Time Series Forecasting of Metropolitan Atlanta Unemployment Rates (2009-2013)

MGS 3100 Project 2

Robin Gillman (100) Darrell Hareford (100)

March 25, 2014

Introduction

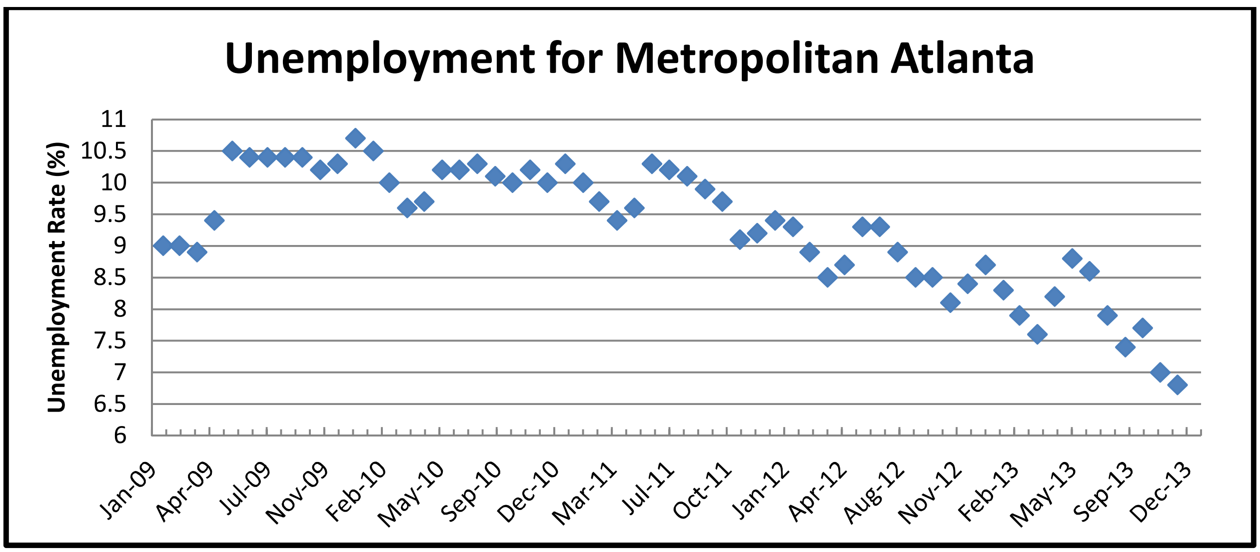
The recent economic crisis requires fast and reliable information to predict economic behavior early, which can be difficult in times of economic change. Using data on unemployment, we use all methods of forecasting to analyze and best predict future changes in unemployment. This information is applicable and pertinent to almost every sector of the economy, and it is invaluable to job-seekers, employers, employees, marketers, economists, and students, among many others. Given the high unemployment rates seen during the recent economic crisis, it is important to gather, analyze, and forecast this data so the stability of the economy and the labor market can be more accurately predicted in this time of economic uncertainty.

Data

Our data was collected online from the Bureau of Labor Statistics (BLS) and includes data on Atlanta, Sandy Springs, and Marietta (Metro Atlanta) unemployment rates. This data was measured by the BLS and collected through monthly sample population surveys of 60,000 homes and approximately 110,000 individuals. We collected monthly unemployment data for the five years of 2009 through 2013, for a total of 60 observations.

Preliminary Analysis

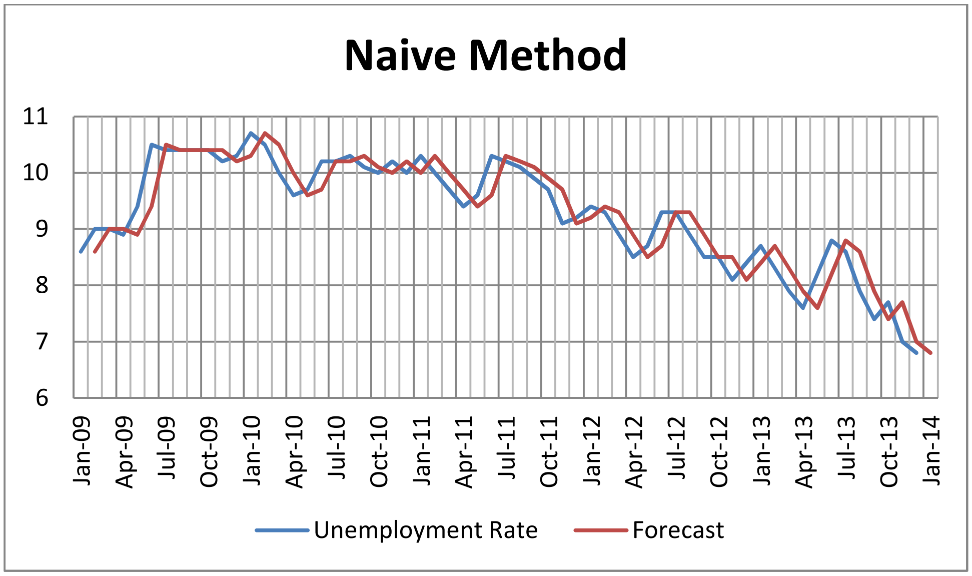
We begin our analysis of Metro Atlanta unemployment by plotting the collected data to see if a trend emerges. The following is a scatter plot of the raw data.



In the chart above, the data is clearly not flat; and since the data is not flat, we believe the Naïve, Moving Average, and Simple Exponential Smoothing methods of forecasting will be inadequate. This leaves us with the Regression Analysis and Classical Decomposition. Regression and classical decomposition are both similar; however, the fact that the data peaks at relatively regular intervals seems to indicate seasonality of the data, meaning that classical decomposition would be the best suited forecasting method because it will allow for the data to be seasonally adjusted, giving a more accurate forecast.

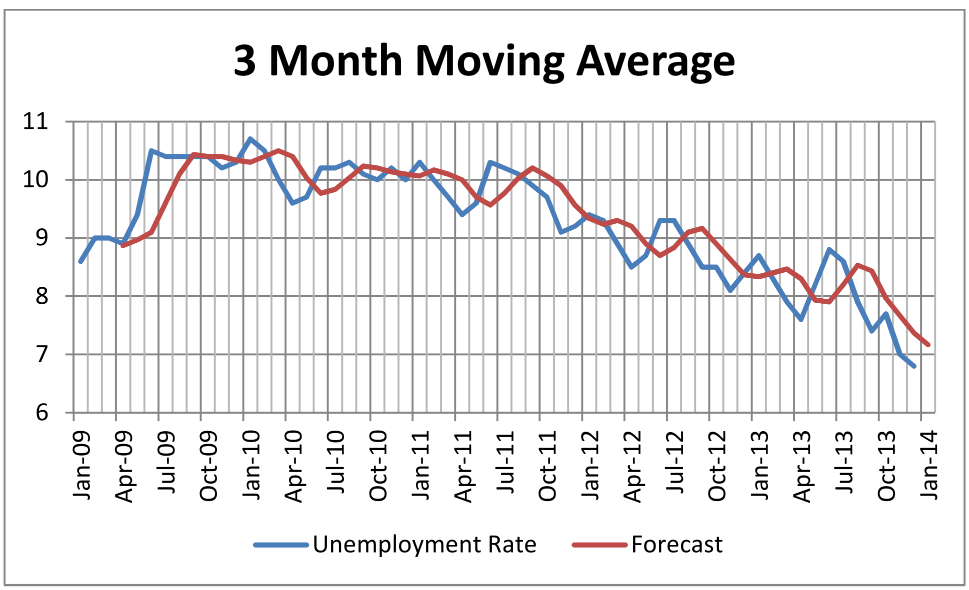
Forecasting

In order to be certain of which forecasting method is best suited for the data, we will analyze each method starting with the Naïve forecasting method.



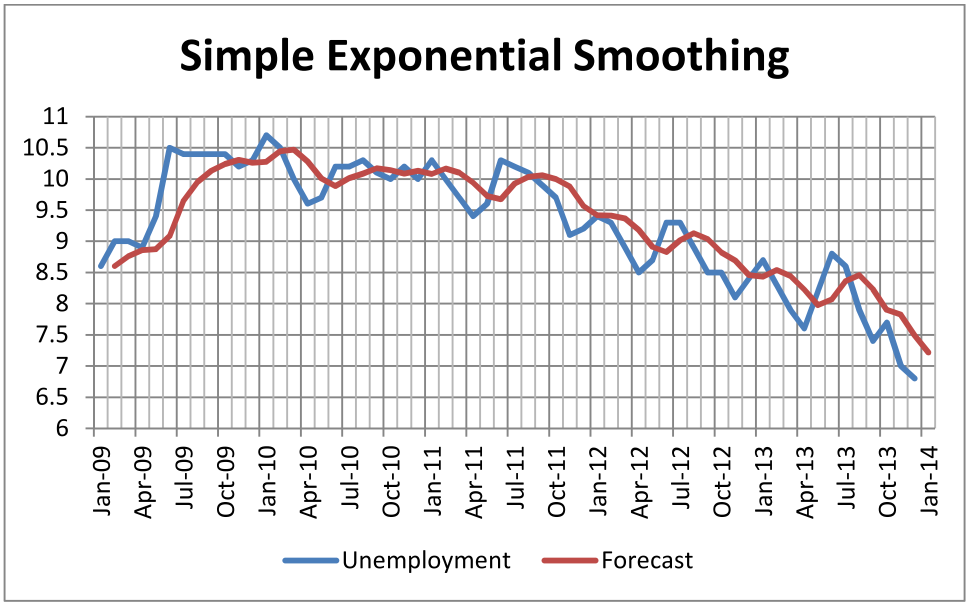
With this method, it is clear that the forecasted values versus the actual values vary in error, with some forecasted values being close to the actual values, and some being far off. By the nature of the predictive methodology it inherits difficulty in tracking rapid oscillations in trend-line direction.

The next forecasting method we will explore is the 3-Month Moving Average forecast.



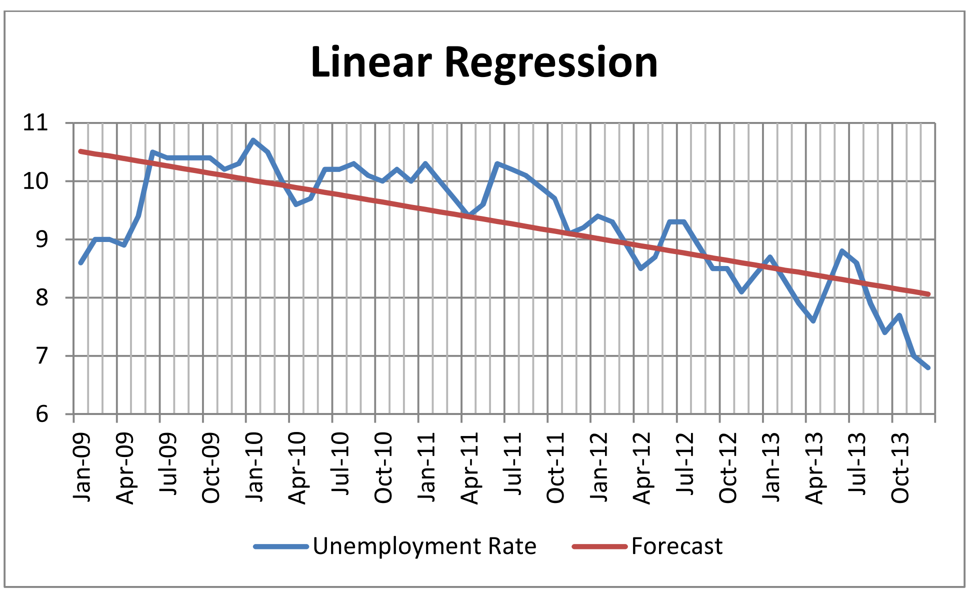
The three month moving average method is supposed to give a forecast based on the average of the previous three months; however, given the volatility of the labor market, this method appears to show ever larger forecasting errors.

Next, we use the Simple Exponential Smoothing (SES) method of forecasting.



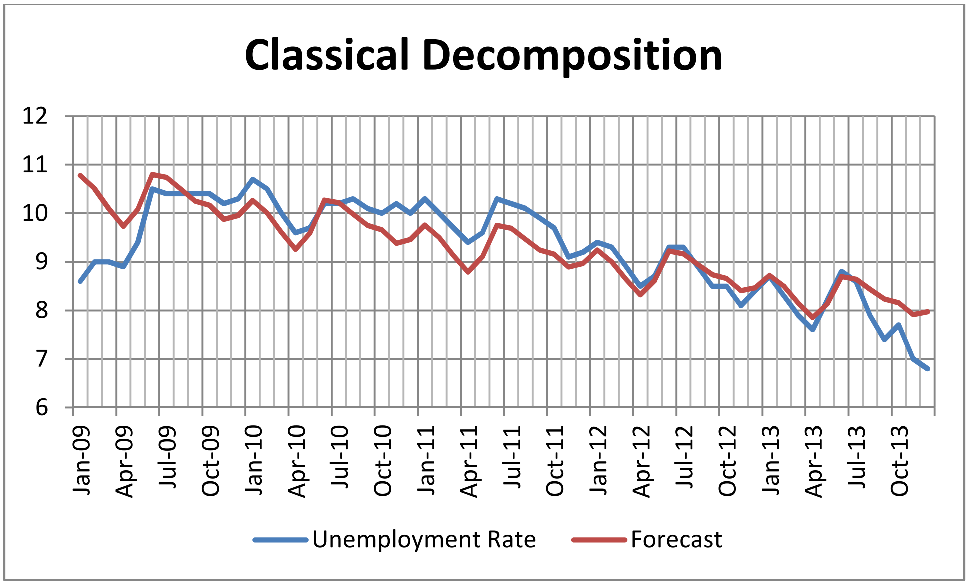
The SES method forecasts by adjusting the previous period's forecast by a factor of the previous period's error. Out of the forecasting methods we have tried so far, this method seems to track the trend direction well; however, the error is still widely varied, indicating that there may be a better method.

As noted in the preliminary analysis, the data shows trend. The next two forecasting methods we try, Regression and Classical Decomposition, tend to be better suited for data that shows trend. We will begin with Regression.



The linear regression here shows that there is a downward trend, but it cannot accommodate the spikes in the trend, and therefore errs even more widely than expected.

The final forecasting method is Classical Decomposition, which accounts for seasonality of trend data.



This method, though accurate in some areas, also has a wide variation in error, overall not fitting the data.

Evaluation

After seeing the above graphs showing the actual values versus the forecasted values for each of the five forecasting methods, we have a more clear perspective on which methods are best suited for our data. However, in order to decide which forecasting method is best, we will look at other measures of fit including bias and error. The chart below shows each of the methods alongside their respective measures of fit.

	Standard Error (SE)	Bias	Mean Squared Error (MSE)	Mean Absolute Deviation (MAD)	Mean Absolute Percentage Error (MAPE)
Naive	0.36869	-0.03051	0.135932	0.2949153	3.28%
3-Period Moving Average	0.484999	-0.06725	0.2352242	0.3912281	4.36%
Simple Exponential Smoothing	0.449954	-0.0586779	0.2024588	0.3643205	4.07%
Regression	0.6503596	0	0.422967664	0.4941244	5.30%
Classical Decomposition	0.565579	0	0.3198799	0.4201973	4.67%

This tables show that despite earlier predictions, it turned out that the forecasting methods intended for data with trend (as this data appeared to have) were the least-best fit. The naïve method, which was expected to be the least-best fit, was actually the best fit of all methods, with simple exponential smoothing trailing behind the best-fit method in every measure by .08 at the most. In the table above, the values shaded red are the lowest numbers for each measure. Clearly, the data shows that the Naïve method is the best forecasting method for this data.

Conclusion

The Preliminary Analysis led us to believe that Classical Decomposition would be the best method for forecasting Metro Atlanta area unemployment; however, further analysis showed that the Naïve method was actually the best method of forecasting for the data. Seasonality of the data was assumed due to the peaking of the apparent trend, but this hypothesis was proven did not lead to a better forecast result. One reason for Naive being the best method that may have been overlooked is that Naive forecasting can provide better forecasts than other methods when the data has a number of rapid transitions from peak to valley. The Naïve Method works remarkably well for economic and financial time series[1]. Using the Naïve method, we conclude that the forecast for January 2014 is a 6.8% unemployment rate.

Factors which may have limited the accuracy of standard regression and classical decomposition methods were the limits imposed upon the time frame of consideration. Recent history having driven the unemployment rate to heights rarely experienced in the US combined with a somewhat tepid job creation rebound may have resulted in significant numbers of previous defined workers now being dropped. The time period studied can have considerable impact on the conclusions drawn, since dramatic events or occurrences may impact the data. In essence, it might result in an aggregate group of outliers which may skew forecasts with more or less impact upon each type of forecast as a function of the methodology of that forecasting method.

---

[1] Rob Hyndman, George Athanasopoulos, *Forecasting: principles and practices*, OTexts, Oct. 17, 2013, http://www.otexts.org/book/fpp, (accessed 3/20/2014)