

US Unemployment Trends

JOSEPH BLUBAUGH*

Statistics

SEAN ROBERSON†

Mathematics, Industrial

AKARSHAN PURI‡

Electrical Engineering

ALISON SHELTON§

Statistics

TRAVIS LILLEY¶

Statistics

BO PANG||

Psychology, Statistics

July 25, 2016

Abstract

The information from above is from the original presentation. The links as to who did what should be modified probably at the end. This is just a starting point. Also the abstract should be written last so I thought it was a good place to put this information.

The writeup below has dummy text so I could set up the sections. I also moved some of the older write-up text to this document to start it all up.

1 Introduction

Unemployment has been a topic of concern throughout the United States in recent years. Graduate and Undergraduate college students alike are concerned over their employment prospects, wondering if their degrees will be enough to gain them a job after graduation. In these times of economic uncertainty, obtaining an income generating position is not the guarantee it has seemed to be in generations past. Therefore, the purpose of our project is to examine trends in unemployment in the United States, focusing on the years from 1992 to 2015, with a goal of forecasting into late 2016 and beyond.

1.1 Goal

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gef-burn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match

the language.

1.2 Data

The unemployment data being examined was obtained from the non-seasonally adjusted, monthly, Civilian Unemployment Rate Series (UNRATENSA), published by the Bureau of Labor Statistics (BLS). This series includes unemployment figures from January of 1948 to May of 2016 (U.S. Bureau of Labor Statistics, 2016). The response variable being analyzed is the unemployment rate defined as the percentage of the labor force that is unemployed. In defining this variable, the BLS restricts this to, “people 16 years of age and older, who currently reside in 1 of the 50 states or the District of Columbia, who do not reside in institutions (e.g., penal and mental facilities, homes for the aged), and who are not on active duty in the Armed Forces”.

2 Exploratory Analysis

Also known as put the initial plots here and then describe it.

As a first step, the data was plotted over time to identify any obvious patterns visually, considering both seasonally adjusted and non-seasonally adjusted versions of the unemployment rate, see Figure 1. The time span included in the study encompasses the presidential terms of Bill Clinton, George W. Bush, and Barack Obama, each serving eight years in office. Initial graphs of the data seem to indicate that, in general, unemployment spiked at the beginning of each president’s term and fell gradu-

*Plots, Data Prep, Code Management

†Presenter

‡Model selection and fitting

§Write-up

¶Diagnostics

||Model fitting and plots

ally over the time he was in office. There are also two noticeable spikes that represent the recessions of 2001 and 2008, respectively. The 2008 recession also follows the burst of a housing market bubble. These are all potential explanatory variables that will be explored in further analysis.

Figure 1: Plot of the original data

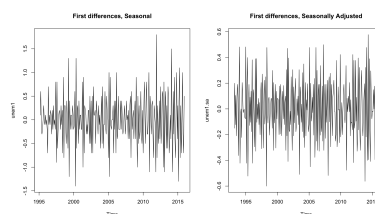
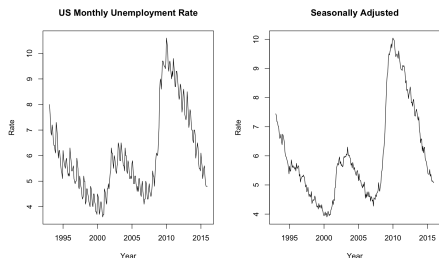


Figure 2: Plots of first differences

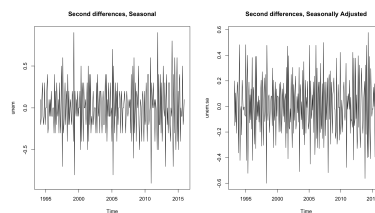


Figure 3: Plots of second differences

3 Achieving Stationarity

After the initial exploration of the time series graphs, the team has chosen to focus on building preliminary models which will serve as the foundation of further analysis. In looking at the initial plots, it appears that the series could benefit from detrending. As such, one of the primary goals has been to transform the data to stationarity.

An Augmented Dickey-Fuller (ADF) test for stationarity was conducted to verify the nonstationarity of the unemployment data. The ADF test tests the null hypothesis that the time series data has a unit root against the alternative that the data are stationary (Shumway and Stoffer, 2006). For the non-seasonally adjusted data the Dickey-Fuller test statistic was -1.4266, with a lag order of 6, and a p-value of 0.8176. Additionally, the seasonally adjusted data had a Dickey-Fuller test statistic of -2.1377, also with a lag order of 6, and a p-value of 0.518. The high p-values suggest that we do not have a stationary model with just the raw unemployment data, whether or not the data are seasonally adjusted.

The first and second differences of the unemployment data were plotted for both the seasonal and seasonally adjusted unemployment data, see Figures 2 and 3. Both the first and second differences appear considerably more stationary when compared to the original data. The associated ADF test results are given in Table 1. Based on the p-values, there is significant evidence of stationarity with each of the 4 differenced models. The consensus in the group was to continue the model building using second differences.

Table 1: ADF Test Results

Model	Statistic	Lag order	p-value
1 st difference	-7.6799	6	< 0.01
1 st difference, SA	-9.3595	6	< 0.01
2 nd difference	-8.4515	6	< 0.01
2 nd difference, SA	-9.3595	6	< 0.01

4 Model Building

4.1 ACF and PACF Plots

The team visually analyzed the ACF and PACF plots within the first season ($h = 1, 2, \dots, 12$), see Figure 4. The PACF appears to decline slowly, while the ACF seems to fall off after 1. Therefore we began by letting $p = 0$, and $q = 1$. Several models were considered by making adjustments to variations resulting in the models found in Table 2.

Figure 4: ACF & PACF Plots

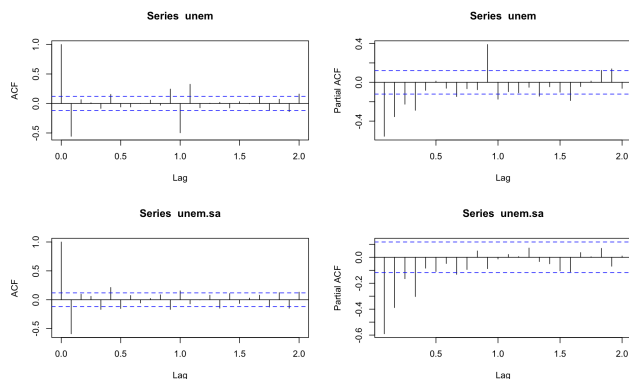
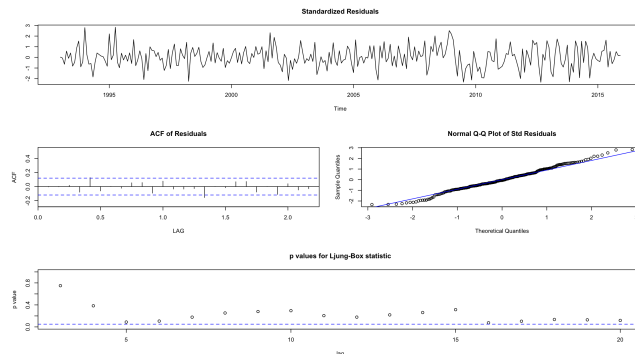


Figure 5: Model 5 Diagnostics



4.2 Model Fit

This will need to be adjusted to make reflect our current decision on the models.

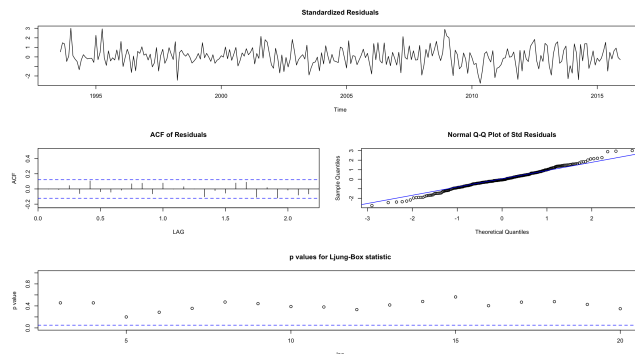
I'll fix it up later but for now I am getting rid of a couple of columns.

Table 2: Model Summaries

Model #	Order	Seasonal Order	AIC	BIC
1	0,2,1	1,1,0	-2.27	-3.23
2	0,2,1	3,1,0	-2.44	-3.37
3	4,2,1	3,1,0	-2.44	-3.32
4	0,2,1	1,0,0	-2.61	-3.58
5	1,2,1		-2.63	-3.60
6	1,0,1		-2.617	-3.578
7	1,0,1		-2.672	-3.565

Based on the AIC values, the two models that show the most promise are models 5 and 7. Model 5 includes only the time series data whereas model 7 also includes some of the predictors of interest. The diagnostic plots are shown in Figures 5 and 6. Both models show a great deal of promise. The standardized residuals show no apparent pattern. The ACF of the residuals show no departure from normality. Although the Normal Q-Q plot of the standardized residuals shows some slight departure from normality in the tails, there is no strong evidence of lack of normality in the residuals. The p-values for the Ljung-Box statistic are high enough at all plotted lags, so there is no indication of lack of fit in the models. Therefore, we will continue to refine these models further as we explore the nature of US Unemployment rate patterns.

Figure 6: Model 7 Diagnostics



4.3 Predictor Variables

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gef-burn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

5 Forecasting

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference

between this text and some nonsense like “Huardest gef-burn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

6 Discussion and Implications

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gef-burn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

References

Shumway, R. H. and Stoffer, D. S. (2006). Time Series Analysis and Its Applications: With R Examples - Third Edition. *Springer*, page 575.

U.S. Bureau of Labor Statistics (2016). *Civilian Unemployment Rate [UNRATE]*.
<https://fred.stlouisfed.org/series/UNRATE>.