Matthew Dahle

Econ 460 Project Report

Forecast of November 2018 Unemployment Rate for 20-24-Year-Old Men

**Description of Data and Time Series:** The economic variable I decided to forecast is the unemployment rate among 20-24-year-old men for the month of November. I found this particularly interesting because this is my age range and sex and I am about to enter the workforce after this spring, so I am curious to know the type of workforce I will be entering. I have obtained this data that dates back to January 1948 on the FRED website with the code LNS14000037. The publication releasing the new unemployment rate comes out on December 7, 2018 on the BLS website at 8:30 a.m. Historically, the unemployment rate for this age and sex started off at 7.5% and stands at 7.4% as of September 2018. Looking at the time series we can see a low point of 3.2% in August 1951, and a high point of 20.0% in April 2010. There seems to be a very slight increase over the time frame, with large fluctuation and swings. Taking another look reveals that this is impacted seasonally and cyclically, with unemployment rates returning to about a 7%-8% every 2-3 years. Historically known times such as the 1980’s oil crisis and 2008 housing market crash reveal high points in the unemployment rate for this group, which is expected. The current rate is very low and has been trending down since 2010, so it seems as though it may rise soon due to the business cycle. The not seasonally adjusted data for this time series reveals that unemployment typically hits a high point during the start of a new year in January and February, and low points in the summer and fall months. Luckily, the BLS provides seasonally adjusted data for this variable, so I was able to use that for my forecast without having to manipulate non-seasonal data.

**Forecasting Method:** For this forecast I decided to project the unemployment rate for the next 12 months, spanning from November 2018 to November 2019. The variables we are looking to forecast are point and 90% interval estimates for each month. I considered several types of models when choosing how I wanted to perform this task and looked at several things such as direct interval forecasts, iterated interval forecasts without trend, iterated interval forecast with trend, and looked at the best lags using AIC scores. First off, I made sure that using lags is even viable in the first place, by looking at the autocorrelation graph, which showed me this series has high autocorrelation up until about the 29th lag. I wished to determine how many lags would produce the best results for this regression, so I regressed unemployment rate on L(1/1), L(1/2), L(1/3), all the way up to L(1/12), and stored each estimate. I then compared each by using AIC scores, and determined that L(1/6) was the best amount of lags to use as the regressor. I decided that using an autoregressive 12 step iterated interval forecast was the method I wanted to use, but it was tough to choose between the model with a trend as a regressor and one without trend. Both models have similar regression outputs, with for the trend model being 0.9514 and the for the model with no trend being 0.9513. Both models produced very low robust standard errors, and the point forecasts themselves were close to each other. Ultimately, I chose to use the method without trend as a regressor, simply because looking at the p-value of the time variable showed that it was only 0.208, so this was not significant enough to produce the best forecast. This essentially produced the STATA command of reg ur L(1/6).ur, r, which means I am simply regressing all the unemployment rate data on the last 6 months of unemployment rate data, with robust standard errors. After that I used the estimates store and forecast solve commands, using 1000 simulations to produce the point estimates and standard errors for each estimate. This process produced me with my best estimate forecast for the unemployment rate among 20-24-year-old men.

**Results:** The main regression equation that is produced from this model is All of the regressors p-values are significant besides L4 and L5 which were 0.217 and 0.652 respectively. All standard errors for the regressors were below 0.05, which showed me that these are good regressors for forecasting. The F-Score for all variables is 0.000, showing me none of them should be equal to zero. The point estimates and intervals are in the table as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| *Month* | *Point Forecast* | *Lower Interval* | *Upper Interval* |
| November 2018 | **7.44** | **6.27** | **8.62** |
| December 2018 | 7.41 | 6.03 | 8.80 |
| January 2019 | 7.49 | 5.84 | 9.14 |
| February 2019 | 7.53 | 5.59 | 9.48 |
| March 2019 | 7.59 | 5.43 | 9.75 |
| April 2019 | 7.65 | 5.23 | 10.08 |
| May 2019 | 7.71 | 5.09 | 10.34 |
| June 2019 | 7.78 | 4.97 | 10.59 |
| July 2019 | 7.84 | 4.90 | 10.78 |
| August 2019 | 7.90 | 4.82 | 11.0 |
| September 2019 | 7.97 | 4.69 | 11.25 |
| November 2019 | 8.02 | 4.60 | 11.46 |

The best model estimate for the November 2018 unemployment rate among 20-24-year-olds is 7.44%, with a 90% confidence interval of [6.27,8.62]. Since the BLS only posts rates with one decimal point, it makes the estimates a little less precise than mine. Ultimately, the 7.44% estimate is higher than the 7.4% rate of September 2018, so it shows a slight increase, which is what I expect to happen. 