Noah Brannon

ECON 305

4/13/2022

**Final Integrated Project**

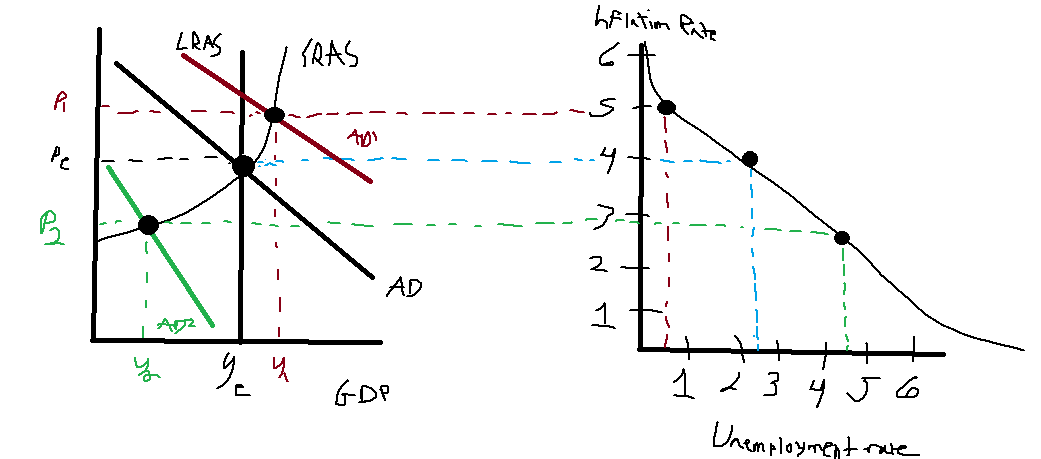
**Abstract**

This paper will be testing whether or not the rate of unemployment and the rate of inflation share a negative relationship in the United Kingdom. These are two variables that are important to understand as they are relevant today due to the COVID-19 pandemic recession. Many people might infer from recent events that once unemployment levels started to rebound or lower a consequence result was that inflation rates increased. The general reasoning behind how the two variables could be negatively related stems from the infamous Phillips Curve, which states that the unemployment rate and the inflation rate share a healthy inverse relationship. However, that is just a theoretical model and does not necessarily prove that the relationship exists in the real world. At first, using more basic statistical methods, especially when you look at the bivariate analysis of the two variables it appears that the two variables do share a weak, but negative relationship. However, as this paper dives into more advanced methods such as linear regression it appears that the two variables do not share any significant relationship, disproving the Phillips Curve’s hypothesis that the two variables share an inverse relationship.

**Introduction**

For this paper, I will be looking into whether or not there is a negative relationship between the unemployment rate and the inflation rate. These two variables are both very important indicators for understanding the general health of an economy. First, the unemployment rate measures the number or percentage of people who are looking for a job but are not currently working and are willing to work. The inflation rate measures how a value of a currency is decreasing, like the US dollar or the British Pound sterling, and in return the general level of prices for goods and services rise as a consequence. Both of these variables are very relevant in the present. The COVID-19 pandemic recession has brought renewed relevance to both of these terms. The unemployment rates in countries around the world rose in the early stage of the pandemic to numbers we have not seen since the Great Depression but started to recover back to pre-pandemic levels once economies started to open back up. Inflation had a different path during the pandemic, at first, inflation rates remained relatively low around the world, but as time went on several factors such as supply chain disruptions caused the inflation rate to rise. To one looking from the outside, it might look like these two variables are negatively related, meaning when one variable’s value goes up, the other goes down. It appears that once the unemployment rates started to lower, the inflation rates started to increase. Perhaps this was true for this instance, but is this a trend that has been there throughout recent history? We will address that in this paper.

**Theory**

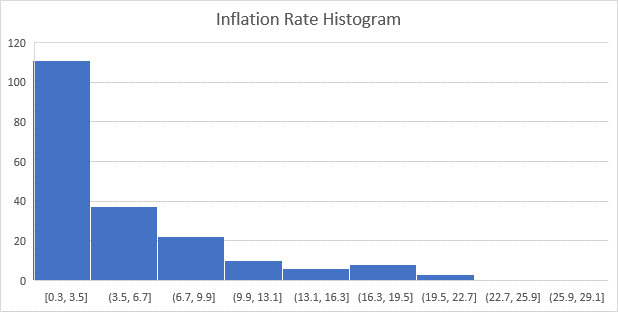
To explain how these two variables could potentially be related we can use a theoretical model which is known as the Phillips Curve, along with an AD-AS graph. The Phillips Curve’s logic is that inflation and unemployment have an inverse or negative relationship. The idea is that when there is general economic growth, inflation will typically rise as more people will have more money to spend which in return makes more money for businesses which then creates more jobs, lowering the unemployment rate. In this model, economic growth or recession will be marked by the change in the value of real GDP in an AD-AS model. The AD-AS model or the aggregate demand-aggregate supply model explains output (real GDP) and price levels through the relationship of AD and AS. As aggregate demand increases so will the short-run aggregate supply which then increases the real GDP and price levels and vice versa as the aggregate demand decreases, the short-run aggregate supply will decrease, and then the real GDP and price levels will decrease. So, to put both of these models together, if there is an increase in aggregate demand, aggregate supply will also increase, which increases the real GDP and price levels (this is inflation), as businesses will have more money to invest, they will then create more jobs, which lowers the unemployment rate. If you look below, you will see that I just explained the red line trend throughout both models. 

**Methodology**

To see if there is any relationship between the two variables we will first have to get an operational definition for both of the variables. The inflation rate in this project will be the dependent variable or the Y variable, which means the inflation rate values will change as a result of the independent variable or the X variable’s value, the unemployment rate changing. For both variables, I have collected data for both the unemployment rate and the inflation rate of the United Kingdom in each quarter from 1971-2020. This data is from FRED or the Federal Reserve Economic Data which is from the Federal Reserve Bank of St. Louis. This sample from FRED has given a sample size of exactly 200 observations for both variables, which is a considerably good size for a project like this. Both variables would be considered to be time series data, as both variables are measured by taking measurements over time, in this case, the measurements will be taken every quarter of a year. Time series data allows us to see how factors influence other factors over periods, which is perfect for this analysis.

Descriptive Analysis

Performing a descriptive analysis of these two variables will start to begin to unravel whether or not these two variables have any statistical relationship between them. To begin we will look at the dependent variable, the inflation rate. Below are the descriptive statistics of the inflation rate. Table

Description automatically generated To start, when looking through this data we will start at the mean and the median. These two values will describe the central location of this sample. The mean in a data set is the average value of the sample and the median is the value in the data set where 50% of the sample occurs in values greater of and the other 50% of values are below the value of the median, or in other words the value in the exact middle of the data. The mean of this sample is 5.38 and the median is 2.89, while it might look looks like these two values are close, 148 observations of these samples lie in between the value of 0% and 6.7%, which can suggest that this data is not normally distributed and since the mean is greater than the median, we can infer that this data is positively skewed, or that the central location of the sample is somewhere on the actual left of the graph. This data is further reinforced that it is not a normal distribution when you look at the standard deviation of this sample. Standard Deviation tells us how dispersed the data is in comparison to the mean, the closer to 0 the standard deviation, the closer the data is to the mean. In this sample, the standard deviation is 5.35, which tells us the data tends to not be close to the mean. The sample variance tells us how each value to what degree differs from the mean. In this sample, the sample variance is 28.64, saying that these values tend to differentiate that much from the mean.The minimum (the lowest numerical value) and the maximum (the greatest numerical value) values of this data also show a pretty wide spread of data when the minimum value in this sample of 0.333 and the maximum of 26.51, especially when you consider the mean of the data to be 5.38. To see whether or not this data is normally distributed or is skewed right (positive skew) we can use a histogram. A histogram will show us a visual representation of the frequency distribution or the number of observations within several given intervals of the sample.  When looking at this histogram of the inflation rate of the United Kingdom, it appears from the histogram that this data is definitely skewed to one side of the graph. This graph confirms that this sample is positively skewed. Now we will do a descriptive analysis of the independent variable, the unemployment value. Below are the descriptive statistics of the independent value. Table

Description automatically generated We will do the same analysis for the independent variable as we did for the dependent variable. Here the mean (6.85) and the median (6.01) are both pretty close to each other, however when you see the minimum value of 3.43 and the maximum value of 11.86, there is a significant difference between the mean and the median. The standard deviation is 2.37, which shows the data is spread out from the mean, but not as much as the independent variables dispersion and the sample variance is 5.615 meaning the data differs that much from the mean. We could predict there will be a positive skew in this sample, but it will be less extreme than the independent variable. We will use a histogram to confirm these results. Chart, histogram

Description automatically generated

This histogram shows that the data does have a positive skew, but the data is much more evenly spread than the independent variable.

Next, we will do a bivariate analysis this will be the first attempt in this project to see whether or not the two variables share any relationship. There are three main parts to a bivariate analysis. The first thing we will do it to present a scatter graph of the Y variable vs the X variable. A scatter graph shows what kind of relationship if any exists between two variables. Below is the scatter graph. This scatter graph does show a negative relationship between the two variables, the trend line (the black solid line) has a slight negative lean. The scatter graph also has a negative coefficient (-0.2317) which would indicate that the two variables have a negative relationship. This does support the Phillips Curve’s hypothesis that says the inflation rate and the unemployment rate share an inverse relationship. The next step in the bivariate analysis is to gather the sample covariance. Sample covariance will tell us how the two variables change together, or in other words, if the two variables are dependent on each other. If you have a negative covariance then you can infer that when your value goes up, the other goes down. Below is the covariance (highlighted in yellow) between the two variables. A picture containing text

Description automatically generated This test shows that the two variables have a negative covariance which could indicate further that these two variables share a negative relationship. The next and last step in the bivariate analysis is to get the sample correlation. The sample correlation will tell us whether or to what degree the two variables are related. These values are limited within the range of -1 to 1, with a -1 meaning these variables are in a perfect inverse relationship and 1 meaning that the two variables share a perfect positive relationship, meaning both variables move in the same direction (when one variable goes up, the other does as well). A picture containing graphical user interface

Description automatically generated

This correlation test that the two variables have a correlation of -.103. This does show there is a negative relationship, albeit a weak relationship between the two variables. This also supports the Phillips Curve’s hypothesis that the two variables share a negative relationship. All three steps of the bivariate analysis showed that the inflation rate and the unemployment rate showed that the two variables might have some kind of negative relationship, however, more will be done to conclude that this is true.

**Inferential Analysis**

This part of the project will start to hammer down whether or not the two variables share any kind of relationship, as we established earlier, we expect the relationship to be negative as predicted by the Phillips Curve. In this part of the project, we will first get a point and interval estimate and interpret them, and then conduct a hypothesis test for both variables. We will start with the Y-variable first. To begin the point estimate of these samples are the sample means. It was established that the mean of the Y-variable was 5.38, so that is the point estimate as well. To calculate the interval estimates, we will take the point estimate of 5.38 and add and subtract the t-stat of what confidence level you choose, and then multiply that value by the standard deviation divided by the square root of the number of occurrences in the sample (known as n). The first confidence interval will be at 95%, which will make the t-stat, 1.96, after plugging that value into the formula you will find that the confidence level value is 0.7463. You then add the CI value to the point estimate of 5.38 to find that the upper confidence interval is 6.12 and to find the lower confidence interval you take the point estimate of 5.38 and subtract the CI value of .7463 to that it is 4.63. This is good as the mean that was estimated at 5.38 falls in between these values. Table

Description automatically generated What this means is that we are 95% confident that the true population mean of the sample is located in between the values of 4.63 and 6.12, this is good to know so we can be confident that the data is reliable for estimation. We can also do a 99% confidence level for this sample, the only thing that will change is the t-Stat in the formula, which will change to 2.576. After plugging the new t-Stat into the formula, the confidence level value is 0.984. We take that value and add and subtract it to the point estimate of 5.38. The upper-level confidence level is 6.37 and the lower confidence level is 4.398, the calculated mean falls in between these values as well. Chart

Description automatically generated with medium confidence This means we are 99% confident that the true population means falls in between the values of 4.398 and 6.37. Now we will do the confidence levels for the X variable or the independent variable. For the 95% confidence interval, a t-Stat of 1.96 will be used. We would find that the confidence level value will be 0.3304. We will then find the upper-level CI to be 6.52 and the lower-level CI to be 7.18. This means we are 95% confident that the true population means falls in between the values of 6.52 and 7.18, with the calculated mean falling in those values. We can also do a 99% CI test, this test will have a t-Stat of 2.576 and will find the CI level to be 7.28 and the lower CI to be 6.41. This means with 99% confidence that the true population means falls in between the values of 6.41 and 7.28, with the calculated mean falling in those values. Next, we will do a hypothesis test of the Y-variable. This will tell us whether or not there is any kind of plausibility in the hypothesis. For this hypothesis test, the null hypothesis will be that the Bank of England has hit an inflation target of 3%, while the alternative hypothesis will be that they did not hit their target (any value that is not 3%). Below is the hypothesis test. Table

Description automatically generated So, for this test what we need to look at is the p-values. A p-value can range anywhere from 0 to 1, 0 means there is we would reject the null hypothesis, or in other words, the Bank of England did not meet its target of 3% and 1 would mean we could not reject the null hypothesis (this does not mean that they did achieve their target, but rather we cannot conclude that they did or not). So, when looking at the p-values, the left or lower-tailed test gave us a p-value of .9999 or 1, which would mean from this test we could not reject the null hypothesis. The other two tests, the upper tail test (right) and the two-tailed test both provided a p-value of 0. This would mean for both tests we can reject the null hypothesis that the Bank of England reached its inflation target of 3% (meaning the value was not 3%). We will also do a hypothesis test for the X variable as well. For this hypothesis test, the null hypothesis would be to see if the unemployment rate of the United Kingdom was greater than 5% (which is the natural rate of unemployment) during this period. Below is the hypothesis test of the X variable. Table

Description automatically generated From this hypothesis test, we have similar results to the Y-variable hypothesis test. The lower-tailed test (left) had a p-value of 1, meaning from that test we can not reject the null hypothesis. While the other two tests, the upper-tail (right) and the two-tailed tests both had a p-value of 0. This means for both of these tests we can reject the null hypothesis. Both of these hypothesis tests produced a result of where the left tail test (the negative side), could not reject the null hypothesis, which could correlate to the actual hypothesis of the project of whether or not the two variables are negatively related to each other. Lastly, we will run a simple regression between the two variables with the inflation rate (Y variable) as a function of the unemployment rate (X variable). This will tell us whether or not these two variables have any correlation with each other or not. Below is the Regression analysis. Graphical user interface, application, table

Description automatically generated To start, we will look at the coefficients, the t-Stat, and the p-values. We can see our coefficients are 6.96 and -.2317. We can plug that into a formula that will look a lot like a regular y=mx+b equation. This equation will look like y = 6.96 - .2317x. What this says for every increase of one percentage point in the unemployment rate the inflation rate will decrease by -.2317%. We could infer that the intercept of 6.96 is the unemployment rate when inflation is at 0%. However, this does not prove that there is any kind of significant relationship between the two variables. To figure out whether or not these two variables share a negative relationship we will need to conduct a test of the slope. So, to figure this out, the slope of these two variables cannot equal 0 if we want to prove any kind of relationship. The null hypothesis will be that the slope = 0 and the alternative hypothesis would be that the slope is less than 0, if the null hypothesis is true, then no linear relationship exists. Since this project is trying to find out whether or not the two variables have a negative relationship, we will do a left-tailed test and find that the rejection value will be if t > -1.65 (or 1.65 in absolute terms), and in this model the t-Stat is -1.45. From this result, we can not reject the null hypothesis that the slope = 0, saying that the two variables are not statistically related to each other. Another indicator that shows that we can not reject the null hypothesis is the p-value and measured that against the significance level of 5%. If the p-value is below 5% then we can consider the relationship to be significant. However, this p-value is .15 or 15%, much higher than the significance level of 5%. R squared in this model is .105, which shows that only about 10.5% of the variability of the inflation rate is explained by changes in the dependent variable again this does not support the hypothesis that the two variables share a negative relationship. Also, if we look at the normality probability plot of the residuals, we can see that the residuals have a skew, which says that the two variables do not have a linear relationship. The histogram of the residuals does not support the hypothesis either as there is a clear positive skew. The other residual plots also show that the data suffers from heteroscedasticity, or the unequal scatter of residuals. There is a general cone shape trend in the other graphs, which means these variables do not have a constant variance, which again proves these variables do not have a good inverse relationship. From all of these tests we can conclude that there is no significant statistical relationship between the unemployment rate and the inflation rate of the United Kingdom.

Conclusion

To conclude, it appeared early on in the research process that these two variables look like they could have some kind of statistical relationship between two, and the Phillips Curve seemed like it would add up. However, once we got into the more advanced statistical methods this proved to not be true. It appears in the data as time went on, especially in the present day the two numbers really did not affect each other at all. You can look during the 1970s and see that the Phillips Curve idea was actually not true, and inflation and unemployment rates were both high and increasing at the same time, a clear contradiction of the model’s hypothesis. For the overall project, the gathering of data was relatively easy, and it was nice working with variables that have a lot of relevance today (at least in my opinion). However, it is clear at the end of this project that there are other variables perhaps such as governmental intervention (such as the stimulus packages that were passed after the Great Recession, etc.) that could be affecting these variables, which would be a cool idea to explore if I could do so.

References

* *Consumer price index: Total all items for the United Kingdom*. FRED. (2022, February 15). Retrieved from https://fred.stlouisfed.org/series/CPALTT01GBQ657N
* *Unemployment rate: Aged 15-64: All persons for the United Kingdom*. FRED. (2022, March 10). Retrieved from https://fred.stlouisfed.org/series/LRUN64TTGBQ156S