Analyzing the Effects of Unemployment on Political Polarization in

New York State

Abstract

In this empirical research I attempt to investigate possible causation between the level of unemployment and the level of political polarization in New York State by county. I hypothesize that the increase in unemployment will lead to higher levels of political polarization following the intuition of political fanaticism grounded on economic distress. The method used in this study to measure the level of political polarization is by dissecting New York State voter enrollment by county and party affiliation, then comparing the number of votes enrolled for polarized parties relative to moderate parties resulting in a polarization index. The changes in the polarization index for all of New York State's 62 counties is examined relative to the unemployment rate over 48 time periods ranging from November 1996 to November 2019. The time it takes for the cumulative unemployment rate to have an effect on the behavior of voters is also analyzed. The outcome of my research repudiates the original hypothesis, as the results show that increased unemployment leads to decreased polarization, citing voter disengagement from smaller polarized parties in favor of larger moderate ones during times of economic distress and financial uncertainty. Results are more statistically significant when incorporating a lag of at least six months between the recorded polarization index and the unemployment rate.

JEL Classification: D72

Mohammed Al-Khulaifi

Georgetown University School of Foreign Service

Advisors: Professors Daniel Westbrook & Zhaoyang Hou

25 April 2020

I. Introduction

It is imperative to begin by differentiating between the different forms of political polarization before proceeding with the intricacies of how economic factors influence it. For the majority of the political discourse regarding polarization in the United States (US), the context concerns polarization across political parties, or *partisan polarization*. Different parties across the US satisfy varying social and moral philosophies ranging from liberalism, moderatism and conservatism; polarization with regards to these three categories refers to *ideological polarization*. David Autor (2017), professor of economics at the Massachusetts Institute of Technology and his co-authors attack the issue driving political polarization by viewing it through the scope of import competition in their 2017 article "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure". Autor et al. (2017) primarily contribute to the realm of economic effects on polarization by confirming that the "vitriolic campaign rhetoric" seen in the 2016 election was as a result of economic strain. Vitriolic in this context can encompass Donald Trump's far right-wing views regarding immigration in the 2016 presidential campaign, not to mention his stark isolationist views on trade illustrated by demonizing China. David Autor, along with his fellow researchers, have shaken the traditionalist views regarding cheap labor and international trade widely welcomed among economists. This is understood when Autor et al. conclude that areas that are greatly affected by Chinese-induced trade shocks and the subsequent higher level of import competition face economic strain and have resultingly become increasingly right-wing Republican. Moreover, congressional voting becomes more ideologically extreme, with the share of moderate Congress members decreasing (Autor, 2017).

Going a step further, Carl Benedikt Frey from Oxford University as well as his fellow researchers Thor Berger and Chinchih Chen follow up on Autor's work with respect to labor markets and political polarization. In their 2018 article "Political machinery: Did Robots Swing the 2016 US Presidential Election?" Frey et al. conclude that Democrats would have been left with the majority in the Electoral College during the 2016 presidential election if not for the adoption of machinery in certain labor markets. Capital intensity, automation and computerization of the global manufacturing sector has increased in an unprecedented manner, signaling new challenges with the structural shift of employment it potentially entails. In conjunction with this increase, we can also see a notable decrease in employment in the manufacturing sector. Michigan, Pennsylvania and Wisconsin represent crucial swing states and Donald Trump obtained the support of voters who are likely to be affected by the onset of machinery in their job prospects (Frey, Berger, & Chen, 2018). As of 2018, approximately 10.5 percent of the US population is employed in manufacturing (West & Lansang). This signals a major portion of the US population that may be concerned with losing their jobs due to automation. When isolating data from New York City, in January of 1990, approximately 279,000 employees were active in the manufacturing sector. Skipping to August 2019 and since then, there has been a steady decline in manufacturing employment down to approximately 68,000 (Federal Reserve Economic Data [FRED], 2019).

Edward J. López, a Distinguished Professor of Capitalism at Western Carolina University, and Carlos D. Ramírez (2004) from George Mason University further support the claim that increased polarization is attributed to economic distress in their joint article “Party Polarization and the Business Cycle in the United States.” Using the standard Downsian spatial competition model, they draw the following conclusions: “(1) inflation should be associated with policy convergence, (2) unemployment should be associated with polarization, (3) the effect of unemployment on polarization should be larger in magnitude than the effect of inflation on convergence, and (4) the effect of unemployment on polarization should be stronger in the House than in the Senate” (Lopez & Ramirez, 2004). Cyclically reoccurring recessions (and subsequent rise in unemployment) due to the business cycle in the US can lead to cyclical rises in polarization. Moreover, López and Ramírez find that a particular polarizing issue for US legislators involves the best way to carry out fiscal stimulus packages. While Democrats prefer greater spending, Republicans prefer tax cuts. This signifies an increase in polarization among the House and the Senate during times of high unemployment (Lopez & Ramirez, 2004). This study provides another dimension from which to examine polarization- namely within the House and the Senate, as polarization sourced in the bicameral system in the US is bound to have an effect on polarization among voters.

Looking internationally, unemployment-induced polarization seems to remain consistently imminent. In her article “On the Determinants of Political Polarization,” Daryna Grechyna (2016) from Middlesex University in London uses a Bayesian model in a sample of 66 countries to deduce the following: lack of trust as well as increased inequality both lead to an increase in polarization. Moreover, the study concludes that “political polarization is a socio-historical and an economic phenomenon” (Grechyna, 2016). Understanding the effect of inequality on polarization will help when drawing out proper controls for the regression later on. Inequality represents another form of economic distress that may polarize voters, as does increased trade exposure, technological advancement in manufacturing as well as the oscillations of the business cycle. To single out the effect of unemployment and remove any biases, controls will have to be incorporated to the regression that can relatively address these various forms of economic distress.

Democratic candidates running for 2020 also cited the risks involved with automation on the ability to find jobs. Andrew Yang, a 2020 Democratic presidential candidate and former executive in the tech industry, stated that automation will replace jobs carried out by people and bring the US to unemployment levels similar to those seen during the Great Depression (Stevens, 2019). It is clear at this point that the notion of employment risk, whether it be due to trade competition or technological advancement, has become a *politically motivated* issue. Moreover, we can deduce that economic distress leads to a higher level of polarization as cited in the literature. This will be the basis for my hypothesis as I claim that increased unemployment should lead to increased polarization following the intuition of political fanaticism grounded on economic distress.

Employment risk and the fear of losing a job is an issue that can translate into a particular political motivation, reflected in voting behavior. Voting (and by definition a polarizing concept) is a behavioral phenomenon that could be induced by a culmination of experiences from a longer period of time. The New York State Department of Labor (2020) outlines in its unemployment code, that successful unemployment claimants in New York State are entitled to unemployment insurance of up to "1/26 of a claimant's highest quarter earnings in all covered employment during the base period used to establish the claim". Does the expiry of these unemployment benefits yield a stronger behavioral response reflected in polarized voting? A recently unemployed individual who is benefiting from unemployment insurance may not necessarily be in a distressed financial position that may induce polarized voting behavior. However, a jobless individual whose unemployment insurance has expired after 26 weeks may exhibit polarizing behavior, which would be reflected in his/her voting as a result. Voting periods in New York State are every six months between November and April and then every eight months from April to November (with the exception of a few periodic irregularities). A six-month period and an eight-month period are both longer than the duration of legally receiving unemployment benefits in New York State. Incorporating this lag variable will be crucial in uncovering if the expiration of unemployment benefits yields a stronger behavioral effect reflected in polarization.

I hypothesize that economically distressed individuals are more likely to vote in a polarizing manner- and the likelihood of economic distress is higher when an individual has no income. The New York State Department's unemployment insurance law makes relevant the notion of lags when looking into the behavioral effects of unemployment. The legal code presented in this case will later influence the generation of lagged variables in the regression.

Multiple economic factors attributed to economic distress can influence politically motivated voting. Autor et al. focus on trade competition, while Frey et al. focus on technological advancement and automation. In my research, I will attempt to aggregate these two economic factors under their presumed impacts on unemployment. Analyzing the effects of polarization through changes in unemployment serves the concerns of the aforementioned economic literature. While trade competition and technological advancement are distinct factors, what unifies them is the notion of job security risk. This is as jobs and job security represent the core fear that voters are unnerved by. As a result, the unemployment rate will serve as the main independent variable for this study.

Previous economic literature on polarization focuses a great deal on partisan polarization, I will go a step further and contribute to the literature by looking at ideological polarization and designating it as my dependent variable for one of the US' most economically influential states, New York. As such, New York State voters will be evaluated on their leanings towards Liberalism, Moderatism and Conservatism (ideological polarization) as oppose to leanings towards Democratic, Independent and Republican political parties (partisan polarization). The latter's prevalence in economic literature necessitates that ideological polarization is equivalently looked into when investigating economic effects on voter behavior.

II. Theoretical Motivation

The aforementioned literature by Autor, Frey and Grechyna look at partisan polarization and the research shows that as economic distress grows, polarization among voters rises amidst financial uncertainty. Increases in inequality, trade exposure and automation are some of the factors that can contribute to economic distress and financial uncertainty at the micro level. In hypothesizing the effects of unemployment on ideological polarization, I believe that the innate political components between partisan and ideological polarization tie them together. In that regard, I hypothesize that as the unemployment rate increases, so does financial uncertainty and economic distress, leading to an increase in ideological polarization.

In attempting to give an adequate control power over the effect of the unemployment rate on the polarization index, I will be incorporating the following controls: the population estimate, the number of adult arrests, the average employment and the average annual wage. The population estimate was chosen because it is important to control for the size of counties to get a better estimate of the regression. To put this into perspective, if one county records 10 votes and another county records 10 million votes, then it would be imperative to control for the population. I also decided to choose the number of adult arrests because crime has a big effect in swinging elections. Hence, it is advantageous to control for the crime level to see an isolated effect of the unemployment rate on the polarization index. The average employment number in counties was chosen as a control because there is a definitional difference between employment and unemployment. The unemployment rate only counts those who are actively searching for jobs but are not able to become employed- as a result, it is vital to control for the number of those employed as well because it will add an extra level of people that are usually not covered in the unemployment count. Finally, I also decided to control for the annual average wage because it is critical to isolate the effect of rich residents from the pattern of their voting.

This paper will explore the case of New York State, regarded as the US' third largest economy that not only excels in finance and banking, but boasts thriving manufacturing and technological sectors as well (Ross, 2019). The reason why New York has been chosen for this study despite its Democratic bias (high level of membership in the Democratic party) is due to its *lack of bias* with respect to political ideology. New York maintains evenness across all three political ideologies ranging from liberal, to moderate, to conservative. A 2014 poll from Pew Research Center shows that across all 50 states in the US, New York possesses an even share of favorability for liberalism, moderatism, and conservatism where approximately thirty percent of the sample is dedicated to each (2014a). This gives us a proper benchmark with which to conduct this research, as the lack of political ideology bias of New York State will limit any biases seeping through into voter registration data. The 62 counties of New York will be analyzed with respect to the unemployment rate and its effect on polarization over 48 time periods ranging from 1996-2019.

III. Empirical Specification

The data collection procedure adopted for this study has been through secondary research, that is, census data that is already available and has been retrieved from online government statistics databases. While some datasets have no real sampling error, such as the case with voter enrollment, other cases involve estimates and averages. The datasets for all variables have been collected for the 62 counties of New York across 48 time periods. Data on the dependent variable has been retrieved from New York State Board of Elections voter enrollment data, later processed into a "polarization index" per time period. Data on the independent variable, the unemployment rate, has been retrieved from US Bureau of Labor Statistics. Data on control variables including the population estimate, the number of adult arrests, the average employment and the average annual wage have been retrieved from the New York State Department of Labor and the New York State Division of Criminal Justice Services (see Table 1).

Table 1: Variables Summary

	Variable	Units	Description
Dependent Variable	Polarization Index	Percentage Change %	Ratio of voter share of polarized parties over the voter share of moderate parties
Independent Variable	Unemployment Rate	Percentage %	Proportion of labor force who are unable to find jobs despite being willing and able to work and despite actively searching for work
Control Variable	Population Estimate	Integer between 1 and 10,000,000	Total number of residents residing in a county
Control Variable	Number of Adult Arrests	Integer between 1 and 5,000,000	Total number of adult arrests in a county
Control Variable	Average Employment	Integer between 1 and 5,000,000	Total number of employed residents of a county
Control Variable	Average Annual Wage	$ value	Average annual wage of a resident in a county

The political aspect of this research necessitates qualitative methodology on top of the quantitative. Before initiating the regression analysis, a thorough review is done on the ideology of all New York States' political parties with automatic ballot access from the years 1996-2019. Each of these political parties are assigned under three categories: "Liberal", "Moderate", or "Conservative". This is done through an analysis of these parties' political ideology. Modern definitions of liberalism, moderatism, and conservatism in the US are greatly understood as the three varying levels of the acceptance towards change. While liberal parties seek progressive agendas far from the reach of traditional and religious institutions, conservative parties seek to conserve these institutions and keep change at a minimum. Moderate parties represent the majority of the US population and possess a mix of both liberal and conservative values. Take the example of the New York State Right to Life Party, whose main goal is to fight back the progressive changes seen across the US with respect to legalizing abortion, deeming it an inherently conservative party (see Table 2).

Table 2: Political Ideology of Parties in New York

Liberal Party	**Moderate Party**	**Conservative Party**
Liberal Party of New York	New York Republican State Committee	Conservative Party of New York State
Women's Equality Party	New York State Democratic Committee	New York State Right to Life Party
Marijuana Reform Party	Independence Party of New York	
Green Party	Reform Party of New York State	
Rent Is Too Damn High Party		
Working Families Party		

As a result of categorizing voters based on the liberal, moderate or conservative party they have voted for, we are effectively analyzing voter behavior based on their ideological leanings rather than their partisan preferences. From there onwards, the changes of the share of liberal, moderate, and conservative voters over time are evaluated. To help illustrate the level of polarization across the voters for a particular period of voter enrollment, a ratio of polarized parties over moderate parties is created, resulting in an index. The higher the value of the index, the higher the level of polarization. For example:

Total number of voters registered for this time period = 3 + 100 + 2 = 105

$Liberals = 2 + 1 = \mathbf{3}$,

$Moderates = 58 + 42 = \mathbf{100}$ $\therefore$ $\frac{polarized\ parties}{moderate\ parties} = \frac{3+2}{100} = \frac{5}{100} = 0.05 = 5\ Polarization\ Index$

$Conservatives = 1 + 1 = \mathbf{2}$

The voter registration data for the main dependent variable (polarization index) and the number of adult arrests enjoy no sampling error as they are census data collected by governmental agencies for the entire population of interest (each county in this case). In other words, no one is omitted from those numbers and hence there is no sampling error. However, there may be some shortcomings associated with the data and methodology when it comes to the unemployment rate, the population estimate, the average annual wage and the average employment. The unemployment rate, for instance, fails to take into account individuals who are discouraged from searching for jobs after months of unfruitful efforts, nor those who take a leave of absence dedicated to taking care of their children, the sick or the elderly. This is not to mention that the unemployment rate is not the best measure in terms of capturing underemployment in the economy. The average annual wage may also present possible anomalies in that it is a measure that involves averaging the state-account-based total wage bill by an average population of employees in New York State. The average employment presents the same issue as an average number is recorded for the entire employed population. Because an accurate count of every individual in a county occurs every 10 years in a decennial census, the population estimate is used instead. This is as the population estimate provides data on the population that is more readily available to use for every recorded voter registration period. However, the limitation associated with this is that estimates may not always be accurate. To address possible anomalies in the data, a method for dealing with outliers is incorporated into the regression analysis process. The UCLA Institute for Digital Research and Education presents a form of the robust regression in Stata statistical software ("rreg" command) that accounts for highly influential observations that are dropped as well as circumstances where observations with high absolute residuals being down-weighted (Statistical Consulting n.d.).

The regression equation must take into account potential biases that are innate within the regression following the discussion on the dependent, independent and control variables. The independent variable (the polarization index) is calculated by producing a ratio of polarized parties over moderate parties for every voting period. The problem that arises with this is defining the units. The literature boasts varying methods of calculating polarization and the various units associated with them. Considering my research, the polarization index observes a point system, ranging from 0 to an undefined upper bound. The method used in my case for measuring polarization is different than that of other political statisticians. In other words, there is no set form or methodology to conform to when attempting to measure ideological polarization. In attempts to standardize the way I measure polarization, I decided to transform the polarization index logarithmically to analyze a percentage change. To "standardize" in this context is to simply to push the emphasis from a point system to a percentage change so that there can be defined upper and lower bounds of 0% to 100% respectively. Secondly, the main independent variable, the unemployment rate, is very much a nonlinear phenomenon. As witnessed by the business cycle in the US, the unemployment rate rises and falls in a nonlinear fashion. In the case of this research, this is true, especially considering the fact that the panel goes through two decades of data. As a result, it is imperative to employ a nonlinear transformation of the unemployment variable to produce a model that will better fit the data. In the majority of cases with this type of research, linear models rarely boast a desirable fit as business cycles render these regressions nonlinear to begin with. As to the manner in which the unemployment variable will be transformed nonlinearly, two methods were considered. The given equation presents an example of a model where the unemployment variable is transformed nonlinearly by raising the term "Unemployment", denoting the unemployment rate, to the third power:

$$\begin{aligned} ln(Polarization\ Index)_{it} \\ &= \beta 0 + \beta 1 * Unemployment^3 + \beta 2 * Population\ Estimate_{it} + \beta 3 * Adult\ Arrests_{it} \\ &+ \beta 4 * Average\ Employment_{it} * \beta 5 * Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

This is an example of a model specification in which the exact nature of the regression curve is being determined by introducing an unconventional nonlinear transformation of the unemployment variable to the third power. Even if it were the case that this model would boast highly significant terms, it would be impossible to justify transforming one unemployment variable to the third as shown on the left. Based on the literature, there simply is not an abundance of data that could theoretically justify transforming one unemployment variable nonlinearly by raising it to the third power (or the fourth, fifth, etc.). In attempting to transform the unemployment term nonlinearly, this would not be the best method to do so. The following model presents an alternative method of transforming the unemployment variable nonlinearly:

$$\ln(Polarization\ index)_{it} = \beta 0 + \beta 1 * Unemployment_{it} + \beta 2 * Unemployment_{it}^2 + \beta 3 * Unemployment_{it}^3 + \beta 4 * Population\ Estimate_{it} + \beta 5 * Adult\ Arrests_{it} + \beta 6 * Average\ Employment_{it} + \beta 7 * Average\ Annual\ Wage_{it} + \varepsilon$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

In this model, rather than having one unemployment variable raised to the third, the model features a third-order polynomial transformation of the unemployment variable whereby all the necessary terms are incorporated. That is, unemployment, unemployment squared, and unemployment cubed. This is more conventional, it is used more widely in the literature and has a theoretically viable basis of the nonlinear transformation of the unemployment variable that can adapt to the deviations reflected in the real world. Going past a fourth-order polynomial presented a number of issues: regression trials produced inconsistent results and deemed the main regressors insignificant which render the models inadequate in estimating the given relationship. Not to mention that going past a fourth order polynomial would render the final regression complicated and too long given the scope of this research. As a result, the nonlinear transformation of the unemployment variable has been capped at the fourth-order.

Finally, the discussion on the importance of lags, made relevant by New York State's unemployment insurance policies, necessitates exploring the incorporation of lags in the unemployment variable when assessing its effect on polarization. The current manner with which the panel is structured is as follows: the voter registration period is associated with the unemployment rate from the month prior. For example, the polarization index observed for the month of April is tied to the unemployment rate from the month of March. This has been done to capture the cumulative effects of unemployment prior to the voting period, where the polarization index is constructed. However, behavior changes reflected in voting may be induced following a greater cumulative effect exposed from imposing a longer lag on the unemployment variable. Consequently, a new lagged variable of unemployment, representing a potential replacement for the original unemployment rate term, has been constructed:

$$Unemployment\ Lag_{it} = Unemployment_{n-1}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

To carry this transformation out, the panel was sorted by county and date. Taking the example of Albany county, the first period of observation was the 1st November 1996 and the second observation was the 1st April 1997. The lag variable for the 1st November 1996 will be blank since there is no period preceding it. The lag variable for the 1st April 1997 will take the observation for the 1st November 1996- and so on this lagged term of unemployment is constructed. In short, by generating a new unemployment lag variable, all the periods are effectively pushed back once. The observations in the panel are recorded in correspondence with the voting registration periods in New York State that typically take place every April and November. However, there are instances of irregular voting registration periods where observations have been recorded in February, March and June. This irregularity in periodicity calls for observations to be treated monthly, as a quarterly or yearly system would result in observations not being counted in the regression. The final two regressions to be tested are fundamentally different by means of the lag associated with the unemployment term and are as follows:

The first regression featuring no lagged term of unemployment. The unemployment variable will later be transformed nonlinearly as an n-order polynomial:

$$\begin{aligned} ln(Polarization\ Index)_{it} \\ &= \beta 0 + \beta 1 * Unemployment_{it} + \beta 2 * Population\ Estimate_{it} + \beta 3 \\ &* Adult\ Arrests_{it} + \beta 4 * Average\ Employment_{it} * \beta 5 \\ &* Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

The second regression featuring the unemployment term being lagged by one period. The unemployment variable will later be transformed nonlinearly as an n-order polynomial:

$$\begin{aligned} ln(Polarization\ Index)_{it} \\ &= \beta 0 + \beta 1 * Unemployment\ Lag_{it} + \beta 2 * Population\ Estimate_{it} + \beta 3 \\ &* Adult\ Arrests_{it} + \beta 4 * Average\ Employment_{it} * \beta 5 \\ &* Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

IV. Data Description

The summary statistics of the data are as follows (see Table 3):

Table 3: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Max
Unemployment Rate	2976	5.761458	1.90204	2.5	18.3
Polarization Index	2976	4.21108	1.88114	0.57946	20.51393
Population Estimate	2852	310510.1	526782.1	4434	2611232
Adult Arrests	2852	8697.256	18392.07	51	114490
Average Employment	2852	135147.3	320380.8	1573	2495088
Average Annual Wage	2852	36864.92	11134.44	18769	123518

The summary statistics table (Table 3) shows that the control variables have 2852 observations which are spread across the 62 counties in New York State and from a period of 1996-2018. For the dependent variable (Polarization Index), there are slightly more observations because of the additional data from 2019 and using the aforementioned modified robust regression deals with any possible outliers, hence, it should not bring forth any issues. Across all counties in New York State, the independent variable (the unemployment rate), averages at around 5.8 percent and varies from the mean with a standard deviation of around 1.9 percent. The highest unemployment rate of 10.3 percent was both in the counties of St. Lawrence and Steuben amidst the 2008 Financial Crisis. The lowest unemployment rate of 2.5 percent was recorded in both Montgomery County and Livingston County in the year 2000. As for the main dependent variable, the polarization index, for it averages at around 4.2 points and varies from the mean with a standard deviation of around 1.9 points. The highest ever polarization index was observed in the far-east county of Rensselaer in 2011 with a value of around 20.5 points. The lowest polarization measure was observed in the southern county of Nassau with an observed value of around 0.6 points.

The following figure shows the scatter plot highlighting the relationship between the independent variable (unemployment rate) and the dependent variable (polarization index). All the observations for these two variables have been plotted as shown (see Figure 1).

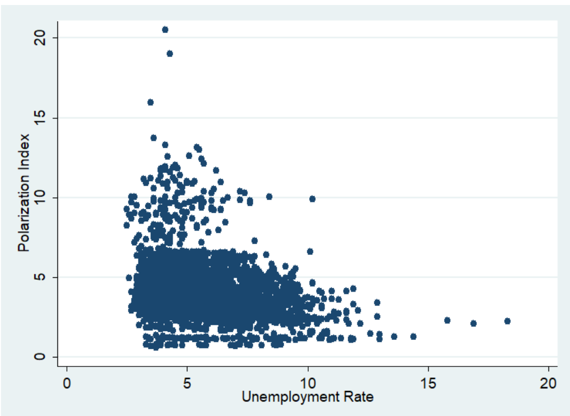

Figure 1. Stata-constructed scatter plot with the unemployment rate on the x-axis and the polarization index on the y-axis.

Figure 1 gives a general indication that as the unemployment rate goes down, the polarization index increases. The negative sign on the unemployment rate variable in the regression results will later confirm this. The unemployment rate appears to not decrease below the 2-3 percent level, this is the natural rate of unemployment observed in New York State. Therefore, the observations get clustered at around 4-8 percent which are observed to be the prevalent unemployment rates that we observe throughout the years 1996-2019. The empirical equations which are to be estimated are as follows:
The first regression featuring no lagged term of unemployment acknowledging that it will later be transformed nonlinearly as an n-order polynomial:

$$\begin{aligned} ln(Polarization\ Index)_{it} &= \beta 0 + \beta 1 * Unemployment_{it} + \beta 2 * Population\ Estimate_{it} + \beta 3 \\ &* Adult\ Arrests_{it} + \beta 4 * Average\ Employment_{it} * \beta 5 \\ &* Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

The second regression featuring the unemployment term being lagged by one period acknowledging that it will later be transformed nonlinearly as an n-order polynomial:

$$\begin{aligned} ln(Polarization\ Index)_{it} &= \beta 0 + \beta 1 * Unemployment\ Lag_{it} + \beta 2 * Population\ Estimate_{it} + \beta 3 \\ &* Adult\ Arrests_{it} + \beta 4 * Average\ Employment_{it} * \beta 5 \\ &* Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

V. Results

My regression analyzes the relationship between the unemployment rate and political polarization and by exclusively looking at ideological polarization. The literature focuses on partisan polarization, for which it cites an increase following a rise in unemployment sourced from various factors including trade exposure and automation. Due to the presumed political ties between ideological and partisan polarization, I hypothesized that an increase in unemployment will lead to an increase in ideological political polarization. The scatter plot shown in Figure 1 as well as the following regression models table (Table 4) reject this hypothesis as they both feature a negative relationship of the unemployment rate on the polarization index. For the first round of regressions seen in Table 4, the models all maintain a logged dependent variable (polarization index) to analyze the percentage change. All models in Table 4 also incorporate all the controls: the population estimate, the number of adult arrests, the average employment and the average annual wage (see Table 4).

Table 4: Regression Models

	Model 1 Ln(Polarization)	Model 2 Ln(Polarization)	Model 3 Ln(Polarization)	Model 4 Ln(Polarization)
Unemployment	-0.743***	-0.496**	-0.705	-2.73
	(-21.30)	(-3.11)	(-0.15)	(-1.94)
Population Estimate	-5.266	-4.511	-3.805	-5.444
	(-0.25)	(-0.21)	(-0.18)	(-0.26)
Adult Arrests	-7.875***	-7.816***	-7.776***	-7.631***
	(-10.80)	(-10.72)	(-10.66)	(-10.45)
Average Employment	1.280**	1.266**	1.223**	1.189**
	(3.17)	(3.13)	(3.03)	(2.94)
Average Wage	-13.26***	-13.37***	-13.34***	-13.41***
	(-15.49)	(-15.58)	(-15.54)	(-15.64)
Unemployment^2		-0.183	-0.751	4.60
		(-1.59)	(-1.25)	(1.67)
Unemployment^3			0.228	-4.12
			(0.95)	(-1.86)
Unemployment^4				0.119
				(1.96)
Constant	-2.298***	-2.371***	-2.467***	-2.014***
	(-64.22)	(-41.14)	(-21.85)	(-8.03)
N	2790	2790	2790	2790
R^2	0.362	0.363	0.364	0.365

t-statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model 1 incorporates one untransformed unemployment term. Models 2, 3 and 4 incorporate a nonlinear transformation of the unemployment term from a second-order polynomial transformation gradually increasing to a fourth-order polynomial transformation observed in Model 4. As expected from Figure 1, the unemployment rate has a negative effect on the polarization index as indicated by the negative sign on the unemployment coefficients observed in Models 1 through 4. Similarly, the population estimate, adult arrests and average wage terms all have negative effects on the polarization index from Models 1 through 4 and they are also significant. Only the average employment variable boasts a positive effect on the polarization index which is in line with the negative effect that the unemployment has on the polarization index. All average employment terms are also significant, maintaining a 1 percent significance level across all models.

In reference to the polynomial models in Table 4, the unemployment terms are not all consistent as we move down from Models 2, 3 and 4. While the adult arrests, average employment and average wage remain significant at least to the 1 percent level, the population estimate remains consistently insignificant throughout all models featured in Table 4. The largest t-statistic of the population estimate is seen in Model 1 and gradually decreases as we go from Model 1 to Model 3. While the insignificance of the population estimate may present limitations to the regression, it is theoretically considered to be the most important control variable, regardless of its significance. Removing the population estimate control from the regression would do more harm than if it were to remain– the reason being that it moderates the size of counties with respect to the number of votes registered. Removing the population estimate will give the same weights to counties with a large number of votes and counties with a small number of votes, relative to their estimated population.

In Table 4, Model 1 seems to show the most promise. All its terms are significant to the 1 percent level, with the exception of the population estimate. T-statistic values are relatively larger than those observed in Models 2, 3 and 4. However, the reluctance to move forward with Model 1 as a final model for the given data is the fact that unemployment is not treated as a nonlinear phenomenon. This presents a danger to the model, since the unemployment rate is never observed as a straight line, meaning that the final model cannot be considered accurate until unemployment is transformed nonlinearly. In Models 2, 3 and 4, the nonlinear transformations of the unemployment rate do not yield statistically significant results. Table 4 shows that without a lag-transformation of the unemployment variable, the regression with a nonlinear unemployment variable is yet to yield statistically significant results.

In the second round of regressions, all of the elements seen in Table 4 remain the same except for the unemployment terms that have been swapped with their respective lagged terms:

$$Unemployment, Unemployment^2, Unemployment^3, Unemployment^4$$

are swapped with,

$$Unemployment\ Lag, Unemployment\ Lag^2, Unemployment\ Lag^3, Unemployment\ Lag^4$$

where,

$$Unemployment\ Lag = Unemployment_{n-1}$$

The new regression table incorporating these lag variables is as shown (see Table 5).

Table 5: Regression Models Incorporating Lagged Unemployment Terms

	Model 1	Model 2	Model 3	Model 4
	Ln(Polarization)	Ln(Polarization)	Ln(Polarization)	Ln(Polarization)
Unemployment Lag	-0.572***	0.108	1.71***	0.0435
	(-15.73)	(0.64)	(3.49)	(0.03)
Population Estimate	16.60	14.84	17.81	17.01
	(0.76)	(0.68)	(0.82)	(0.78)
Adult Arrests	-9.259***	-8.987***	-9.027***	-8.926***
	(-12.29)	(-11.94)	(-11.99)	(-11.85)
Average Employment	1.446***	1.422***	1.386**	1.346**
	(3.42)	(3.36)	(3.27)	(3.18)
Average Wage	-12.94***	-13.21***	-13.24***	-13.24***
	(-14.28)	(-14.50)	(-14.55)	(-14.54)
Unemployment Lag2		-0.516***	-2.77***	0.575
		(-4.27)	(-4.39)	(0.20)
Unemployment Lag3			0.974***	-1.74
			(3.89)	(-0.75)
Unemployment Lag4				0.742
				(1.17)
Constant	-2.405***	-2.600***	-2.945***	-2.662***
	(-64.75)	(-43.48)	(-24.87)	(-10.03)
N	2727	2727	2727	2727
R2	0.330	0.335	0.336	0.337

t-statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 5, the same models are run as in Table 4, however this time the unemployment term is replaced with an unemployment lag term where the recorded periods of unemployment get pushed back one voter registration period. This seems to yield better results across all models when comparing to Table 4 where no lags were involved. The negative relationship between the unemployment rate and the polarization index is maintained even with the lags in the unemployment variables. The population estimate becomes positive across all models in Table 5 and this seems to be the more correct sign of the coefficient as the t-statistic values are larger when comparing to Table 4. The average employment maintains the positive coefficients across all models in Table 5 as expected due to the fact that as unemployment is correlated negatively with polarization, then employment would correlate positively. Average wage and adult arrests maintain the same negative sign on their coefficients. The population estimate, adult arrests, and average employment controls all boast an increase in t-statistic vales across all models in Table 5 when comparing to Table 4.

Again, when referencing Table 5, Model 1 cannot be considered as unemployment is treated linearly, making it an unrealistic intepretation of unemployment. In Table 5, when looking at Models 2, 3 and 4 in isolation, Model 3 stands out as the only model where all the unemployment terms in the third order polynomial are signficant to the 0.1 percent level. This is not the case in Models 2 and 4, as they show unemployment lag terms that are insignificant, meaning they cannot be considered.

Considering all the models in Tables 4 and 5, Model 3 in Table 5 shows the most promise. In this model, the three controls, adult arrests, average employment and average wage are all significant at least to the 1% percent level. The population estimate in Model 3, though insignficant, boasts the largest t-statistic value across all models in both tables. Model 3's R^2 value of 0.336 is well above average meaning that the model adequately explains the variability of the response data around the given mean. As a result, the final regression equation for this empirical research chosen based upon the overwhelming statistical significance seen in Model 3 from Table 5 and the equation is as follows:

$$\begin{aligned} \ln(Polarization\ index)_{it} &= \beta0 + \beta1 * Unemployment\ Lag_{it} + \beta2 * Unemployment\ Lag_{it}^2 + \beta3 \\ &* Unemployment\ Lag_{it}^3 + \beta4 * Population\ Estimate_{it} + \beta5 * Adult\ Arrests_{it} + \beta6 \\ &* Average\ Employment_{it} + \beta7 * Average\ Annual\ Wage_{it} + \varepsilon \end{aligned}$$

Where i represents the individual counties in New York State and t represents the time in voter registration periods in New York State from 1996 to 2019

Following the selection of the most suitable regression model, a series of cross-sectional data has also been analyzed (See Figure 2).

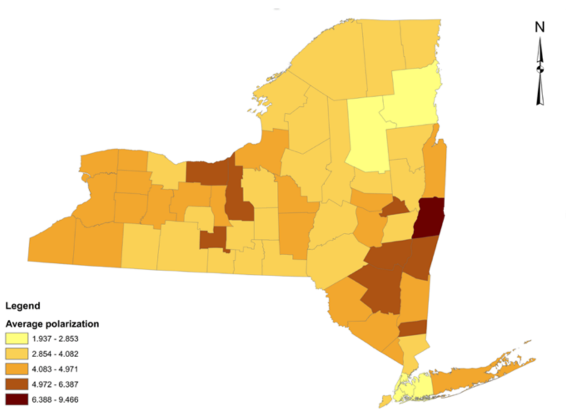

Figure 2. ArcMap-constructed geographic heatmap of New York State portraying the average polarization index from 1996-2019.

Figure 2 shows that counties with relatively higher levels of polarization are clustered around the far west and far east regions of the state on average. Northern counties boast relatively lower polarization when comparing to southern counties on average. Additional heatmaps were constructed to analyze the fluctuations of the polarization index over the course of the 2008 Financial Crisis. New York was at the epicenter of this crisis. The depreciation in the mortgage market on the national level soon led to the collapse of one of the most influential investment banks in the US, Lehman Brothers based in New York. The collapse of the bank arguably catapulted the crisis to the international level. An analysis of the polarization index throughout the period of the 2008 Financial Crisis reveals the following (see Figure 3):

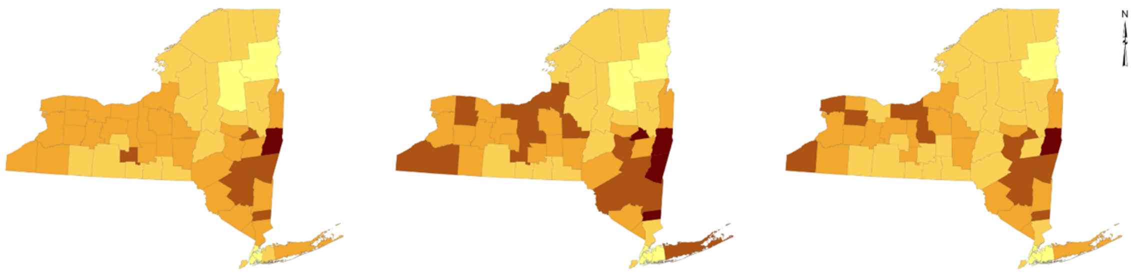

Figure 3. ArcMap-constructed heatmaps of polarization indices recorded in New York State counties throughout the period of the 2008 Financial Crisis

In Figure 3, the first heatmap on the left records polarization indices in April 2007, the second heatmap shown in the middle of the figure records polarization indices in November 2009 and the third heatmap shown on the right of the figure records polarization indices in April 2016. Figure 3 brings forth a major inference, the increase in polarization seen between the first two maps indicates that the events that occurred in the 2008 Financial Crisis resulted in the opposite observed effect on the relationship between unemployment and polarization. According to the data used for this empirical study, polarization is observed to decrease following an increase in unemployment. The period between the first two maps in Figure 3 saw record-levels of unemployment across counties in New York State following the events of the 2008 Financial Crisis. However, rather than a decrease in polarization, when comparing the second map to the first map in Figure 3, there appears to be a general increase in polarization in the mid, midwestern and southeastern regions of New York State. The cross-sectional data portrayed in Figure 3 deduces that national-level crises such as that observed in the 2008 Financial Crisis lead to a positive effect of unemployment on the polarization index. Moreover, the last map, recorded in April 2016, appears to show certain regions with permanent increases of polarization 8 years following the crisis- this is true for the mid, midwestern and southeastern regions of New York State.

In a cross-sectional analysis of the polarization index by adult arrests, there appears to be a 19 percent drop in polarization when comparing counties with the bottom 10 percent of adult arrests to counties with the top 10 percent of adult arrests (see Figure 4):

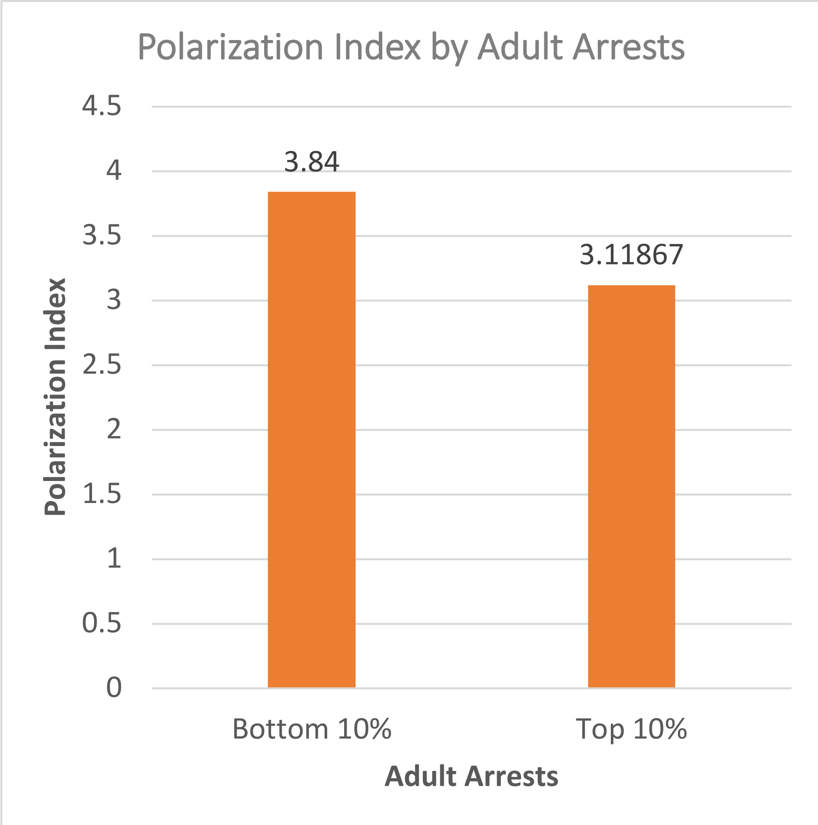

Figure 4. Polarization index by the top and bottom 10 percent of adult arrests observed in New York State counties.

The drop in polarization observed in the top 10 percent of adult arrests in all of New York State's counties aligned with the data observed in Table 5, which boasts negative signs on the adult arrests coefficient across all five models.

Singling out New York County, the polarization index is observed from 1996-2019 (see Figure 5).

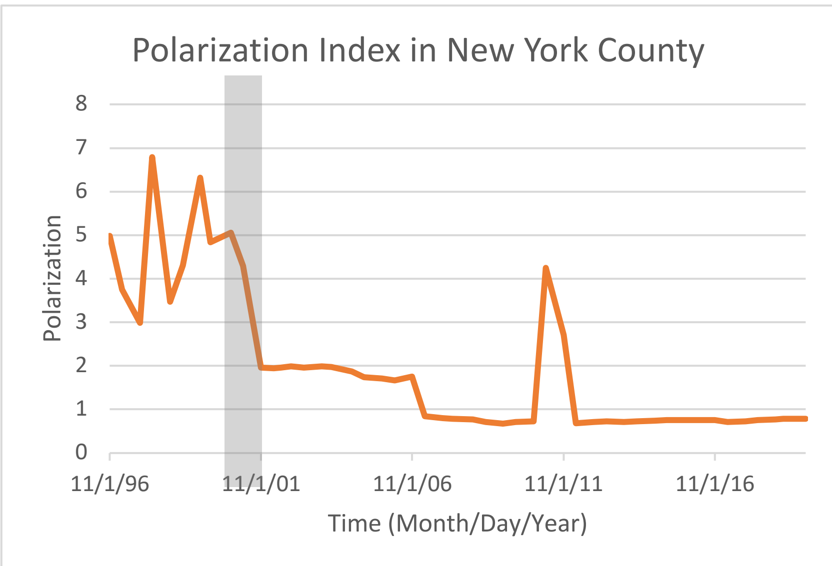

Figure 5. Polarization index in New York County from 1996-2019 where the shaded region marks the time period beginning and following the September 11 attacks.

The shaded region in the figure marks the beginning of September 11 attacks, with the months following. Within the shaded region, there appears to be a significant drop in the polarization index. A number of reasons could clarify this. Firstly, policy consolidation among liberal and conservative parties and their subsequent reconciliation with the policies of moderate parties can be a possible explanation. Narratives among political parties amalgamate under the notion of shock, recovery and mourning amidst an attack from a common foreign enemy. Moreover, the voters, who are at the core of this sudden drop of polarization observed in the shaded region in Figure 4, are unified as victims of this attack, led by a common purpose in overcoming a horrifying event that shook the nation.

VI. Discussion and Conclusion

In conclusion, the empirical research presented in this paper deduces that the relationship between the unemployment rate and the polarization index is a negative one, with brief instances of inflection as observed in the 2008 Financial Crisis. This repudiates the original hypothesis, which predicted that the relationship would be positive based on the literature and the assumed political links between partisan and ideological polarization. The results indicate that the differences between partisan polarization (observed in the literature) and ideological polarization (used for this empirical study) are greater than what was presumed prior to conducting this research. Moreover, a greater behavioral response reflected in voting is observed when analyzing polarization incorporating a lag period that surpasses the expiration of unemployment insurance benefits. This places emphasis on the fact that unemployed individuals in New York State are not necessarily incomeless, contingent on the qualification for unemployment insurance. An unemployed individual with unemployment insurance is not as inclined to vote differently. This is why when unemployment insurance income ceases after 26 weeks, as portrayed by the involvement of a lag variable within the regression, results appear to be more statistically significant.

While this empirical study looks at the case of New York State, there is still much to be explored with other states in the US as well as other countries that possess different unemployment policies, work cultures and political structures. Polarization appears to have effects that extend beyond the parameters of unemployment as seen in Figure 4. The 2008 Financial Crisis shows that ideological polarization can go against the trends observed in the data- which emphasizes the need to build upon trends seen in international-level economic phenomena such as trade exposure observed by Autor et al. In these international-level circumstances, ideological and partisan polarization move positively in unison. However, for the rest of the time periods observed in the subject of this analysis, ideological polarization disassociates itself from partisan polarization.

Policy implications of polarization have historically been viewed relative to the struggle between the Democrats and the Republicans- and even more so today. Rather than observing fluctuations in policies due to changes in partisan politics, this empirical research ultimately aims to shed light on the importance of looking into ideological fluctuations. To develop and improve upon this research, ideological categorization of political parties must be dissected further. Moderate parties represented in this study are assumed to have a combination of liberal and conservative qualities, acting as a compromise between the two through moderatism. Pew Research Center challenges this by deducing from a 2014 survey conducted in the US that “being in the center of the ideological spectrum means only that a person has a mix of liberal and conservative values, not that they take moderate positions on all issues” (2014b). Consequently, voter behavior must be examined in more detail in order to more accurately categorize voter results if one were to effectively capture the relationship between unemployment and polarization. While partisan polarization boasts a negative relationship with economic prosperity, ideological polarization appears to be positively correlated. The disassociation observed between the two forms of polarization have made conducting this research and estimating the model worthwhile, for not only does this study illuminate the differences between partisan and ideological polarization, but empirically substantiates it.

References

Autor, D., Dorn, D., Hanson, G. & Majlesi, K. (2017). Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure. *MIT*, https://economics.mit.edu/files/11665.

Federal Reserve Economic Data (FRED). (2019). All Employees: Manufacturing in New York City, NY, *FRED*, https://fred.stlouisfed.org/series/SMU36935613000000001SA.

Frey, C. B., Berger, T., & Chen, C. (2018). Political Machinery: Did Robots Swing the 2016 US Presidential Election? *Oxford Review of Economic Policy*, 418–42. https://doi.org/10.1093/oxrep/gry007.

Grechyna, D. (2016). On the Determinants of Political Polarization. *Economics Letters 144*, 10–14. https://doi.org/https://doi.org/10.1016/j.econlet.2016.04.018.

López, E. J., & Ramírez, C. D. (2004). Party Polarization and the Business Cycle in the United States. *Public Choice 121*(3/4), 413–30. https://www.jstor.org/stable/30026547

New York State Board of Elections. (2020). *Enrollment by County.* Albany, New York.

New York State Department of Labor. (2019). *Quarterly Census of Employment and Wages Annual Data: Beginning 2000*. Albany, New York.

New York State Department of Labor. (2020). *Unemployment Insurance Division, Unemployment Insurance Benefits - An Employer's Guide* §. Accessed April 22, 2020. https://labor.ny.gov/formsdocs/ui/IA318.2.pdf.

New York State Department of Labor. (2020). *Annual Population Estimates for New York State and Counties: Beginning 1970.* Albany, New York.

New York State Division of Criminal Justice Services. (2019). *Adult Arrests by County: Beginning 1970*. Albany, New York.

Pew Research Center. (2014a). *Political Ideology by State*. Pew Research Center's Religion & Public Life Project. Pew Research Center. https://www.pewforum.org/religious-landscape-study/compare/political-ideology/by/state/.

Pew Research Center. (2014b). Section 4: Political Compromise and Divisive Policy Debates. Pew Research Center - U.S. Politics & Policy. https://www.pewresearch.org/politics/2014/06/12/section-4-political-compromise-and-divisive-policy-debates/

Ross, S. (2019). "New York's Economy: The 6 Industries Driving GDP Growth." Investopedia. Dotdash, https://www.investopedia.com/articles/investing/011516/new-yorks-economy-6-industries-driving-gdp-growth.asp.

Statistical Consulting. (n.d.). "Robust Regression | Stata Data Analysis Examples." UCLA Institute for Digital Research and Education. UCLA. Accessed April 23, 2020. https://stats.idre.ucla.edu/stata/dae/robust-regression/.

Stevens, M. (2019). "Why Andrew Yang Says Automation Is a Threat to the Country." *The New York Times*. https://www.nytimes.com/2019/06/27/us/politics/andrew-yang-automation.html.

U.S. Bureau of Labor Statistics. (n.d.). "Local Area Unemployment Statistics Map." Washington, District of Columbia.

West, D. M., and Lansang, C. (2018). "Global Manufacturing Scorecard: How the US Compares to 18 Other Nations." Brookings. The Brookings Institution, https://www.brookings.edu/research/global-manufacturing-scorecard-how-the-us-compares-to-18-other-nations/.