US Unemployment Trends

Joseph Blubaugh[*] Sean Roberson[†] Akarshan Puri[‡] Alison Shelton[§]

Statistics Mathematics, Industrial Electrical Engineering Statistics

Travis Lilley[¶] Bo Pang[‖]

Statistics Psychology, Statistics

July 26, 2016

Abstract

US unemployment rates follow a complex countercyclical pattern, exhibiting smooth, sharp rises toward the beginning of presidential terms followed by much choppier, slower declines. Multiple ARIMA, SARIMA, and VAR models were developed and compared both to describe temporal unemployment trends from 1993-2015 and ultimately to predict unemployment rates from 2016-2018. The ARIMA(1, 2, 1) models with no exogenous predictors very accurately forecasted unemployment for six months and offered the greatest level of parsimony, but their predictions in the long term were unfeasible and prone to explosively large error bounds. A VAR(1) model with the three most useful exogenous predictors—construction, retail sales, and recession presence—performed comparably to the ARIMA(1, 2, 1) in the short run, and more accurately and precisely predicted unemployment in the long run by accounting for the sharp upswings in unemployment. Overall, our analysis suggests that unemployment trends require a layer of multivariate model complexity in order to be fully described and forecasted.

1 Introduction

Unemployment has been a topic of concern throughout the United States in recent years. The Great Recession of 2007 was accompanied by the worst unemployment crises seen since the 1930s (Wanberg, 2012). The results have been enduring. In 2010 the US job deficit was estimated to be over 10 million (Katz, 2010). 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. These worries are well-founded as full-recovery of college graduate employment rates and earning is expected to be a slow process (Carnevale and Cheah, 2015). In these times of economic uncertainty, obtaining an income generating position is not a guarantee, as it has seemed to be in past generations.

Unemployment has far-reaching consequences that extends beyond financial security. Unemployment is linked to psychological difficulties, including depression, physical deterioration, and even suicide (Wanberg, 2012; Kim and von dem Knesebeck, 2015; DeFina and Hannon, 2015). A study of Greek students found a relationship between parental unemployment and PTSD symptoms related to bullying (Kanellopoulos et al., 2014). In Nigeria, unemployment has been linked to insurgency and terrorism (Akanni, 2014). Given the impact that unemployment has on fiscal, mental, and physical health, research into unemployment patterns an important part of developing policies to improve the welfare of the local, national, and global populace.

1.1 Goal

The purpose of our project is to examine trends in unemployment in the United States. We will focus on the years surrounding the Great Recession of 2007, 1992 to 2015. Our primary goal is to forecast future unemployment rates.

1.2 Data

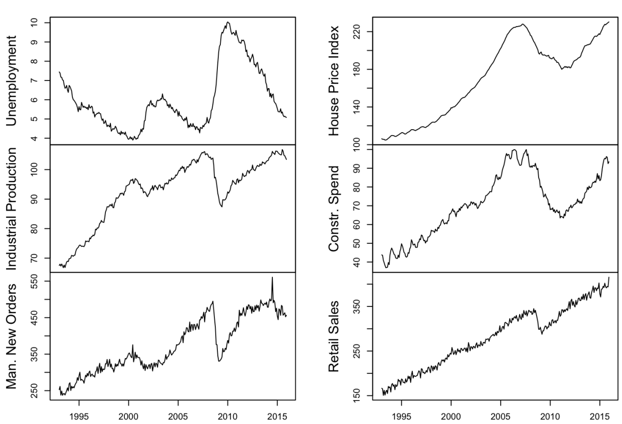
The unemployment data being examined was obtained from the seasonally adjusted, monthly, Civilian Unemployment Rate Series (UNRATE), published by the Bureau of Labor Statistics (BLS). This series includes unemployment figures from January of 1948 to May of 2016 (U.S. Bureau of Labor Statistics, 2016). 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

[*] Plots, Data Prep, Code Management
[†] Presenter
[‡] Model selection and fitting
[§] Write-up
[¶] Diagnostics
[‖] Model fitting and plots

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".

Unemployment tends to follow a counter-cyclical pattern, increasing quickly during times of economic slowdowns and decreasing slowly in times of growth (Montgomery et al., 1998). To address this we have chosen to include a recession indicator as a possible predictor of unemployment. Recession dates were obtained from the National Bureau of Economic Research (NBER) (The National Bureau of Economic Research, 2016). The NBER identifies recessions and US business cycles based upon a variety of economic indicators. These include Gross Domestic Product (GDP), Gross Domestic Income (GDI), and a variety of lesser known indicators such as aggregate hours of work in the total economy.

Figure 1: Timeplots of included variables



We also explored several predictor variables related to unemployment. Industrial production measures enterprise output of the U.S. establishments (The Board of Governors of the Federal Reserve System, 2016). Value of manufacturers' new orders for all manufacturing industries refers to manufacturer's sales and inventory, except for new orders from the semiconductor industry (US. Bureau of the Census, 2016c). The purchase only house price index for the United States follows sales for a specific set of single-family homes (US. Federal Housing Finance Agency, 2016). We also included Retailers Sales (US. Bureau of the Census, 2016a) and Total Construction Spending (US. Bureau of the Census, 2016b). Each of these predictors shows an overall increasing trend over time, see Figure 1.

2 Exploratory Analysis

Figure 2: Monthly unemployment, 1948-2015

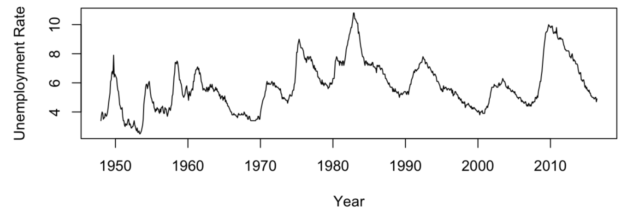
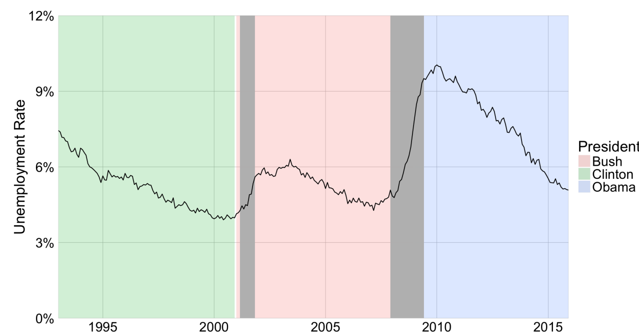
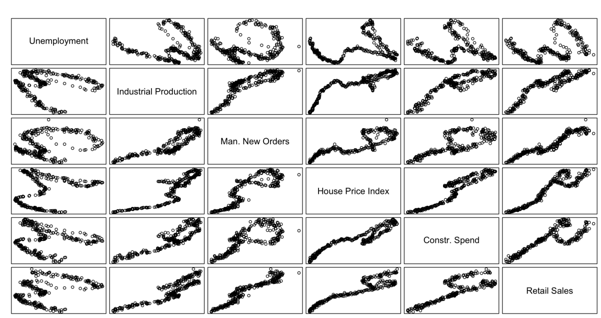


Figure 3: Monthly unemployment, 1993 - 2015



As indicated in Figure 2, raw untransformed unemployment appears relatively volatile. There are several time periods of sudden spikes in the unemployment rate, followed by a slower recovery period. This counter-cyclical movement is consistent with the descriptions of unemployment data found in the literature (Katz, 2010; Montgomery et al., 1998; Shimer, 2012).

Figure 4: Scatterplot matrix of unemployment and potential predictors

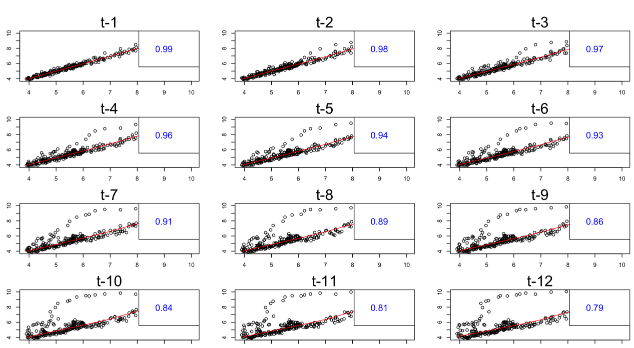


Because of marked potential differences in the trend surrounding times of economic downturn, such as those that occurred after World War II and in the 70s and the 80s, we have chosen to limit our analysis to a more recent set of unemployment data. Ultimately, we decided to focus the time preceding and following the Great Recession of 2007. We limited our initial analysis to 1993 to 2015, which 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 gradually over the time he was in office, see Figure 3. There are also two noticeable spikes that represent that recessions of 2001 and 2008, respectively. The 2008 recession also follows the burst of a housing market bubble. These are all explanatory variables that can potentially inform unemployment patterns. A scatterplot matrix of these predictors can be seen in Figure 4.

3 Achieving Stationarity

In analyzing the initial plots, it appears that the series could benefit from detrending. A graph of various potential lagged values for unemployment can be seen in Figure 5. The high values of the correlation coefficients, particularly through lag 6, further suggest a high degree of autocorrelation within the unemployment dataset. An Augmented Dickey-Fuller (ADF) test for stationarity was conducted to verify the nonstationarity of the unemployment data. The ADF 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). The high p-values suggest that we do not have a stationary model with just the raw unemployment data.

Figure 5: Autocorrelation of unemployment data



The first and second differences of the unemployment data were plotted for the seasonally adjusted unemployment data, see Figure 6. Both sets of differencing, bring the data closer to stationarity with a consistent mean and more constant variance. 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 differenced models. Visually, the second differences best approximate a white noise series. Therefore, the consensus in the group was to continue the model building process using second differences.

Figure 6: Timeplots of lagged unemployment

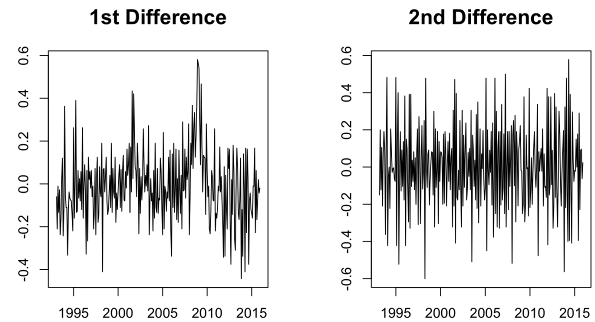


Table 1: ADF Test Results for unemployment

Model	Statistic	Lag order	p-value
No difference	-1.2646	0	0.8859
No difference	-2.1377	6	0.8859
1st difference	-9.3595	6	< 0.01
2nd difference	-9.3595	6	< 0.01

The predictor variables were also detrended using second differences. The timeplots of these second differences can be seen in Figure 7. Although the housing prices still retain some non-constant variance, overall the differencing improves the stationarity of all the predictor variables. Furthermore, the ADF test of the differenced data provides evidence of stationarity for each of the variables, see Table 2.

Figure 7: Timeplots of differenced predictors

An attempt to stabilize the variance of the housing prices, with a log transformation, does not improve this stationarity much (ADF changes from -9.104 to -9.5211) and a scatterplot of the differenced logs still shows evidence of heteroscedasticity in the variance over time, see Figure 8.

Figure 8: Timeplot of the housing price log transformation, d=2

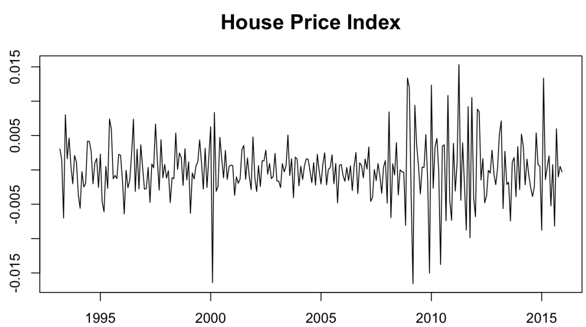


Table 2: ADF Test Results for Predictors, $d = 2$

Variable	Statistic	p-value
Industrial Production	-9.2333	< 0.01
New Orders	-8.391	< 0.01
House Prices	-9.104	< 0.01
Construction Spending	-10.447	< 0.01
Retail Sales	-10.72	< 0.01

4 Model Building

The correlogram (ACF plot) and partial correlogram (PACF plot) of the unemployment data is shown in Figure 9. The ACF seems to tail off and the PACF seems to cut off at either 1 or 3. A tailing ACF function with a PACF that cuts off at p suggests an AR(p) model (Box et al., 2008). Therefore, these initial plots suggest a possible AR(1) or AR(3) model. When looking at the ACF and PACF of the second differences, we have evidence of a possible integrated model with $d = 2$. For example, an ACF of difference d that decays exponentially after lag 1 with a PACF that is dominated by an exponential decay pattern after lag 1 would be evidence of an ARIMA(1,d,1) model. Therefore, it is worthwhile considering ARIMA models such as ARIMA(1,2,1). Of course predictor variables may help to improve the predictive strength of our models; therefore, models with regressors and vector autoregressive Models (VAR) were considered as well.

Figure 9: ACF & PACF Plots

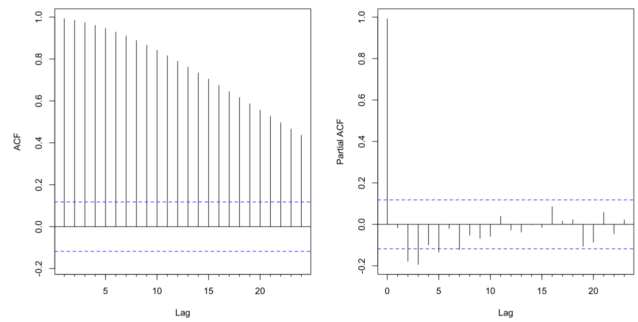
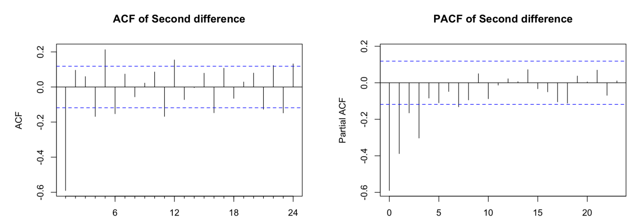


Figure 10: ACF & PACF Plots of Second Differences



4.1 Models Considered

4.1.1 ARIMA Models

Table 3: ARIMA models considered

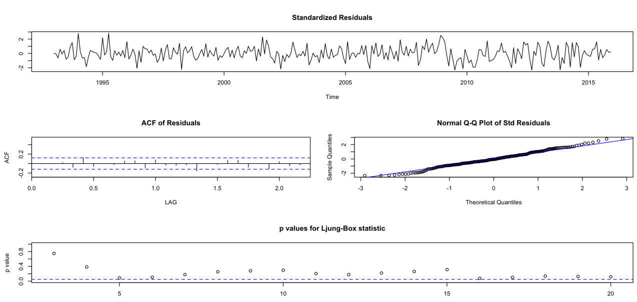
Model	Order	Reg	AIC	BIC	Best
1	1,2,1	NA	-212.30	-201.46	BIC
2	2,2,2	NA	-211.81	-193.74	
3	3,2,3	NA	-215.48	-190.19	
4	1,2,1	X	-211.56	-182.65	
5	2,2,2	X	-209.83	-177.32	
6	3,2,3	X	-215.10	-171.74	
7	1,2,1	LagX	-222.45	-193.69	AIC
8	2,2,2	LagX	-220.70	-188.35	
9	3,2,3	LagX	-217.89	-174.76	

We began by exploring three potential ARIMA models without regressors, ARIMA(1,2,1), ARIMA(2,2,2), and ARIMA(3,2,3). Although model 3, ARIMA(3,2,3), has the lowest AIC of the three models, model 1, the ARIMA(1,2,1) model, has the lowest BIC. Model 1 is also the most parsimonious model of the three. So of the three initial models, without regressors, we chose to retain model 1.

Although their parsimony is attractive, the ARIMA models are too simplistic to accurately portray the asymmetric nature of unemployment data and have a tendency to underestimating unemployment during economic slowdowns (Montgomery et al., 1998). Therefore, we repeated the above analysis using multivariate ARIMA models. The variables Industrial Production, Value of Manufacturers' New Orders, Purchase Only House Price Index, Retailers Sales, and Total Construction Spending were included as potential predictors of unemployment.

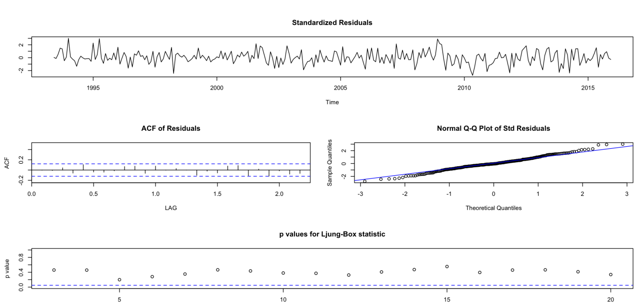
Models 4 through 6 were ARIMA(1,2,1), ARIMA(2,2,2), and ARIMA(3,2,3) respectively. These predictors had lower AIC and BIC values than their original counterparts without regressors, ash shown in Table 3. Since these models had predicted a lagged response variable using data that was potentially nonstationary, we chose to repeat the process using lagged regressors. Models 7, 8, and 9 refer to the ARIMA(1,2,1), ARIMA(2,2,2), and ARIMA(3,2,3) models with lagged predictor variables. Of these three new models, model 7 has both the smallest AIC and the smallest BIC values. Of the 9 original models, model 7 has the lowest AIC overall, as shown in Table 3.

Figure 11: Model 1: Residual diagnostics



ARIMA(1,2,1) with no regressors

Figure 12: Model 7: Residual diagnostics



ARIMA(1,2,1) with lagged regressors

Based on the AIC and BIC values, the two ARIMA models that show the most promise are models 1 and 7. Model 1 includes only the time series data, whereas model 7 also includes some lagged versions of the predictors of interest. The diagnostic plots for these models are shown in Figures 11 and 12. Both models show a great deal of promise. The standardized residuals show no apparent pattern. The ACF of the residuals shows no departure from normality. Although the Normal Q-Q plot of the standardized residuals shows some slight departure from normality in the tails, for both models, there is no strong evidence of lack of normality in the residuals The p-values for the Ljung-Box statistic are well above 0.05 at all plotted lags, so there is no indication of lack of fit in the models.

4.1.2 VAR Models

Table 4: VAR models considered

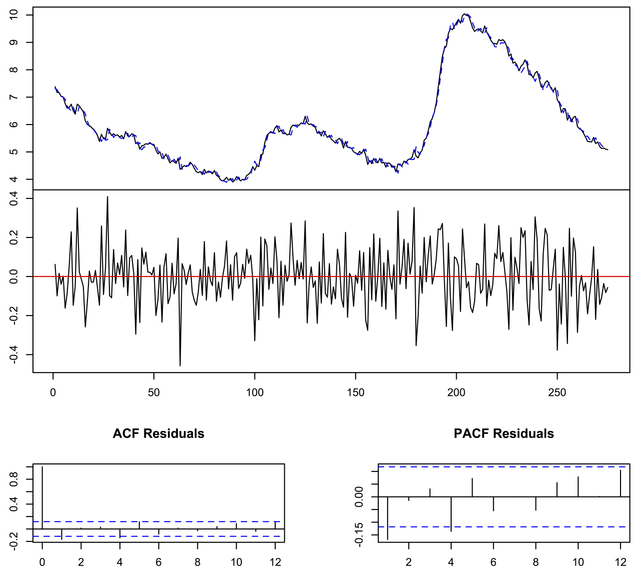
Model	P	Type	AIC	BIC	Best
1	1	NA	-223.67	-201.97	
2	2	NA	-217.83	-185.31	
3	1	Ind	-256.77	-231.45	BIC/AIC
4	1	LagX	-216.65	-195.06	
5	2	LagX	-212.53	-180.17	
6	1	Both	-245.72	-220.53	

Much of the recent literature on modeling unemployment trends suggests that vector autoregressive models (VAR) have the capacity to outperform ARIMA models and are widely used by professional forecasters (Meyer and Tasci, 2015; Tasci and Treanor, 2015; Barnichon and Garda, 2016). VAR models provide a mechanism for modeling complex, multivariate times series in the absence of a moving average term (Chatfield, 2001) . The ACF and PACF plots shown in Figures 9 and 10 do not conclusively demonstrate that the moving average term is necessary in this case; therefore we have decided to explore the potential in fitting VAR models to the unemployment data in order to improve the performance of our predictions.

We started with 6 initial VAR models to compare. Models 1, 2, and 3 use the predictors of construction spending and retail sales, without differencing. Model 1 is a VAR(1), model 2 is a VAR(2), and model 3 is a VAR(1) with the regression indicator included as well. Models 4, 5, and 6 repeat the analysis using the differenced version of the predictors. Table 4 shows the AIC and BIC values for each of these models.

Model 3, the unlagged model with the regression indicator, has the lowest AIC and BIC values. The diagram of the fit and residuals for model 3 is provided in Figure 13. The blue line indicates that the actual and predicted values of unemployment are similar in this model. A timeplot of the residuals is consistent with a white noise series. The ACF and PACF of the residuals give no indication of lack of fit. Therefore, we have chosen to retain model 3 to compare with the ARIMA models developed earlier.

Figure 13: Model 3 fit and residuals for unemployment



4.2 Initial Model Comparisons

Table 5: Comparison of ARIMA and VAR models

Model	Type	AIC	BIC
ARIMA #1	Univ ARIMA(1,2,1)	-212.29	-201.45
ARIMA #7	Mult ARIMA(1,2,1)	-222.45	-193.69
VAR #3	VAR(1)	-256.76	-231.45

In the previous model building process, we retained 3 models for further comparison. ARIMA model 1 is a univariate ARIMA(1,2,1) model without predictors, ARIMA model 7 is an ARIMA(1,2,1) model with exogenous predictors, and the VAR model 3 is a VAR(1) model with no regressors and an indicator variable for recession among its predictors.

At first glance the VAR(1) model appears to be the best model. It has the lowest values of both AIC and BIC. Of the two ARIMA models the ARIMA(1,2,1) model with predictors has a lower AIC but a higher BIC. However, being a multivariate model it allows us to leverage the additional information provided by indicators of the nature of the economy to refine our predictions about future unemployment rates.

Since the rate of increasing unemployment is so different from the rate of decreasing unemployment, forecasting can be very difficult without some indicator of whether we are currently in a increasing or decreasing portion of the cycle (Montgomery et al., 1998). For this reason, we have chosen to continue our model comparisons with only the two models that include predictors of economic strength. In the next section, we compare the forecasting performance of the best multivariate ARIMA model with best VAR model. These are the multivariate ARIMA(1,2,1) model and the VAR(1) with an indicator variable for recession among its predictors.

5 Forecasting

In the initial data setup, the 2016 values were initially excluded to provide a dataset with which to evaluate the performance of our predictions. Short-term predictions from the multivariate ARIMA(1,2,1) model and VAR(1) models were then compared compared with the actual unemployment rates for January 2016 through May 2016. These results were used to create our final model.

5.1 Multivariate ARIMA(1,2,1)

Figure 14: 5 month forecast with ARIMA(1,2,1)

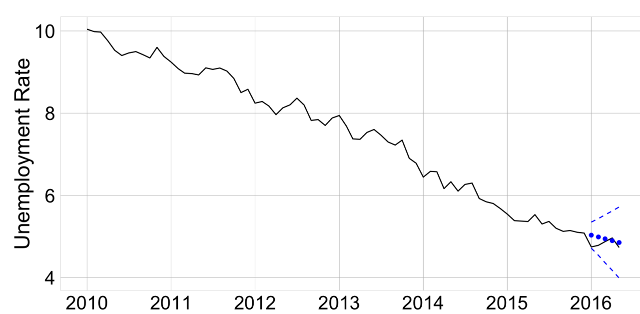


Table 6: 2016 Unemployment Rate Predictions from Multivariate ARIMA(1,2,1)

Month	Observed	Predicted	95% CI	Residual
Jan	4.74	5.03	(4.71 , 5.35)	-0.29
Feb	4.78	4.99	(4.54 , 5.43)	-0.20
Mar	4.87	4.94	(4.36 , 5.53)	-0.07
Apr	4.96	4.90	(4.17 , 5.62)	0.07
May	4.73	4.85	(3.99 , 5.72)	-0.12

The unemployment rates for January 2016 to May 2016 were forecast using the Multivariate ARIMA(1,2,1) model, see Table 6. This model did a good job overall, since all predicted values were relatively close to the actual unemployment rates and entirely inside the confidence bands. Figure 14 shows the predicted values graphed against the actual unemployment rates. The model predicts a steady decrease in the unemployment rate over the 5 month period. In general, the ARIMA model provided an overestimation of unemployment rate, except in April-16 when the rate spiked slightly. The confidence bands spread outward rapidly, making a definitive long term trend hard to predict.

5.2 VAR(1)

Figure 15: 5 month forecast with VAR(1)

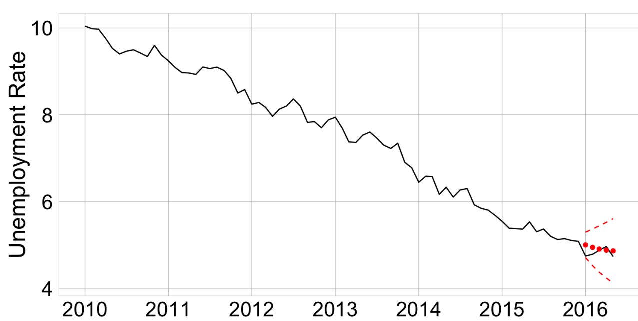


Table 7: 2016 Unemployment Rate Predictions from Multivariate VAR(1)

Month	Observed	Predicted	95% CI	Residual
Jan	4.74	5.00	(4.71 , 5.29)	-0.26
Feb	4.78	4.94	(4.52 , 5.36)	-0.16
Mar	4.87	4.90	(4.37 , 5.44)	-0.03
Apr	4.96	4.88	(4.24 , 5.52)	0.09
May	4.73	4.86	(4.12 , 5.60)	-0.13

The unemployment rates for January 2016 to May 2016 were also forecast using the Var(1) model, see Table 7. At first glance the results look very similar to the ARIMA(1,2,1) model. All predicted values were within 0.3% of the actual unemployment rates and entirely inside the confidence bands. Figure 15 shows the predicted values graphed against the actual unemployment rates. The model predicts a nonlinear decrease in the unemployment rate over the 5 month period, with the rate of decrease slowing over time. In general, the VAR model also provided an overprediction of unemployment rate, except in April-16.

5.3 Forecast comparisons

Figure 16: ARIMA and VAR Model comparison of 3 year forecast

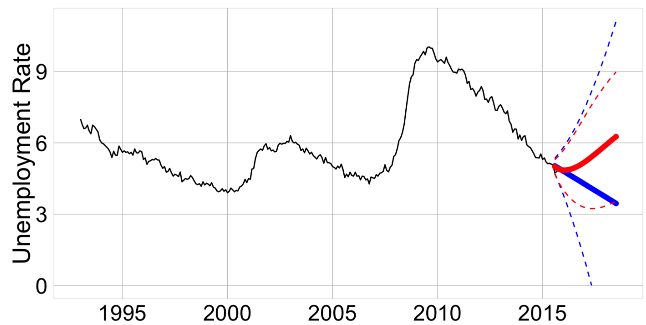


Figure 16 provides a graphical comparison of the two candidate models. The ARIMA(1, 2, 1) shows a steady, linear decrease of the unemployment rate over time, which is unrealistic in the long term. The Var model predicts a decrease in unemployment followed by an increase, which is more consistent with actual unemployment patterns. The mean square error from the VAR(1) model (0.0097) is much lower than the mean square error of the ARIMA(1,2,1) model(0.0151). The implication of this is that the confidence interval of the ARIMA model quickly explodes, suggesting that it may not a good choice for longer term forecasts.

In the long run the VAR(1) model seems to provide a much more accurate forecast than the ARIMA(1, 2, 1) model and with much narrower prediction intervals. The complex, countercyclical nature of unemployment necessitates the use of a more advanced model for longer time periods. However, in the short term VAR(1) is not dramatically more effective than ARIMA(1, 2, 1), suggesting that an VAR(1) model might be over-fitting the data in this region.

6 Final model: VAR(1)

6.1 Model estimates

The final model chosen was the VAR(1) with construction spending, retail sales, and the recession indicator as predictors. The estimates of the coefficients for this model can be found in 8. For this model, all predictors have p-values below .001, indicating that they add significantly to the model with the other variables included. The positive coefficient for trend suggests that, overall, unemployment has increased over the last 23 years, which is consistent with the overall trend in the timeplot of 2.

The coefficient for retail spending is negative, which makes practical sense. If people are spending more money, we would expect that business would have more capital to hire employees. Conversely, people tend to spend less when they are unemployed. The positive coefficient for construction is a little harder to explain. Higher construction costs tend to be associated with higher unemployment rates, keeping all other variables constant. This may be just an artifact of the multivariate nature of the dataset. For a given amount of retail spending, more construction spending is associated with more unemployment. Perhaps this is an indication of the higher cost of resources, when construction materials cost more there is less capital to spend on human resources. Or maybe it is an indication of income inequality, when unemployment is high builders may cater to wealthier clients to recoup costs. Finally, there is a high positive coefficient for the recession indicator, suggesting that during times of recession unemployment is expected to be higher than in times of economic strength.

Table 8: Estimation results for equation VAR(1) model

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
unem.l1	0.9752	0.0010	97.677	<2e-16
constr.l1	0.0043	0.0012	3.412	0.000744
retail.l1	-0.0059	0.0011	-5.241	3.24e-07
recession.l1	0.1924	0.0318	6.051	4.80e-09
const	0.8438	0.2015	4.188	3.82e-05
trend	0.0045	0.0009	4.759	3.17e-06

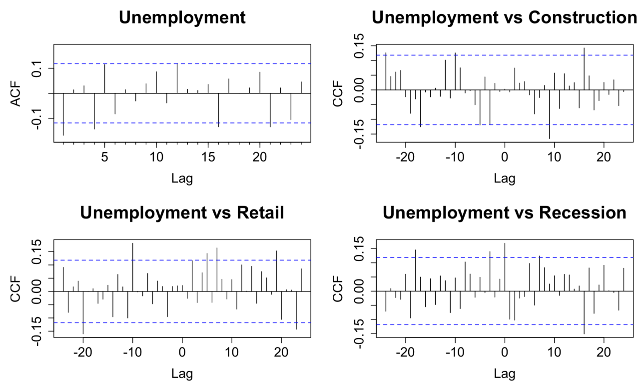
The final equation for the VAR(1) model is:

$$\begin{aligned}\widehat{\text{Unemployment}} = 0.8438_{(0.2015)} &+ 0.0045(t)_{(0.0009)}\\ &+ 0.9752\text{Unemployment}_{t-1(0.0010)}\\ &+ 0.0043\text{ConstructionSpend}_{t-1(0.0012)}\\ &- 0.0059\text{RetailSales}_{t-1(0.0012)}\\ &+ 0.1924\text{Recession}_{t-1(0.0012)}\end{aligned}$$

6.2 Model Diagnostics

Figure 17 shows the ACF and the CCF plots for the VAR(1) model of unemployment. The ACF for the unemployment rate does not indicate any lack of fit for the model. However, the CCF for unemployment vs retail and unemployment vs recession suggest that there may be some remaining dependency not captured by the VAR(1) model. The Portmanteau Test, used to detect model mispecification for multivariate time series (Davies and Newbold, 1979), was conducted to test for stationarity in the residuals. Since the Portmanteau statistic was large

Figure 17: VAR Residual Plots



($\chi^2 = 529.42$, $df = 176$, $p < .001$), we rejected the null hypothesis. There is strong evidence of nonstationarity in the VAR(1) residuals. This small p-value, along with the graphs above, suggests there is still some dependency left in our model.

7 Discussion and Implications

The VAR(1) model given in section 6.1 is a strong model overall. However, given the unexplained dependency in the residuals, further investigation may be necessary to create a more reliable model in the long term. Furthermore, there are many economic predictors we did not investigate in this project. Further research should explore these potential explanatory variables in more depth.

We were not able to analyze the effect that presidential transitions had on unemployment. Cursory visual inspection of both Figures 2 and 3 reveal a sharp increase in unemployment, followed by a slower decline. This could be investigated using an indicator variable for election year. Furthermore, it would be interesting to evaluate the impact that party affiliation has on unemployment over time. Using measures for presidential party and variables indicating the number of democrats and republicans in Congress, we may be able to detect a relationship between unemployment rate and political climate.

Additionally, a mixed model might be beneficial when modeling complex variables like unemployment rates. "Because of the evidence of fractional integration in the unemployment, stationarity and non-linearity issues" researchers in Croatia utilized a "multivariate singular spectrum model (MSSA) for modeling unemployment" (Skare and Buterin, 2015). It would be interesting to explore its potential for use in the U.S. which has a significantly larger economy and population. Or, perhaps we could create a model that somehow combines the simplicity of the ARIMA models for short term predictions with the more complex structure needed for the long term. After further development, the model should be tested on the other unemployment rates, such as industry specific or local unemployment rates.

7.1 Conclusion

Unemployment in the United States is an issue that is complex and socially meaningful. Recent research has explored the use of VAR models and worker flow data to predict the direction of unemployment in the short term. This project has attempted to predict unemployment rates using other economic indicators, including spending levels and recession time period. While the model created was strong, with relatively precise short term accuracy, it could benefit from further refinement. Future research in this area should include more variables, spectral analysis, and a mixed model approach.

References

Akanni, A. (2014). History of terrorism, youth psychology and unemployment in nigeria. *Journal of Pan African Studies*, 3:65.

Barnichon, R. and Garda, P. (2016). Forecasting unemployment across countries: The ins and outs. *European Economic Review*, 84:165–183.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2008). *Time series analysis : forecasting and control. 4th ed. George E.P. Box, Gwilym M. Jenkins, Gregory C. Reinsel.* Wiley series in probability and statistics. Hoboken, N.J. : J. Wiley and Sons, [2008].

Carnevale, A. P. and Cheah, B. (2015). From Hard Times to Better Times: College Majors, Unemployment, and Earnings. *Georgetown University Center on Education and the Workforce.*

Chatfield, C. (2001). *Time-series forecasting. Chris Chatfield.* Boca Raton : Chapman and Hall/CRC, [2001].

Davies, N. and Newbold, P. (1979). Some power studies of a portmanteau test of time series model specification. *Biometrika*, 66(1):153–155.

DeFina, R. and Hannon, L. (2015). The changing relationship between unemployment and suicide. *Suicide & Life-Threatening Behavior*, 45(2):217 – 229.

Kanellopoulos, A., Tsiantis, A., Tsiantis, J., Beratis, I., Sygelakis, E., Stamati, G. D., and Psarras, R. (2014). EPA-1642–Parental unemployment and post-traumatic stress disorder symptoms. a study through the fog of greek financial crisis. *European Psychiatry*, 29:1.

Katz, L. F. (2010). Long-term unemployment in the Great Recession.

Kim, T. J. and von dem Knesebeck, O. (2015). Is an insecure job better for health than having no job at all? a systematic review of studies investigating the health-related risks of both job insecurity and unemployment. *BMC Public Health*, 15:985.

Meyer, B. and Tasci, M. (2015). Lessons for Forecasting Unemployment in the United States: Use Flow Rates, Mind the Trend. *Working Paper Series (Federal Reserve Bank of Atlanta)*, 2015(1):1.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., Tiao, G. C., Montgomery, A. L., Tsay, R. S., and Tiao, G. C. (1998). Forecasting the U.S . Unemployment Rate. *Journal of the American Statistical Association*, 93(442):478–493.

Shimer, R. (2012). Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*, 15(2):127–148.

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

Skare, M. and Buterin, V. (2015). Modelling and Forecasting Unemployment Non-linear Dynamics Using Spectral Analysis. *Engineering Economics*, 26(4):373–383.

Tasci, M. and Treanor, C. (2015). Forecasting Unemployment in Real Time during the Great Recession: An Elusive Task. *Economic Commentary*, 2015(15):1.

The Board of Governors of the Federal Reserve System (2016). Industrial Production and Capacity Utilization - G.17.

The National Bureau of Economic Research (2016). Us business cycle expansions and contractions.

U.S. Bureau of Labor Statistics (2016). Civilian Unemployment Rate (UNRATE).

US. Bureau of the Census (2016a). Retailers Sales.

US. Bureau of the Census (2016b). Total Construction Spending.

US. Bureau of the Census (2016c). Value of Manufacturers' New Orders for All Manufacturing Industries (AMTMNO).

US. Federal Housing Finance Agency (2016). Purchase Only House Price Index for the United States.

Wanberg, C. R. (2012). The individual experience of unemployment. *Annual review of psychology*, 63:369–396.

Appendix

Project contributions

- Joseph Blubaugh
  - Plots
  - Data Prep
  - Code Management
  - Model refinement
- Sean Roberson
  - Presenter
  - Key talking points
  - Model clarification
  - Literature
- Akarshan Puri
  - Model selection
  - Model diagnostics
  - Model fitting
- Alison Shelton
  - LaTeX code management
  - Writing
  - Literature
- Travis Lilley
  - Model building
  - Model fitting
  - Model diagnostics
  - Abstract
- Bo Pang
  - Model building
  - Model diagnostics
  - Model fitting

Please note, this project was done in a collaborative fashion using GitHub as a tool for collaborative coding and writing. Therefore, the above list of contributions is only a rough estimate of how the workload was distributed. Each member worked on every component. While the project was small we used Overleaf for collaborative writing, as it grew changes were pushed to GitHub. Literature references were shared on GitHub and using Mendeley.

It is particularly important to note that once an initial file was placed on the repository, all members had access to and modified each other's code. We believe this brings strength to our project but it makes it difficult to definitively state who contributed to each portion in a concise manner.

For the R-code used in this project and more information about our project management process please see `https://github.com/JestonBlu/STAT626_PROJECT`.