Measuring County-Level Deindustrialization in the United States

Emily Lucinda Kurtz

University of Minnesota, Twin Cities

kurtz217@umn.edu

Abstract

Existing measures of county-level deindustrialization in the United States are often single numbers, typically capturing manufacturing job losses over a short time. This misses some nuance because job loss trajectories are not always linear and the starting point is consequential. I implement a time series clustering algorithm to categorize counties based on trends in manufacturing jobs over 53 years for the 2,316 counties with available data. Six distinct clusters represent the major deindustrialization trajectories. Declines have been quite sharp in some clusters, giving those counties less time to transition to a post-industrial economy. These cross-cluster differences are consequential to theories and past findings. For example, I find that TAA claims, a traditional measure of local deindustrialization, meaningfully impact individuals' views on investment only when they live in non-deindustrialized counties. In deindustrialized counties, views are homogeneous across demographics, except across partisanship, where Republicans are more supportive of investment.

Introduction

Deindustrialization is one of the, if not the, most discussed drivers of local economic downturns in the United States, and it can impact individuals' political and societal attitudes. The loss of local manufacturing jobs led to a decreased vote share for Democrats in the 2016 election (Baccini and Weymouth, 2021), a loss of faith in institutions like labor unions (Russo and Linkon, 2009), and an increased perception of the value of formal education (Winters et al., 2011). Feng, Kerner and Sumner (2021) found that those living in communities most exposed to trade-related job loss were more likely to oppose foreign, especially Chinese, business investment.

Deindustrialization is hard to measure, though, in part because it is hard to concretely define. Milner (2018) defines deindustrialization as simply "the relative decline of jobs in manufacturing." While this is fundamentally the definition I use throughout this paper, three things are worth noting. First, "manufacturing" itself can be hard to define, and there has been some push to focus less on the manufacturing label and more on the set of activities performed at a firm to determine whether the jobs done there qualify as manufacturing jobs (Bernard, Smeets and Warzynski, 2017). Second, while research often focuses directly on jobs lost (Baccini and Weymouth, 2021; Feng, Kerner and Sumner, 2021; Kerner, Sumner and Richter, 2020), the term originated as a way to describe the shuttering of factories, mills, etc., not directly as a state of job losses (High, 2020). Third, measuring the "relative" decline of jobs necessitates choosing a starting and ending point, and it does not give room to describe any patterns that occurred within the range. Particularly on this last point, current measures of deindustrialization are inadequate. Political economy research typically uses a single quantitative measure of the current state of the local economy, often a change in manufacturing jobs over a set time period, to summarize the level of deindustrialization a community has experienced. For example, Kerner, Sumner and Richter (2020) used per capita Trade Adjustment Assistance (TAA) claims for trade-related job losses over 16 years (2000-2016), and Baccini and Weymouth (2021) used Quarterly Workforce Indicators (QWI) to calculate the total number of layoffs in the manufacturing sector at the county level between 2004 and 2016.

Although this research has clearly produced valuable knowledge and these measures have been useful in answering important questions, these approaches have some limitations. TAA claims theoretically cover all jobs, not just manufacturing jobs, though in practice because these are jobs lost to trade, manufacturing jobs likely make up a large proportion of those covered. More notably, according to the U.S. Department of Labor, "a petition for TAA may be filed by a group of three or more workers, their union, or other duly authorized representative" (U.S. Department of Labor, N.d.). Thus, this measure likely misses situations in which very few jobs in a firm are lost to trade, workers are not aware of the fact that they can file such claims, or workers are not represented by unions. This last point could even lead to misleading geographical patterns to deindustrialization since states in the North typically have higher proportions of workers represented by labor unions than states in the South (U.S. Bureau of Economic Analysis, 2024). The QWI data used by Baccini and Weymouth (2021) are quite detailed, covering not just information on the jobs but also demographic information on the workers themselves (invaluable for the research questions they sought to answer). However, the data unfortunately cover a short time period and miss the first 25 years of the downward trajectory of American manufacturing jobs (Harris, 2020).

Additionally, calculating a single rate of change or net loss of jobs over time, even if that time period is quite long, misses a lot of intermediate detail. Communities with similar current economic states or changes over even more than 50 years can have vastly different histories overall. For example, Martin County, NC and Union County, NC have similar histories regarding manufacturing jobs according to traditional measures. Between 1979, the peak for manufacturing jobs in the United States, and 2010, when manufacturing jobs were at their post-WWII lowest (Harris, 2020), manufacturing jobs in Martin County fell from being 33% of total jobs to 14%, and from 32% to 13% in Union County. This represents an identical 19-point shift in both counties over that time period, and a very similar 43%

reduction in Martin County compared to Union County's 41% reduction. Figure 1 makes clear that these counties have not had quite the identical paths that these statistics imply. Union County started at a slightly higher rate of manufacturing in 1969, and Martin County had a notable, approximately decade-long rally in manufacturing jobs. However, by 2021, both had declined to 11%. In brief, Union County's jobs were more steadily lost, while Martin County's were lost more piecewise.[1] While Martin County's rally in the late 1980s and early 1990s was surely beneficial to the local economy while it was happening, it also comes with a drawback. Specifically, residents need only look back one generation to the "good times" when 4 in 10 people were employed in manufacturing jobs, whereas Union County residents had a more gradual loss of jobs to adjust to.

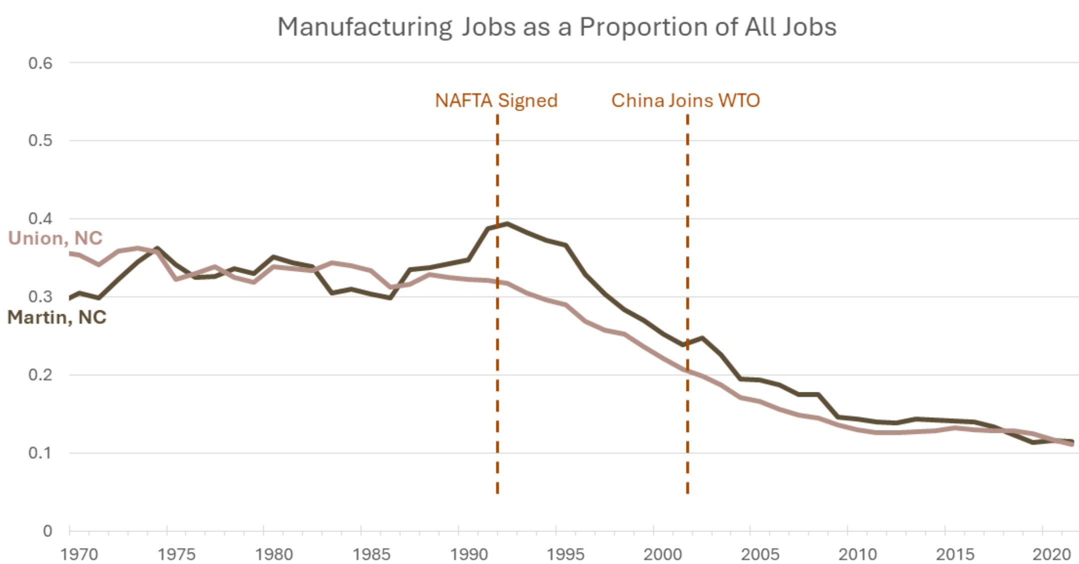

Figure 1: The proportion that jobs in the manufacturing sector make up of total jobs in Martin County, NC and Union County, NC.

As shown in the previous example, the actual pattern/shape to job loss matters. A

[1]These two counties are in fact clustered into two different groups by my algorithm; Martin County is an example of a "Temporary Holdout then Rapid Decline" county whereas Union County is a "Steady Decline" county. Steady decline should give counties time to adjust to a post-industrial economy, whereas piecewise decline, with brief, local peaks, could give counties a sense of false hope.

start and end point could be connected by a straight line. If this is the truth in reality, the community's job losses are evenly distributed over time, which could allow the community to acclimate and predict into the future. However, the start and end point could be connected in infinitely many other ways. The line may stay relatively level before plummeting, giving the community very little time to adjust to the drastic loss of manufacturing jobs. The proportion of jobs may even grow before again shrinking, giving younger workers a taste of what America's manufacturing heyday was like and making the loss of jobs even larger than it appears with the choice on arbitrary start and end points. These are but a few of the possible paths that communities' manufacturing jobs could take over time, and the impacts these few different paths could have on the economy of the community are already highly variable.

Looking at patterns over a longer historical time period may also help unearth new political theories. Not all communities are hit by deindustrialization in the same way, and this difference is not just quantitative (i.e. how many jobs were lost or firms were closed). Keeping a note of approximately when deindustrialization was severe in a community highlights the probable cause of the deindustrialization. A downward trend in the 1990s, for example, could imply that the community may be more inclined to blame their economic woes on NAFTA, whereas a decrease in the 2000s could be blamed on China joining the World Trade Organization. Feng, Kerner and Sumner (2021) evaluate skepticism of both foreign and Chinese investment, and with additional data, my method of measuring deindustrialization, and their general theory, future studies may find that counties that only experienced a substantial decline in manufacturing jobs in the 2000s are not necessarily antagonistic towards foreign investment broadly, but only towards Chinese investment. Counties with declines in the 1990s, on the other hand, may be skeptical of all foreign investment. In addition, these two major trade liberalization agreements began to have an impact at the beginning of the terms of two presidents of different parties, Bill Clinton and George W. Bush. Baccini and Weymouth (2021) study the impact that race and exposure to deindustrialization have on presidential vote shares. Residents of counties that began to lose manufacturing jobs in the 1990s may be more inclined to blame Democrats for local economic downturns broadly, whereas those from counties with substantial job losses beginning in the 2000s may blame Republicans.

To address the shortcomings of past methods of measuring local deindustrialization, I develop a categorical measure of US counties' manufacturing job trajectories in this paper. Instead of summarizing county-level deindustrialization with a single number, I categorize counties into groups based on the general shape of the rise and/or fall of manufacturing jobs in these individual counties over a period of over 50 years. I do this using hierarchical clustering of time series, with some modifications to deal with outliers. This process offers a longer and more detailed history of communities' reliance on manufacturing jobs over time that can, nonetheless, be represented by a single categorical variable for the purposes of quantitative analysis.

Data and Measures

I use data from the US Bureau of Economic Analysis (BEA). This organization collects jobs data categorized by Standard Industrial Classification (SIC) codes prior to 2001 (U.S. Bureau of Economic Analysis, N.d.*b*), and North American Industry Classification System (NAICS) codes from 2001 to present (U.S. Bureau of Economic Analysis, N.d.*a*). In total, I have data from 1969 to 2021, for a total of 53 individual years.

The tables include total job counts for each county (though data are missing in some cases) and for each year from 1969 to 2021. The data are also split by sector. For example, total jobs in mining, service, retail trade, farming, and more are included. While the data set includes total jobs for each county-year for a number of industries, I focus only on the rows containing manufacturing job counts and total job counts since this paper focuses on manufacturing job trends specifically. I create a new data set that is the number of manufacturing jobs for each county-year divided by the total number of jobs for that given county-year.[2]

There are 3,132 unique counties in the SIC dataset and 3,118 in the NAICS dataset. Once the tables are merged (combining only counties that are present in both the NAICS and SIC datasets), there are 3,110 counties. There are 3,142 counties currently in the country (United States Census Bureau, 2020). Counties do change over time; for example, the footnote of the NAICS dataset notes that the newest county, Broomfield County, CO, was established only in 2001 (U.S. Bureau of Economic Analysis, N.d.*a*). Given the combined data set has just shy of 99% of the counties in the country today, though, the data and merging process reflect today's county makeup well.

Once dividing the manufacturing jobs by the total jobs, I have 3,110 rows in my dataset. However, for the time series clustering step, only rows that have data for each of the 53 years can be used; after removing county-industry combinations with missing data, I have data on 2,316 counties, which is just shy of 75% of the total counties. Data are missing from BEA reports for two main reasons. Rarely, the county in question not does not exist for some time frame within the 1969-2021 range, like the aforementioned Broomfield County, CO. Far more often, counties with missing data either do not have a large enough number of workers in the specific manufacturing industry or have their data suppressed intentionally. At times, a sufficiently substantial number of workers are employed in the industry and the county's data are still suppressed to "avoid the disclosure of confidential information of business establishments that could be identifiable in metropolitan area estimates" (U.S. Bureau of Economic Analysis, 2008). Unfortunately, this implies the data are unlikely to be missing at random, and instead missing data are likely to be from counties with small populations

[2] There is an argument to be made for making total population, not total employed population, the denominator here. In essence, my approach makes "importance of manufacturing jobs on the overall job market" over time the measure, whereas using total population as the denominator would be measuring the importance of manufacturing jobs on all individuals in the community. I choose the former because of the large variety of age distributions from state to state and even county to county. Manufacturing jobs may seem less important to the community than they are in actuality in a county where, for example, much of the population is over 65 and retirement rates are high.

or with single, dominant employers. Fortunately, while these 794 counties represent about a quarter of all counties by raw count, they were home to 5.1% of Americans as of July 2022. This clustering method thus still captures the manufacturing jobs patterns of the counties that about 95% of Americans live in.

Methods

The goal of this paper is to group US counties based on their manufacturing job trends. I do this by tracking the proportion of the job force in each county that is employed in the manufacturing industry over several decades. Time series clustering enables this to be done relatively quickly and usefully for this high number of counties. Time series clustering, and clustering more widely, has a number of moving parts that one needs to get right in order to have sensible results. One must choose a reasonable clustering algorithm overall (including selecting a distance measure), have a strategy for determining the optimal number of clusters, deal with outliers if it is relevant, and have a way to evaluate model fit (Kotsakos et al., 2018).

Clustering Algorithm

This paper uses hierarchical clustering, specifically agglomerative hierarchical modeling, to sort counties into clusters based on the historical trajectories of their manufacturing jobs. Hierarchical clustering has a number of advantages. It is robust to outliers (Karna and Gibert, 2022), though purposefully identifying and accounting for outliers is still helpful, especially for interpretation's sake. Hierarchical clustering does not assume any knowledge about the data before being run (Madhulatha, 2012), which is helpful in this case because obtaining in depth knowledge of the 3,143 counties in the Unites States is quite impractical. This fact can be a double edged sword, making it easy to run the algorithm without understanding the process and the results, but used responsibly, these drawbacks can be avoided. Hierarchical clustering also can be quickly iterated to test different parameters. Calculating a distance matrix is a necessary step and can be slow if the number of observations is large, but once complete, making new clustering models with different numbers of clusters, different subsets of the data, etc. is quick. Finally, hierarchical clustering is popular because it is one of the best algorithms for understanding the underlying structure of the data being clustered (Karna and Gibert, 2022).

Agglomerative hierarchical clustering works by assuming each observation is in its own cluster and merging the two "closest" clusters at each step of the algorithm, until the last step, when all observations are merged into a single cluster (Murtagh and Contreras, 2012). The entire process can be visually represented using a "dendrogram," where two observations are more closely related if they are connected by a branch lower in the dendrogram and less closely related if they are connected by a branch higher in the dendrogram. When there is a relatively small number of observations, the dendrogram by itself is typically helpful in showing how the observations relate to each other. However, when observations are numerous, it is all but impossible to visually evaluate the dendrogram. Figure 2 is a dendrogram of manufacturing jobs in South Carolina counties. This number of observations is manageable, but the figure illustrates how difficult determining an underlying structure can become as the number of observations increases.[3]

Distance Measure

As mentioned, hierarchical clustering relies on a distance matrix. Most clustering algorithms can cluster based on many distance metrics beyond Euclidean distance, and hierarchical clustering is no different. The overall goal of clustering is to group observations so that there is minimal dissimilarity (or distance) among the observations within groups but maximal

[3]The figure also features two horizontal lines. These are cut points. If the dendrogram is cut at a height of 0.9 (represented by the gray dotted line), then the counties are grouped into three clusters. If the dendrogram is cut at a height of 1.2 (represented by the black dotted line), then the counties are grouped into two clusters. This is noted not as an argument for either of those numbers to be chosen as the number of clusters, but instead to illustrate further how the dendrogram can be used.

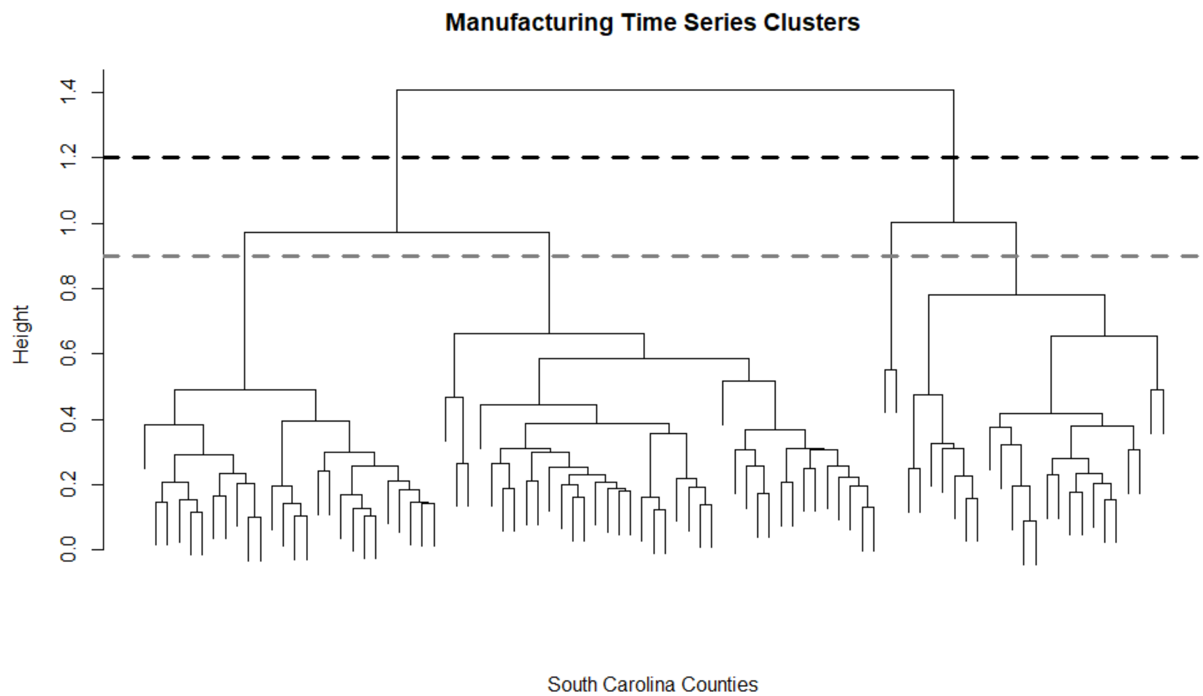

Figure 2: A dendrogram showing hierarchical clustering of time series of manufacturing jobs proportions in all South Carolina counties

dissimilarity between groups (Latifi-Pakdehi and Daneshpour, 2021). Dissimilarity can be measured in many ways. A widely used dissimilarity measure in time series clustering is dynamic time warping (DTW). DTW is popular for methodological and subject-specific reasons. First, DTW is somewhat robust to outliers (Brankovic et al., 2020).[4] Compared to the traditional Euclidean distance metric, DTW is likely to classify two time series with the same general shape into the same cluster even if there is a time lag between the two (Müller, 2007). This is appealing in many practical settings, including for this project. When we are interested in the long term political impacts of manufacturing jobs trends at the county level, it matters little whether the decline in manufacturing jobs began in 1976 or 1977. Instead, what matters in the general pattern overall: Were jobs steady over a period of time or did the county experience a decade of extreme volatility year to year? How

[4]I stress the word "somewhat" here, though, as the sensitivity to outliers is the driving force behind the development and use of the iterative algorithm of this paper.

quickly were jobs lost - i.e. how steep was the slope of the time series - when the decline began? Two counties that experienced extremely steep declines, albeit in different years, likely have more in common in response to those losses than two counties that experienced the losses over the exact same years, but at vastly different rates.[5] For example, a quick decline may be a sign that the county did not have time to transition to another source of jobs. In addition, a drastic, instead of steady, decline means that the "good old days" when manufacturing jobs were plentiful and the local economy was good are more recent and relevant in residents' memories. Because of these methodological and substantive strengths, I use DTW to evaluate inter-observation similarity here.

Determining the Number of Clusters

The decision of how many clusters to include is often more of an art than a science when using an unsupervised clustering technique like hierarchical clustering. Techniques range from fully qualitative processes like expert judgement, to relatively quantitative processes like the elbow method, to fully quantitative methods like those proposed by (Patel, Sivaiah and Patel, 2022). All of these strategies are valid, though some may be more applicable than others in certain contexts. Moreover, the strategies can be used in combination.[6] My algorithm relies heavily on the elbow method, which calculates the Within-Cluster-Sum of Squared Errors (WSS) for a range of numbers of clusters (Jaslam et al., 2022). As the number of clusters grows, the WSS will necessarily shrink. However, at a certain point, each additional cluster results in a minuscule decrease in the WSS, and the loss of simplicity for the model and risk of overfitting outweigh the marginal improvement in WSS. The point at which this occurs is called the "elbow," and the number of clusters corresponding to the elbow is chosen as the optimal number of clusters for the model overall.[7]

[5] At least to an extent. Certainly there is a difference between experiencing manufacturing job losses in 1920 vs. 2020.

[6] In fact, I use a mix of the elbow method and some substantive evaluation to choose the final model in this case.

[7] This is not as "objective" as the measures proposed in Patel, Sivaiah and Patel (2022) because the process typically involves the researcher "eyeballing" the WSS plot. However, there is a way to automate

Dealing with Outliers

Since clustering aims to minimize intra-group dissimilarity and maximize inter-group dissimilarity, it is common for algorithms to identify clusters with very few observations. This is understandable, but it is not always useful. To the first point, if some number of observations are extreme outliers, it is not useful to try to sort them into a cluster that they really don't belong to, and so it is mathematically and substantively preferable to create a new cluster that has very few, and possibly only one, observation. This is, however, at odds with the general goal of simplifying the data by reducing them down to as few clusters/general patterns as possible. In the context of realizing the underlying trends in manufacturing jobs in US counties, it may be useful to note that a single county has a very different pattern than the rest, but if it is in a cluster of its own, or a cluster with very few others, it is not particularly useful to include it in an overall model of American counties. To avoid this, I build a step into the hierarchical clustering process that checks if the model built contains a cluster with very few counties.[8] The process then adds the members of these clusters to a list of outlier counties and reruns the clustering algorithm without them.

Evaluating Model Fit

Finally, I must determine whether the end model is a good fit for the data or not. Since there is no true categorization for the counties, I am not able to evaluate the fit of the model against some "known truth." Instead, the model must be used to evaluate itself. There are a number of objective measures that exist to evaluate clustering models, two of which include the Silhouette coefficient (Rousseeuw, 1987) and the cophenetic correlation (Saraçli, Doğan and Doğan, 2013). Both of these statistics are correlations, and like any other correlation, are

the process, and I build this into the algorithm. Specifically, I preset some plausible values by which the WSS can decrease and still be considered meaningful. The elbow occurs when the WSS stops dropping by that amount.

[8] "Very few" is a parameter that can be modified. I run many models, with minimal cluster sizes set to 5, 10, 15, 20, 25, and 30, and pick a "best" model based on criteria discussed in the next section. More details about how this step is built into the algorithm can be found in the Appendix.

measured between -1 and 1, with 1 being an ideal fit. These values could artificially increase if too many outliers are removed, though, so a best model could be found by balancing the number of rows removed as outliers with the rise in these correlations. Because the most time-intensive step of this method, calculating the distance matrix, is necessary whether one or 1,000 models are fit, I can create many models and compare them against each other. I thus create many models and pick the one that best balances fit (as measured by the cophenetic correlation), number of data points removed as outliers, and interpretability.

Results

Technical Results

I choose the model that uses a minimum cluster size of 15, a maximum drop in WSS with added cluster of 0.1, and a sequence of just 1 of added clusters for which the drop must be seen. This model is appealing because it has a cophenetic correlation of 0.615 (representing a moderately good fit - a somewhat difficult feat with thousands of observations), six understandable and relatively sizeable clusters, and only 25 outliers removed, which is only about 1% of the data overall.

The six clusters come in a range of sizes, with the smallest containing 36 counties and the largest containing 804. Figure 3 is a plot of the time series of all counties, divided by cluster. This figure is helpful in showing the general number of member counties in each cluster and illustrating similarities of manufacturing job trajectories within cluster. The y axis is consistent across all clusters so that the clusters are directly comparable. It is thus easy to see that manufacturing jobs make up a larger proportion of total jobs in counties of clusters 1, 3, and 4 compared to clusters 2 and 6, for instance.

To make the overall pattern of each cluster easier to visualize, Figure 4 shows each cluster's "centroid" or "prototypical" county. This is the county for which the sum of distances to all other counties in the cluster is smallest (Sarda-Espinosa, 2019). In a sense, it repre-

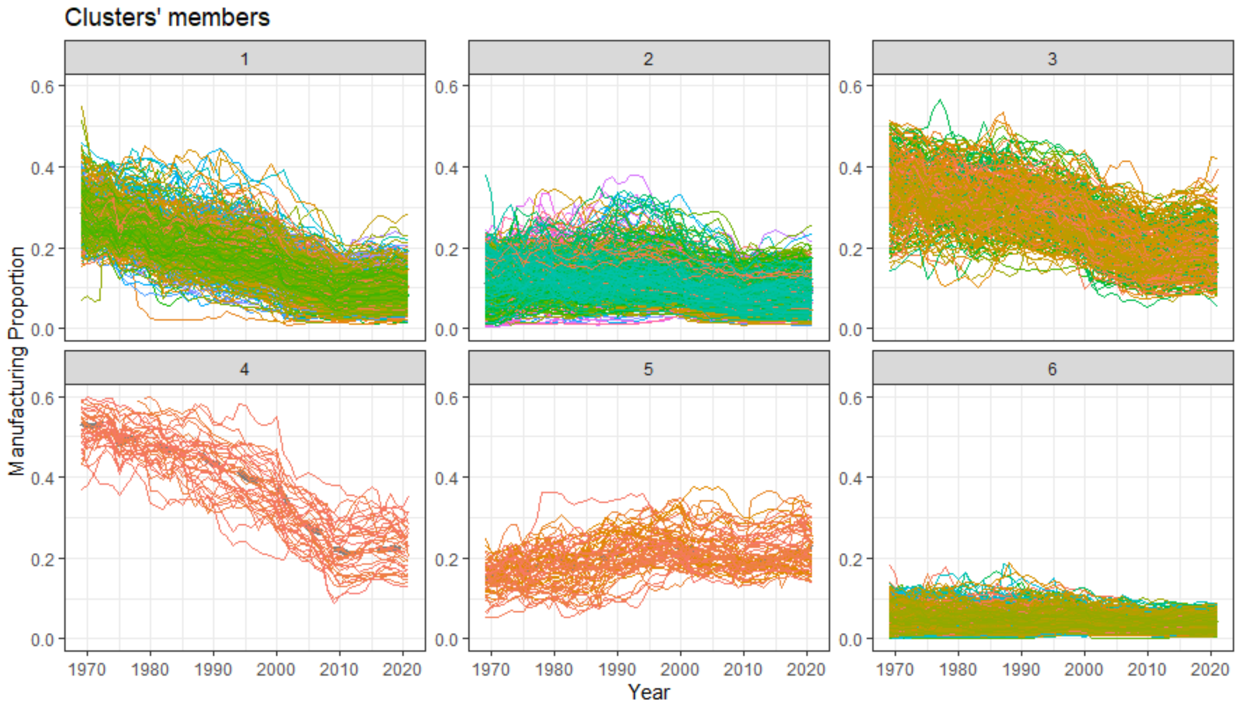

Figure 3: Time series of all counties in each cluster

sents the "typical" county's path of manufacturing jobs over the past 50 or so years in that cluster. In this case, I do not make the y axes consistent for all clusters so that finer patterns are not visually hidden by scaling. However, it is again worth noting that manufacturing jobs, as a proportion of all jobs, are relatively more numerous in clusters 1, 3, and 4 than they are in clusters 2 and 6. Manufacturing jobs in cluster 5 have only somewhat recently risen to the proportions historically seen in clusters 1, 3, and 4.

For these six clusters, the general shape is primarily of note, though the numbers on the side, signifying the proportions of jobs attributable to manufacturing in the county, also matter. The vast majority of counties have experienced loss of manufacturing jobs, with three of the six clusters very explicitly showing that pattern. One cluster, Cluster 5, has actually gained manufacturing jobs over time, though it did experience a temporary downturn in the recent past.[9] Unfortunately for the American manufacturing sector and those employed by

[9]The prototypical county plotted here, Preble County, Ohio, experienced a temporary lack of manufacturing jobs right around the peak of the Great Recession. While it is likely that most other counties in this cluster experienced their downturns at the same time, it is technically possible that the downturns happened

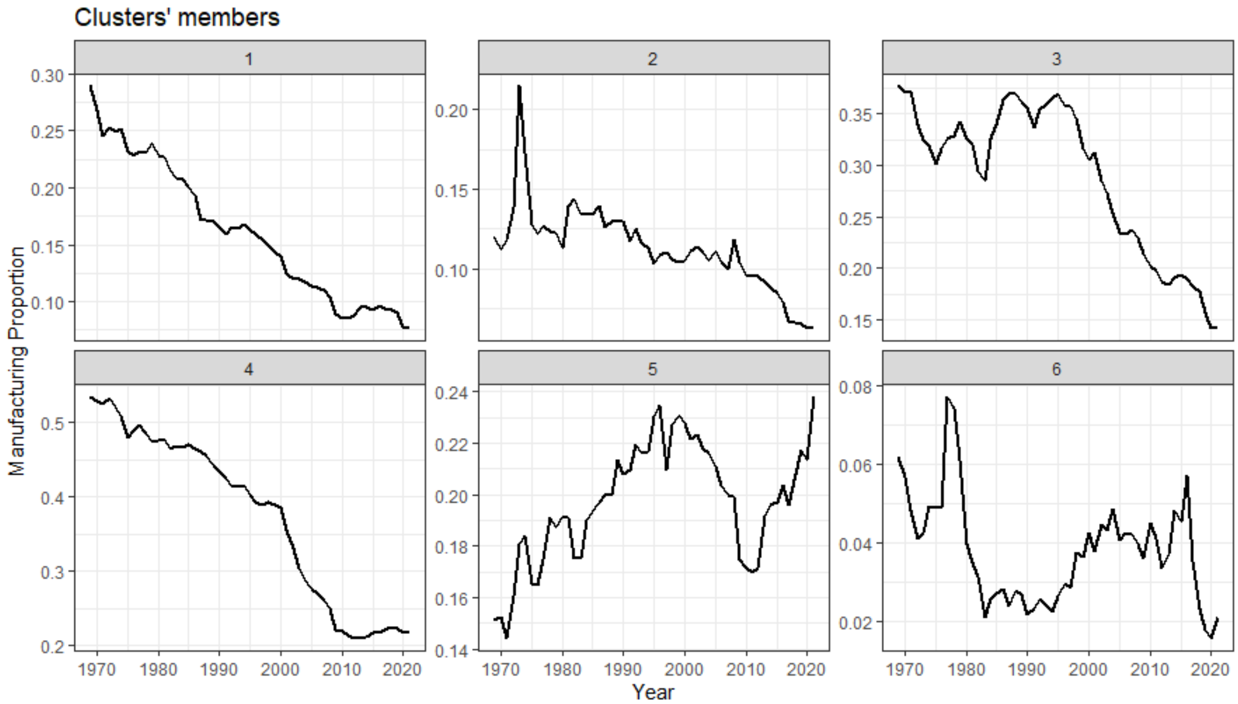

Figure 4: Time series of centroid counties for each cluster

it, this cluster is the second smallest, with only 67 member counties. I call this cluster the "Increase, Temporary Recent Bust" cluster.

The centroid counties for the other five clusters have all lost manufacturing jobs since 1969. Counties in Cluster 6 could be said to have roughly maintained their numbers, but there are two things worth noting. First, there are cycles, with short periods of relatively low proportions of manufacturing jobs and periods with relatively high proportions of jobs. Second, the overall proportions of manufacturing jobs in general is quite low, with only 8% of jobs attributed to the sector at the high point for the prototypical county in the cluster, Mingo County, West Virginia. I call this cluster the "Cycles, Regularly Low" cluster.

Cluster 1 has the simplest, and perhaps most expected, shape. With only the most minor of deviations, this cluster represents a consistent decline in manufacturing jobs. It is all but a straight line with a negative slope over the entire range of years. Manufacturing jobs were relatively important here in the early years, with its centroid county, Licking County, Ohio,

at other times, because of how dynamic time warping distance calculations are somewhat time-agnostic.

owing about 30% of its jobs to the sector in 1969. Owing to its consistent negative pattern, I call this cluster the "Steady Decline" cluster.

Cluster 2, like Cluster 1, shows a noticeable decline in manufacturing jobs. However, it is not too drastic. In the early years of the range of study, counties in this cluster did not have a particularly large number of manufacturing jobs. They soon saw a sharp jump in manufacturing jobs, and just as quickly as that jump occurred, it ended. After that point, the proportion of jobs in manufacturing slowly diminished. That temporary jump may look like a data error, but as I discuss in the next section, there is a reasonable explanation, namely automation, for such a sharp but short peak in manufacturing jobs around this time. I call this group the "Short-lived Peak, Slow Decline" cluster. This cluster is by far the largest, with almost a third of all counties, 804, belonging to it.

Cluster 3 is relatively large, with 329 member counties. Manufacturing jobs here are more crucial than for all other clusters but Cluster 4. Counties in Cluster 3 experienced far more manufacturing job volatility in the earlier portion of our range of interest before experiencing a steady decline. While there was volatility, over the first 25 years or so, these counties saw occasional rebounds in manufacturing jobs before their ultimate decline. I call this cluster the "Temporary Holdout then Rapid Decline" cluster.

Cluster 4, by far the smallest, is somewhat similar to cluster 3 and contains counties that have seen on average the most recent and drastic declines in proportions of manufacturing jobs. From 1969 to 2021, these losses started out somewhat slow, but soon became much quicker, until around 2009 when the proportion of jobs in manufacturing leveled out. Jobs in manufacturing have always been relatively abundant in these counties, with the centroid county, Catawba County, North Carolina owing about half of its jobs to the sector in 1969 and still about 20% today. Due to the trajectory of jobs over time, I call this cluster the "Slow then Fast Decline" cluster. Table 1 summarizes these findings.

Cluster	Members	Centroid	Example County
1. Steady Decline	578	Licking, OH	Wayne, MI
2. Short-lived Peak, Slow Decline	804	Dewitt, TX	San Mateo, CA
3. Temporary Holdout then Rapid Decline	329	Patrick, VA	Lorain, OH
4. Slow then Fast Decline	36	Catawba, NC	Chester, SC
5. Increase, Temporary Recent Bust	67	Preble, OH	Carver, MN
6. Cycles, Regularly Low	477	Mingo, WV	Teton, WY

Table 1: Cluster names, size of membership, and typical and example county for each cluster

Clusters in more Detail

In addition to summarizing the previous section, Table 1 also includes a better-known "example county" for added context. For example, while Licking County, Ohio is the centroid, or prototypical, county for the Steady Decline counties, this descriptor represents many counties home to larger cities, especially Rust Belt cities. One of the best understood examples of a steadily deindustrializing county is Wayne County, Michigan, home to Detroit. Counties that are home to relatively similar cities like Cleveland, Pittsburgh, Chicago, and Buffalo (Cuyahoga, Allegheny, Cook, and Erie, respectively) also fall into this cluster.

For the two smallest clusters by member size, I pick example counties that may not be particularly well known but that have illustrative histories that help to explain the patterns unique to their clusters. Carver County, Minnesota (directly southwest of Minneapolis's Hennepin County) is heavily reliant on manufacturing today, with 24% of its jobs in the sector, the most of any industry. The county is home to many small advanced manufacturing companies, creating goods for medical, pharmaceutical, aerospace, robotics, and electronics uses (Finance Department of Carver County, 2021). While traditional manufacturing jobs like milling steel or building cars may generally be at risk, the manufacturing jobs building newer technologies, and the communities in which these jobs are abundant, are thriving relatively. Chester County, South Carolina, meanwhile, historically has had a mixed agricultural and industrial economy. The county began to thrive around the turn of the 20^{th} century thanks to the emergence of textile mills and sawmills, and as these jobs have left the economy turned to a reliance on cash crops (Choose Chester, N.d.).

In many counties across the country, manufacturing has never been particularly important to the local economy. The prototypical county for which this is the case is Mingo County, West Virginia. A city in this county, Williamson, is home to the Coal House Museum, highlighting the importance of a single industry, mining, to the county (Communities LEAP, N.d.). I choose to include Teton County, Wyoming, home to Grand Teton National Park and part of Yellowstone National Park, as the example here. There is, perhaps surprisingly, a large enough number of manufacturing jobs here to register on the BEA data set, but, unsurprisingly, these jobs are not terribly important, and never have been, in an area where the protection and preservation of the natural landscape is all but inarguably the foremost goal (Data USA, N.d.).

Manufacturing jobs in Cluster 2 are somewhat more important than they are in the previous cluster, but are less important than they are in all other clusters. These counties did not supply many manufacturing jobs prior to the 1970s, when they experienced a short lived peak and a slow decline. The 1970s was the decade when automation ramped up most quickly, and while manufacturing jobs long term were harmed by this innovation, in the shorter term they benefited (Bong, 2022). Specifically, creating machines to aid in automation in factories in and of itself created jobs, even if those very machines then took jobs away (Grillot, 2022; Wu, 2022). While this cluster is by far the largest, and no single explanation can describe the general experience of these 804 counties, this automation explanation describes the example of San Mateo County, California well. Today, this county (one county south of San Francisco) is heavily reliant on professional, scientific, and technical services, a very modern industry that Silicon Valley is well-known for. Before this industry could take off, though, manufacturing was important to the area (U.S. Bureau of Economic Analysis, N.d.*a*).

Finally, Cluster 3 may best represent standard view of American manufacturing over time. Counties in this cluster are often home to smaller communities that were historically very reliant on manufacturing jobs. While they held onto their manufacturing jobs for a while, they ultimately, and quickly, followed the pattern that most of the rest of the country had already started experiencing. A good example of this pattern is Lorain County, Ohio, one county west of Cleveland. Manufacturing has long been crucial to its economy, and arguably still is today. Compared to Cleveland, though, which is, like Detroit, in the "Steady Decline" cluster, Lorain County lost manufacturing jobs, as a proportion of total jobs, later and more quickly. Lorain County, like many in this cluster, is outside a major city and thus has a smaller population and less diversified economy than its neighbor county (U.S. Bureau of Economic Analysis, N.d.*a*). This leads to volatility, which can be seen even before the drastic, permanent decline in manufacturing jobs that is indicative of the cluster; manufacturing jobs are volatile, experiencing the most drastic and frequent swings of all counties, before completely plummeting around the turn of the century.

Geographical Patterns

The example counties chosen in the previous section have a benefit beyond being relatively well-known and having interesting and straightforward histories that make their memberships in their respective clusters relatively clear. They are also useful in highlighting the geographical pattern to the clustering. Specifically, the Plains and Mountain West states are home to many of the "Cycles, Regularly Low" counties (though Florida also noticeably has many), and the choice of Teton County, Wyoming reflects that pattern. The "Increase, Temporary Recent Bust" counties are mostly clustered in the Upper Midwest, whereas the "Slow then Fast Decline" counties are clustered in the South. The "Temporary Holdout then Rapid Decline" counties are mostly suburban/exurban counties of major cities in the Rust Belt, though many of these counties are also clustered around Southern cities. The "Steady Decline" counties also include some Rust Belt suburbs but also most major cities in the Northeast and Midwest and many counties spread across the South and Pacific Coast. There does not seem to be much pattern to the "Short-lived Peak, Slow Decline" counties; every continental US state has at least one such county. Figure 5 shows the cluster membership of every county, with the outlier counties and those with missing data marked as not belonging to any cluster (NA).

Application

In this section, I demonstrate the value added of my measure by revisiting Feng, Kerner and Sumner (2021), adding to their findings and showing how both their and my measure of deindustrialization can be used together to build a more robust theory of how deindustrialization impacts all Americans' investment views. In their paper, they find that recent TAA claims per capita - their way of measuring local exposure to trade-related deindustrialization (TRD) - impact Americans' views on foreign direct investment (FDI) and investment more broadly. Specifically, Americans' views on investment grow more positive as their exposure to TRD goes up, but this effect is effectively negated when the investment is specified to be from a Chinese source.

I rerun the model in Feng, Kerner and Sumner (2021), but I split their data set in two. One subset includes residents of broadly "deindustrialized" counties, as determined by my measure, and the other includes residents of broadly "non-deindustrialized" counties. My results support their general findings but show how using these clusters can provide a more appropriate test of their theory. Specifically, while their main independent variable of interest, TAA claims per capita, is significant in predicting investment disapproval in the country overall, it is not significant in the counties within the clusters most heavily impacted by deindustrialization. Additionally, different results between the two models for some of the demographic control variables (namely age and partisanship) can also illuminate how the politics of deindustrialization may play out differently based on historical trajectory.

Deindustrialization can impact a community over a long period of time, and its impacts may take a while to be fully realized. Measuring a county's history of deindustrialization using data beginning only in 2000 does not capture the impacts that might take a generation to unfold, nor does it account for the workers impacted before the data were collected. US manufacturing peaked in the late 1970s (Harris, 2020), so decades of job losses, whether they would qualify for TAA claims or not, are not captured when data collection only begins in 2000. My measure covers manufacturing jobs from about a decade before their peak to almost present day.

This isn't to say that TAA claims are not useful to this area of research. They do capture something. Specifically, in many counties in the United States, manufacturing has never been particularly important to the local economy. That said, some even in those communities do hold manufacturing jobs, and recent losses of those jobs, even if not particularly numerous, could impact these individuals and immediate members of their social networks. While the time scale is not large enough for deindustrialization to permeate the core identity of the county, individuals can still be impacted. These counties stand in contrast with those counties that *have* experienced long term deindustrialization. In those areas, I would expect effectively all residents to understand the impact of manufacturing, trade, investment, etc. on their communities and local economies. In essence, I would expect a higher convergence of opinion on these issues among all residents, no matter their age, race, direct work experience in manufacturing, etc. Thus, local, recent TAA claims and short term drops in manufacturing employment would not play as large of a role in these Americans' investment views. Put simply, people living in these communities have seen what globalization has done, for better or worse, over generations - why would individual job losses just within the current generation shift opinion much?

For these reasons, I expect that TAA claims per capita will remain significant and meaningful in predicting negative opinions towards investment in the *non-deindustrialized counties*, but it will not be significant for the deindustrialized counties. I have six clusters, though, and this expectation only discusses two groups. This is intentional for two reasons. The first is purely practical. Feng, Kerner and Sumner (2021) contains 1,029 observations - splitting these data into six groups could (and indeed does) make some groups very small - too small to have the degrees of freedom to test all variables that the original article did. Thus, the value of a replication and extension like this would be limited. The second reason is for the sake of parsimony. While I stand by the value of six clusters in a general sense, I believe they can be collapsed in this particular application. Cluster 3 may be "more deindustrialized" than Cluster 1, but *both* should behave similarly (or at least their estimates should have the same signs and should both either be or not be significant) in regards to whether or not TAA claims per capita impact views on investment. Other researchers may justifiably argue for different ways in which to collapse these six clusters both in this application and in other work, but for this specific example, condensing the six clusters into two in this way is both pragmatic, given the available, and illuminating. The ability to flexibly choose whether and how to condense the clusters is a feature of including and discussing the least collapsed, six cluster version - individual researchers can choose whether and how to consolidate the clusters based on their specific research project and ability to gather sufficient data. To be specific about how I group the clusters for this replication, I will consider clusters 1, 3, and 4 to be "deindustrialized" and 2, 5, and 6 "non-deindustrialized." The collapsing of counties is shown in Figure 6.

Figure 7 gives more insight into the logic of consolidating the clusters in the way I have proposed for this demonstration. This figure is identical to Figure 4, but with the y axis standardized across all six clusters. Visualizing each of these counties on the same y axis gives added perspective to the relative importance of manufacturing jobs in these counties. It shows how important these jobs were in the last century in especially clusters 3 and 4, but also in Cluster 1, and how these jobs have been lost in these counties, thus justifying their status of "deindustrialized" counties. In contrast, manufacturing jobs have never been proportionally numerous in Cluster 6, only saw a temporary surge in Cluster 2, and were not very numerous but have actually been *rising* in Cluster 5. I do not consider these counties deindustrialized because cluster 2 and 6 never truly industrialized, and Cluster 5 is industrializing more over time.

I run the same logistic regression model that Feng, Kerner and Sumner (2021) run, but I subset the data as just described. Table 2 contains results for, in order, the original data set, the set that contains deindustrialized counties, and the set that contains non-deindustrialized counties.[10] As expected, TAA claims per capita is no longer significant for the deindustrialized counties, with an estimated effect of -0.055. However, this variables is also insignificant for non-deindustrialized counties, contrary to the expectation. That said, its magnitude is larger for non-deindustrialized counties than for the whole sample, so its loss of significance at the 0.1 level is likely the result of a smaller sample size compared to the full sample.

The results of the three models in Table 2 show that there is meaningful reason to split counties into these two groups for analysis - the original results pooled two distinct groups. The estimates for the main variables of interest, namely the three treatment levels, are very similar in the entire sample and non-deindustrialized county models. Respondents from deindustrialized counties, however, are more accepting of investment when no information of the source of the investment is given, but the difference between their feelings of no information investment and both foreign and Chinese investment is larger than that of respondents overall and from non-deindustrialized counties. Of the control variables, two stand out. Age, as previous mentioned, was significant and positive (meaning older respondents were more skeptical of investment in general) in the original model and remains so for the non-deindustrialized model, but there is no significant difference in investment views among respondents of different ages in deindustrialized counties. This may further support the argument that views on investment are somewhat more homogeneous in deindustrialized

[10] 104 of the original 1,029 counties are not in any cluster. Of these, 14 are due to the original study not having a county for the respondent, 4 are due to the county being identified as an "outlier" by my algorithm, and the rest are due to missing data in the BEA data set (i.e. the data were suppressed because the county is so small or manufacturing is so unimportant that individuals or their employers would be identified if the data were included). With this in mind, one could argue that the observations for which a county is identified but where no cluster is assigned should be considered non-deindustrialized since manufacturing jobs are not numerous. I, for a few reasons related to the causes of the missing data, do not agree, but I do include a model in the Appendix where those counties are included along with clusters 2, 5, and 6, and there are no meaningful differences in the results.

counties as nearly everyone is aware of the impact of globalization on these communities.

Table 2:

	Dependent variable:		
	investmentbad		
	All Counties	Deindustrialized	Non-Deidustrialized
treatmentforeign	2.351***	3.356***	2.363***
	(0.394)	(1.051)	(0.597)
treatmentChinese	3.197***	4.498***	3.281***
	(0.394)	(1.020)	(0.598)
TAApclogscale	−0.465*	−0.055	−0.548
	(0.256)	(0.181)	(0.376)
unemployedpercapita	14.067	32.360	−1.963
	(21.598)	(46.186)	(27.280)
man	−0.265	−0.364	−0.174
	(0.186)	(0.311)	(0.257)
white	−0.059	0.575	−0.474
	(0.219)	(0.414)	(0.299)
agescale	0.313***	0.163	0.428***
	(0.095)	(0.142)	(0.137)
republican	−0.110	−0.607*	0.178
	(0.186)	(0.311)	(0.248)
manufacturing	−0.076	−0.072	−0.314
	(0.206)	(0.348)	(0.303)
college	−0.294	−0.443	−0.220
	(0.182)	(0.322)	(0.252)
treatmentforeign:TAApclogscale	0.349	0.315	0.078
	(0.282)	(0.402)	(0.389)
treatmentChinese:TAApclogscale	0.485*	0.124	0.544
	(0.292)	(0.356)	(0.427)
Constant	−3.609***	−5.305***	−3.273***
	(0.682)	(1.635)	(0.911)
Observations	851	324	451
Log Likelihood	−371.900	−137.846	−189.549
Akaike Inf. Crit.	769.799	301.691	405.098
Note:			*p<0.1; **p<0.05; ***p<0.01

While people are more similar along demographic characteristics in deindustrialized counties, the same cannot be said along partisanship. The results show that Republicans in deindustrialized counties are less likely to view investment as bad than non-Republicans, with an estimated effect of -0.607, or about 35%. Superficially, this finding may seem at odds with the narrative that there are many Trump supporters in the Rust Belt who are skeptical of investment (at least when it is foreign or Chinese). These results could be simply due to the sample or the geography. For example, Republicans in the Rust Belt may indeed be more skeptical of investment, but this could be outweighed by Republicans outside the Rust Belt being more pro-investment. There is also a sensible political explanation. According to Keser et al. (2024), Republicans have been more supportive of trade than Democrats until very recently. This pattern flipped in 2017, the year Feng, Kerner and Sumner (2021) fielded their experiment. This switch in economic preferences may have only started or not fully developed by the time the experiment occurred, and by using their data from 2017, I may have captured a finding about Republicans in deindustrialized counties that was true at the time but is no longer. In addition, Keser et al. (2024) attribute this flip in trade preferences by party to the Trump effect, with those who approve of Trump being far more likely to have increasingly negative views of trade. Given that not all Trump supporters identify as Republicans, and those who identify as Democrats are likelier to be located in the geographic "Trump Belt" of the Upper Midwest and Northeast (Muravchik and Shields, 2020), it is also possible that my deindustrialized subset of counties captures a lot of anti-trade "Trump Democrats."

While the specific findings here are interesting in their own right, the main takeaway from this replication is that deindustrialization happens over multiple decades, or even generations, not just years. Short term shocks to the supply of manufacturing jobs may illicit strong responses from those immediately impacted, but longer term, drastic loss of jobs, especially in areas where those jobs had been so important for so long, permeates through the entire community and shapes community views. TAA claims per capita in the county *are* measuring something meaningful, but it seems to be at a more personal, individual scale than tracking the course of manufacturing jobs in the county over decades.

Discussion and Conclusion

On the whole, the proportion of the United States labor force employed in manufacturing has been declining over time from its peak in the 1970s. However, the story of shuttered factories and abandoned steel mills is not true for every community across the country. Manufacturing is a highly heterogeneous industry; turning iron pellets into steel, mass producing children's toys, and combining metals, plastics, and electronics into cell phones are all examples of manufacturing. Foreign firms have become more competitive than American firms in producing and trading goods that require cheap, low-skilled labor, and so certain manufacturing jobs, especially those that involve more routineness, are more prone to offshoring (Owen and Johnston, 2017; Feenstra, Hanson and Swenson, 2000). We are a long way away from completely automating manufacturing (if total automation is even possible), though, and cannot or will not move all manufacturing jobs overseas (Wu, 2022; Owen and Johnston, 2017; Acemoglu and Autor, 2011; Markusen, Deardorff and Irwin, 2005). This is especially true as the manufacturing process globalizes further and a single manufactured good contains components produced in several different countries (Chase, 2008; Grossman and Rossi-Hansberg, 2008). Thus, claiming that American manufacturing has been declining without exception ignores the very real success stories that some areas have been experiencing, especially in advanced manufacturing.[11]

This is not the first paper to argue that some communities have suffered from deindustrialization more than others, but work done before now has essentially just provided a before and after picture. This is the first project I am aware of to use year-by-year trends in every individual county for which data were available. While this project did not focus on political implications of deindustrialization, it did provide a small but meaningful extension to existing theories on the role of exposure to globalization on views of investment. Thus, in addition to highlighting the nuance that measuring deindustrialization using time series

[11] There isn't really a simple, universal definition of what constitutes "advanced manufacturing," but this type of manufacturing often produces innovative, new technologies and goods and goes beyond turning a raw material into a finished good, which itself is known as traditional manufacturing (Thareja, 2005).

data for each county can bring, this project has created a new categorical variable that can be incorporated into future works. This variable can be useful when revisiting old works, as demonstrated in this paper, or whenever a measure of local deindustrialization is needed more broadly. In this way, researchers can benefit from the fuller and more detailed data while easily incorporating that information into their projects.

Future work can improve and expand on this project. For example, multivariate time series clustering, i.e. including time series of jobs not just in the manufacturing sector, may unearth more information that would be particularly valuable in work done in political economy. This would be valuable as the political consequences of two counties with a similar pattern of manufacturing jobs over time could vary drastically based on how and if those jobs are being replaced. A community that transitions from manufacturing to high-paid sectors (for example, "Professional, Scientific, and Technical Services," "Information," or "Finance and Insurance" in NAICS classification) would be very likely to have far different political and social outcomes than one that transitions from a manufacturing-dominated community to one with high unemployment or otherwise lower paid jobs (workers in the "Accommodation and Food Services" NAICS classification, for example, made over $12 less per hour than manufacturing workers in August 2022). Time series for more generic demographic data, like county populations overtime, populations of working age adults, proportions of men or women, and total jobs/unemployment may also be useful. This would give a more robust picture of what is occurring in the counties, as the loss of these traditionally male-dominated, well-paid roles can have major demographic impacts. This would also highlight differences among counties whose proportions of manufacturing jobs have declined due to a loss of those jobs, a consistency in those jobs but an abundance of other jobs, and simply a quickly growing population overall.

Finally, the algorithm developed in this paper could be applied to political science questions outside political economy or even completely different subject areas. For example, using county-level election data over several decades would likely give a fuller picture of the political trends of those counties than simply recording which candidates the counties supported over the course of two elections, and clustering global cities based on their pollution histories would be more informative than recording how levels of pollution have changed between two points in time. While the focus on this paper is on manufacturing labor statistics, the algorithm could be useful to clustering time series for any case where the researcher has many observations and would like a way to flexibly handle outliers.

While many picture major Rust Belt cities like Detroit, Cleveland, Pittsburgh, and Buffalo when thinking about deindustrialization in the United States, this study has shown the phenomenon is not limited to these areas, and it arguably is not even the most severe there either. The loss of manufacturing jobs is more drastic and less consistent and predictable in smaller counties in the Rust Belt and some Southern counties. Recognizing this and the differences in deindustrialization patterns over a long period of time among all of these counties is the first step in drawing more attention to those who may have up until now felt ignored and left behind, but who have nonetheless been among the most severely affected by the loss of manufacturing jobs.

References

Acemoglu, Daron and David Autor. 2011. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*. Vol. 4 Elsevier pp. 1043–1171.

Baccini, Leonardo and Stephen Weymouth. 2021. "Gone for good: Deindustrialization, white voter backlash, and US presidential voting." *American Political Science Review* 115(2):550–567.

Bernard, Andrew B, Valerie Smeets and Frederic Warzynski. 2017. "Rethinking deindustrialization." *Economic Policy* 32(89):5–38.

Bong, Nathan. 2022. "The Evolution of Automation.".

Brankovic, Milutin, Kevin Buchin, Koen Klaren, André Nusser, Aleksandr Popov and Sampson Wong. 2020. (k, l)-Medians Clustering of Trajectories Using Continuous Dynamic Time Warping. In *Proceedings of the 28th International Conference on Advances in Geographic Information Systems*. pp. 99–110.

Chase, Kerry A. 2008. "Moving Hollywood abroad: Divided labor markets and the new politics of trade in services." *International organization* 62(4):653–687.

Choose Chester. N.d. "South Charlotte's Innovation Corridor.".

Communities LEAP. N.d. "Mingo and Logan Counties, WV.".

Data USA. N.d. "Teton County, WY.".

Feenstra, Robert C, Gordon H Hanson and Deborah L Swenson. 2000. Offshore assembly from the United States: Production characteristics of the 9802 program. In *The impact of international trade on wages*. University of Chicago Press pp. 85–125.

Feng, Yilang, Andrew Kerner and Jane L Sumner. 2021. "Quitting globalization: trade-related job losses, nationalism, and resistance to FDI in the United States." *Political Science Research and Methods* 9(2):292–311.

Finance Department of Carver County. 2021. The Carver County Annual Comprehensive Financial Report. Technical report Finance Department of Carver County.

Grillot, Ben. 2022. "The Industrial Landscape of the 1970s.".

Grossman, Gene M and Esteban Rossi-Hansberg. 2008. "Trading tasks: A simple theory of offshoring." *American Economic Review* 98(5):1978–1997.

Harris, Katelynn. 2020. "Forty years of falling manufacturing employment.".

High, Steven. 2020. "Deindustrialization and its Consequences." *Routledge international handbook of working-class studies* pp. 169–179.

Jaslam, PK, Nitin Bhardwaj, Monika Devi, Vivek Kumar Singh and Anu Anu. 2022. "Diversity structure analysis based on hierarchical clustering method." 2357(1).

Karna, Ashutosh and Karina Gibert. 2022. "Automatic identification of the number of clusters in hierarchical clustering." *Neural Computing and Applications* 34(1):119–134.

Kerner, Andrew, Jane Sumner and Brian Richter. 2020. "Offshore production's effect on Americans' attitudes toward trade." *Business and Politics* 22(3):539–571.

Keser, Ceren, James C Garand, Ping Xu and Joseph Essig. 2024. "Partisanship, Trump favorability, and changes in support for trade." *Presidential Studies Quarterly* 54(1):46–64.

Kotsakos, Dimitrios, Goce Trajcevski, Dimitrios Gunopulos and Charu C Aggarwal. 2018. Time-series data clustering. In *Data Clustering*. Chapman and Hall/CRC pp. 357–380.

Latifi-Pakdehi, Alireza and Negin Daneshpour. 2021. "DBHC: A DBSCAN-based hierarchical clustering algorithm." *Data and Knowledge Engineering* 135:101922.
URL: *https://www.sciencedirect.com/science/article/pii/S0169023X21000495*

Madhulatha, T Soni. 2012. "An overview on clustering methods." *arXiv preprint arXiv:1205.1117* .

Markusen, James R, Alan V Deardorff and Douglas A Irwin. 2005. Modeling the Offshoring of White-Collar Services: From Comparative Advantage to the New Theories of Trade and Foreign Direct Investment [with Comments and Discussion]. In *Brookings trade forum*. JSTOR pp. 1–34.

Milner, Helen V. 2018. Globalization and its political consequences: the effects on party politics in the West. In *Annual Meeting of the American Political Science Association, Boston, August*.

Müller, Meinard. 2007. "Dynamic time warping." *Information retrieval for music and motion* pp. 69–84.

Muravchik, Stephanie and Jon A Shields. 2020. *Trump's Democrats*. Brookings Institution Press.

Murtagh, Fionn and Pedro Contreras. 2012. "Algorithms for hierarchical clustering: an overview." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2(1):86–97.

Owen, Erica and Noel P Johnston. 2017. "Occupation and the political economy of trade: Job routineness, offshorability, and protectionist sentiment." *International Organization* 71(4):665–699.

Patel, Punyaban, Borra Sivaiah and Riyam Patel. 2022. Approaches for finding optimal number of clusters using k-means and agglomerative hierarchical clustering techniques. In *2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP)*. IEEE pp. 1–6.

Rousseeuw, Peter J. 1987. "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis." *Journal of computational and applied mathematics* 20:53–65.

Russo, John and Sherry Lee Linkon. 2009. "The social costs of deindustrialization." *Manufacturing a better future for America* pp. 183–216.

Saraçli, Sinan, Nurhan Doğan and İsmet Doğan. 2013. "Comparison of hierarchical cluster analysis methods by cophenetic correlation." *Journal of inequalities and Applications* 2013(1):1–8.

Sarda-Espinosa, Alexis. 2019. "Comparing time-series clustering algorithms in R using the dtwclust package." *The R Journal* .

Thareja, Priyavrat. 2005. Manufacturing paradigms in 2010. In *Proceedings of National Conference on Emerging trends in Manufacturing Systems, JMIT, Radaur*.

United States Census Bureau. 2020. "Most of the Counties with the Largest Population Gains Since 2010 are in Texas.".

U.S. Bureau of Economic Analysis. 2008. "What does a (D) mean in place of an estimate for GDP by Metropolitan Area?".

U.S. Bureau of Economic Analysis. 2024. "Union affiliation of employed wage and salary workers by state, 2022-2023 annual averages.".

U.S. Bureau of Economic Analysis. N.d.*a*. "CAEMP25N: Total full-time and part-time employment by NAICS industry.". (accessed Monday, August 28, 2023).

U.S. Bureau of Economic Analysis. N.d.*b*. "CAEMP25S: Total full-time and part-time employment by SIC industry.". (accessed Monday, August 28, 2023).

U.S. Department of Labor. N.d. "Trade Act Programs.".

Winters, Lisa, Wanda Rushing, Martin Levin, Troy Blanchard et al. 2011. "Deindustrialization, Class, and Adolescents: Changing Gender Attitudes in Middletown." *Sociology Mind* 1(03):114.

Wu, Nicole. 2022. "Misattributed blame? Attitudes toward globalization in the age of automation." *Political Science Research and Methods* 10(3):470–487.

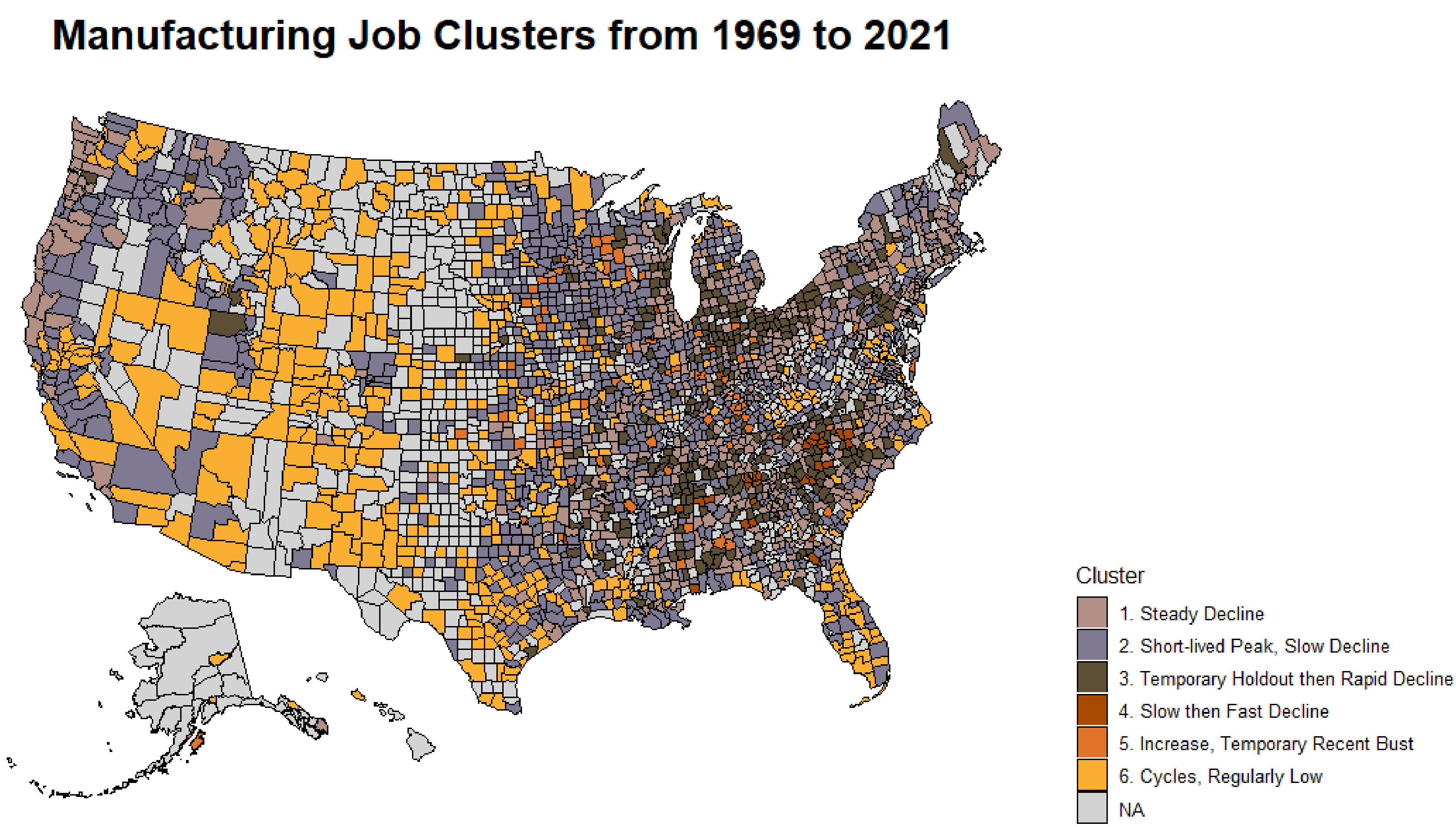

Figure 5: County-level map showing membership in one of six clusters

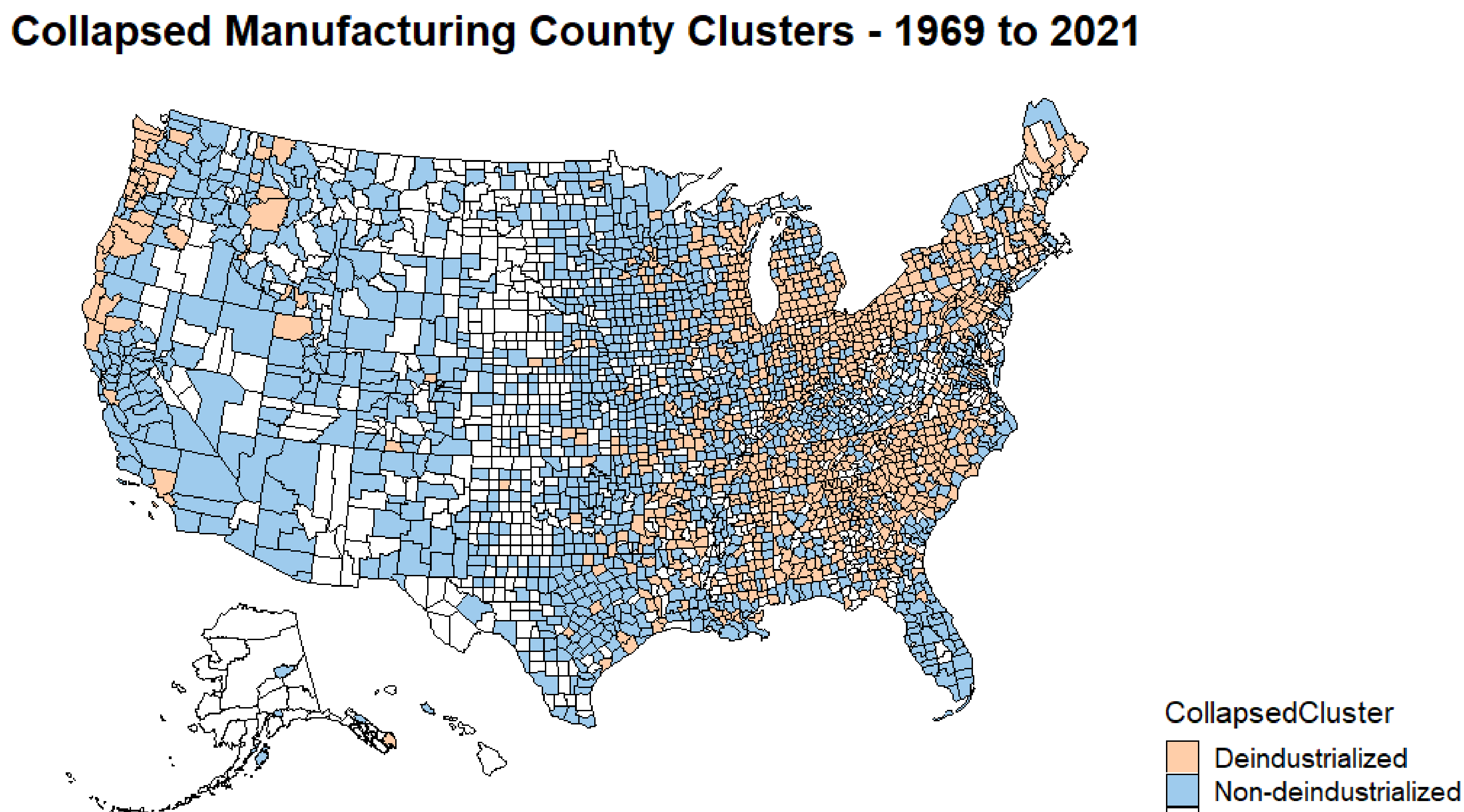

Figure 6: Map of US Counties with Six Manufacturing Clusters Condensed Down to Two

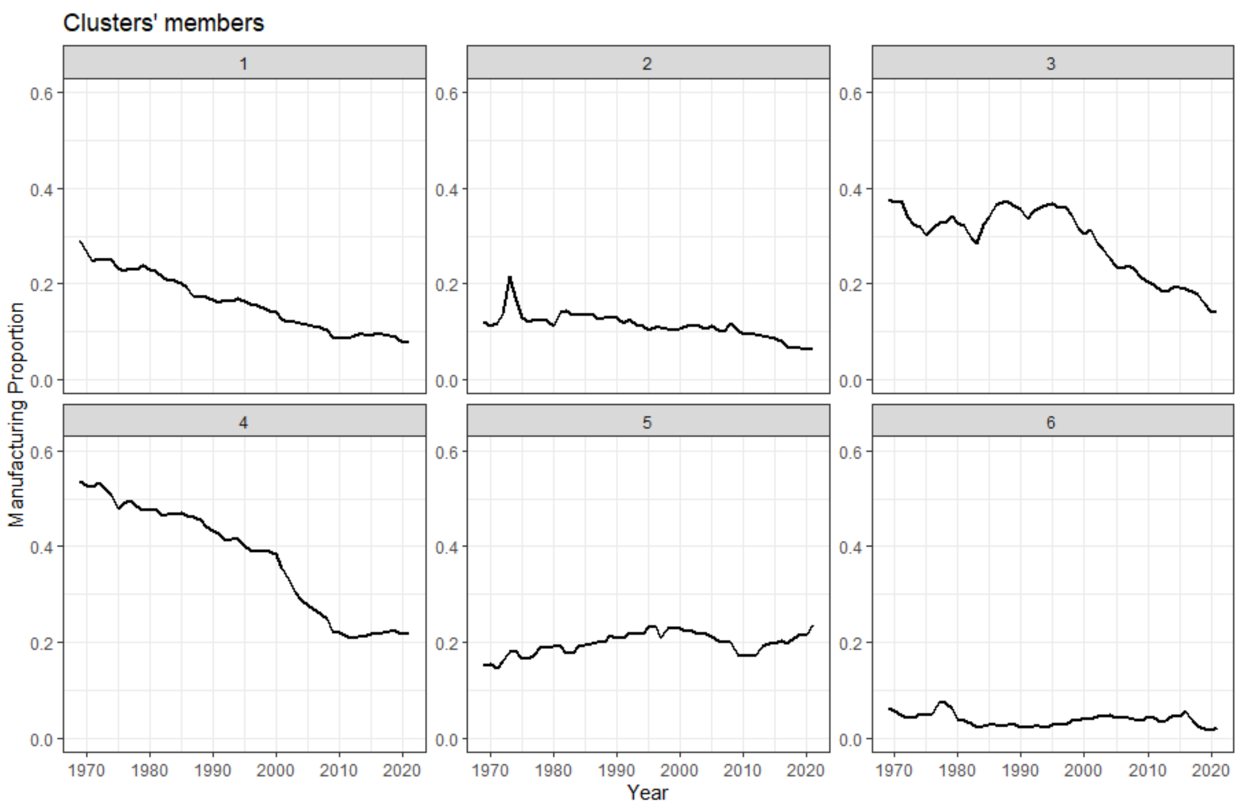

Figure 7: Time series of the each cluster's centroid county with standardized y axis