ECON 102: First Assignment

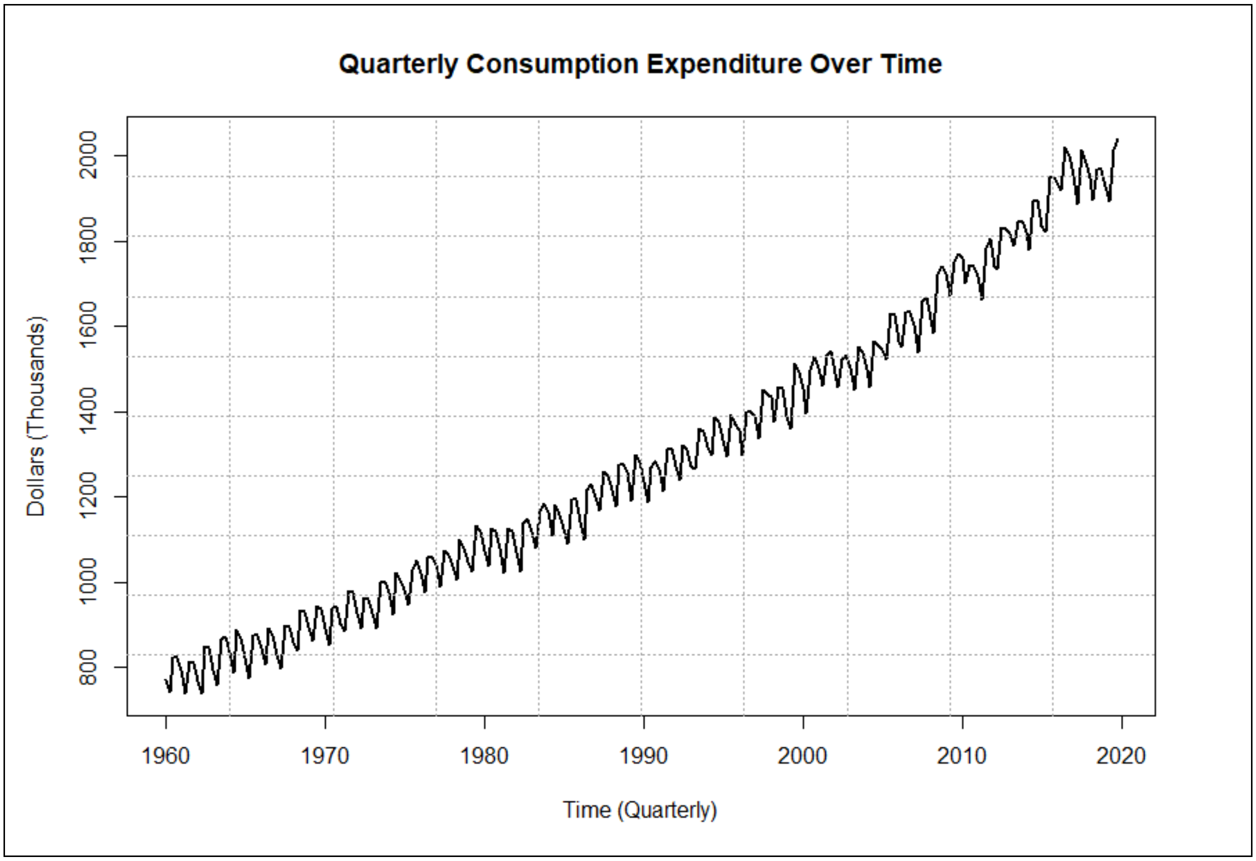
Saturday, October 17, 2020

Robert Knowles (20878339)

Dat7.csv

Part A: Visualization

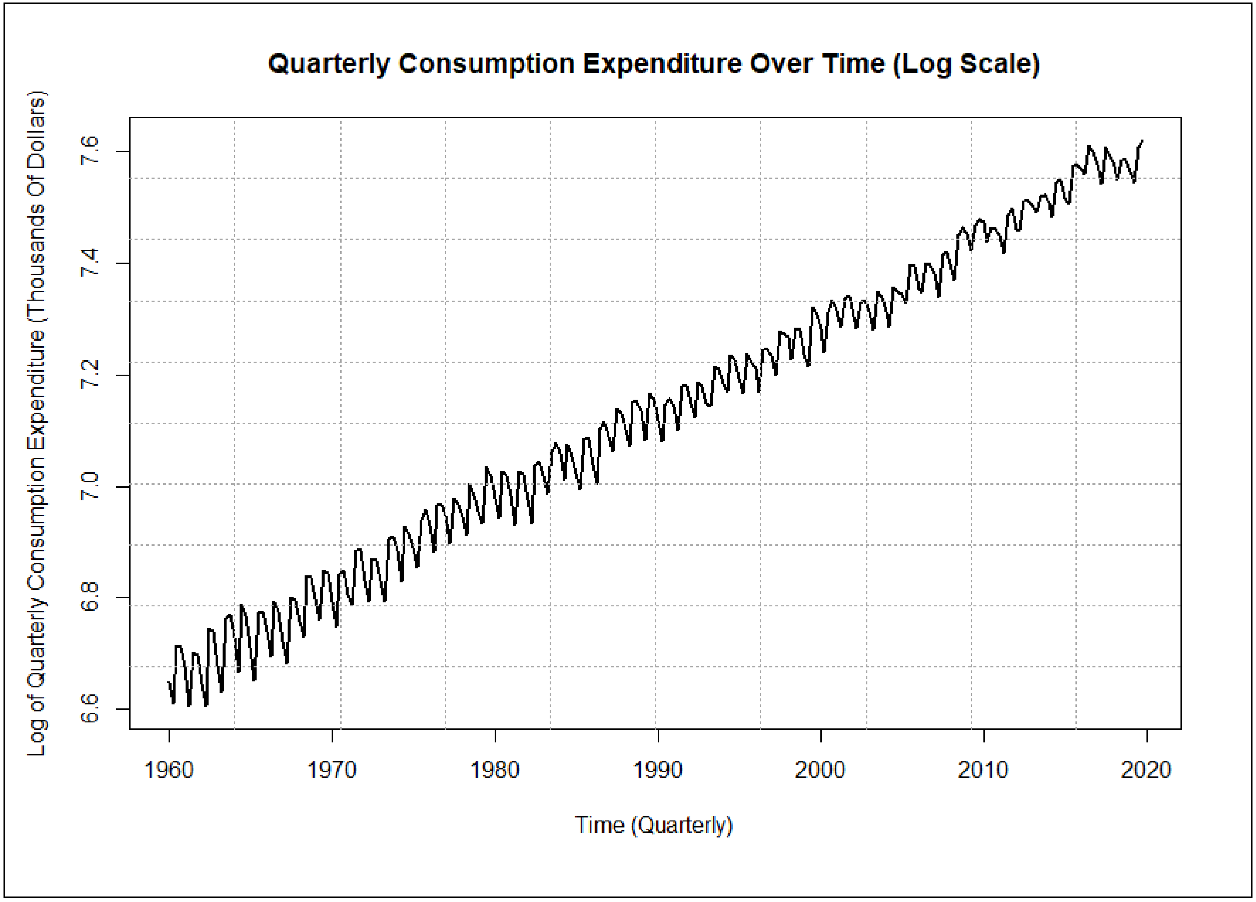
Q1) Plot the series using a line chart. Briefly describe what you see: Is it a positive or negative trend? Is the trend increasing? What kind of short-term fluctuations do you observe?



The upward slope of the graph shows that consumption expenditure is increasing; the average of the last four documented values (covering 2019) is 1969.25 thousand dollars and this is 249% greater than the 790.25 average for the first four documented value (covering 1960).

The plot also shows a large amount of volatility in the quarterly results, with peaks generally occurring in the 3rd or 4th quarter of every year and troughs in the 1st and 2nd quarter. This implies the existence of a seasonal component in the data. However, the fact that the peaks do not happen at uniform times (for example in 2010 the highest value was recorded in the first quarter) suggests that there may be a residual component in the data. Therefore, the short-term fluctuations reflect both a seasonal component and a residual component the data.

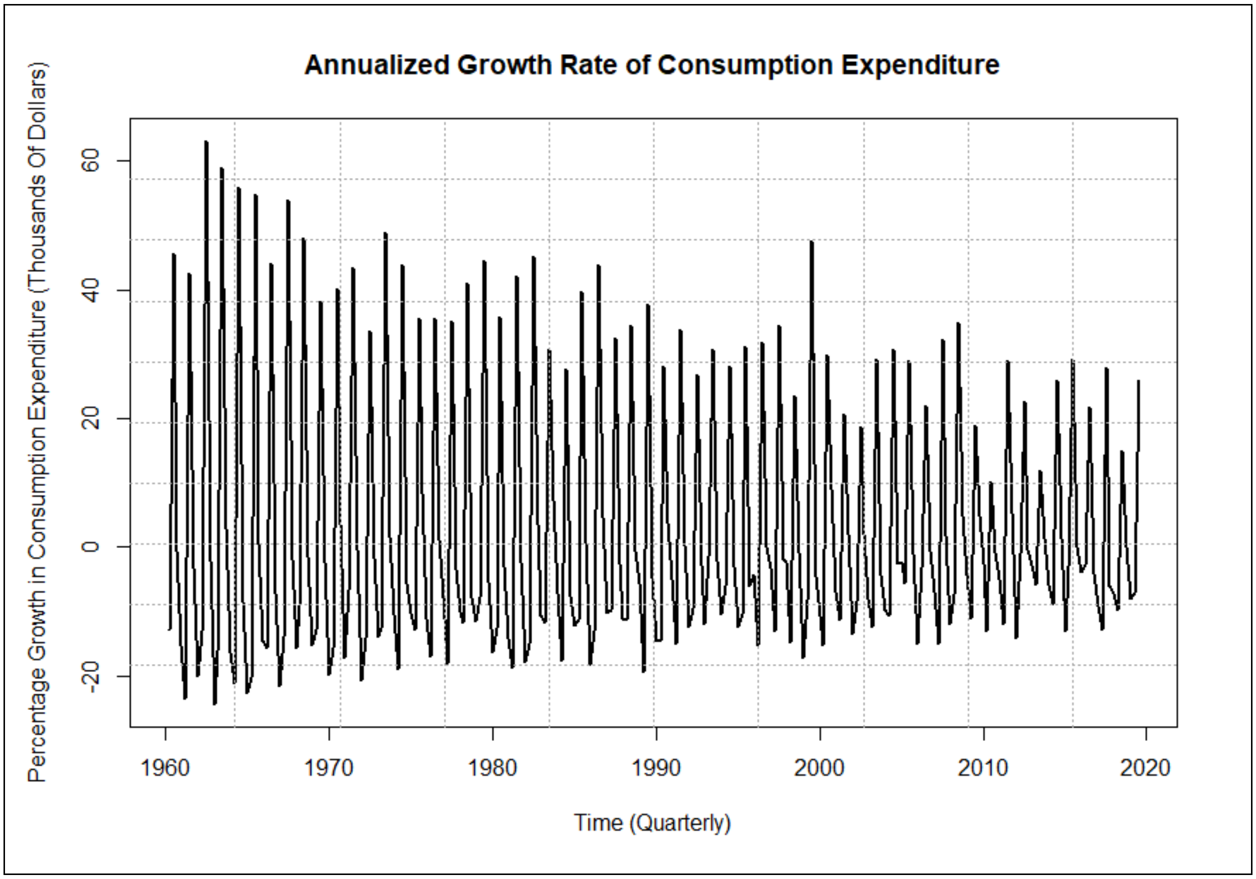
Q2) Answer the previous question using the log-scale. Can you tell if the growth rate is increasing or decreasing on average over the period?



On the log scale we can see that quarterly consumption expenditure is increasing, this means that the trend is positive, we can verify this as the average documented value for the first year in the data (the four quarters of 1960) is 6.67 and this is less than the 7.58 value for the final year of the data (representing the four quarters of 2019). The difference between each quarter and the equivalent quarter in the previous year stays roughly constant, meaning that the overall growth rate is staying the same over time.

It is hard to discern the growth rate trend accurately as there's high volatility from both the seasonal and residuals component of the data series. These components create the short-term fluctuations which makes it hard to decipher the movement of the trend in the log scale form.

Q3) To better see how the growth rate evolves through time, plot the annualized growth rate of consumption expenditure. Describe what you see. Is it constant on average?

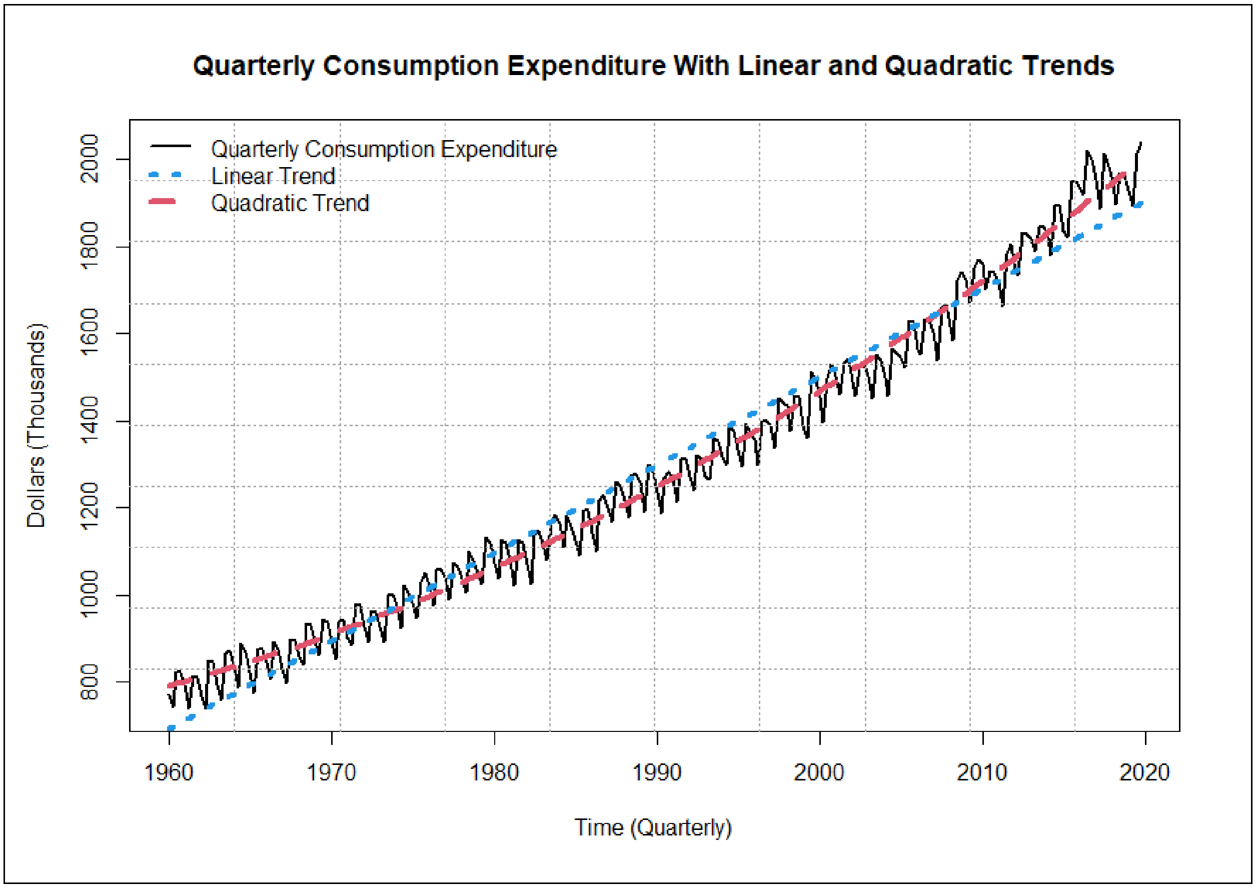


To begin we see a highly volatile trend that is centered around very roughly around 2.5%. We also notice that the peaks are much larger then the troughs, we also know that seasonally quarters 1 and 2 are most likely negative and quarters 3 and 4 are most likely positive (shown in the pervious graphs as quarters 1 and 2 were troughs and quarters 3 and 4 were peaks).

We also can see that both the peaks and troughs decrease in size as we go across the period, in 2019 they are roughly half the size of what they would have been in 1962. This indicates the volatility is decreasing overtime, this could be because as the trend becomes larger, all the short-term fluctuations become less significant in comparison.

Part B: Time Series Decomposition

Q1) Fit a linear and quadratic trends to your series. Then, create a line chart with your original series and the two trends. Which trend seems to best fit the series? Explain.



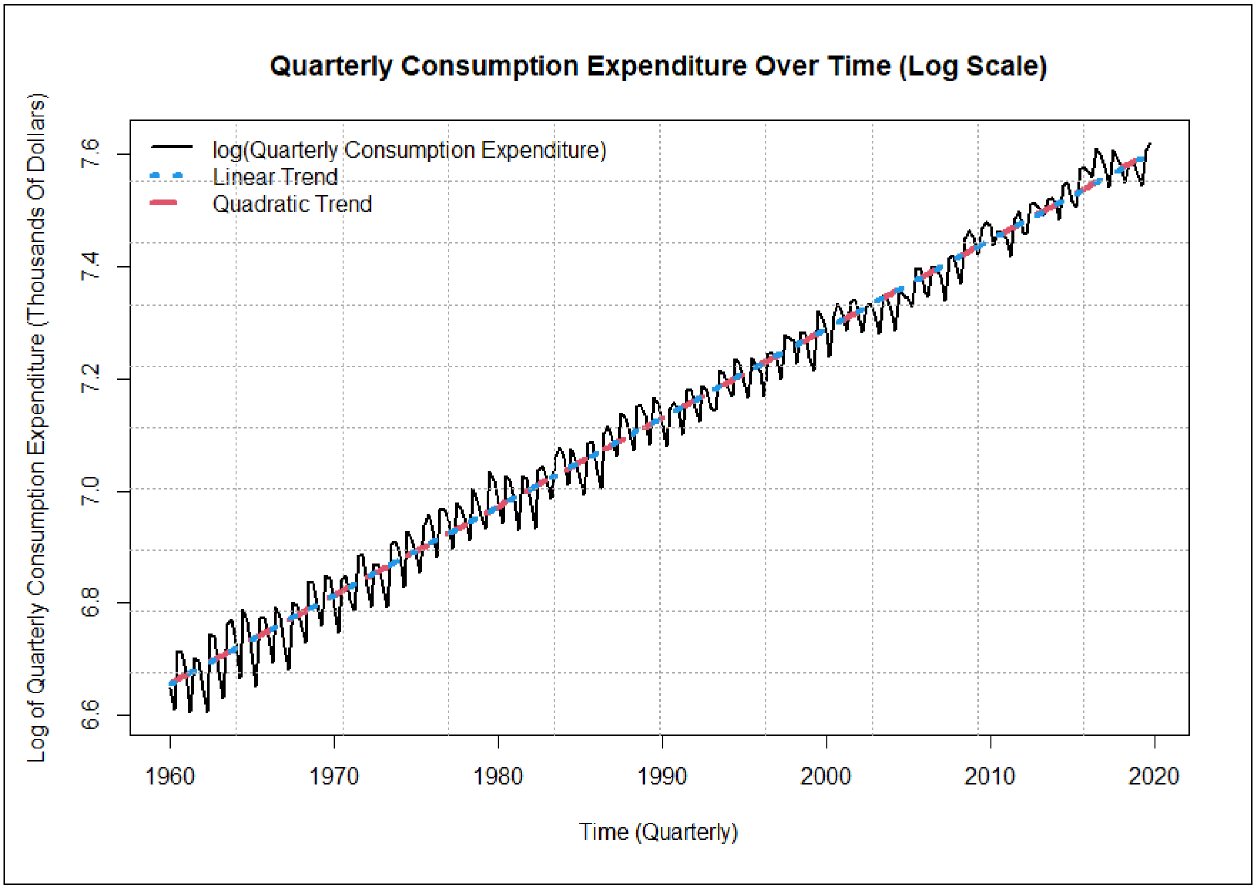
The quadratic trend line fits the data better as the calculated value falls within the range of the actual quarterly data series in every annual period, whereas the linear trend only in certain periods bisects the data series.

This finding suggests that the data series reflects the impact of a geometric/compounding influence (such as inflation) rather than a simple arithmetic increase (such as a constant rate of population growth). If the growth in the data were simply arithmetic, then the linear line would demonstrate the better fit.

Instead we see that at the start (1960) the linear trend is below the data series, near the middle (1993) the linear series is above the data series and near the end the linear trend is below the data series (2019).

Therefore, the quadratic trend is a better fit for our data.

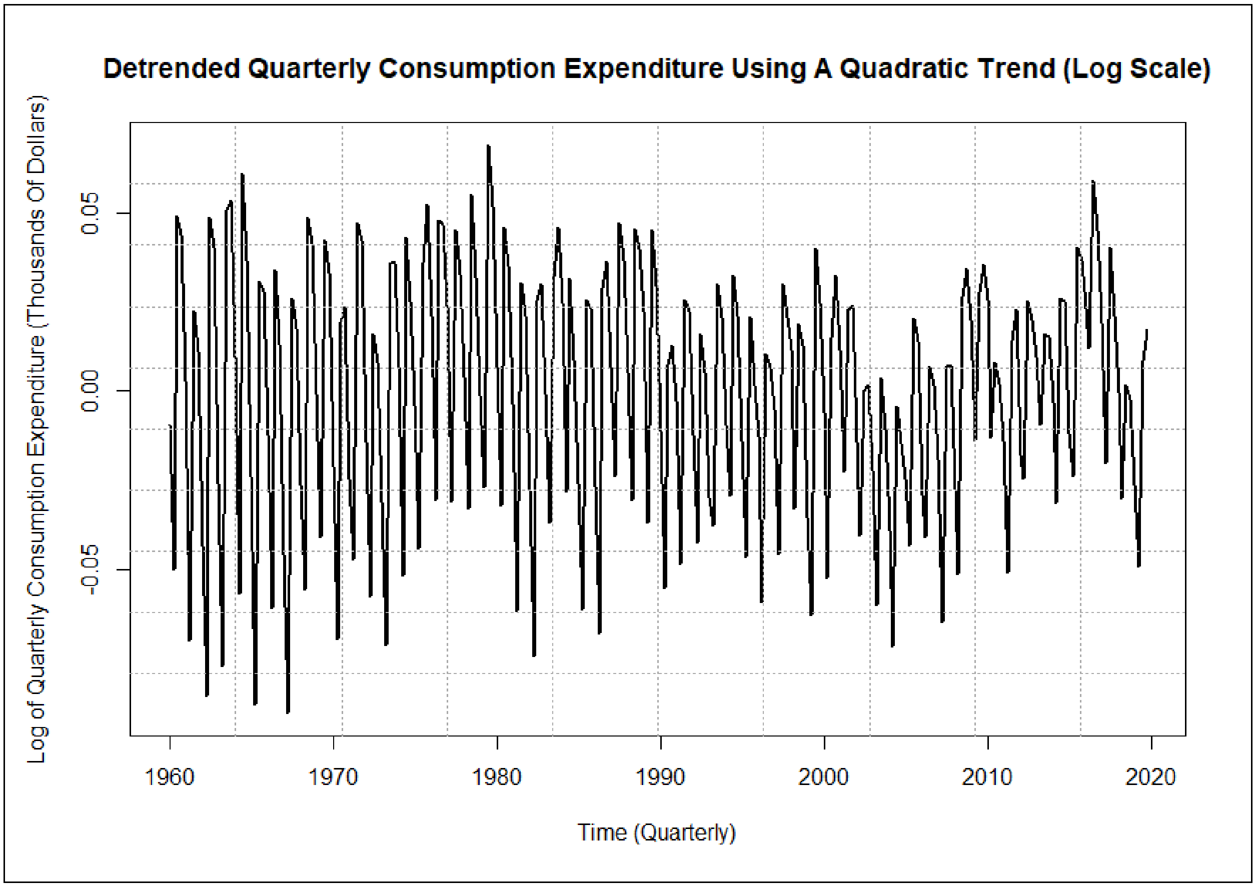
Q2) Answer the previous question for the series expressed in logs. Does the linear trend fit the log of the series better? Explain why.



In the log approximations the linear and quadratic trend lines both demonstrate a good fit with the actual data. This finding is consistent with a rate of growth that is both constant and compounding (such as a constant rate of inflation). In this scenario the linear log line and the quadratic log line both produce good fit since the rate of change is constant over time.

However, we know that the quadratic trend has an extra degree of movement (the x^2 term) in comparison to the linear trend, which means that it will have an equal deviation from each value then the linear trend so would give a clearer indication of change. Therefore, based on that logic I believe that the quadratic trend would fit the series better, although with this data set the difference is marginal.

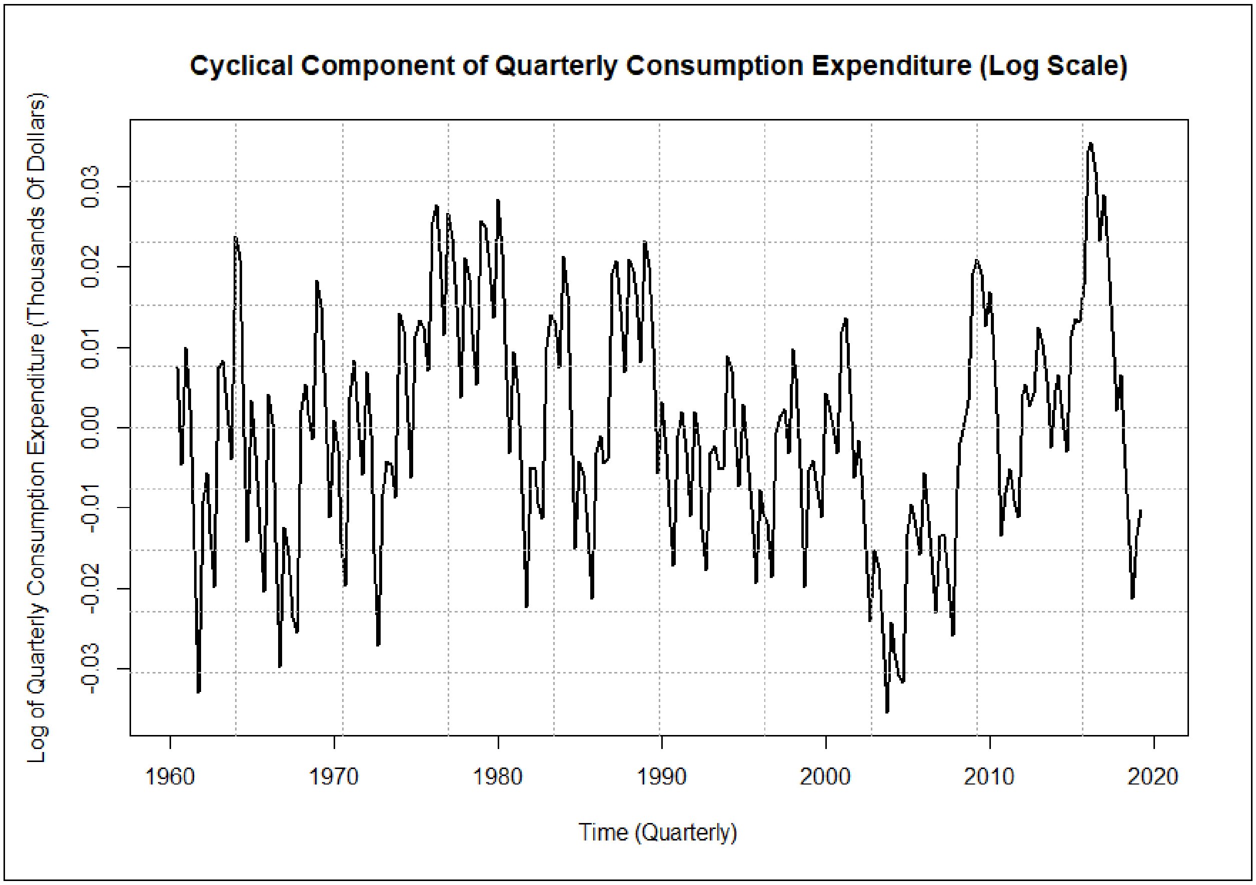
Q3) Plot the detrended series using the trend that best fit the series. Briefly describe what you see: Do you better detect short term fluctuations?



The detrended data series exhibits high volatility, however we see that on average the peaks get smaller over time and the troughs also get smaller over time. We can also see that the that the average of all observations is slightly very slightly positive ($2 * 10^{-15}$). We can also see that the peaks of quarter on quarter growth occur in the third and fourth quarter and the troughs occur in the first and second quarter.

The maximum observed value for the detrended data is 0.07 and the minimum is -0.09. Despite this, the overall growth trend is slightly positive because the peak values are less volatile and have a slightly higher average than the negative average for the troughs.

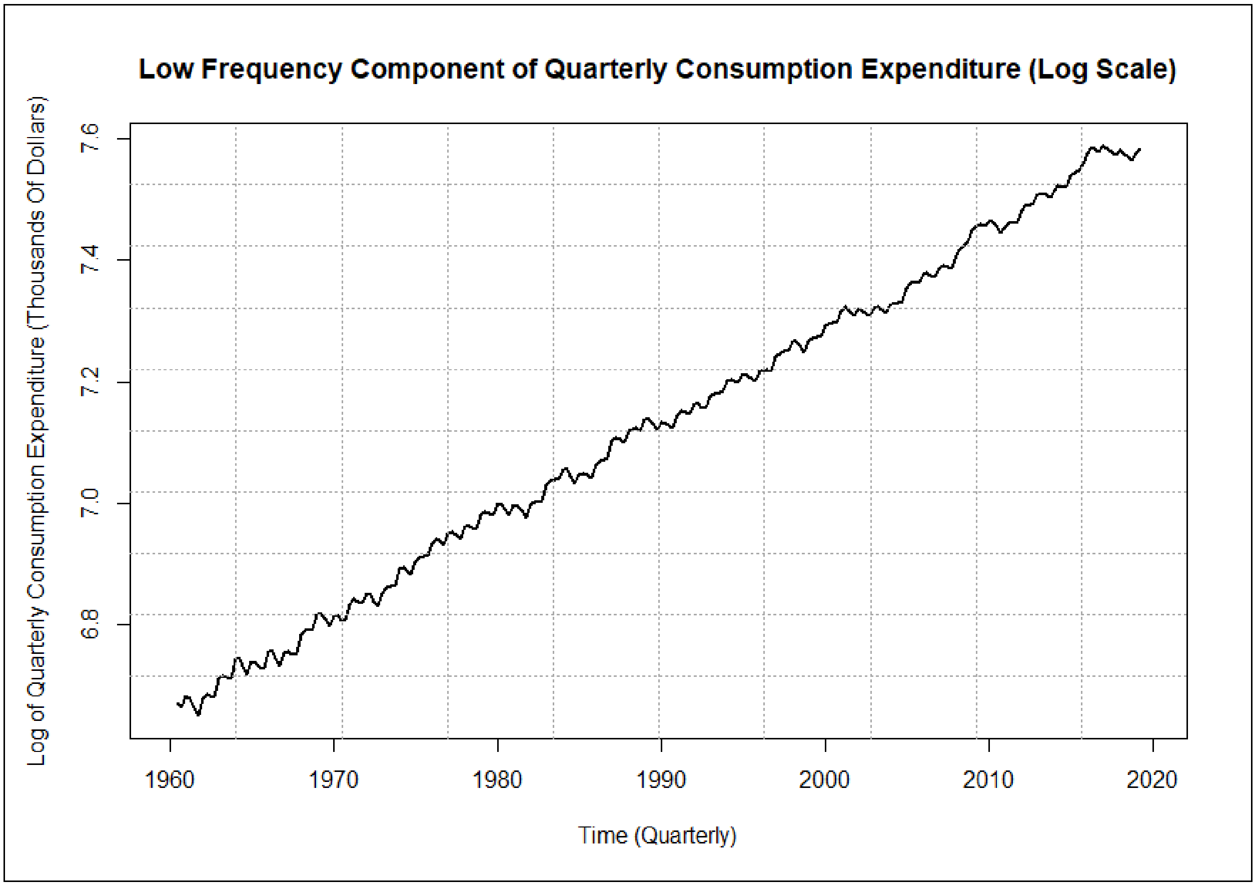
Q4) Using a moving average of order 5, compute the cyclical component of your series. Then, plot the cycle and briefly describe what you see: interpret the values of some peaks and troughs.



We can see that the cyclical component oscillates between being positive or negative roughly every 10 years. The lowest value is -0.0354, the highest value is 0.0354. Over the entire 60 years, the average for the cyclical component is essentially zero and because the peaks and troughs are bounded similarly this means that very little bias noise exists and that the cyclical trend will have a minimal net impact on the data series.

As well the 0.0354 is 0.5% of the average term in the log scale data series, which means that it while the cyclical component represents the biggest component in the detrended series it still is much smaller than the seasonal component.

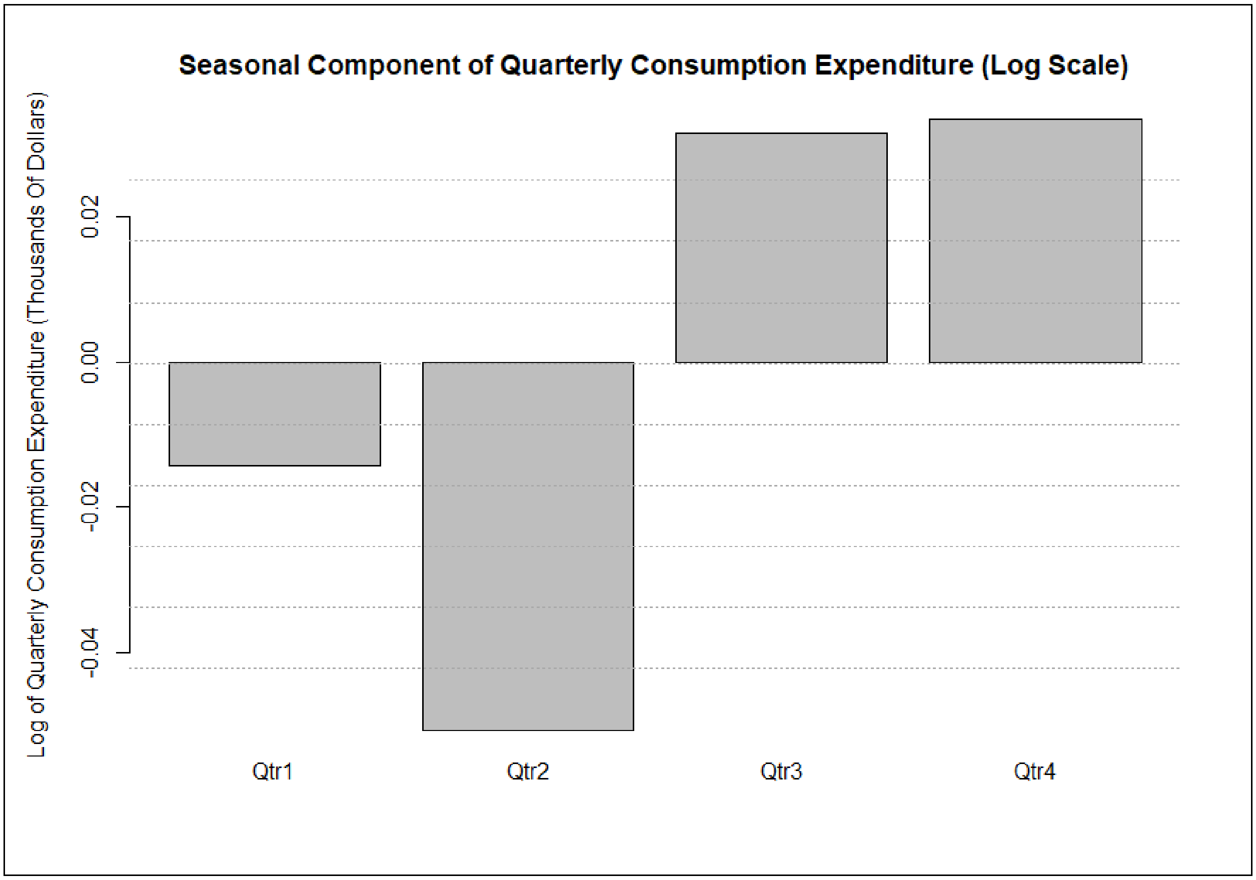
Q5) Plot the low frequency of your series and briefly describe what you see.



In comparison to the log data series we see a much smoother data trend which represents the combination of the cyclical component and the trend component. We see significant drops in 1962, 1972, 1982, 1990, 1998, 2011, which correspond roughly the cyclical trend having a frequency of once ever 10 years.

The average of the low frequency component only differs by $4.5 * 10^{-5}$ in comparison to the original log scale date series, this number is relatively very small and shows that the low frequency component of the data series on average mimics the data very well so there should be relatively little bias error in the residual component.

Q6) Compute the seasonal component and represent it on a bar chart (only the 4 quarters). Interpret the four seasonal values.



We see that the seasonal values are negative in the first and second quarter, but positive and roughly equal in the third and fourth quarter. Remember in the detrended series we could see that larger troughs existed but there were more consistent peaks. This is echoed in the data as the seasonal Q2 quarter (responsible for the large troughs) is much larger then the individual positive quarters (Q3 and Q4), however the average of Q3 and Q4 equals the average of Q1 and Q2, which shows that the seasonal peaks and troughs on average will be equal but the troughs will have more extreme values.