

Social and Economic Factors Leading to Differences in US State Unemployment

Taylor Neighbours^a

Abstract:

This paper dives into the socio-economic factors that lead to differences in U.S. State unemployment rates. The study will look at and evaluate prior economic literature, including regression analysis, comparative statistics, and raw data from the BLS and Census Bureau. Further, it will take select parts from past models to create and analyze an economic regression model. The data to be found below is a cross-sectional interpretation of each of the 50 states for the years 2000 & 2010. The root causes for fluctuations in unemployment can help economists better understand how to monitor and manage unemployment levels on a state by state basis.

JEL Classification: [E24, J64](#)

Keywords: Unemployment

^a Taylor Neighbours, 315 Hopmeadow St., Simsbury, CT 06089. Phone: (860) 392-9797.
Email: tneighbo@bryant.edu

1.0 INTRODUCTION

The Unemployment Rate is a lagging economic indicator. It measures the effects or results of specific events (i.e. recession, war, political change, etc.); however, the effects do not show until after the event has occurred. Similar to the idea of global warming, once the events have taken place there is not much that can be done to fix the problems immediately. Sometimes the best way to reduce unemployment is to allow it to naturally recede. Many economists believe that tinkering with the system could cause more damage than help. With this paper, I hope to be able to come to conclusions on how to use the independent variables in the regression model as either leading or lagging indicators for unemployment levels. This will allow policy makers to take action when they start to see the leading indicators change in an unfavorable way and it will also allow them to see when the policy is no longer necessary by watching the lagging indicators return to favorable levels.

U.S. Unemployment is a broad topic that has many economic implications and has the potential to affect each and every one of us. However, the amount it affects each of us can change based on which state we live in. Why would that be? Each state has their own set of laws and the people in each state are made up of an assortment of demographics. Any 1 to 100 of these factors could be the cause(s) of these differences. This study will look into the potential factors that can be measured, which the data is available, and that are relevant for the two observed years.

This study is a Cross-Sectional regression analysis on the social and economic factors that contribute to fluctuations in state unemployment rates. The data for the study was collected for each of the 50 states and for the years 2000 and 2010. These years were chosen based on the

availability of information due to censuses being distributed every 10 years. Between this ten year gap the United States experienced significant shocks to the economy, by placing the study on opposite ends of these shocks allows us to take a closer look at the true impact they had on both Unemployment and the independent variables. However, the negative implications of choosing these dates are the fact that the 2010 data may not be as relevant to compare to the 2000 data due to the massive impact the 2008 recession had on unemployment rates.

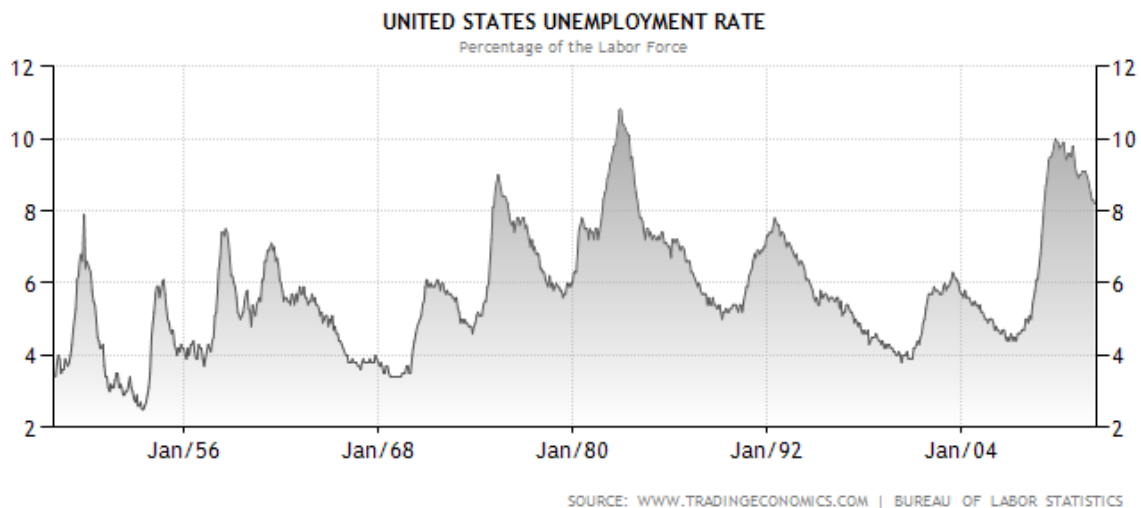
Ways in which this study differs from those done previously on the topic: First, this study's main focus is a cross-sectional regression analysis of Unemployment on a regional basis. This has not been done before; the only analysis that has been done regionally is on unemployment benefits and insurance (which has no relevance here). The literature reviews and prior studies done on the U.S. were nationwide analyses or descriptive statistic studies. Second, it is done on the United States; other studies have been done on regional unemployment for both Spain and Italy in the past decade however, nothing on factors influencing U.S. regional unemployment. Third, it uses a combination of independent variables from prior studies, as well as two variables that I determined through my own research could show to be important. Although these differences made data collection much more challenging, it will hopefully fill a gap of missing research.

The rest of the paper will be presented as follows: 2.0 Unemployment Trends, 3.0 Literature Review, 4.0 Data and Empirical Methodology, 5.0 Empirical Results, and 6.0 Conclusion. Included in the conclusion are policy recommendations and data limitations.

2.0 UNEMPLOYMENT TRENDS

Since American Industrialization, and even earlier, we have seen wide fluctuations in unemployment levels as a nation as a whole, and on a state-by-state basis as well. To get a grasp on the context in which the paper works within, unemployment trends and relevant current events will be discussed. According to Trading Economics (their information gathered from the Bureau of Labor Statistics graphed below), “The unemployment rate in the United States was last reported at 8.2 percent in March of 2012. From 1948 until 2010 the United States' Unemployment Rate averaged 5.70 percent reaching an historical high of 10.80 percent in November of 1982 and a record low of 2.50 percent in May of 1953.⁹”

Figure 1: U.S. Unemployment Rate: 1948-2012

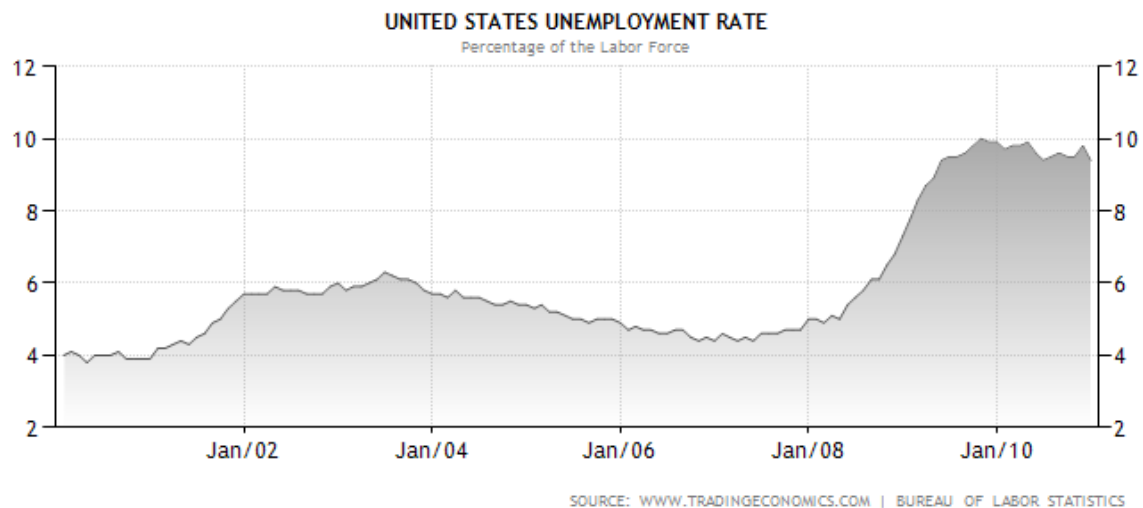


Source: Trading Economics⁹

With a low of 2.5 and a high of 10.8, from an outsiders view it shows that the wide fluctuations in unemployment are hardly being managed by the government, not that it is an easy indicator to manage. However, its more than just policy that goes into managing the

unemployment rate. There are unmeasurable factors that can greatly influence employers decisions of whether or not to hire any new employees or even how many employees must they lay-off. In the graph below you can see the effects of two of these types of factors. The 9/11 terrorist attacks that fueled the Iraqi war triggered a spike in unemployment beginning in the end of 2001 and lasting until the end of 2006. Shortly after that time period the United States entered into recession starting in 2008 that triggered the highest levels of unemployment since 1982.

Figure 2: U.S. Unemployment Rate: 2000-2010



Source: Trading Economics⁹

Another factor yet to be taken into account is the ever increasing amount of global competition. More and more firms seem to be relocating their operations overseas or outsourcing branches or divisions of their company elsewhere. Firms do this clearly for financial purposes (occasionally for regulatory reasons), however it hurts our economy as a whole by taking jobs away from people here in America and sending them to other countries to do the

work for less pay. This shows that there are a tremendous amount of factors that have influenced the levels of unemployment.

3.0 LITERATURE REVIEW

Overall eight papers were used to get an understanding of studies previously done on unemployment. Of these papers some were regression analyses, some were descriptive statistics, and two were more for understanding the implications of unemployment persistence. The data driven papers were used collaboratively and combined with a base unemployment knowledge to develop the adapted model used for this study.

Clemente, et al (2005) used a time series study of unemployment rates in the United States to determine the structure of unemployment. Their results showed a rejection of a presence of a unit root in the variables and showed how cyclical U.S. unemployment is. The paper was useful in determining how to avoid my data being affected by frictional or seasonal unemployment. Prakash (2012) wrote on the relationship between unemployment, mortgage rates, and housing prices. This paper was used to evaluate possible variables to use for the model. However, the paper used unemployment as a factor for housing prices in their test for Granger Causality, whereas my paper is more interested in what influences unemployment (not what unemployment influences). Shimer (1998) used a structural interpretation to the relationship between demographics and aggregate unemployment in the U.S. while studying why the U.S. unemployment rate is so much lower than other countries. His paper supports the idea that the increase in the aggregate unemployment rate has mostly been due to the baby-boom generation getting older. He evaluates several variables such as education, race, and age to support his theory that this generation caused a general 2% rise in the unemployment rate.

Hornstein, et al (2011) wrote about the causes and implications of increasing long-term unemployment. Some of the main causes include extensions of unemployment insurance, low unemployment exit rates, and the effect of unemployment duration on job prospects. This study was used to understand some of the critical influences on extended unemployment to better know how to evaluate unemployment data obtained.

Romero-Avila (2008) set out to evaluate the hysteresis hypothesis for regional unemployment in Spain. Hysteresis hypothesis suggests that in periods of economic downturn those who lose their jobs and become unemployed become discouraged and lose some of their skills. Once the economy turns around these people are left behind and not included in the prosperity of everyone else and remain in long-term unemployment. His study supports this hypothesis of unemployment hysteresis in Spanish regions. However, Lanzafame (2010) did a similar study on Italy and produced opposite results; his results reject the pure hysteresis unemployment hypothesis. This shows the volatility of unemployment based on not only country, but region within the country. This confirmed to me that I could not just trust the results of a prior study, nor expect to see the same results in my study.

Lopez-Bazo, et al (2005) studied the inequalities of unemployment in regions of Spain and this is where I obtained my original model from. The regression analysis looked into variables including: employment growth, % employment in industries, male and female participation rate, migration, and more. Results showed that in all models that were run the male and female participation rates were both significant, as well as employment growth. This study was crucial to obtaining a full understanding of a few of the factors affecting unemployment. Murphy and Payne (2003) wrote a similar study on a regional approach to the U.S.

unemployment rate; however, it was for the years from 1960 to 1996. This showed me that my topic was the perfect study to fill the gap of information on Unemployment. Murphy and Payne (2003) showed the relationship between unemployment rates by regions of the country and by state. Their study was used collaboratively with Lopez-Bazo, et al.'s study to piece together the model for this study. Murphy and Payne (2003) used variables like relative wage, % with 4yr college degree, incarceration rate, unionization rate and more. All eight of these papers were used together to develop an extensive model and to grasp an understanding of the previous work in the field of unemployment.

4.0 DATA AND EMPIRICAL METHODOLOGY

4.1 Data

This study uses annual U.S. data for the years 2000 and 2010. The data was obtained from the U.S. Census Bureau and the U.S. Bureau of Labor Statistics. The data used was gathered on a state-by-state basis for the two years under evaluation; no states were omitted. These years were chosen based on the availability of information due to censuses being distributed every 10 years. Summary Statistics for the data are provided in Table 1 and Table 2 for 2000 and 2010, respectively.

Table 1 – Summary Statistics (2000 Data)

Variable	Obs.	Mean	Median	Min.	Max.
UR	50	3.80	3.88	2.30	6.20
STTXS	50	5.00	4.65	0.00	7.00

COLLEGE	50	15.6%	15.4%	8.9%	22%
FEMLF	50	62.3%	61.8%	51.3%	70.3%
MALLF	50	75.6%	74.9%	63.5%	82.0%
POPPOV	50	11.2%	11.9%	6.5%	19.9%
POPDEN	50	82.28	153.50	0.95	964.80

Table 2 – Summary Statistics (2010 Data)

Variable	Obs.	Mean	Median	Min.	Max.
UR	50	8.76	8.65	3.90	14.90
STTXS	50	5.0	5.7	0.0	8.3
COLLEGE	50	27%	26%	17%	38%
FEMLF	50	60%	59%	48%	68%
MALLF	50	71%	72%	61%	78%
POPPOV	50	13%	13%	8%	21%
POPDEN	50	164.2	89.3	1.1	1008.1

4.2 Empirical Model

Due to the nature of the study, the exact same model was used for both 2000 and 2010.

The model for the two regressions is as follows:

$$UR_{(per\ state)} = B_0 + B_1(STTXS) + B_2(COLLEGE) + B_3(FEMLF) + B_4(MALLF) + B_5(POPPOV) + B_6(POPDEN) + \epsilon$$

Here the dependent variable is the Unemployment Rate of each state and is represented by $UR_{(per\ state)}$. The Unemployment Rate is measured through dividing the number of people unemployed by the number of people in the labor force. There are a few flaws by the nature of this measurement. The UR doesn't take into account those people who have given up looking for jobs due to long-term unemployment and choose to leave the labor force. It also doesn't consider those people deemed as underemployed. Underemployed classifies the people who wish to work full-time, however only currently work part-time.

The model consists of six independent variables that were determined through both the literature reviews and research of available data. See Appendix A and B for descriptions, sources, and expected signs of both the dependent and independent variables. The first variable STTXS represents state sales taxes, this variable was used because the lower the tax the more incentive there is to buy goods. The rationale behind this is that with more goods being purchased business should be able to afford to hire more people and then in-turn lowering the unemployment rate. Second, COLLEGE represents the percent of state population with a four year college degree. Third, FEMLFP stands for female labor force participation rate. It was calculated through dividing total females in the labor force by total female population per state.

The fourth variable, similar to the third, is MALLF, which stands for male labor force participation rate and was calculated the same way. Fifth, POPPOV represents the percent of population living below the poverty line. This was obtained from the Census website directly for 2010, however for 2000 it had to be calculated based on the 2000 poverty income level and comparing to 2000 census data on number of people based on household income breakdown. And sixth, POPDEN represents the population density per state and was calculated by dividing the total state population by the number of acres² in that state.

5.0 EMPIRICAL RESULTS

The main objective of this study was to determine which factors, both social and economic, influence the levels of unemployment between U.S. States. The regression results for each year are listed separately below.

Table 3 - Model Regression 2000 (R-squared = 0.4228)

<u>Independent Variables</u>	<u>Coefficient</u>	<u>T-Stat</u>
STTXS	-0.0572	-0.855
COLLEGE	-3.3646	-0.516
FEMLF	-0.2798	-0.051
MALLF	1.1981	0.199
POPPOV	15.8204**	2.418
POPDEN	0.50648**	2.026
C	1.9604	0.422

Note: ***, **, and * denotes significance at the 1%, 5%, and 10% respectively.

Shown above, in Table 3, are the results from the regression run on data from the year 2000. Due to the complex nature of unemployment, and the magnitude of factors truly determining unemployment levels, the results are somewhat conflicting. According to the data for 2000, labor force participation, education, and taxes did not have a significant influence on the levels of unemployment. This is shown through the t-stat values being less than 2 and the probability of error being much higher than 10%. However, results show that poverty levels and population density both being statistically significant, with both variables having significance at 5%, with t-stats over 2.0 and positive coefficients. This shows that as poverty levels and population density increases, it will cause an increase in the unemployment rate. Due to the significantly greater coefficient of poverty levels, the unemployment rate is more sensitive or subject to variation with a change in poverty, more than a change in population density.

Table 4 - Model Regression 2010 (Rsquared = 0.459)

<u>Independent Variables</u>	<u>Coefficient</u>	<u>T-Stat</u>
STTXS	0.6856**	2.0160
COLLEGE	-2.2394	-0.2517
FEMLF	-33.8603***	-2.9795
MALLF	10.0254	0.7652
POPPOV	-7.4295	-0.4683
POPDEN	0.0017	0.9306
C	22.2909	2.5812

Note: ***, **, and * denotes significance at the 1%, 5%, and 10% respectively.

Shown above, in Table 4, are the results from the regression run on the data from the year 2010. This was a standard OLS regression of a cross-sectional analysis and done using the same exact method as the regression in Table 3. Although the regressions were run identically with the same independent variables, the results came out completely different from one another. This could be due to many reasons, including the massive economic shocks our market endured between the September 11th terrorist attacks in 2001 or the economic recession ranging between 2007 and 2009.

The results show that the independent variables of education, poverty, population density, and male labor force participation were statistically insignificant for 2010. However, female labor force participation had a significant influence on decreasing unemployment for 2010. The data says that the more females that participate in the labor force between states then the lower unemployment would become there. And this is statistically significant at 1%, which shows the strength of the influence the independent variable has on the dependent one. State Taxes were also significant at 5% which shows that the level of taxes between the states has a strong influence on unemployment levels. These results are extremely interesting in comparison to the 2000 data due to the discrepancy in what is statistically significant. Not to mention the fact that one would believe that educational attainment would have a significant impact on the unemployment levels. However, after the analysis of results and further research one can see that just because you don't have a higher education doesn't mean you have a higher chance of being unemployed. There would be a strong connection to salary, however there are always jobs out there for different levels of skill, which is why educational attainment would appear to be statistically insignificant.

6.0 Conclusion

Unemployment is a complex indicator that has far reaching effects throughout an economy. Many scholars have tried to determine the factors that influence unemployment; however, the results seem to fluctuate depending on certain variables. Prior studies contrasted on country of analysis, time period in question, and variables tested. In a previous study done on Regional Spanish Unemployment, Lopez-Bazo, et al. (2005) determined that female labor force participation rates and job growth had high statistical significance. However, Murphy and Payne (2003), in their study on U.S. unemployment, found significance in the incarceration rate and percent of population with college degree. This shows that although significance in data can be established, unemployment studies have historically shown contrasting results.

Interpreting appropriate policy recommendations after obtaining contrasting results can be quite challenging. However, it can be said that these discrepancies were caused by the severe market shocks that the U.S. experienced; including 9/11 and the global recession. The results however are extremely useful in determining how to target unemployment in good or poor economic times. The first regression run shows that in times of growth the unemployment levels are most sensitive to population density and poverty levels. Knowing this, individual states can target poverty and over-crowding to reduce their unemployment. The second regression shows that, in times of large recession, each state could lower their sales tax and that will in turn slowly reduce unemployment. Also, each state can focus on increasing the amount of women participating in the workforce to help lower the rate further.

This study experienced several data limitations with respect to including more independent variables. Due to an unavailability of data for specific variables (such as: job

growth by state, average state corporate tax rate, % long-term unemployment is of total unemployment per state, and more), the study was unable to include these factors in the regression.

This study took into account six variables for determining influences leading to differences in regional U.S. unemployment. The variables being evaluated were state taxes, male and female labor force participation rates, population levels in poverty, population density, and percent of population with 4-year college degrees. For the year 2000, percentage of population in poverty and population density had significant influence on unemployment. Then in 2010, state taxes and female labor force participation rates had significance at 5% and 1%, respectively. The discrepancy of the results in the cross sectional analysis is a testament to how complex and unpredictable unemployment is.

Appendix A: Variable Description and Data Source

Variable	Description	Source
UR	U.S. Unemployment Rate by State	U.S. Dept. of Labor Statistics
STTXS	State Sales Tax	http://www.usa-sales-use-tax-e-commerce.com
COLLEGE	% of population with 4 year college degree	U.S. Census Bureau
FEMLF	% of female population in the labor force	U.S. Dept. of Labor Statistics
MALLF	% of male population in the labor force	U.S. Dept. of Labor Statistics
POPPOV	% of population under the poverty line	U.S. Census Bureau
POPDEN	Population per square acre by state	U.S. Dept. of Labor Statistics

Appendix B: Variables and Expected Signs

Variables	Description	Expected Sign	Rationale
STTXS	State Sales Tax	+	The higher the tax on goods the less people will buy and the less employees companies will hire
COLLEGE	% of population with 4 year college degree	-	The more people with 4 year degree the lower the unemployment rate should be
FEMLF	% of female population in the labor force	-	The more women in the L.F. the lower the U.R. should be
MALLF	% of male population in the labor force	-	The more men in the L.F. the lower the U.R. should be
POPPOV	% of population under the poverty line	+	The more people below the poverty line the more unemployed there will be
POPDEN	Population per square acre by state	+	The denser the population gets the harder it should be to get a job therefore increasing unemployment

Bibliography

1. Clemente, J., Lanaspa, L. and Montanes, A. (2004) The Unemployment Structure of the US States, *The Quarterly Review of Economics and Finance*, Vol. 45, Issues 4-5,pg. 848-868
2. Hornstein, A., Lubik, T., and Romero, J. (2011) Potential Causes and Implications of the Rise in Long-Term Unemployment, *Richmond Fed Economic Briefs*, Sep2011, Vol. 11, Issue 9, pg. 1-5
3. Lanzafame, M. (2010) The Nature of Regional Unemployment in Italy, *Empirical Economics*, Dec2010, Vol. 39, Issue 3, pg. 877-895
4. Lopez-Bazo, E., Del Barrio, T. and Artis, M. (2005) Geographical Distribution of Unemployment in Spain, *Regional Studies*, Vol. 39, Issue 3, p.305-318
5. Murphy, K. and Payne, J. (2003) Explaining Change in the Natural Rate of Unemployment: A Regional Approach, *Quarterly Review of Economics & Finance*, Vol. 43, Issue 2, p.345-369
6. Prakash, D. (2012) The Association Between Unemployment, Mortgage Rates, and House Prices. *Advances in Management*, Feb2012, Vol. 5, Issue 2, pg. 22-27
7. Romero-Avila, D. (2008) On the Persistence of Spanish Unemployment Rates, *Empirical Economics*, Aug2008, Vol. 35, Issue 1, pg. 77-99
8. Shimer, R. (1998) Why is the U.S. Unemployment Rate So Much Lower?, *NBER Macroeconomics Annual 1998*, Vol. 13, pg. 11-74
9. Trading Economics, April2012 <<http://www.tradingeconomics.com/united-states/unemployment-rate>>