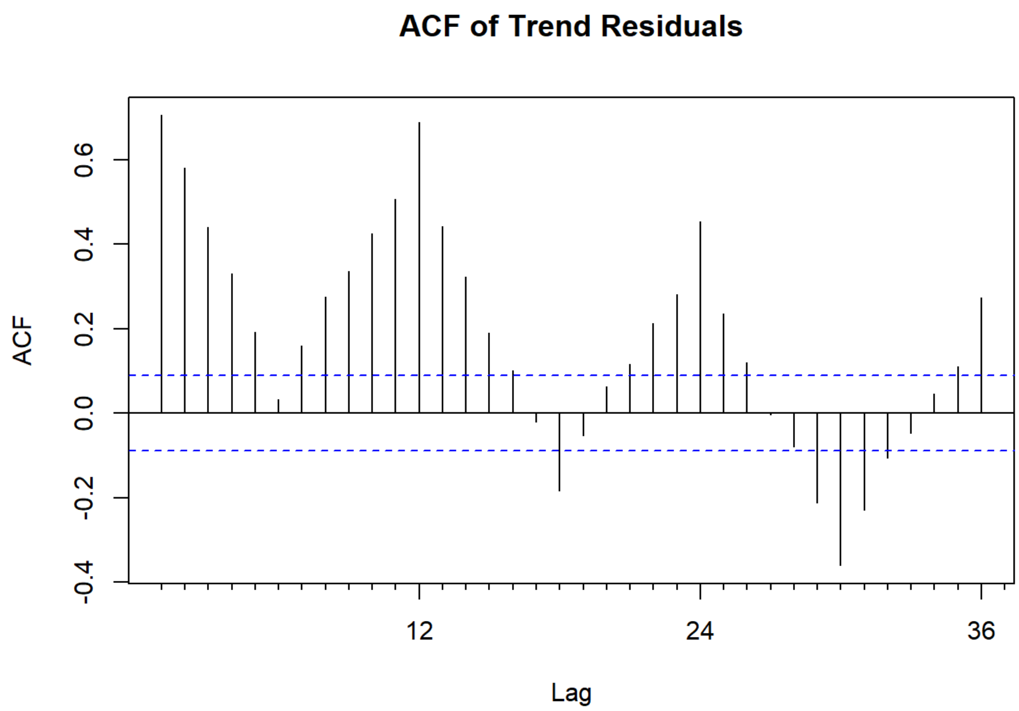
Lastly, since there are still spikes in the ACF and PACF of the residuals, we will add a seasonal component. This will be an S-ARMA, and once again we looped through to figure out the optimal set. We found that the lowest AIC and BIC values were found at cell (2,2) which corresponds with S-ARMA(1,1). Unfortunately, even though both seasonal variables are statistically significant, and help the model, we still see that we did not capture all the dyanamics due to the spikes in the ACF and PACF.

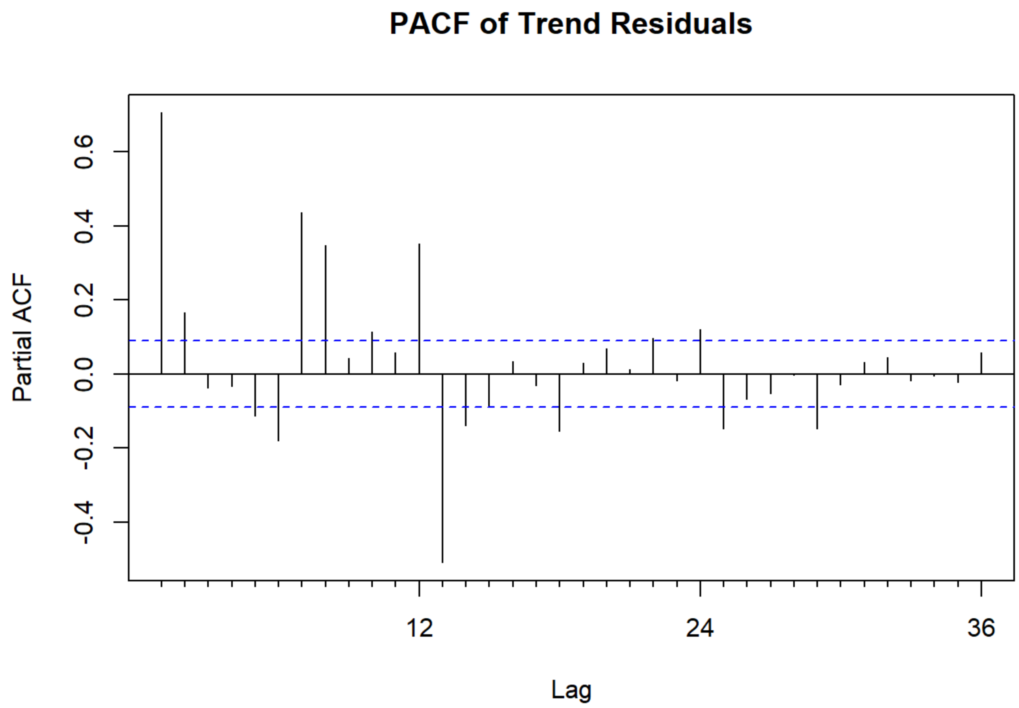
Model for Beef

```
##
## Call:
## tslm(formula = beef ~ t.b + t2.b)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5970  -3.6866  -0.1676   3.8886  15.1307
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.017e+05  8.573e+03  -11.87   <2e-16 ***
## t.b          1.011e+02  8.573e+00   11.79   <2e-16 ***
## t2.b        -2.510e-02  2.143e-03  -11.71   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.6 on 477 degrees of freedom
## Multiple R-squared:  0.716,  Adjusted R-squared:  0.7148
## F-statistic: 601.3 on 2 and 477 DF,  p-value: < 2.2e-16
```




Trend Residuals
% of Pork Production
-10
-5
0
5
10
15
1980
1990
2000
2010
2020
Years


ACF of Trend Residuals
ACF
-0.4
-0.2
0.0
0.2
0.4
0.6
12
24
36
Lag



For the beef data, we chose to also use a quadratic function, since when inputting a cubic polynomial, the p-value becomes insignificant than if we only use a quadratic function. Also, when graphing the time series function, it seems that the quadratic fit fits better than the cubic function fit. As we can see from the residual plot, it seems that the residual is quite evenly distributed around 0, and this is one prove that the quadratic fit is a good fit for the beef data, though there is still some room for improvement since the residual plot still shows some pattern.

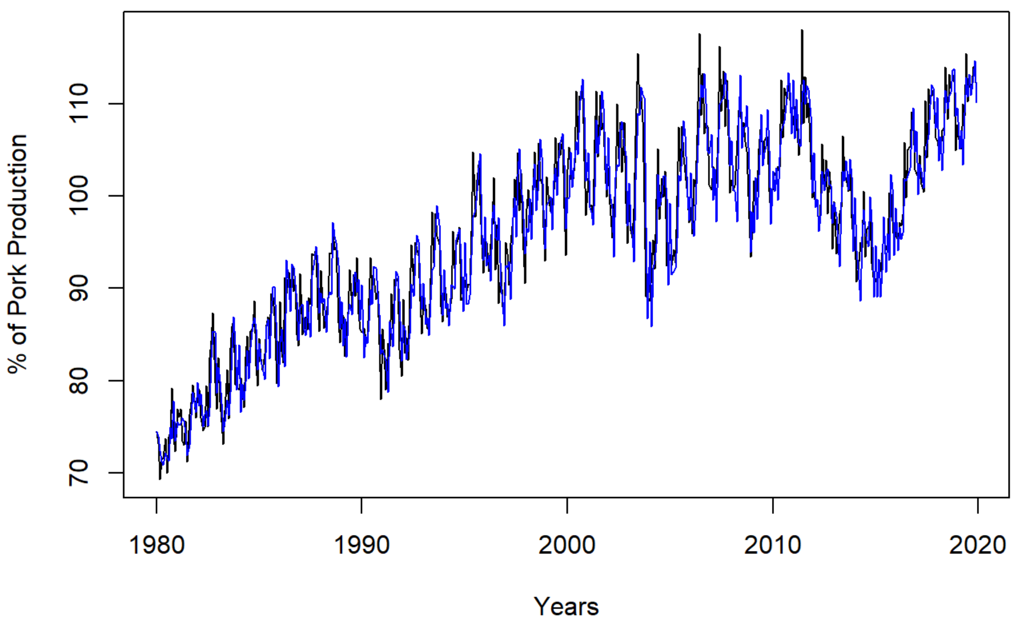
```
##          [,1]     [,2]     [,3]     [,4]     [,5]
## [1,] 2713.485 2715.251 2712.958 2694.983 2686.461
## [2,] 2715.191 2701.814 2694.674 2696.046 2682.616
## [3,] 2689.567 2626.853 2608.965 2535.587 2585.502
## [4,] 2718.567 2590.382 2580.723 2535.422 2653.944
## [5,] 2716.387 2564.033 2564.049 2466.850 2530.849
```

```
##          [,1]     [,2]     [,3]     [,4]     [,5]
## [1,] 2730.180 2736.120 2738.001 2724.199 2719.852
## [2,] 2736.060 2726.857 2723.890 2729.436 2720.180
## [3,] 2714.609 2656.070 2642.355 2573.151 2627.239
## [4,] 2747.783 2623.772 2618.287 2577.160 2699.855
## [5,] 2749.777 2601.597 2605.787 2512.762 2580.935
```

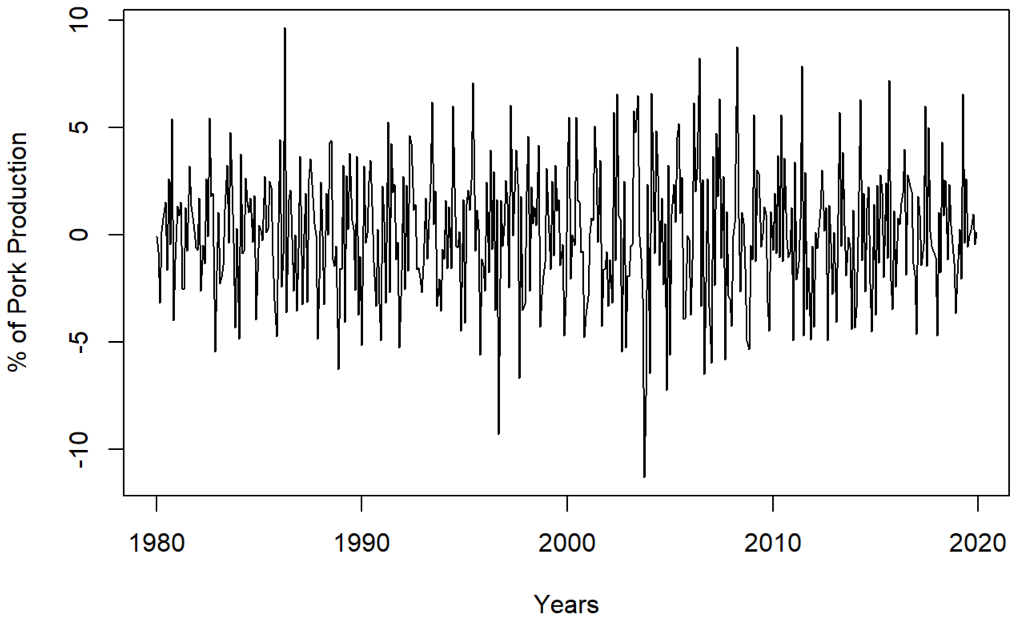
From the AIC and BIC model, we can see that the model with the lowest possible value of AIC and BIC is when the model follows ARMA(5,4). Different from the chicken data, it seems that the AIC and BIC for the beef data is congruent in determining the lowest AIC and BIC for the model. Thus from here onwards we will use ARMA(5,4) fit for our beef model.

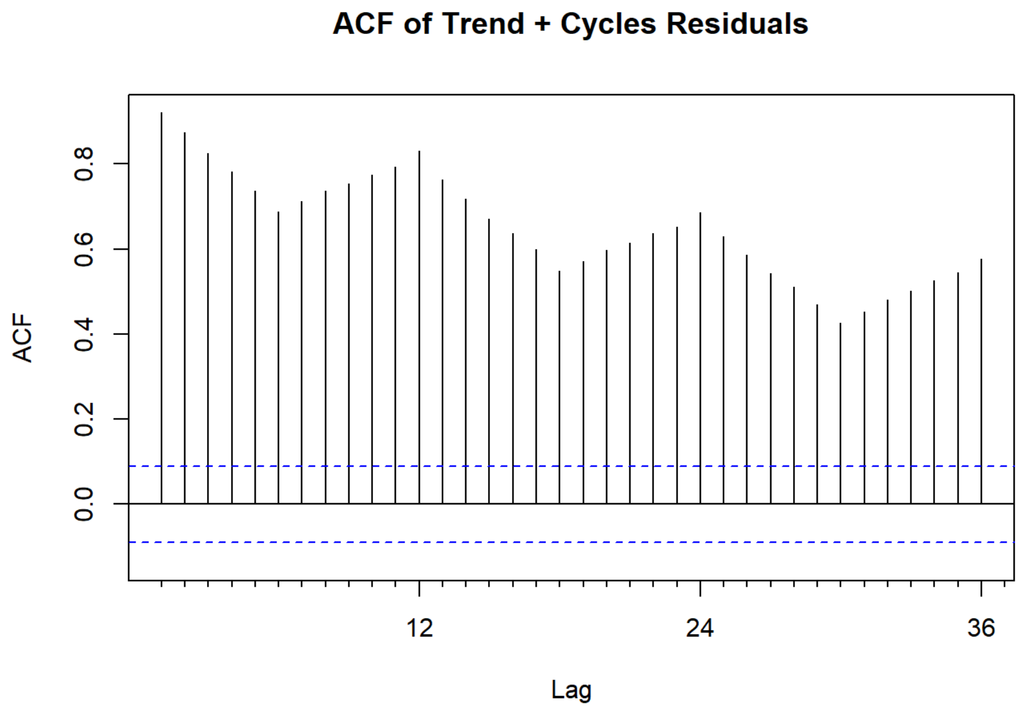
```
## 
## Call:
## arima(x = beef, order = c(5, 0, 4))
## 
## Coefficients:
##           ar1      ar2     ar3     ar4     ar5     ma1     ma2     ma3     ma4  intercept
##       -0.7241  -0.2595  0.2673  0.7212  0.9951  1.6583  1.9017  1.6581  0.9998    95.5211
## s.e.   0.0036   0.0039  0.0033  0.0042  0.0029  0.0115  0.0184  0.0221  0.0144        NaN
## 
## sigma^2 estimated as 9.139:  log likelihood = -1222.43,  aic = 2466.85
## 
## Training set error measures:
##                      ME     RMSE     MAE         MPE     MAPE      MASE        ACF1
## Training set 0.08259696 3.023123 2.37367 0.03538388 2.491126 0.6867597 -0.04411211
```

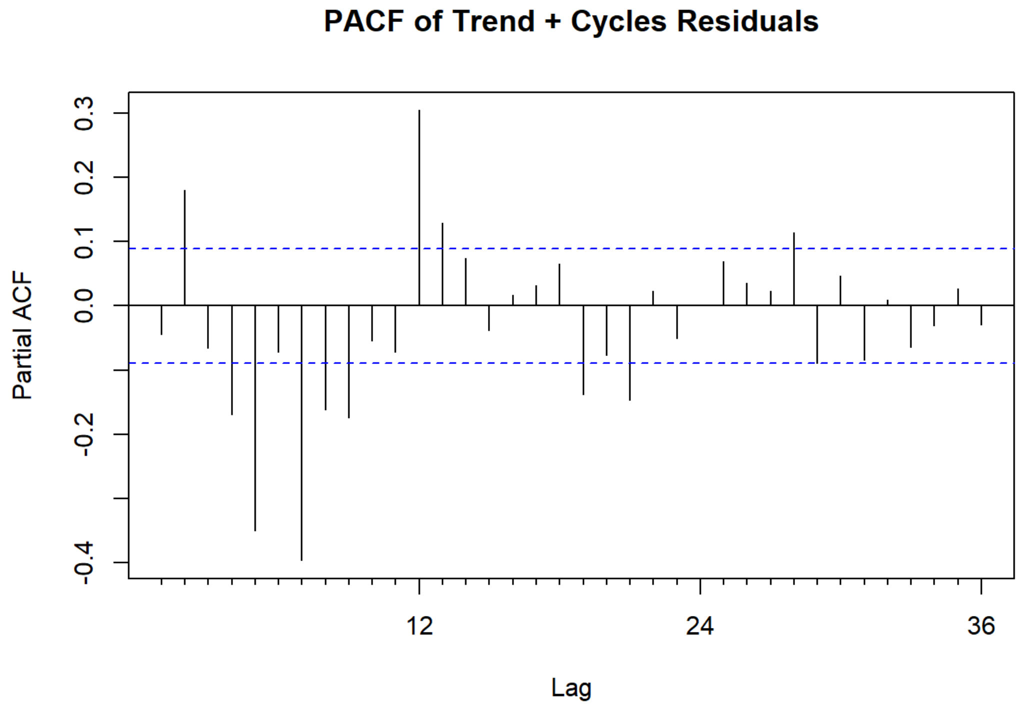
Trend+Cycle Model of Pork Production



Trend + Cycles Residuals







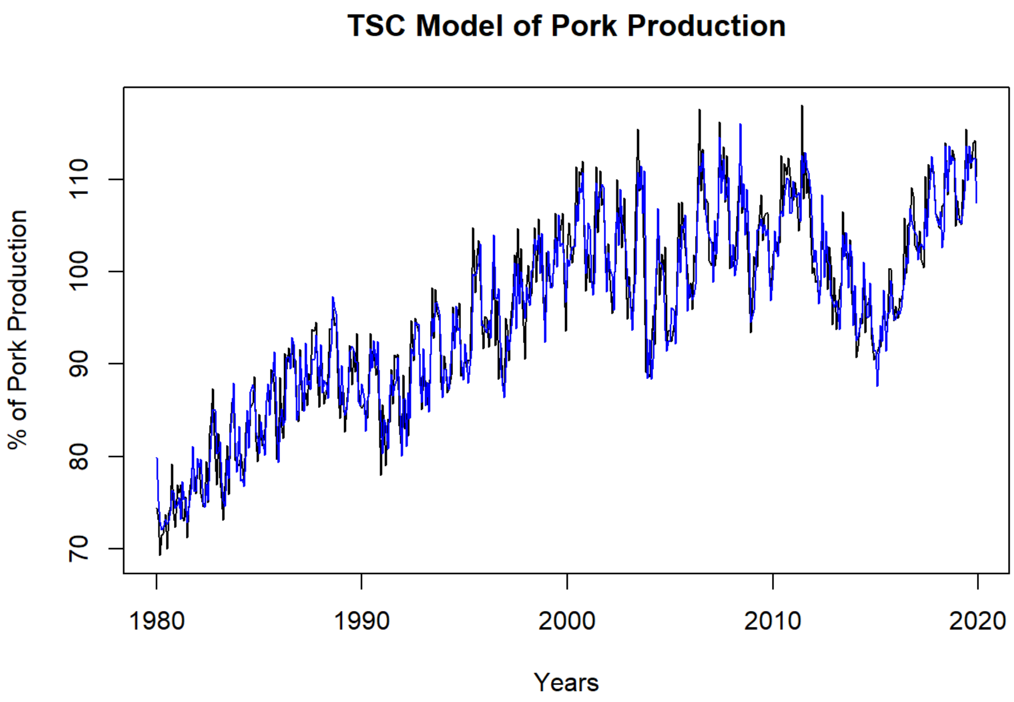
Then, we add the cyclical component to our model using the aforementioned ARMA(5,4) fit for our model. As we can see from the residuals, the trend + cyclical fit for the beef data is a better representation for the data since the residuals are more evenly distributed around 0 and resembles a white noise more than the trend only fit.

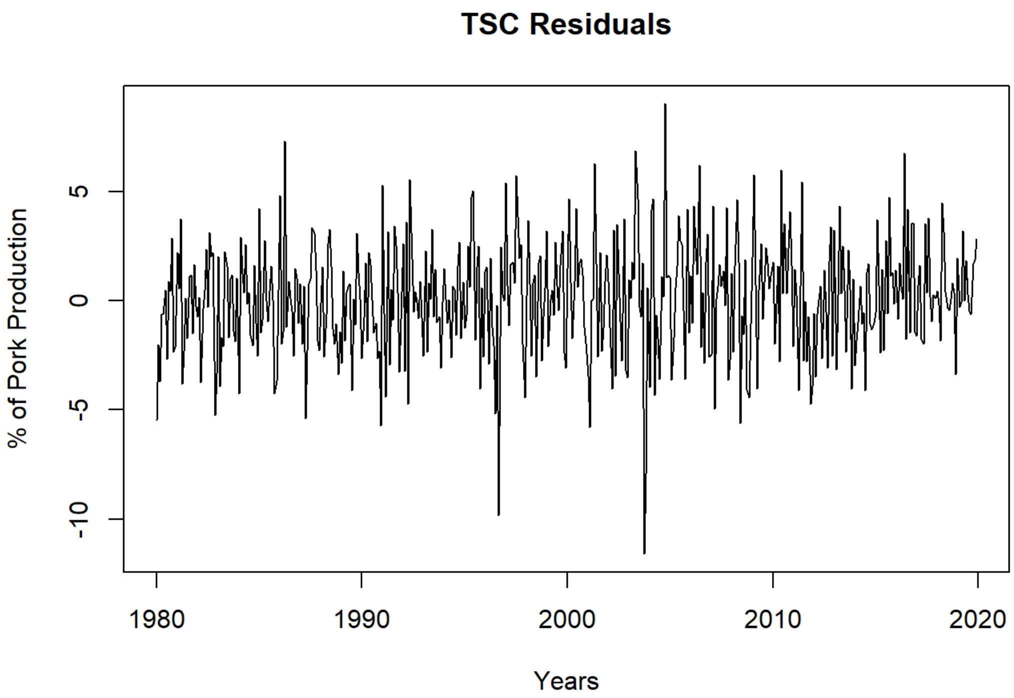
```
## [1] 2280.702
```

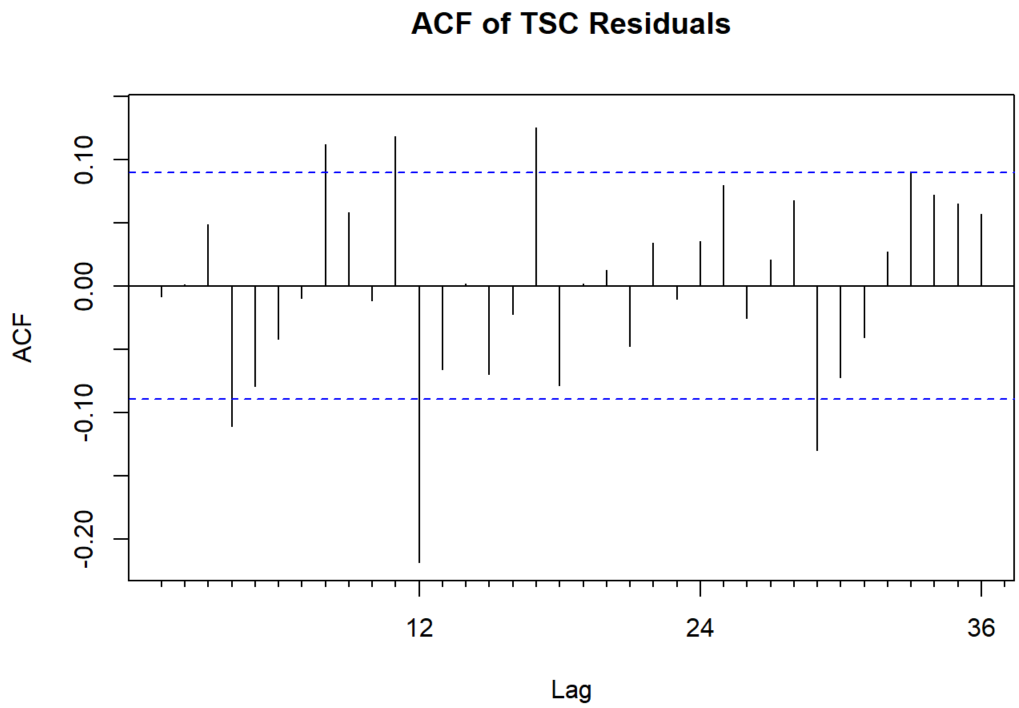
```
## [1] 2330.787
```

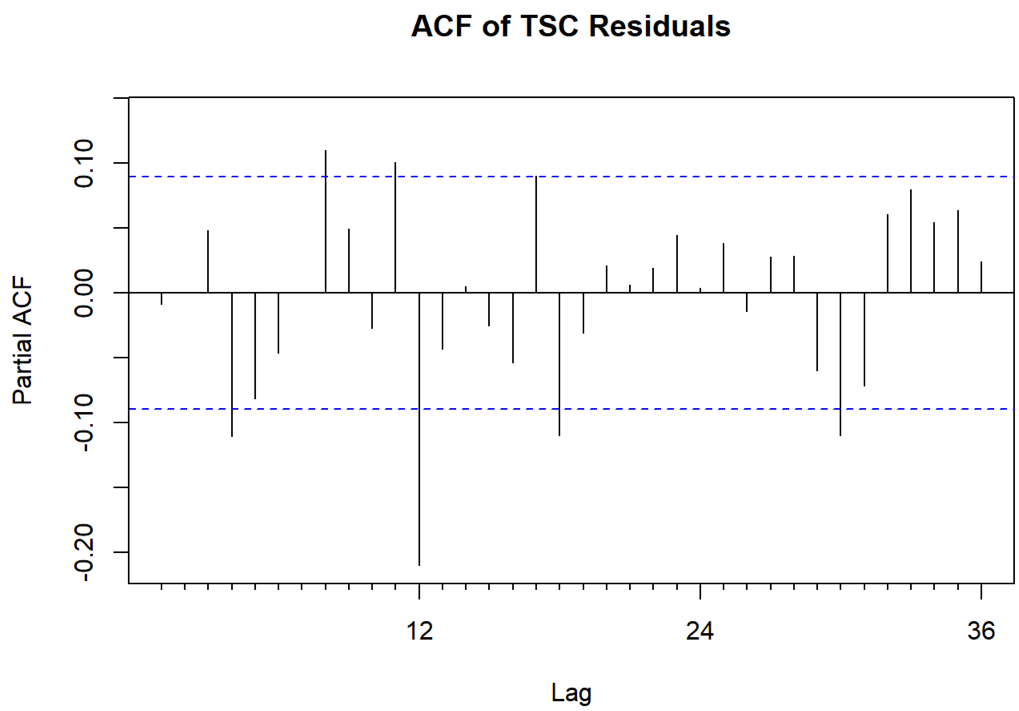
It seems that the only possible value of seasonality is when the seasonal is (1,0). We tried doing the same looping process as the chicken data, but the result seems to return an error message. This model corresponds to an S-ARMA model fit.

```
##
## Call:
## arima(x = beef, order = c(2, 0, 1), seasonal = list(order = c(1, 0, 0)))
##
## Coefficients:
##          ar1     ar2      ma1    sar1  intercept
##       0.7385  0.1623  -0.1540  0.7714    95.1036
## s.e.  0.1370  0.1175   0.1356  0.0292     4.0171
##
## sigma^2 estimated as 6.748:  log likelihood = -1145.71,  aic = 2303.43
##
## Training set error measures:
##                     ME     RMSE      MAE        MPE     MAPE      MASE         ACF1
## Training set 0.1099518 2.597707 2.021747 0.03823115 2.134383 0.5849397 -0.008322337
```





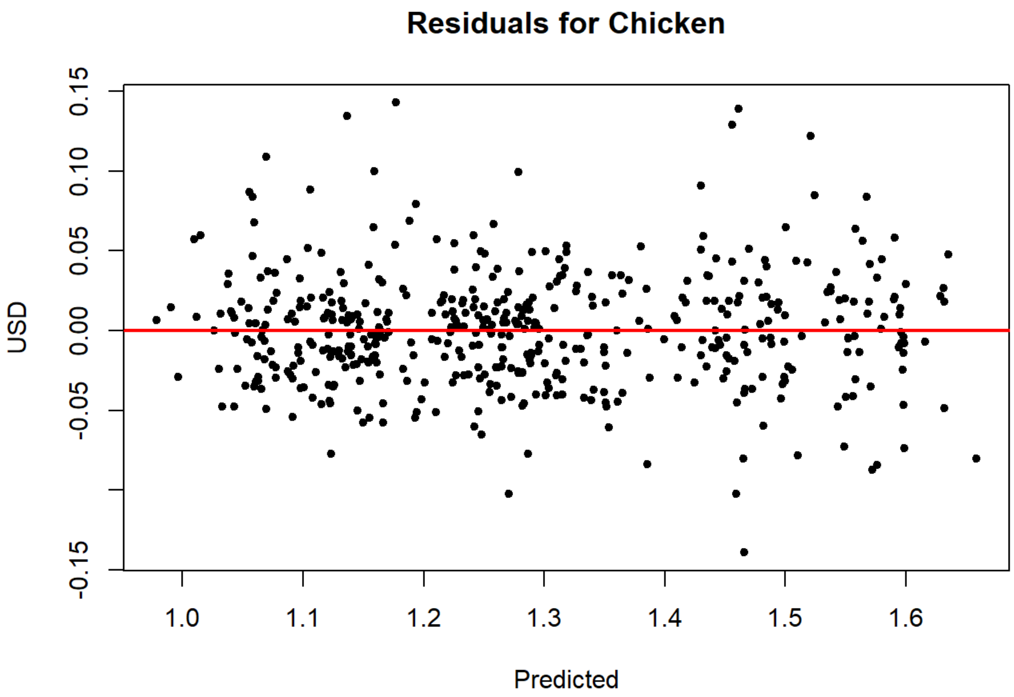




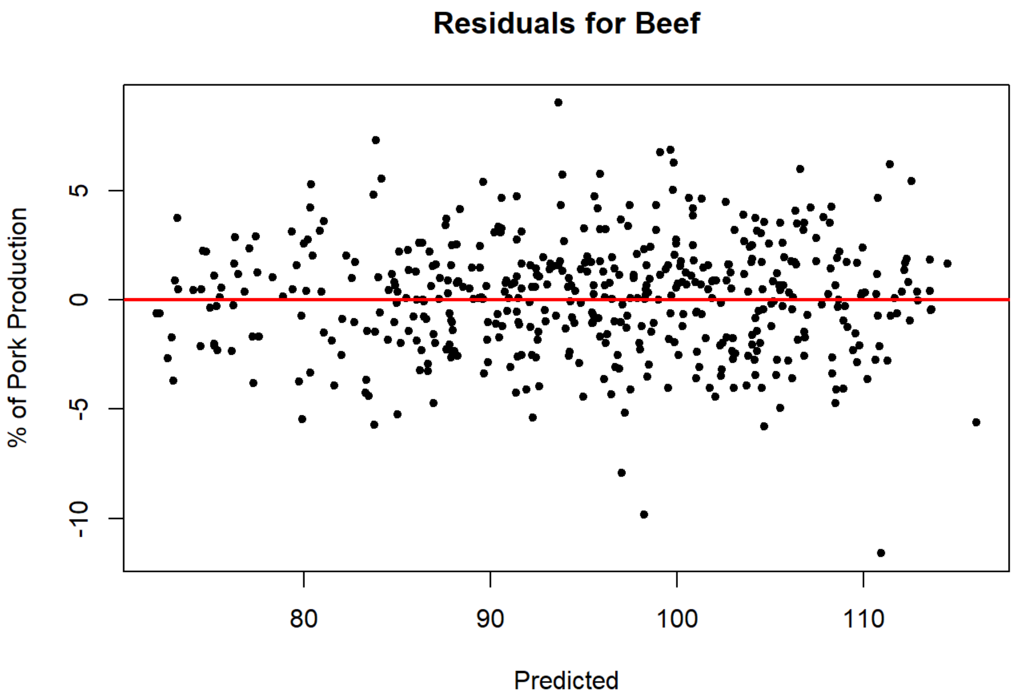
As we can see from the fitted plot, we have done a quite good job in fitting the model to the actual data. In addition to that, the residual plot seems to follow a white noise process, and the ACF and PACf plot of the residuals seem to be consistent with my argument of how the residual is following a white noise process.

c.

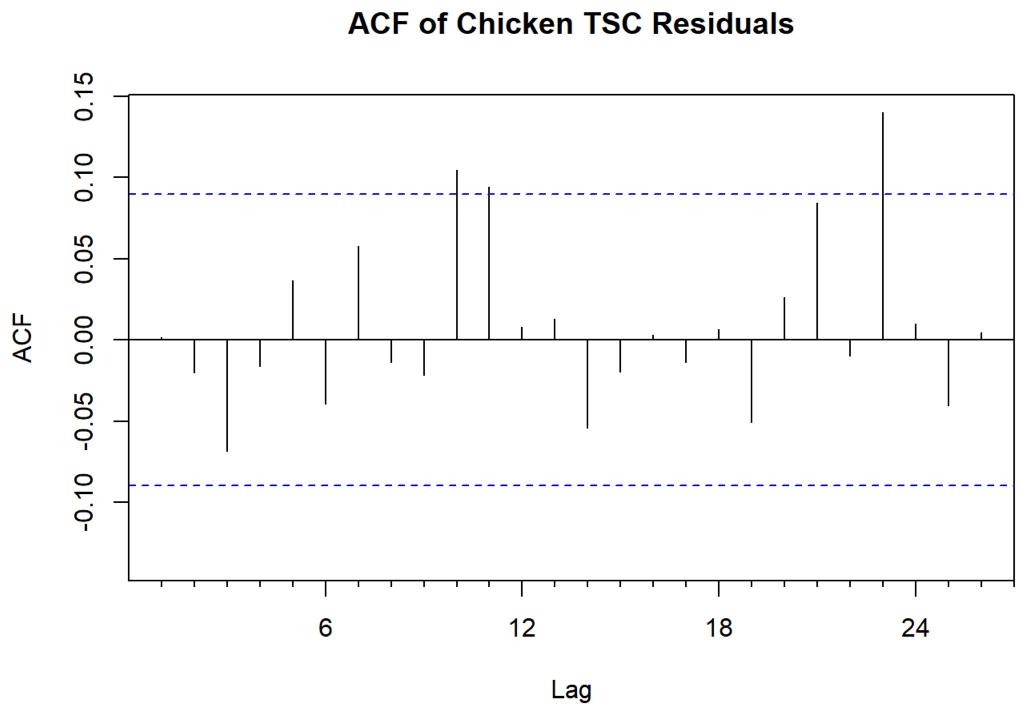
Below is the residuals vs. fitted values of the chicken data. As we can see, though is it quite a good fit, the points are not evenly distributed around 0. However, this represenatation of the data is still good enough.

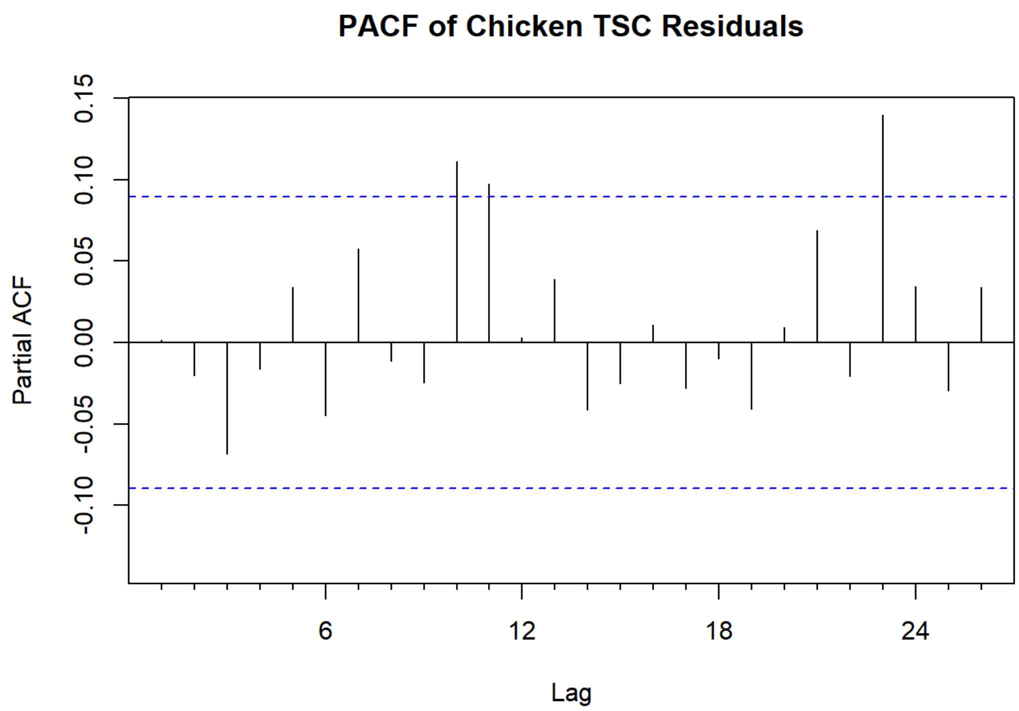


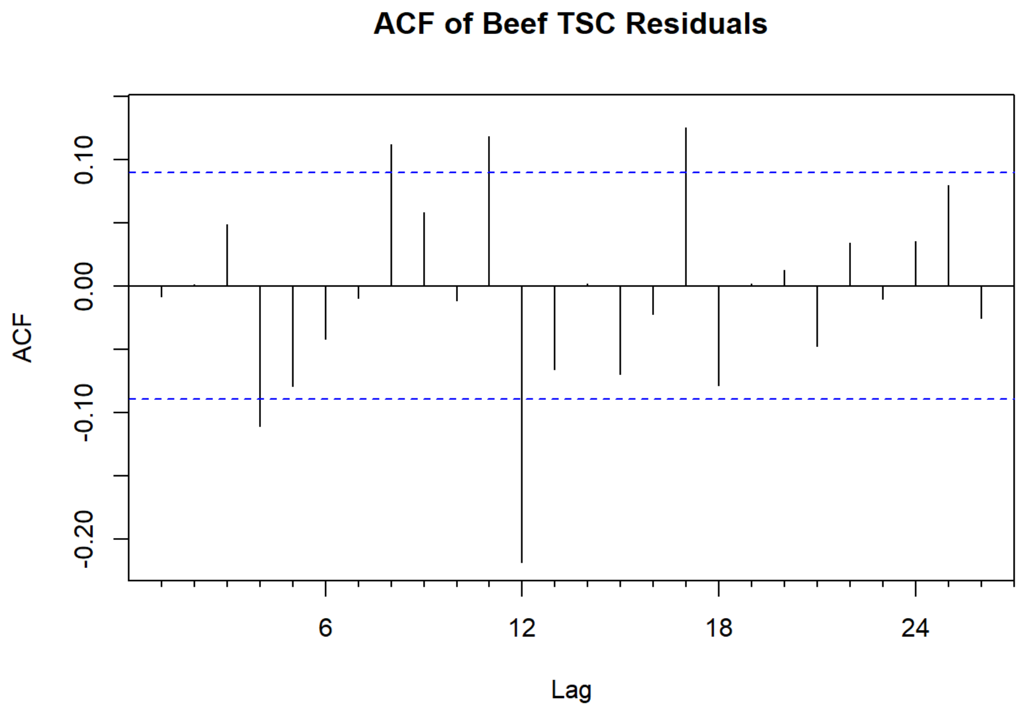
And below is the residuals vs. fitted values of the Beef data. As we can see, the residuals vs. fitted value of the data seems to show an even distribution around 0. This shows that there is a correlation between the residuals and the fitted value.

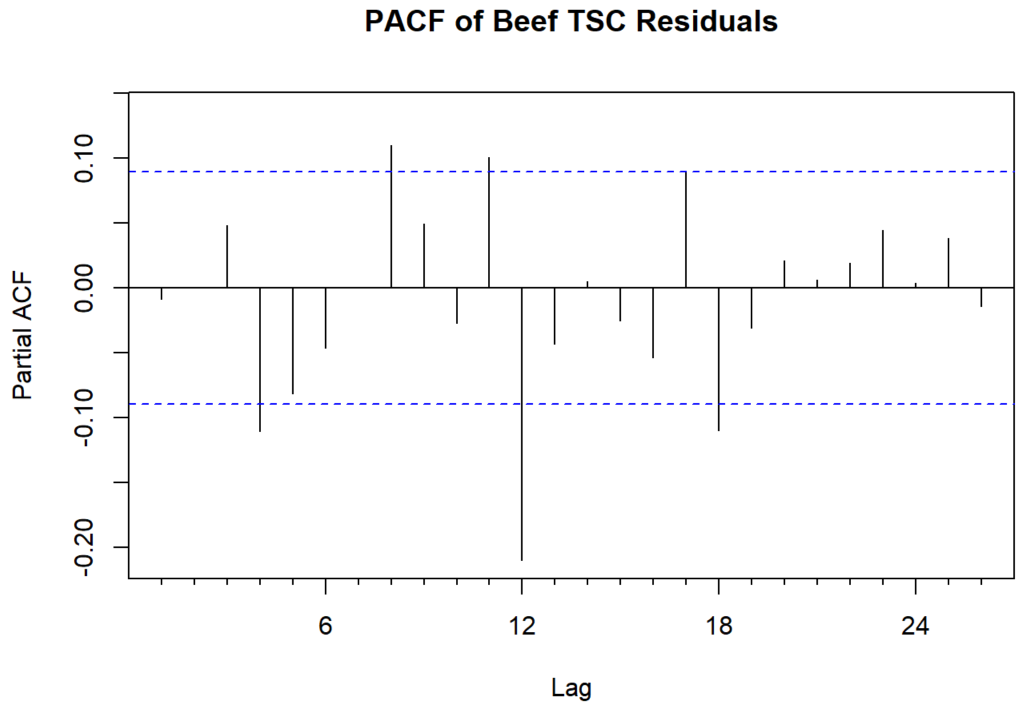


e.


ACF of Chicken TSC Residuals
ACF
0.15
0.10
0.05
0.00
-0.05
-0.10
6
12
18
24
Lag


PACF of Chicken TSC Residuals
Partial ACF
0.15
0.10
0.05
0.00
-0.05
-0.10
6
12
18
24
Lag

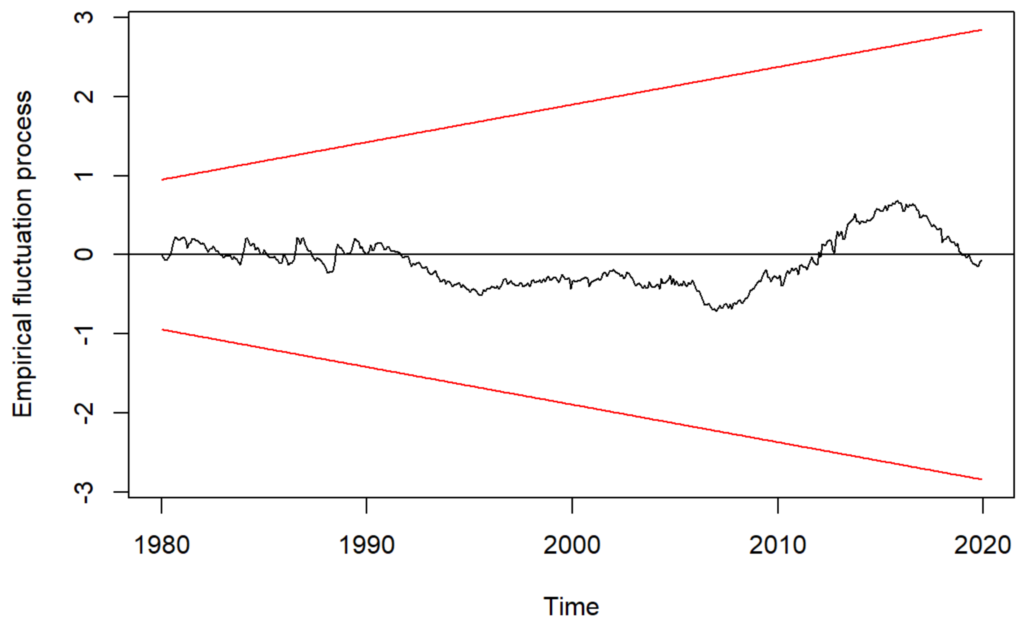




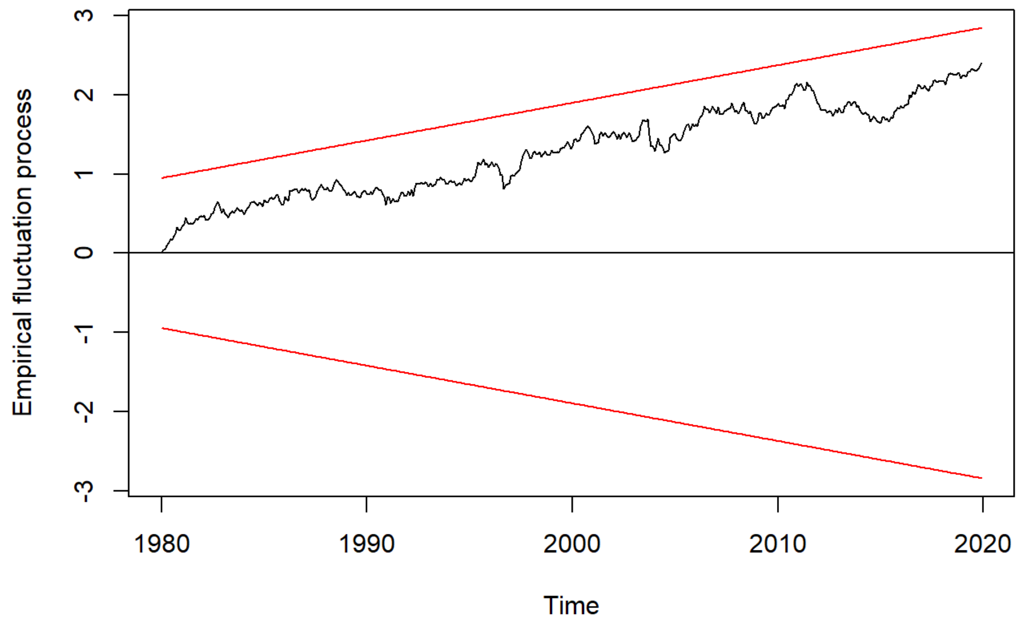
As we can see from both the ACF and PACF plot from the residuals of the data, both residuals seems to follow a white noise process, thus showing that our fit of the data is a good fit.

f.

Recursive CUSUM for Chicken TSC

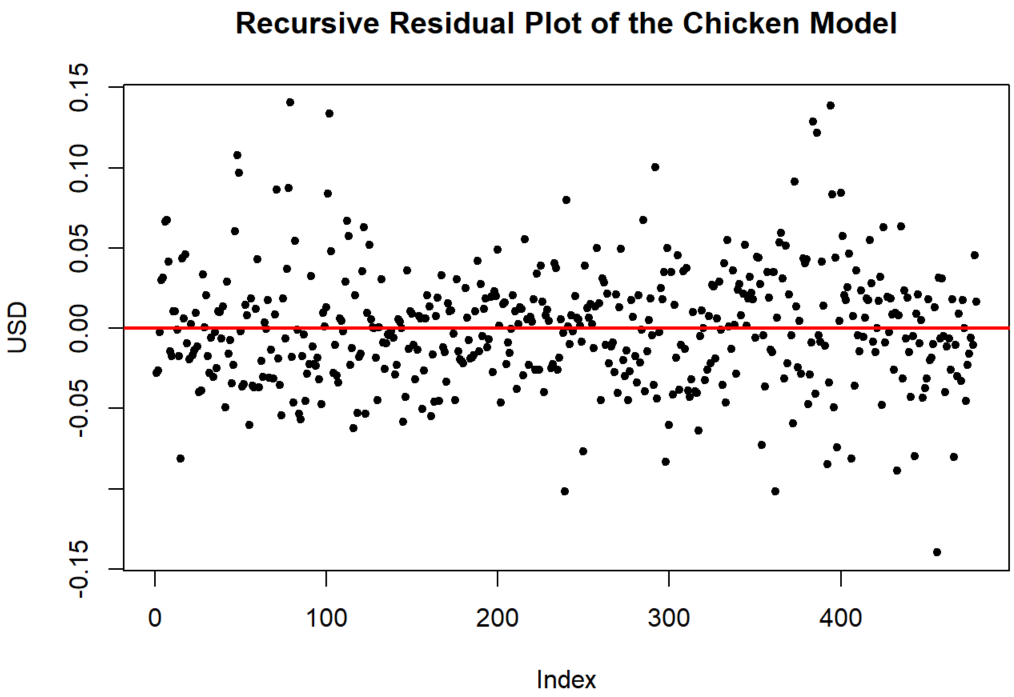


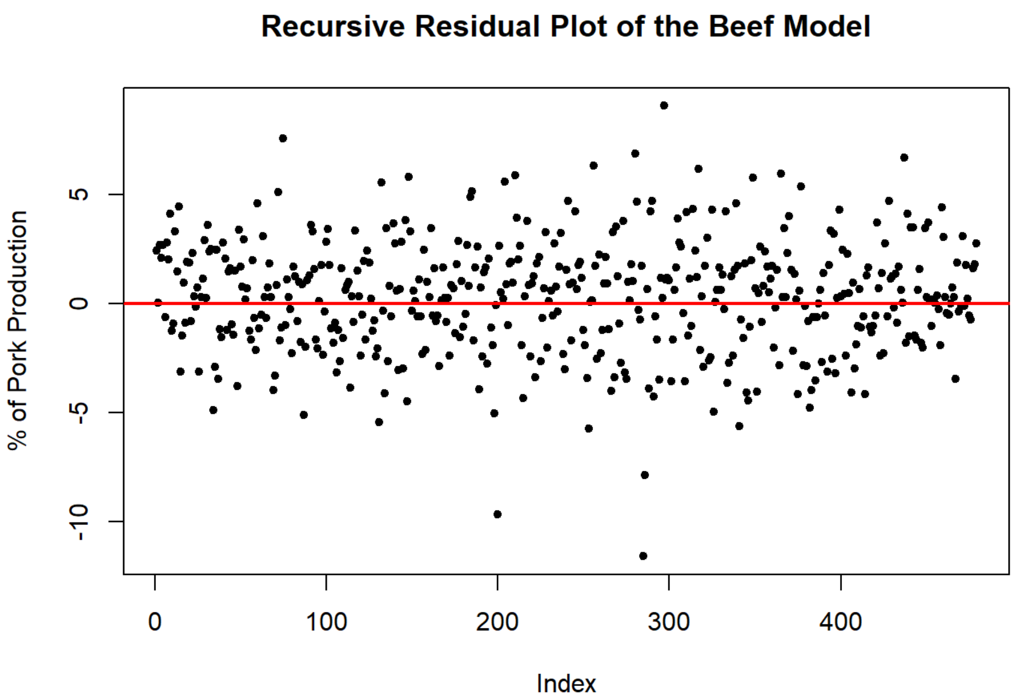
Recursive CUSUM for Beef TSC



From the Recursive CUSUM plots, we know that both models have passed the test; as we add more observations to our data, the variance stays inside the band, thus making the models stay consistent as we add more data. What we also know, is that the model for the Chicken data seems to be more consistent as we add more data than the Beef data, since the variance of the Beef Recursive CUSUM test seems to be increasing, though still staying inside the band, while the variance of the Chicken Recursive CUSUM test seems to be more persistent around 0.

g.





The recursive residual plot of both models seems to show even distributions around 0, which makes them normally distributed around 0. This is a good sign that our model is a good fit for the actual data.

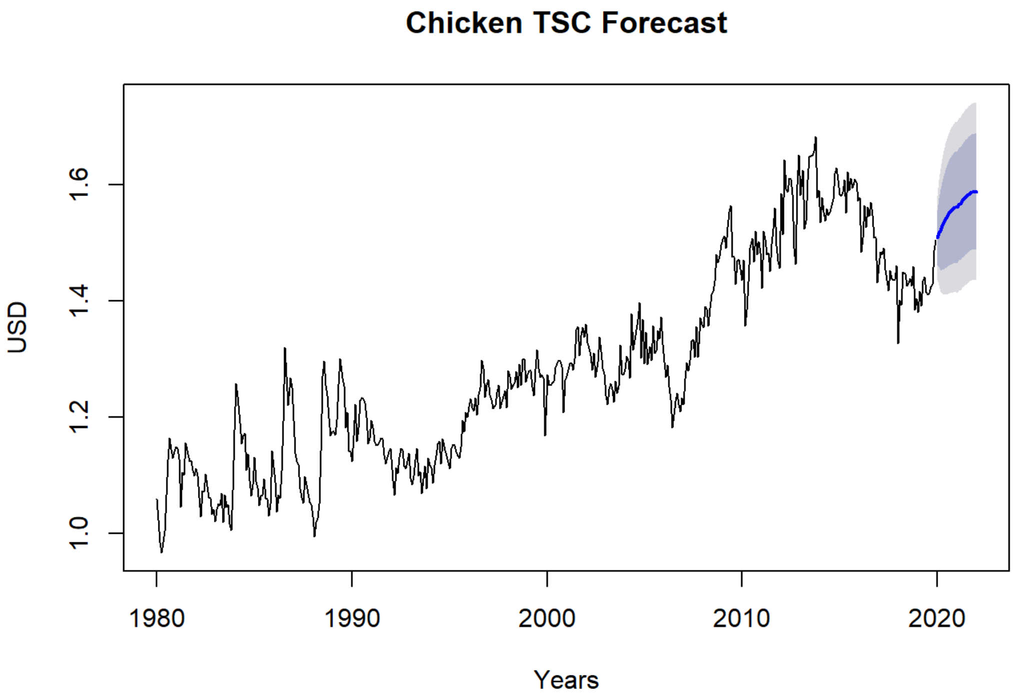
h.

So, in conclusion, we have used a TSC model for chicken, with a trend of t and t^2, a seasonality component of S-ARMA(1,1), and a cyclic component of ARMA(2,1). We can see that most of the values in the ACF and PACF are not statistically significant, aside from a few random spikes at late lag intervals. These we will attempt to remove by using a GARCH model in a later section. The recrusive residual plot looks relatively normally distributed, and the CUSUM shows no subset of the data that is model breaking. Thus this model, at least utilizing just ARMA and trend, is the best we can do.

So, in conclusion, we have used a SC model for chicken, a seasonality component of S-ARMA(1,0), and a cyclic component of ARMA(5,4). We can see that most of the values in the ACF and PACF are not statistically significant, aside from a few random spikes at late lag intervals. These we will attempt to remove by using a GARCH model in a later section. The recrusive residual plot looks relatively normally distributed, and the CUSUM shows no subset of the data that is model breaking. Thus this model, at least utilizing just ARMA and trend, is the best we can do.

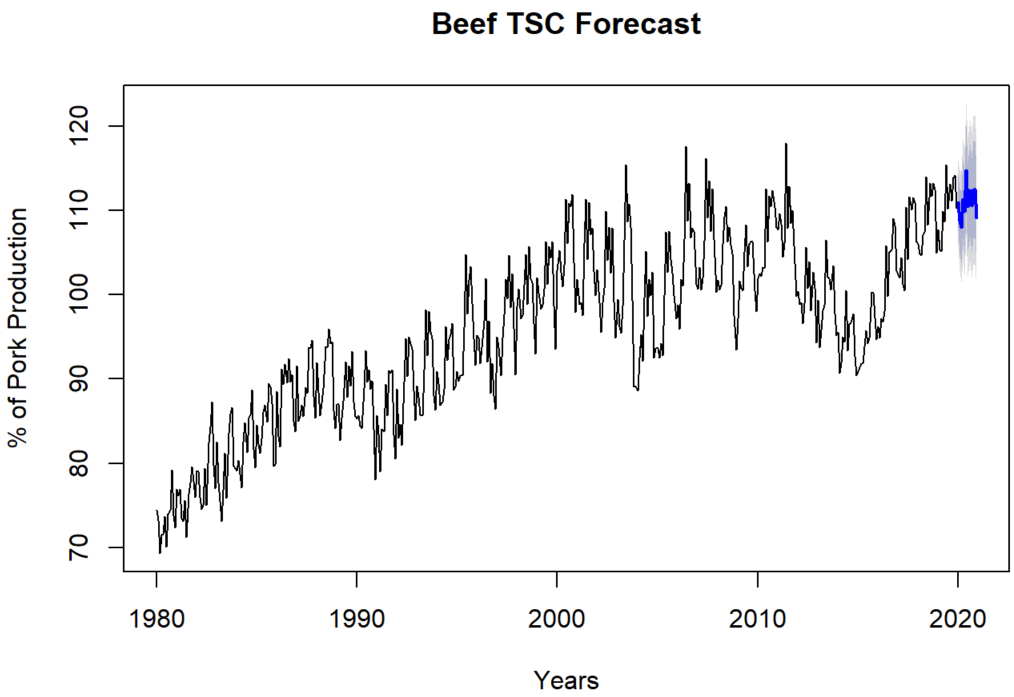
i.

The model below shows the h=12 steps forecast of the Chicken data. From the function `accuracy()`, we know that the MAPE for this particular forecast is 2.126451.




```
##                         ME      RMSE       MAE         MPE     MAPE      MASE        ACF1
## Training set -2.594935e-05 0.0363295 0.0272436 -0.08041339 2.126451 0.4285716 0.001303348
```

The model below shows the h=12 steps forecast of the Beef data. From the function `accuracy()`, we know that the MAPE for this particular forecast is 2.134383.

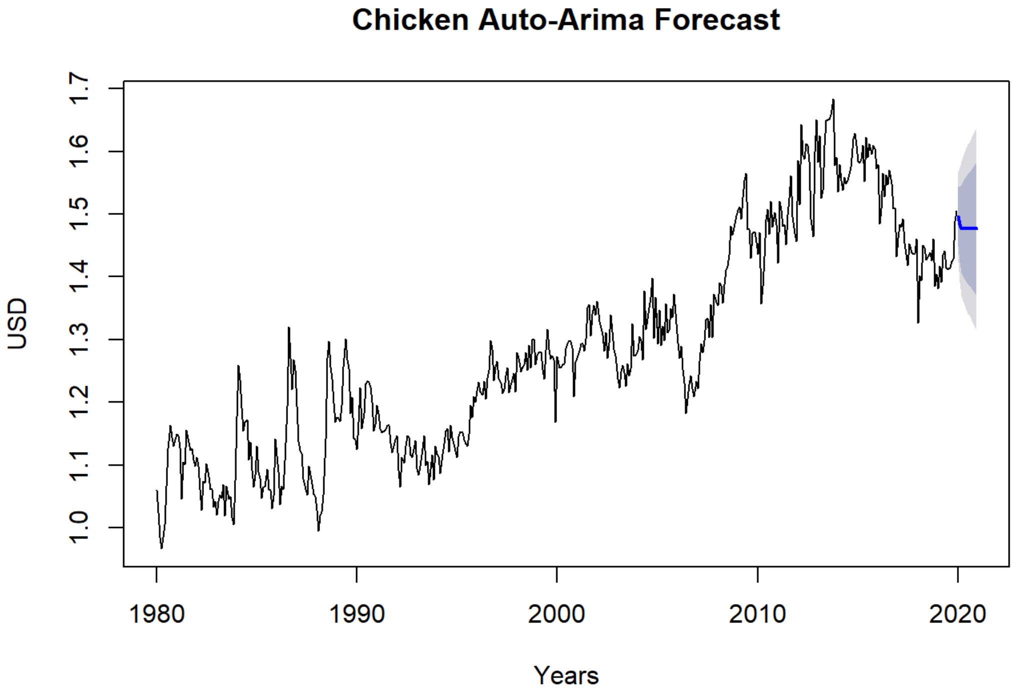


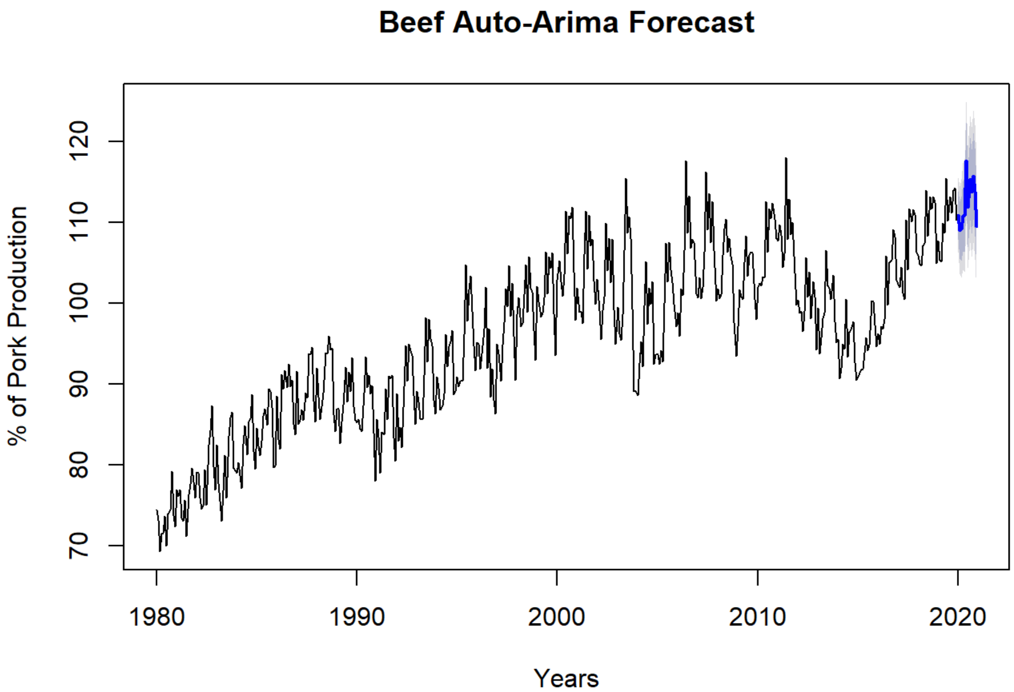

```
##                     ME     RMSE      MAE        MPE     MAPE      MASE         ACF1
## Training set 0.1099518 2.597707 2.021747 0.03823115 2.134383 0.5809309 -0.008322337
```

j.

Before doing the forecast, we need to prepare each model fit for the data, and we need each AutoArima, ETS, and Holt-Winters model fit for both the Chicken and Beef data.

Below are the Autoarima forecasts for both Chicken and Beef data:

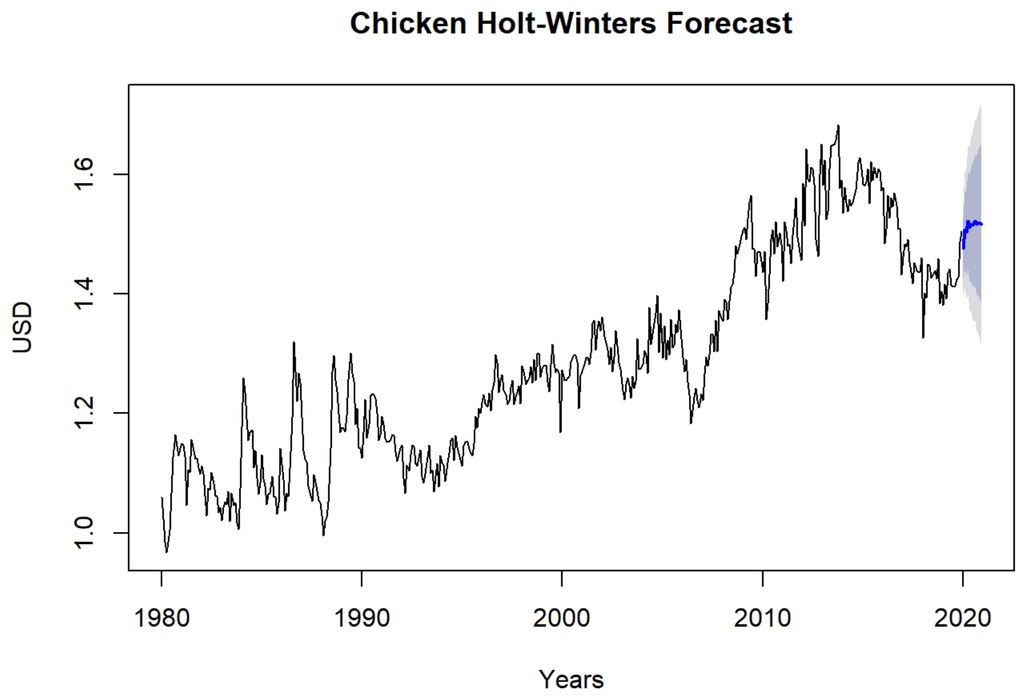


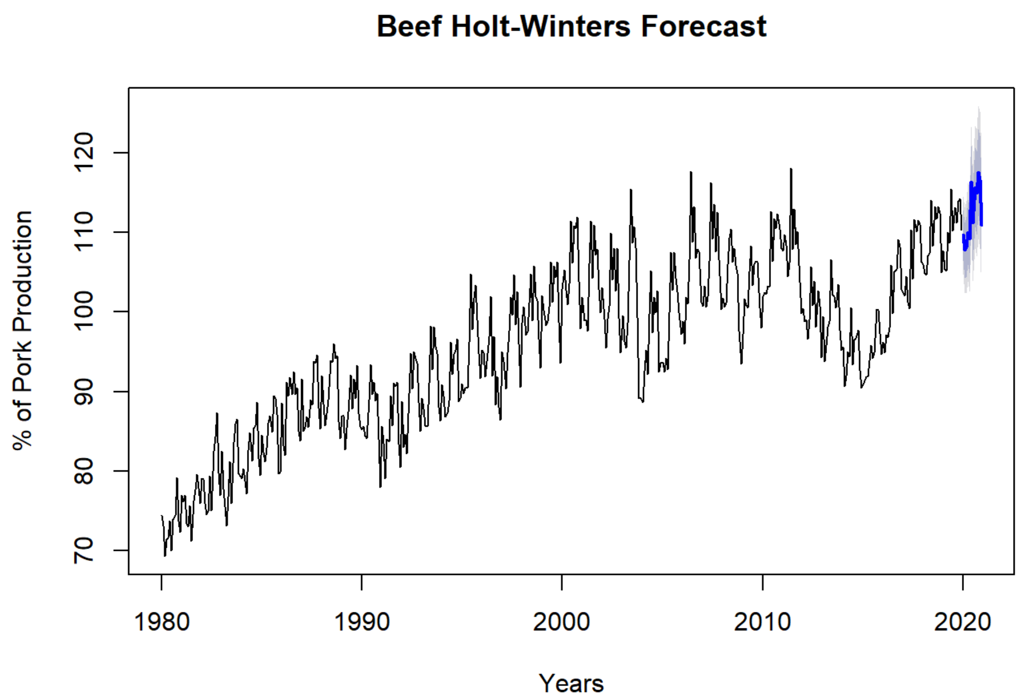



```
##                       ME       RMSE       MAE        MPE     MAPE      MASE        ACF1
## Training set 0.001569703 0.03660246 0.0272018 0.06577892 2.124362 0.427914 0.00484088
```

```
##                      ME     RMSE      MAE         MPE     MAPE      MASE         ACF1
## Training set 0.03036151 2.299271 1.763083 0.006314319 1.849402 0.5066062 -0.002691989
```

Below are the Holt-Winters forecasts for both the Chicken and Beef data:

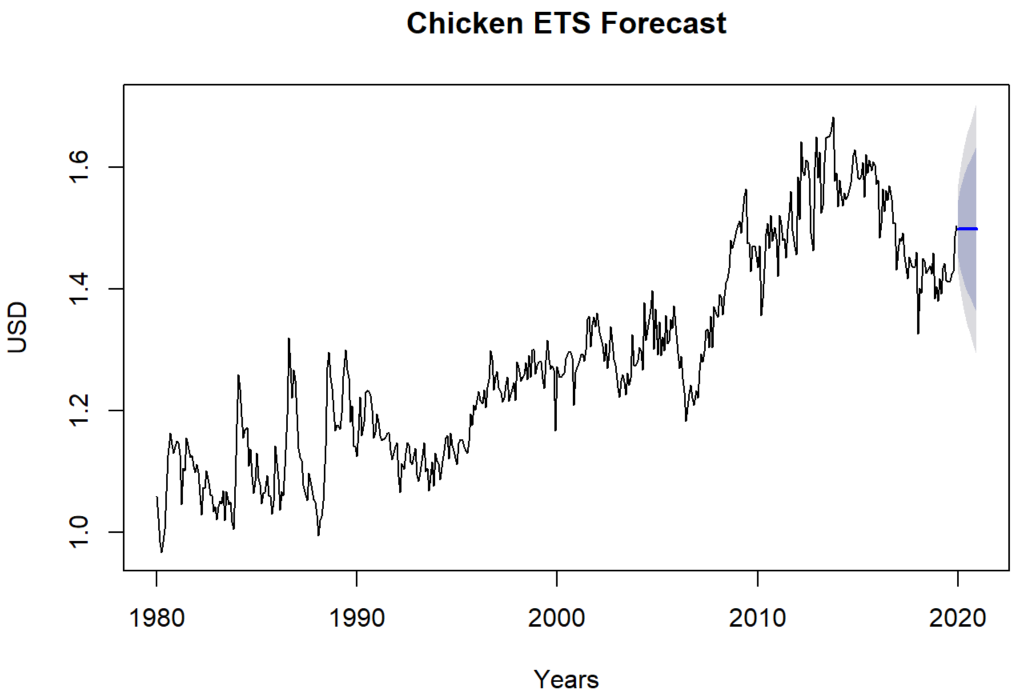


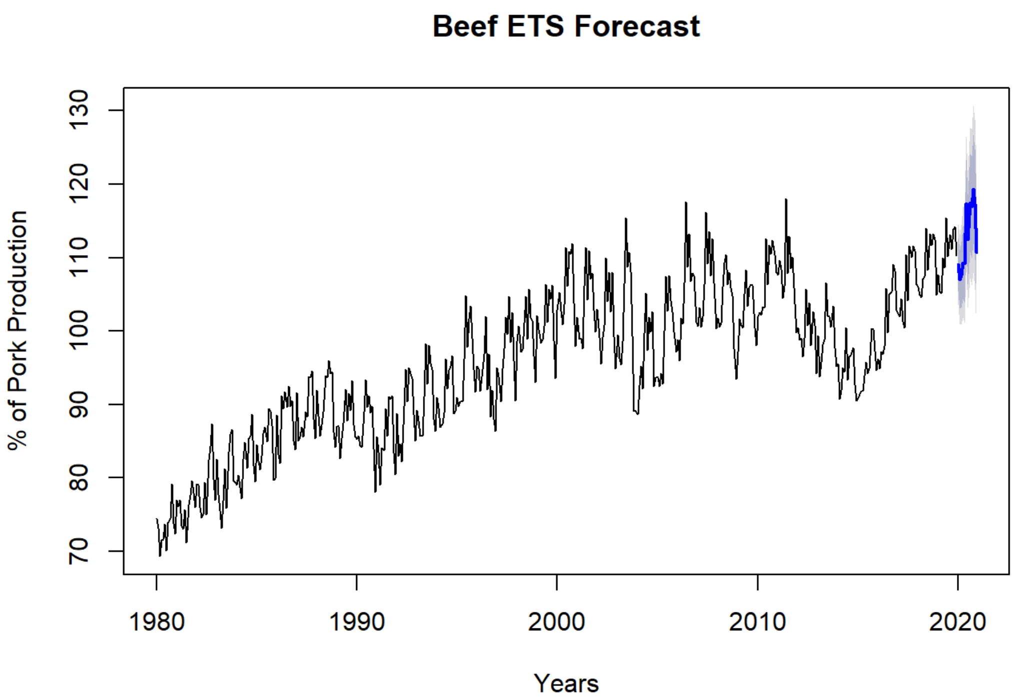



```
##                        ME       RMSE        MAE        MPE     MAPE      MASE       ACF1
## Training set -0.002508601 0.03972812 0.03013484 -0.2511285 2.364465 0.4740539 0.06531961
```

```
##                      ME     RMSE     MAE        MPE     MAPE      MASE      ACF1
## Training set -0.1468704 2.478606 1.89811 -0.1962871 1.990644 0.5454052 0.1832855
```

Below are the ETS forecasts for both Chicken and Beef data:





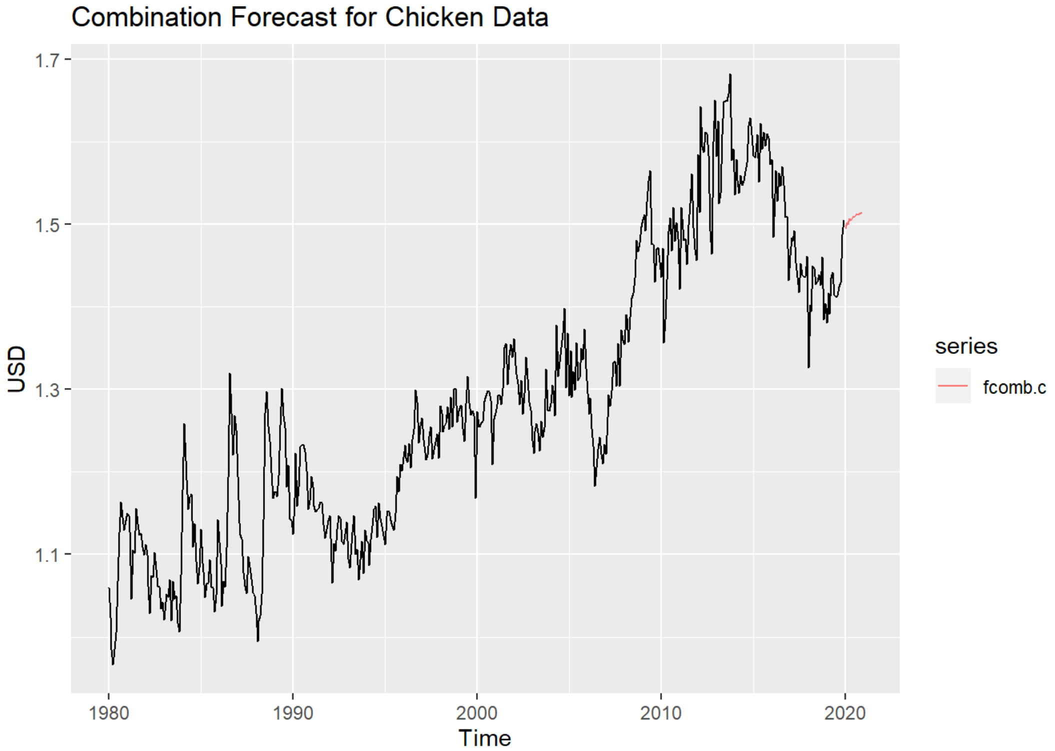

```
##                       ME       RMSE        MAE        MPE    MAPE      MASE        ACF1
## Training set 0.001166513 0.03725769 0.02778719 0.04192302 2.17316 0.4371228 0.005847022
```

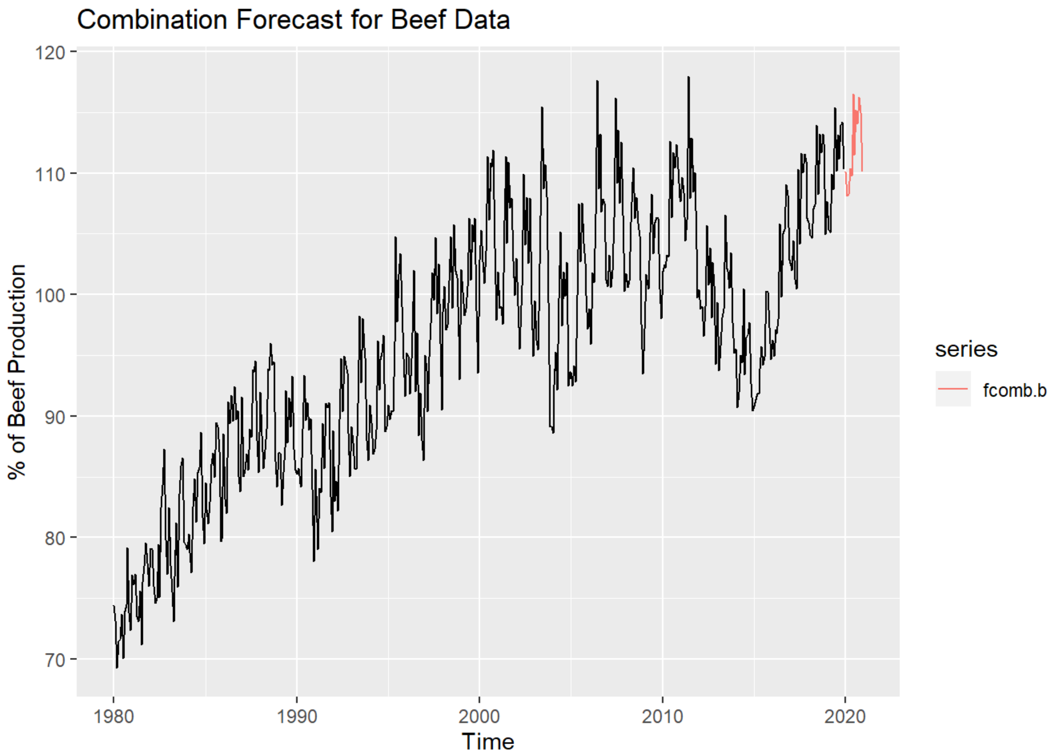
```
##                       ME     RMSE      MAE         MPE     MAPE      MASE    ACF1
## Training set -0.05065987 2.422001 1.872273 -0.08780547 1.965051 0.5379809 0.12667
```

using the `accuracy()` function, we know that the MAPE from the Beef model has a value of 2.134383, while from the Autoarima, Holt-Winters, and ETS have 1.849402, 1.990644, and 1.965051. judging from the MAPE alone, the best performing forecast comes from the Autoarima model, since it has the lowest value of MAPE among the three other forecast models.

k.

For the purpose of this question, we chose to use take the average of the 4 forecasts combined in combining our forecast for both datasets. In doing so, we take the average of each forecast, sum them, and divide them by 4, which is the amount of forecast that we are going to combine: S-ARMA(p,q), AutoArima, ETS, and Holt-Winters.






```
##                 MAPE of Chicken MAPE of Beef
## Auto-Arima          0.019265514  0.006250020
## Holt-Winters        0.005232833  0.005554077
## ETS                 0.006042746  0.012742726
## TSC Model           0.021706712  0.013997115
```

I.

As we can see from our VAR model result for equation `chicken`, only the first lag of chicken, or `chicken.l1`, that affected `chicken`. This is showed by the p-value of each lags, in which only `chicken.l1` that is statistically significant. `chicken.l2` and `chicken.l4` may have some effect, though they are less statistically significant then `chicken.l1`.

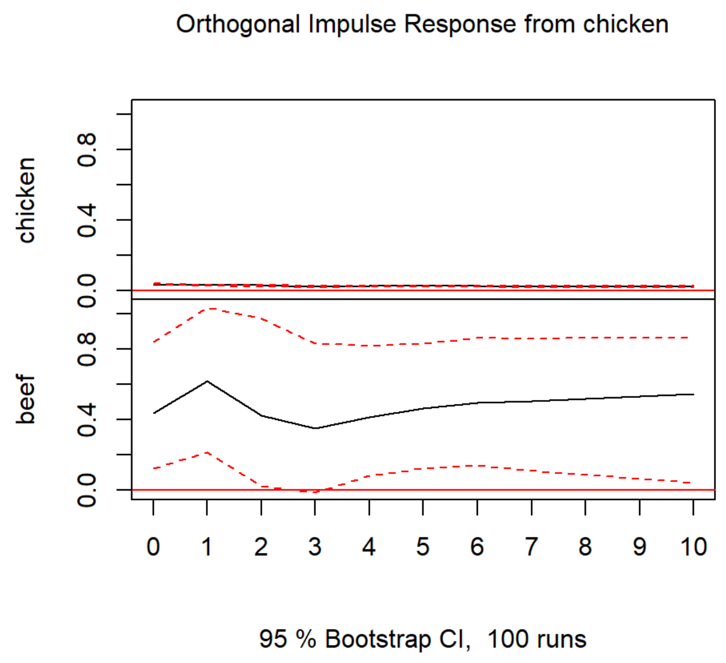
For our estimation results for equation `beef`, only the first and second lag of beef, `beef.l1` and `beef.l2` respectively, affects the changes in `beef`. Consistent with our previous argument, this is showed by the small p-value, hence their statistical significance.

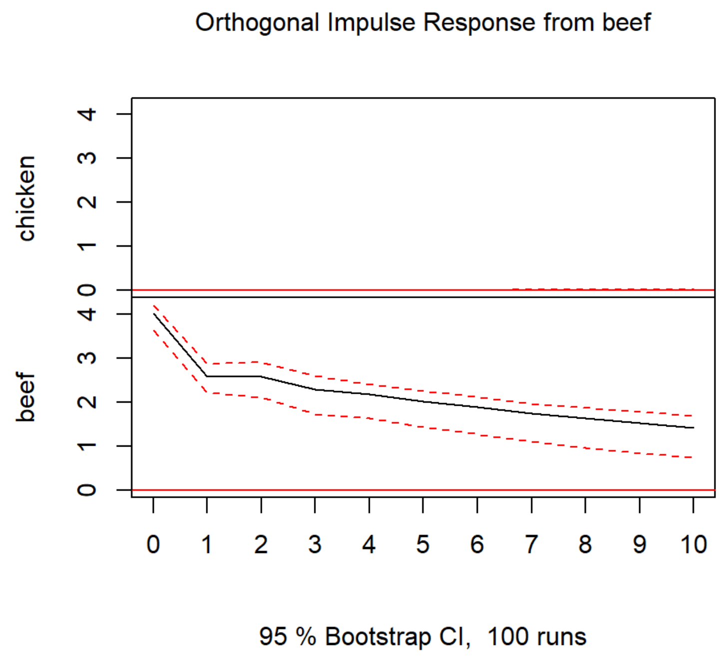
```
##
## VAR Estimation Results:
## =========================
## Endogenous variables: chicken, beef
## Deterministic variables: const
## Sample size: 476
## Log Likelihood: -432.912
## Roots of the characteristic polynomial:
## 0.9852 0.9142 0.5614 0.5022 0.5022 0.346 0.1736 0.1736
## Call:
## VAR(y = comb.data, p = 4)
##
##
## Estimation results for equation chicken:
## ========================================
## chicken = chicken.l1 + beef.l1 + chicken.l2 + beef.l2 + chicken.l3 + beef.l3 + chicken.l4 + beef.l4 + con
st
##
##              Estimate Std. Error t value Pr(>|t|)
## chicken.l1  0.7906313  0.0461626  17.127  < 2e-16 ***
## beef.l1     0.0001813  0.0004264   0.425  0.67085
## chicken.l2  0.1574018  0.0586711   2.683  0.00756 **
## beef.l2    -0.0004318  0.0005059  -0.853  0.39386
## chicken.l3 -0.1117649  0.0587631  -1.902  0.05779 .
## beef.l3     0.0004011  0.0005057   0.793  0.42803
## chicken.l4  0.1285910  0.0460328   2.793  0.00543 **
## beef.l4     0.0002245  0.0004265   0.526  0.59883
## const       0.0108401  0.0166936   0.649  0.51643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.03695 on 467 degrees of freedom
## Multiple R-Squared: 0.9531,  Adjusted R-squared: 0.9523
## F-statistic:  1186 on 8 and 467 DF,  p-value: < 2.2e-16
##
##
## Estimation results for equation beef:
## =====================================
## beef = chicken.l1 + beef.l1 + chicken.l2 + beef.l2 + chicken.l3 + beef.l3 + chicken.l4 + beef.l4 + const
##
##            Estimate Std. Error t value Pr(>|t|)
## chicken.l1  9.21117    5.04027   1.828  0.06826 .
## beef.l1     0.64170    0.04655  13.784  < 2e-16 ***
## chicken.l2 -9.24363    6.40601  -1.443  0.14970
## beef.l2     0.22967    0.05524   4.158 3.83e-05 ***
## chicken.l3 -1.77955    6.41606  -0.277  0.78163
## beef.l3     0.01330    0.05521   0.241  0.80979
## chicken.l4  4.83295    5.02610   0.962  0.33676
## beef.l4     0.01878    0.04657   0.403  0.68691
## const       5.45842    1.82269   2.995  0.00289 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 4.034 on 467 degrees of freedom
## Multiple R-Squared: 0.8493,  Adjusted R-squared: 0.8467
## F-statistic:   329 on 8 and 467 DF,  p-value: < 2.2e-16
##
##
##
## Covariance matrix of residuals:
##          chicken     beef
## chicken 0.001365  0.01612
## beef    0.016117 16.27476
##
## Correlation matrix of residuals:
##         chicken   beef
## chicken  1.0000 0.1081
## beef     0.1081 1.0000
```

m.

From the first IRF plot, which shows the response from the chicken data, it is clear to see that the own-variable response is very close to zero, while the effect on the cross-variable response is fairly constant at a level of around 0.4, without the plot surpassing the confidence band.

on the second IRF plot, which shows the response from the beef data towards beef itself and chicken, it is evident that the result is almost similar to the IRF plot of the chicken data. It seems that beef itself has a response on its own variable, in which it starts high and then decarys slowly to zero, whie the cross-variable response seems to be constant at a level of 0.





n.

From the granger causality tests on the effect of each beef and chicken on each other, we know that we fail to reject the null hypothesis that beef causes chicken and chicken causes beef by their high p-values. Thus from this test we conclude that both chicken and beef might not even affect each other's movements.

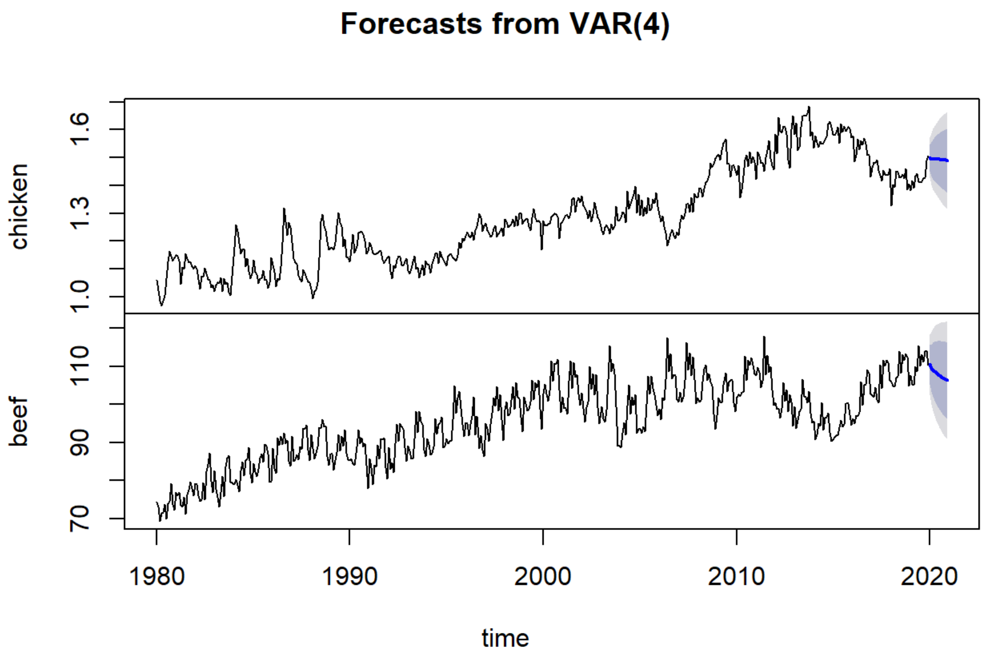
```
## Granger causality test
##
## Model 1: beef ~ Lags(beef, 1:4) + Lags(chicken, 1:4)
## Model 2: beef ~ Lags(beef, 1:4)
##   Res.Df Df      F Pr(>F)
## 1    467
## 2    471 -4 1.8785  0.113
```

```
## Granger causality test
##
## Model 1: chicken ~ Lags(chicken, 1:4) + Lags(beef, 1:4)
## Model 2: chicken ~ Lags(chicken, 1:4)
##   Res.Df Df      F Pr(>F)
## 1    467
## 2    471 -4 0.9817 0.4171
```

n.

The forecasts that are found using the VAR model is a lot less complex / detailed than that of the other models. For the chicken time series, it seems to be a completely flat line, devoid of the complications found in our model (that had it increasing), or the automatically generated by Holt-Winters (that had it slightly increasing the slowly leveling off). This is more in line with what we found in auto.arima and ETS however.

For the beef time series, our VAR model predicts something completely different. It predicts that the price of beef will begin to drop quite drastically, while the other models seem to predict an increase first then drop during baseline levels. This is due to the seasonality component, which is very strong in the beef component. Since the var doesn't take this into account, it just predicts a general decrease.

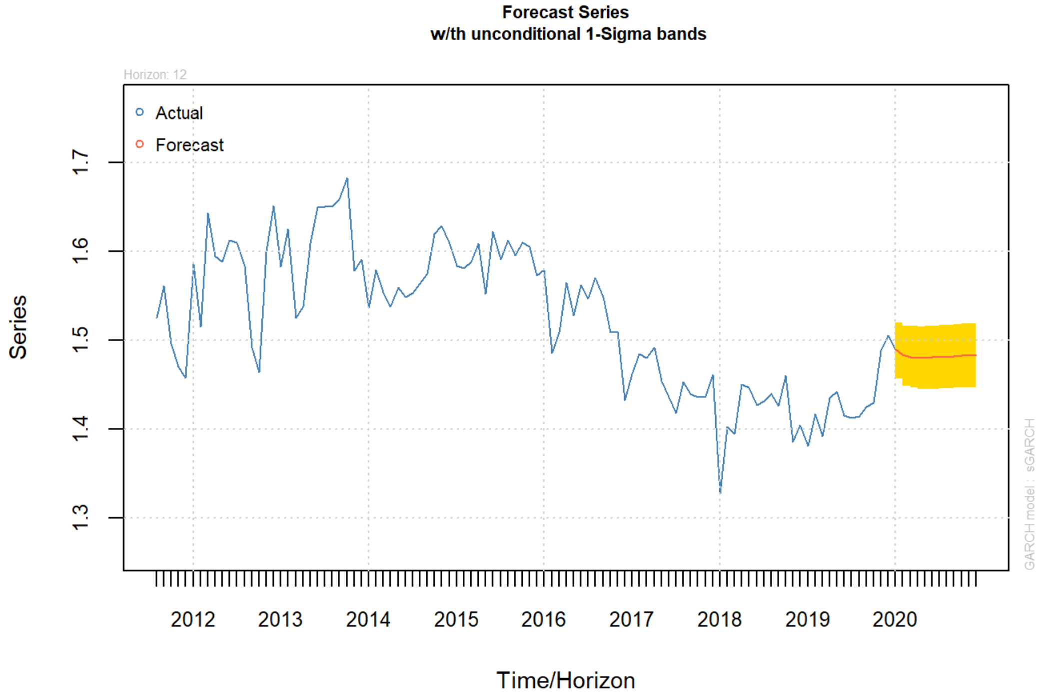


o.

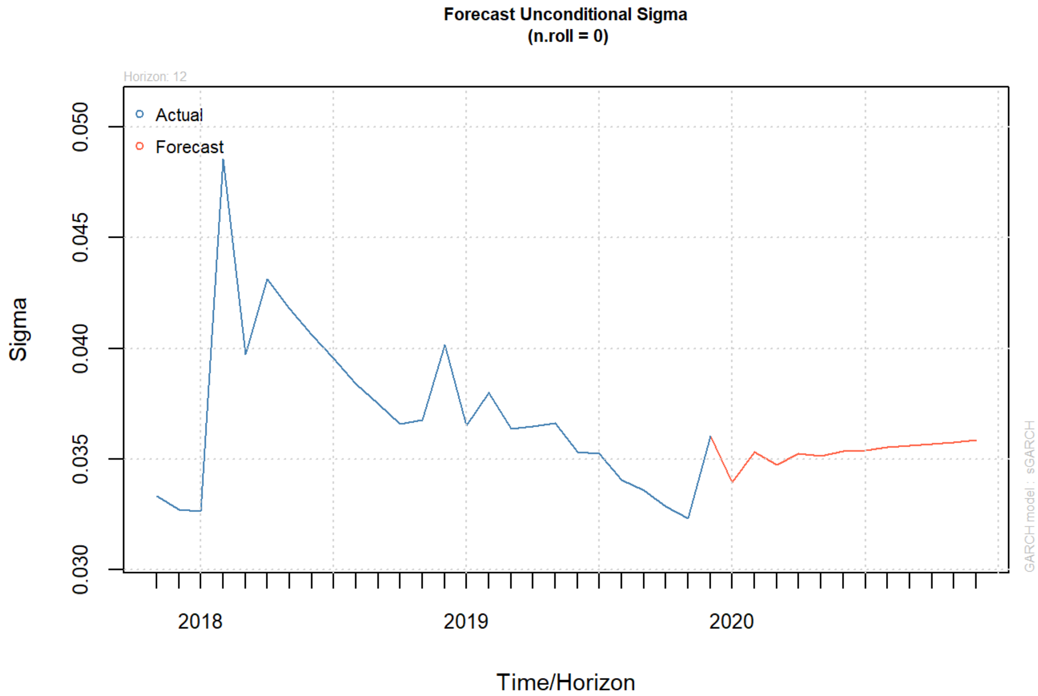
For the purpose of this question , we will choose our model's ARMA models. Note that the trend and seasonality component are not present.

Chicken:

```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```



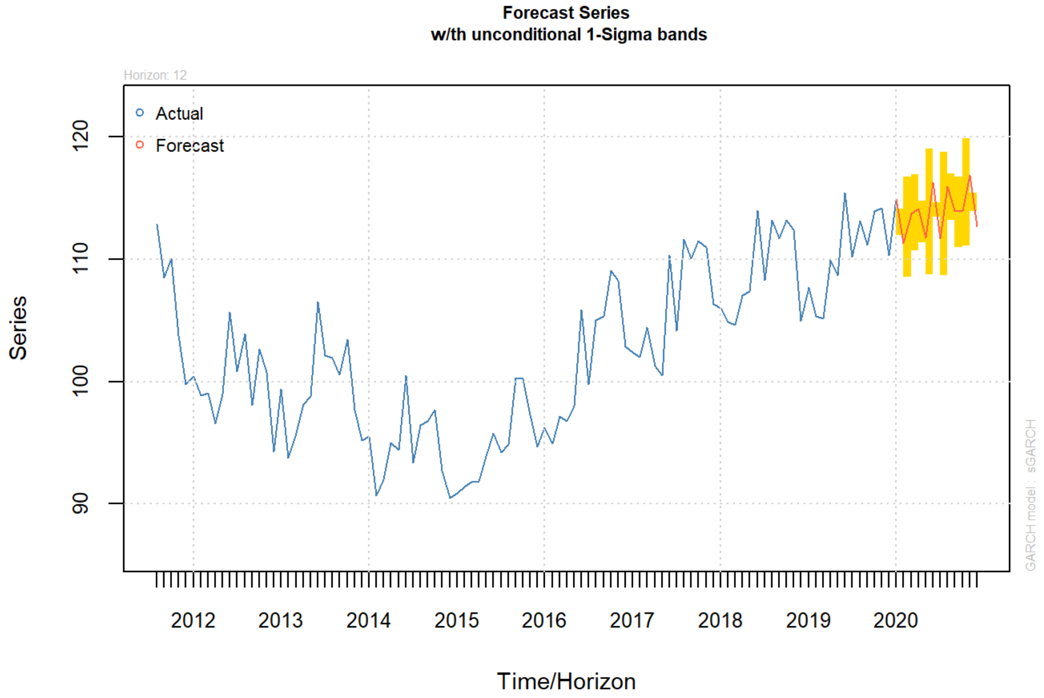

```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```



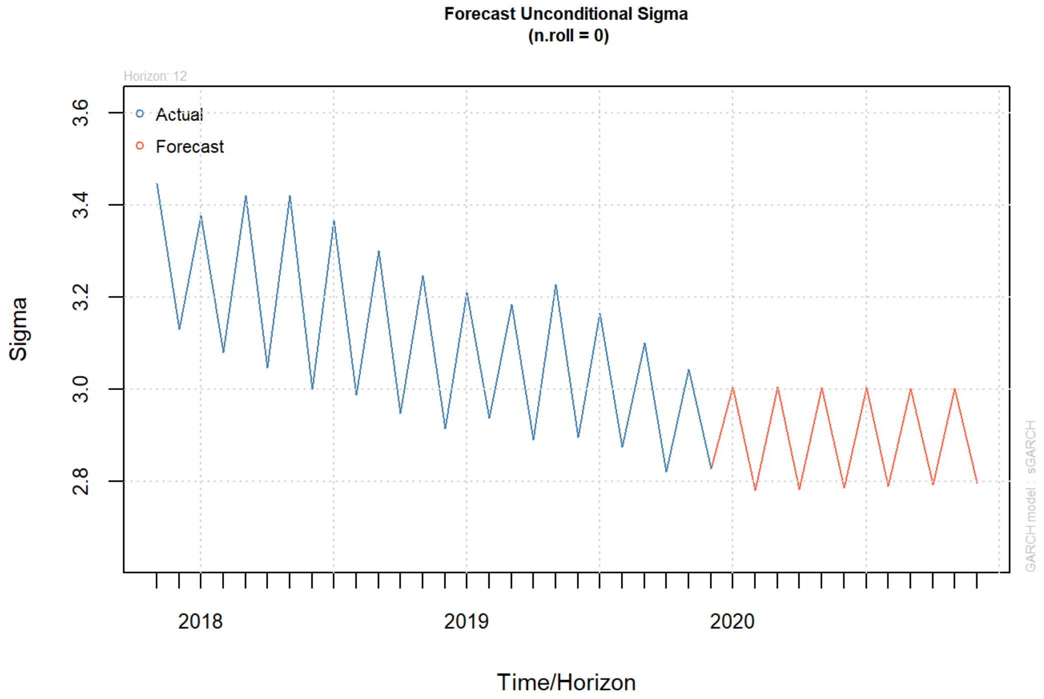

```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```

Beef:

```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```




```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```




```
##
## Make a plot selection (or 0 to exit):
##
## 1:   Time Series Prediction (unconditional)
## 2:   Time Series Prediction (rolling)
## 3:   Sigma Prediction (unconditional)
## 4:   Sigma Prediction (rolling)
```

Part III

In conclusion, we have come up with numerous different models for chicken leg price per pound and monthly beef production: a TSC model, ARIMA, Holt-Winter, ETS, VAR, and GARCH. We also combined the first four of these forecasts into a combination model. We note that many of these forecasts state different things, albeit some are roughly the same. For the Chicken dataset, our model interpreted a trend variable, and thus forecasted an increase. The other forecasts did not, and thus the forecast for chicken is either basically flat or a slight decrease and then a flattening out. The Beef dataset had much more varied forecasts. The first four forecasts had similar forecasts, where seasonality played a very large role in dictating the forecast, where it would spike up during the summer months and then die back down. VAR on the other hand predicted a downward trend, and GARCH predicted a relative leveling out with much less drastic peaks and valleys. We also note that there is little evidence of any causality or relation between the lags of chicken and beef with beef and chicken respectively.

There are multiple ways that we could improve upon our current models. We ran into a large amount of issues using our xreg and higher order ARMAs, and if these issues were to be solved (perhaps through more data) it is possible that we could get a better fit. We could also add more data / features that could explain away some of the peaks and valleys found in our models. Additionally, we could find datasets that are more cross-correlated, and thus would have a functional VAR that would help our forecasts.

Part IV

http://sub1.economagic.com/em-cgi/data.exe/frbg17/n311611t3b_ipnsa (dataset)

http://sub1.economagic.com/em-cgi/data.exe/blsap/apu0000706212 (dataset)

https://cran.r-project.org/web/packages/forecast/forecast.pdf

https://www.cmegroup.com/education/courses/introduction-to-agriculture/livestock/understanding-seasonality-livestock.html