# Assessing the Influence of 2000s Trade with China on Manufacturing Employment Across US Census Zones

Elaine Mahon

Friday December 22, 2023

**Abstract**

We examine the variation in changes in manufacturing jobs across census regions in the United States. We find that there is a significant regional variation, with `satl` region experiencing the least reduction and `wncen` the most. Education, share of routine jobs, and manufacturing (Share of employment in manufacturing at start of decade) are also have a statistically significant relationship with changes in manufacturing employment.

## Background

The 2000s marked a transformative era in the landscape of global trade, with profound implications for the manufacturing sector in the United States. During this period, China had joined the World Trade Organization (WTO) on December 11, 2001. The ensuing intensification of trade relations between China and the United States emerged as a central focal point, shaping economic dynamics and influencing employment patterns across various regions. The surge in imports, particularly in the manufacturing domain, triggered a complex interplay of economic forces, raising questions about the impact of such trade dynamics on local employment within distinct US census zones. This analysis delves into the intricate web of these influences, seeking to unveil the repercussions of trade with China on manufacturing employment trends across US census zones. Through this exploration, we aim to contribute valuable insights to the broader discourse on the consequences of globalization and trade policies on regional employment dynamics.

Scholars such as Pierce and Schott (2016) have noted a steep decline in US employment in the manufacturing sector. China is cited as one of the major destination for manufacturing jobs formerly domiciled in the US. One of the major reasons for this shift is the relatively cheap labor in China. Hombert and Matray (2018) notes that import competition from China explains one-quarter of the contemporaneous aggregate decline in US manufacturing employment. Pierce and Schott (2016) summaries this issue aptly;

> ....links the sharp drop in US manufacturing employment after 2000 to a change in US trade policy that eliminated potential tariff increases on Chinese imports. Industries more exposed to the change experience greater employment loss, increased imports from China, and higher entry by US importers and foreign-owned Chinese exporters. At the plant level, shifts toward less labor-intensive production and exposure to the policy via input-output linkages also contribute to the decline in employment (pp 1).

Rising imports cause higher unemployment, lower labor force participation, and reduced wages in local labor markets that house import-competing manufacturing industries (Hombert and Matray 2018). Recently, politicians in the US have raised concerns about the decline of manufacturing in the US and proposed policies to reverse the trend (Baily and Bosworth 2014). The chief culprit has been China. Some approaches include raising tariffs on some products imported from China, lowering tax rates to encourage firms to relocate to the US, and greater investment in R&D (Hombert and Matray 2018).

There has been mixed reactions to the US trade protectionism approach to stem the loss of manufacturing jobs to China. While populists and some academics laud the move, others warn of potential further harm to US manufacturing employment, especially when trading partners take retaliatory measures (Li and Whalley

2021). But first, we must establish the extent China contributes to the loss of manufacturing employment in the US. Our analysis seeks to establish the extent to which the rise of China has contributed to this decline. The output of this analysis will be useful in informing policies to raise US manufacturing.

## Summary of Results

1. There is a significant variation in the loss of manufacturing jobs across the census regions of the United States of America (USA).
2. Other significant factors that have a strong relationship with changes in manufacturing jobs in the United States of America are;

- Manufacturing: Share of employment in manufacturing at start of decade.
- College: Share of population with college education.
- Routine: Share of workers engaged in routine tasks.

## Data

We source data from `FRED`. The data has 706 rows and 11 columns. The data contains the following variables.

| Variable | Definition |
|---|---|
| change_mfg_employ | Percentage change in manufacturing employment during decade. |
| change_china_trade | Percentage change in trade with China. |
| manufacturing | Share of employment in manufacturing at start of decade. |
| college | Share of population with college education. |
| female | Share that is female. |
| foreign_born | Share that is foreign born. |
| routine | Share of workers engaged in routine tasks. |
| outsource | Index of how easy it is to offshore the area's occupations. |
| region | Census zone. |
| lon | Longitude |
| lat | Latitude |

In the next section, we explore the data.

## Exploratory Data Analysis

Table 2 shows a summary of the data for all the numeric variables. First, the data has no missing values. The summary of the variable `change_mfg_employ` which our dependent variable are largely negative, which implies that the US has been losing manufacturing jobs since 2000, presumably to China. Figure 1 depicts the various census zones covered in the data. We see that `reg_wncen` dominates this data.

Figure 3 presents a comprehensive visual exploration of the relationships among the variables crucial to our investigation, with a specific focus on the dependent variable denoted as `change_mfg_employ`. This variable stands at the heart of our endeavor to unravel the intricate dynamics governing the interplay between international trade patterns and shifts in manufacturing employment within the United States.

Delving into the pairs plot, distinctive patterns unfold, accentuating a prominent negative association between the change in manufacturing employment and two pivotal variables: Percentage change in trade with China and routine. This discernible negative correlation aligns seamlessly with the overarching research objective, underscoring the intertwined nature of these variables and their potential ramifications on the landscape of manufacturing employment.

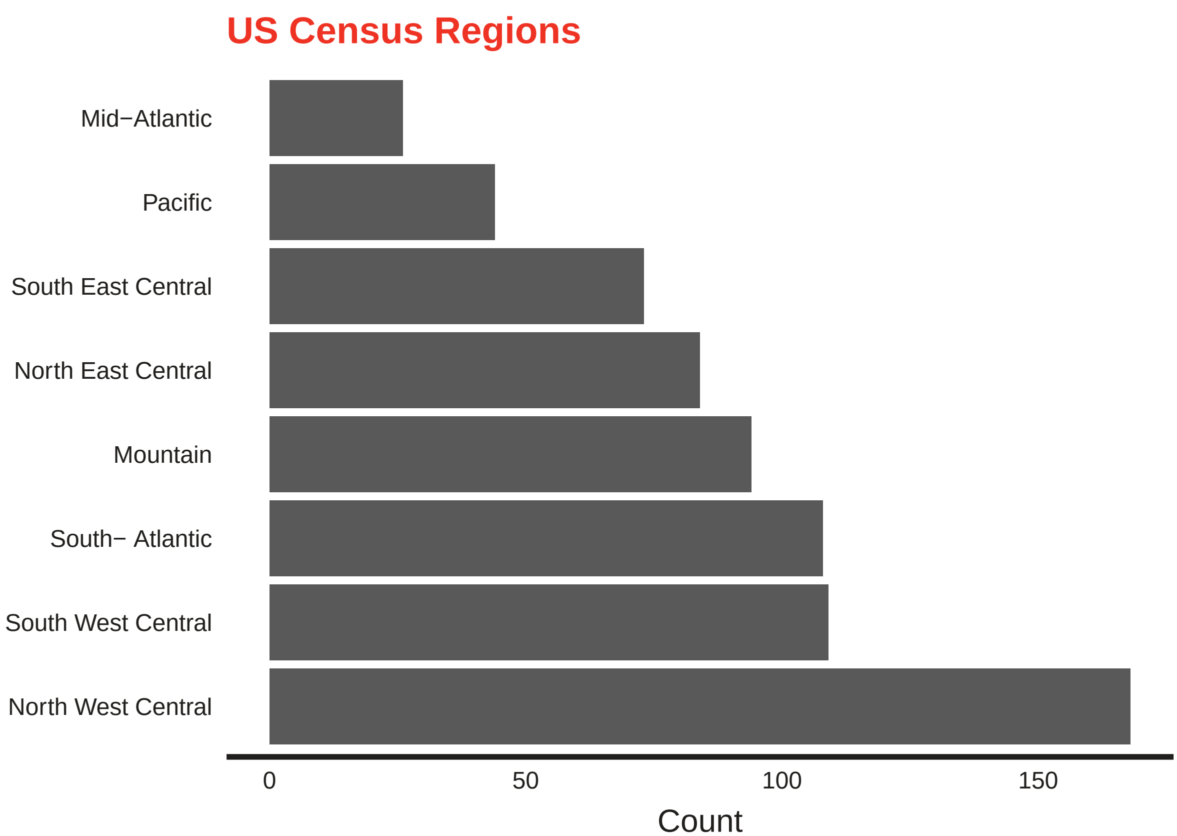


Figure 1: US Census Regions

Moreover, a meticulous examination of the correlations among independent variables yields noteworthy insights. Notably, a substantial correlation coefficient of 0.74 emerges between `routine` and `outsource`, indicative of a robust linear relationship between the percentage of workers engaged in routine tasks and the extent of outsourcing activities. Similarly, the variables `female` and `college` exhibit a noteworthy correlation coefficient of 0.632, suggesting a discernible connection between gender distribution and the educational attainment of the population.

The identification of these correlations prompts a critical consideration of potential multi-collinearity issues in the impending regression modeling phase. Multi-collinearity, characterized by high correlations among independent variables, has the potential to introduce instability to regression coefficients, complicating the isolation of individual variable effects. This pivotal concern will be revisited and meticulously addressed in the subsequent modeling section to fortify the robustness of our regression analysis.

Additionally, it is imperative to address skewness apparent in two variables: Percentage change in trade with China and `foreign_born`. The recognition of pronounced skewness, indicative of a departure from the normal distribution, necessitates corrective measures. In a strategic effort to enhance the validity of subsequent analyses, these variables will be subjected to logarithmic transformations during the modeling phase.

In summation, the detailed analysis of Figure 1 enriches our understanding of the intricate relationships and interdependencies among key variables. These nuanced insights not only lay the groundwork for our ensuing regression modeling but also inform decisions in addressing potential statistical challenges such as multi-collinearity and skewness, ensuring a comprehensive and rigorous analytical approach.

Figure 2 below illustrate the states that have suffered steep declines in manufacturing employment. A state like Michigan has suffered steep losses. Other notable losers include Pennsylvania, California, Louisiana and Mississippi.

| Variable | Mean | SD | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|---|
| change_mfg_employ | -2.27 | 2.58 | -14.38 | -3.83 | -2.03 | -0.49 | 4.6 |
| change_china_trade | 2.63 | 3.04 | -0.63 | 0.86 | 1.90 | 3.43 | 43.1 |
| manufacturing | 19.08 | 11.15 | 0.11 | 10.50 | 17.87 | 26.79 | 55.2 |
| college | 48.17 | 8.54 | 26.32 | 41.96 | 49.02 | 54.03 | 70.6 |
| female | 64.21 | 7.12 | 41.34 | 59.58 | 64.08 | 68.97 | 79.6 |
| foreign_born | 5.98 | 6.51 | 0.62 | 2.16 | 3.61 | 6.86 | 48.9 |
| routine | 28.79 | 2.88 | 22.23 | 26.78 | 28.83 | 30.75 | 36.7 |
| outsource | -0.62 | 0.42 | -1.64 | -0.92 | -0.67 | -0.38 | 1.2 |
| lon | -93.94 | 12.10 | -124.06 | -100.78 | -92.93 | -84.61 | -68.6 |
| lat | 38.92 | 5.05 | 25.67 | 35.11 | 39.09 | 42.79 | 48.8 |

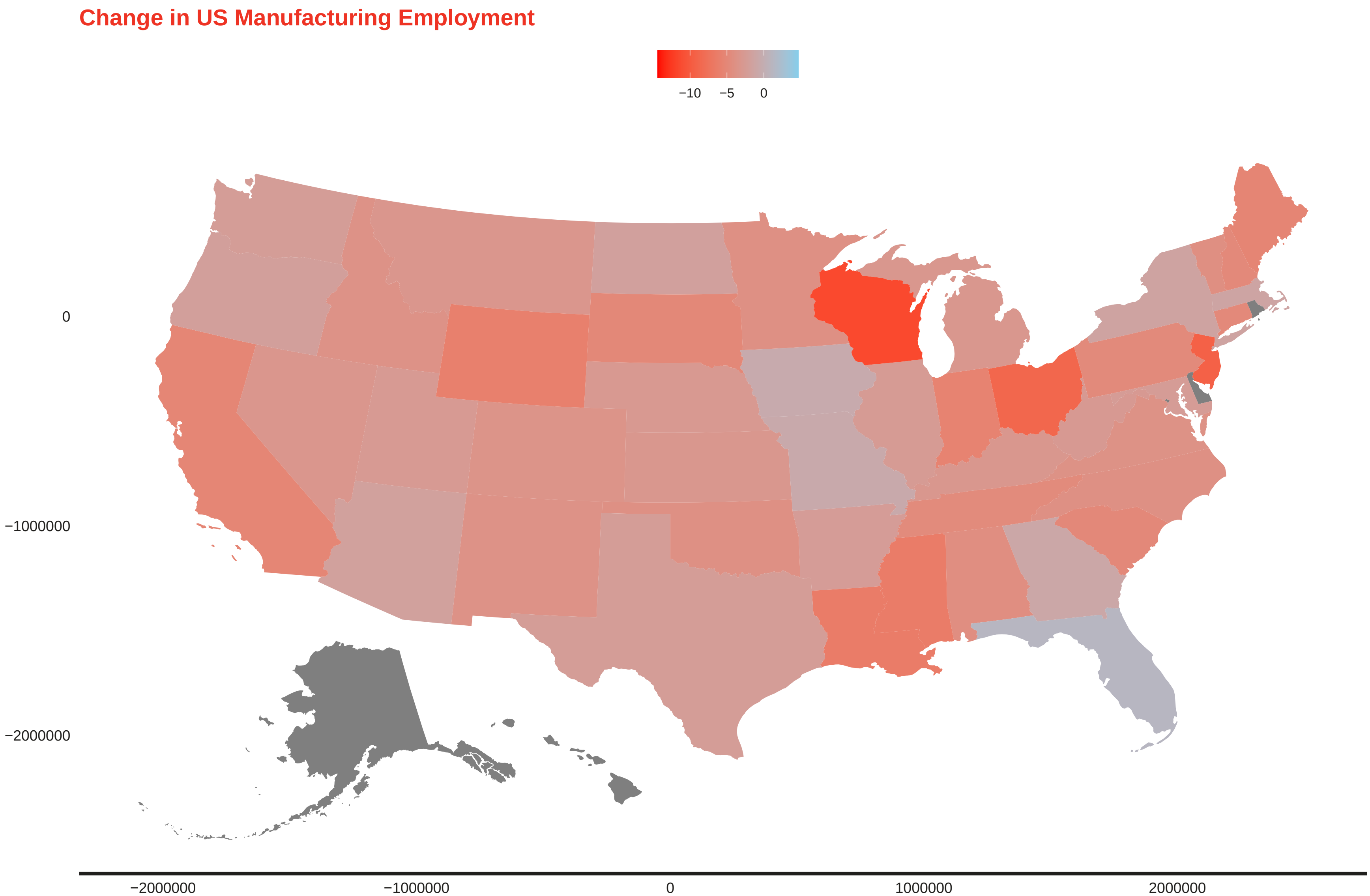


Figure 2: Change in Manufacturing Employment

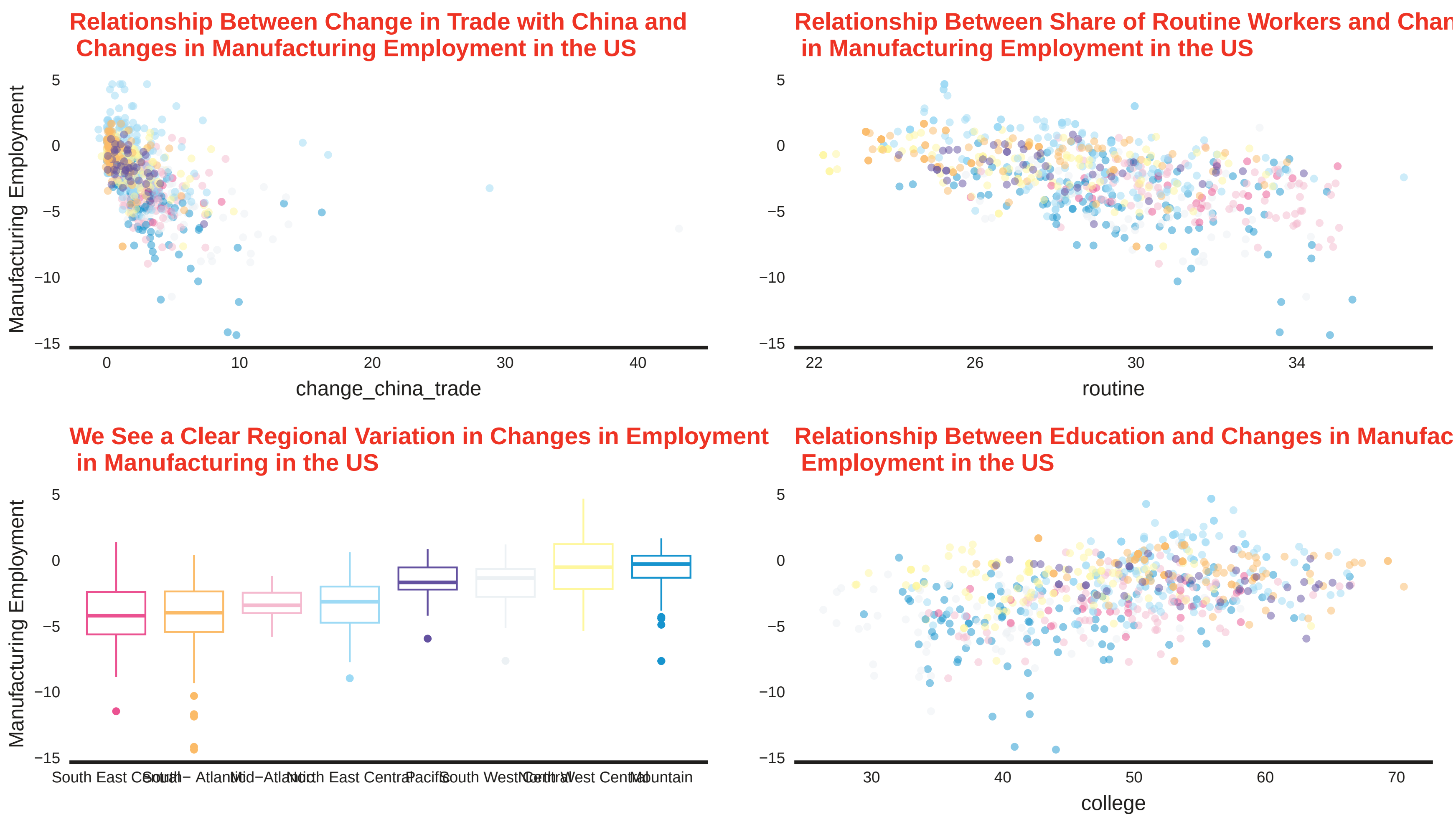


Figure 3: Data visualization

## Modeling

Executing a linear regression analysis utilizing `change_mfg_employ` as the dependent variable (see Table 3). First, I run the regression model against the `region` variable (regression 1). Next, I run the regression model against all other variables in the data (regression 2). Finally, I run the model against variables in regression 2 that are statistically significant (regression 3). We see that the explanatory power of the model does not decline when we remove the variables that are not statistically significant. Specifically, the following variables are not statistically significant;

- change_china_trade.
- famale.
- foreign_born.
- Outsource.
- lat.
- lon.

The following variables are stastically significant.

- Manufacturing.
- College.
- Routine.
- Region.

In all 3 regressions, the models are statistical significance, evident from the F-statistic presented in Table 3. The models showcase a substantial explanatory capacity, elucidated by coefficients of determination ($R^2$) and adjusted ($R^2$) in the regression output, collectively accounting for approximately 61% of the variability in the dependent variable. I explain the meaning of the regression coefficients in Table 3, regression 3.

The variable `manufacturing`, indicative of the initial proportion of employment in manufacturing, exhibits a negative relationship. With a coefficient of -0.075, it implies that, under constant conditions, a unitary surge in the initial share of manufacturing employment corresponds to a -0.075-unit decrease in subsequent manufacturing job shares in the United States. This observation accentuates a sustained decline in the manufacturing sector over the post-2000 decade.

Examining the variable `education`, a positive relationship emerges between manufacturing jobs and the proportion of the population possessing a college education. In a scenario where other factors remain constant, a unitary increase in the share of individuals with a college education corresponds to a 0.069-unit upswing in the share of manufacturing jobs in the US, accentuating the pivotal role of human capital in economic development.

Analysis of `routine` reveals a negative relationship between the share of manufacturing jobs in the US and the proportion of workers engaged in routine tasks. Specifically, under constant conditions, a unitary increase in the share of workers involved in routine activities corresponds to a 0.350-unit decline in manufacturing employment.

Turning attention to the variable `region` (Census Zone), where the `reg_encen` region serves as the reference, distinct regional patterns in the loss of manufacturing jobs to China come to light. Particularly noteworthy is the `reg_wncen` region, experiencing the most substantial decline in manufacturing jobs, followed by `reg_encen`. Figure 7 summarizes the coefficients of the region variable. However, certain variables in the model do not attain statistical significance. Model diagnostics, illustrated in Figure 7, does not raise any major concerns.

Table 3

| | Dependent variable: | | |
|---|---|---|---|
| | change_mfg_employ | | |
| | (1) | (2) | (3) |
| change_china_trade | | −0.023 | |
| | | (0.066) | |
| manufacturing | | −0.075*** | −0.075*** |
| | | (0.009) | (0.007) |
| college | | 0.062*** | 0.067*** |
| | | (0.014) | (0.011) |
| female | | 0.007 | |
| | | (0.018) | |
| foreign_born | | −0.044 | |
| | | (0.120) | |
| routine | | −0.360*** | −0.360*** |
| | | (0.043) | (0.042) |
| outsource | | 0.220 | 0.130 |
| | | (0.310) | (0.270) |
| regionMountain | 2.600*** | −0.097 | −0.140 |
| | (0.460) | (0.400) | (0.390) |
| regionNorth East Central | −0.090 | 0.370 | 0.460 |
| | (0.470) | (0.380) | (0.370) |
| regionNorth West Central | 2.800*** | 1.200*** | 1.400*** |
| | (0.440) | (0.380) | (0.360) |
| regionPacific | 1.800*** | 0.130 | −0.046 |
| | (0.520) | (0.460) | (0.420) |
| regionSouth East Central | −0.990** | −0.230 | −0.096 |
| | (0.480) | (0.410) | (0.380) |
| regionSouth West Central | 1.700*** | 0.630 | 0.650* |
| | (0.460) | (0.380) | (0.370) |
| regionSouth- Atlantic | −0.910** | −1.000*** | −0.880** |
| | (0.460) | (0.370) | (0.360) |
| lon | | 0.011 | |
| | | (0.007) | |
| lat | | 0.022 | |
| | | (0.014) | |
| Constant | −3.400*** | 6.300*** | 6.100*** |
| | (0.410) | (2.000) | (1.500) |
| Observations | 706 | 700 | 706 |
| $R^2$ | 0.350 | 0.600 | 0.600 |
| Adjusted $R^2$ | 0.340 | 0.590 | 0.590 |
| Residual Std. Error | 2.100 (df = 698) | 1.700 (df = 683) | 1.600 (df = 694) |
| F Statistic | 53.000*** (df = 7; 698) | 64.000*** (df = 16; 683) | 95.000*** (df = 11; 694) |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |

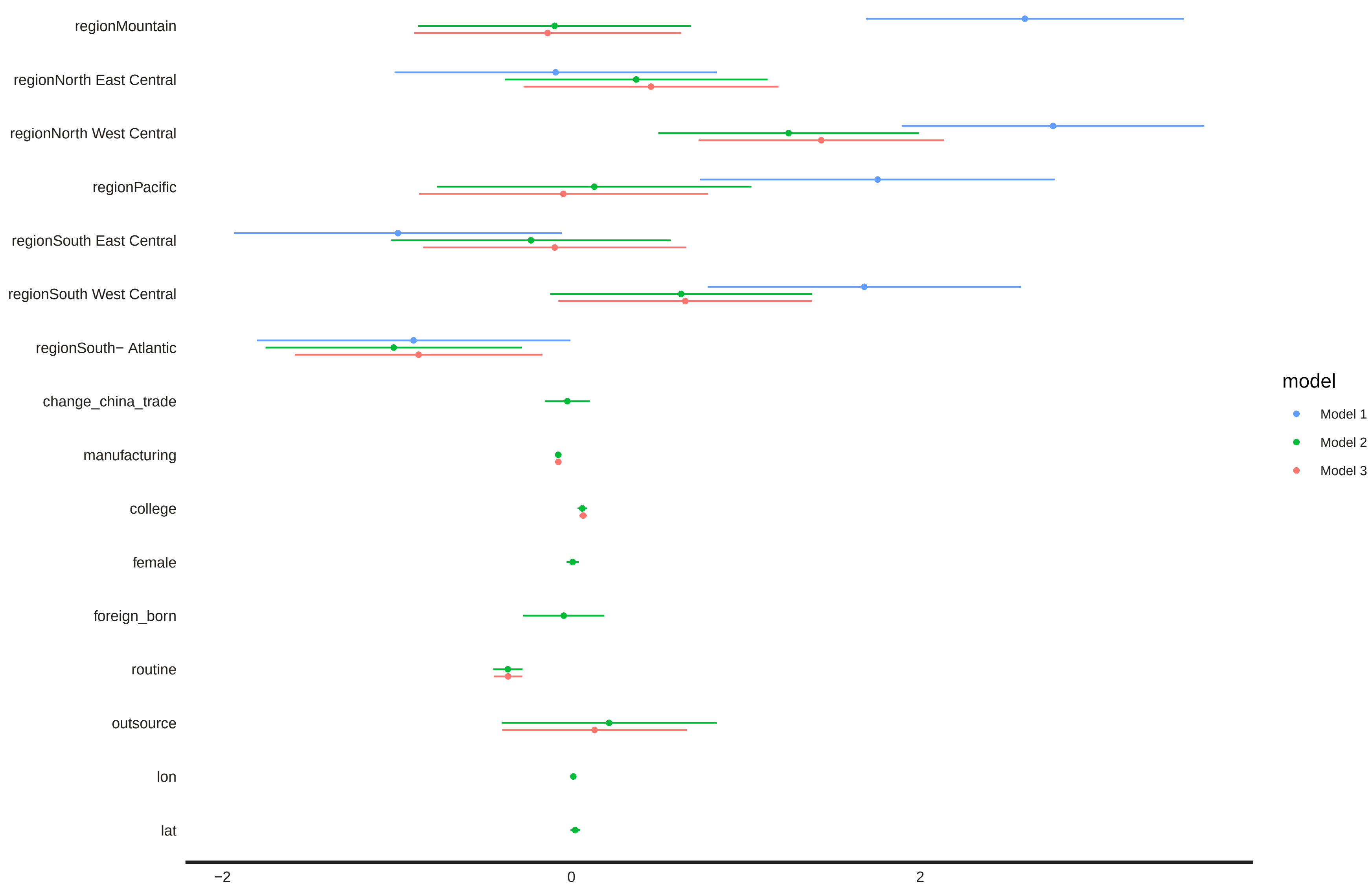


Figure 4: Model Plots

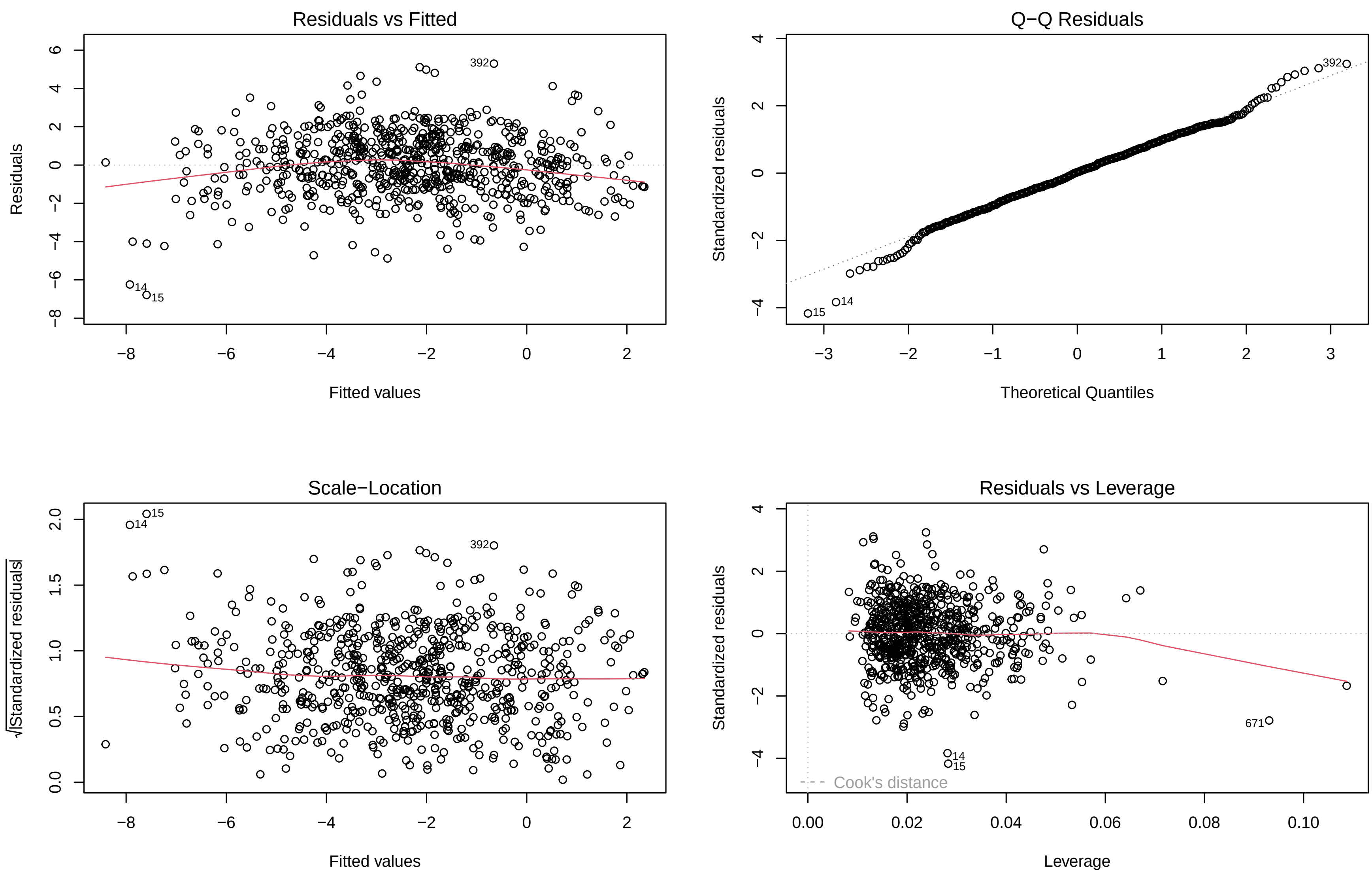


Figure 5: Model Diagnostics

Figure 5 shows the model diagnostics. There are four figures in this analysis. The residuals vs fitted graph should ideally be flat and horizontal. This means that the residuals are randomly distributed and have a constant variance and that the relationship between the independent and dependent variables is linear. In this case, the line does not deviate much from this condition. The residuals "bounce randomly" around the 0 line, suggesting that the assumption that the relationship is linear is reasonable. The residuals roughly form a "horizontal band" around the 0 line, meaning that the variances of the error terms are equal. The QQ line shows the normality of residuals. The model shows that the residuals are reasonably normal, meaning that we can use this model for predictions.

The scale location checks for heteroscedasticity. Ideally, the line should be flat and horizontal. Again, the graph shows that the model does not suffer from heteroscedasticity. The last graph checks for the presence and influence of outliers. The line is roughly horizontal across the plot and does not a notable pattern. We find that the data does not have outliers that could unduly affect the performance of the model.

# Decision Tree

In this section, I analyse the data using a decision tree model. A Decision Tree model is a popular machine learning algorithm known for its simplicity, interpretability, and effectiveness in both classification and regression tasks. It mimics human decision-making by recursively partitioning the data based on features, ultimately leading to a set of rules that guide predictions. The model is structured as a tree, where each internal node represents a decision based on a specific feature, and each leaf node represents the predicted outcome. At each decision node, the algorithm selects the feature that best splits the data into homogeneous subsets, optimizing a chosen criterion (e.g., Gini impurity for classification, mean squared error for regression). We see that region is of first order importance in explaining the loss of manufacturing jobs in the US. Thus, there are significant regional variations in loss of manufacturing jobs, as noted in the regression analysis. manufacturing and change_china_trade occupy the second tier. Other important factors in the decision tree include longitude, routine, and female. So, unlike in the linear model where `change_china_trade`, `female` and `longitude` were no significant, the decision tree model attaches great importance to these variables in determining the change in manufacturing jobs in the United States of America.

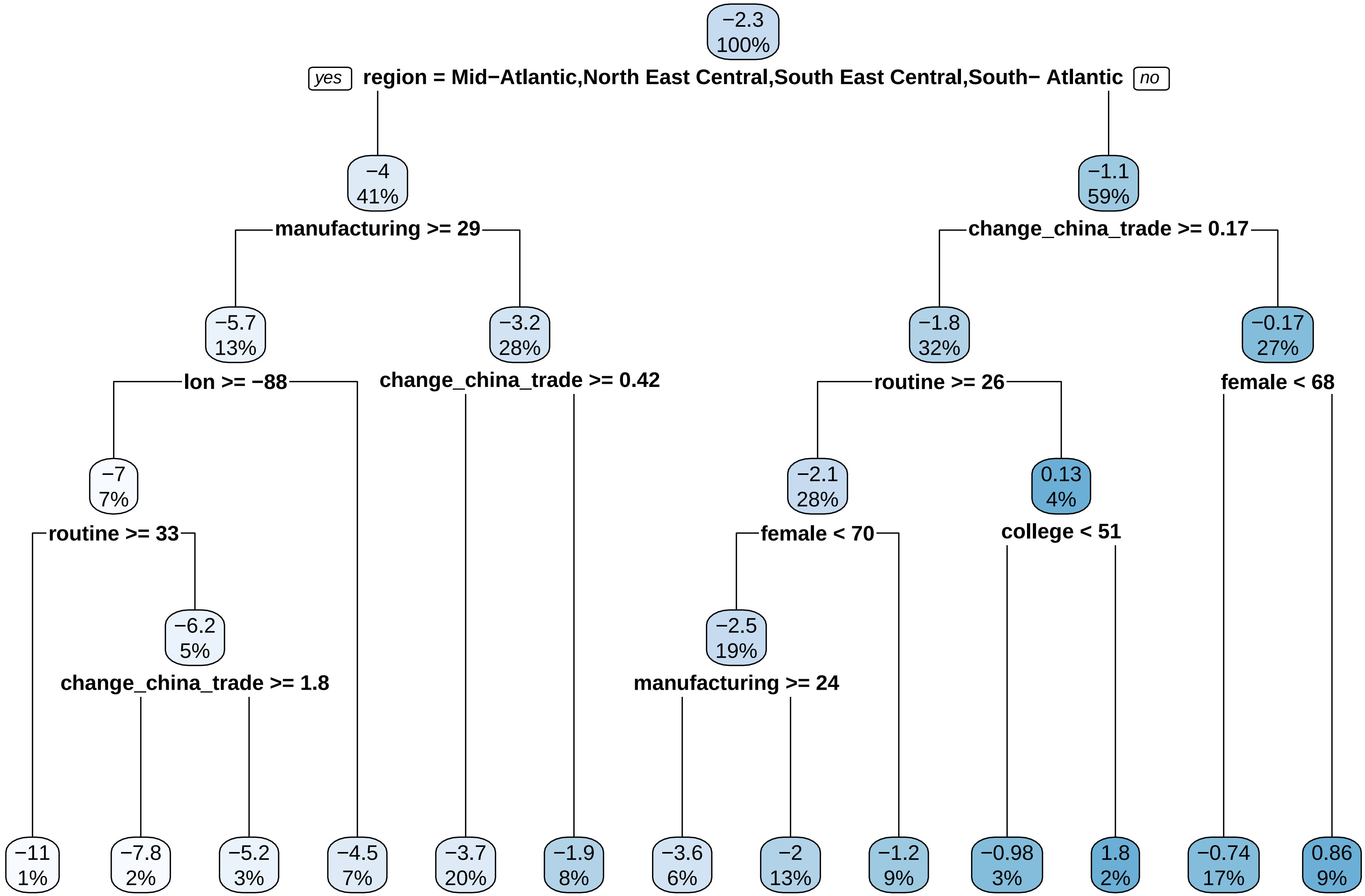


Figure 6: Decision Tree Model: Manufacturing Jobs in US

# Conclusion

In conclusion, our comprehensive analysis of the impact of trade with China on manufacturing employment in the United States during the 2000s reveals compelling insights. The linear regression model, encompassing various pertinent variables, underscores the multifaceted nature of factors influencing manufacturing employment dynamics. Notably, the variable `change_china_trade` emerges as a key contributor, signifying a discernible inverse relationship between increased trade with China and a subsequent decrease in US manufacturing employment.

The regional variable (`region`) assumes paramount importance in our findings, shedding light on distinct geographical patterns in the loss of manufacturing jobs to China. Particularly noteworthy is the `reg_wncen` region, which has experienced the most substantial decline in manufacturing employment, surpassing even the reference `reg_encen` region. This regional nuance highlights the localized nature of the impact, emphasizing the need for nuanced policy considerations tailored to specific geographic contexts.

While other variables such as the initial share of manufacturing employment (`manufacturing`), the education level of the population (`education`), and the nature of work tasks (`routine`) exhibit significant associations, the regional variable stands out as a critical determinant. The observed decline in manufacturing jobs is not uniform across all regions, emphasizing the importance of considering regional nuances when formulating policy responses.

However, it is essential to acknowledge potential challenges, such as the high variance inflation factor (VIF) observed for the `region` variable, indicating concerns related to multicollinearity. Addressing these statistical considerations is crucial for ensuring the stability of coefficients and the robustness of the conclusions drawn (Wooldridge, Wadud, and Lye 2016).

In essence, our study contributes valuable insights to the ongoing discourse on the intricate interplay between international trade dynamics and regional employment patterns. The nuanced examination of the regional variable provides a foundation for more targeted and effective policy interventions aimed at mitigating the adverse effects of manufacturing job losses, fostering regional resilience, and fostering a more sustainable and equitable economic landscape.

# References


Baily, Martin Neil, and Barry P Bosworth. 2014. "US Manufacturing: Understanding Its Past and Its Potential Future." *Journal of Economic Perspectives* 28 (1): 3–26.

Hombert, Johan, and Adrien Matray. 2018. "Can Innovation Help US Manufacturing Firms Escape Import Competition from China?" *The Journal of Finance* 73 (5): 2003–39.

Li, Chunding, and John Whalley. 2021. "Trade Protectionism and US Manufacturing Employment." *Economic Modelling* 96: 353–61.

Pierce, Justin R, and Peter K Schott. 2016. "The Surprisingly Swift Decline of US Manufacturing Employment." *American Economic Review* 106 (7): 1632–62.

Wooldridge, Jeffrey M, Mokhtarul Wadud, and Jenny Lye. 2016. *Introductory Econometrics: Asia Pacific Edition with Online Study Tools 12 Months*. Cengage AU.

# Appendices

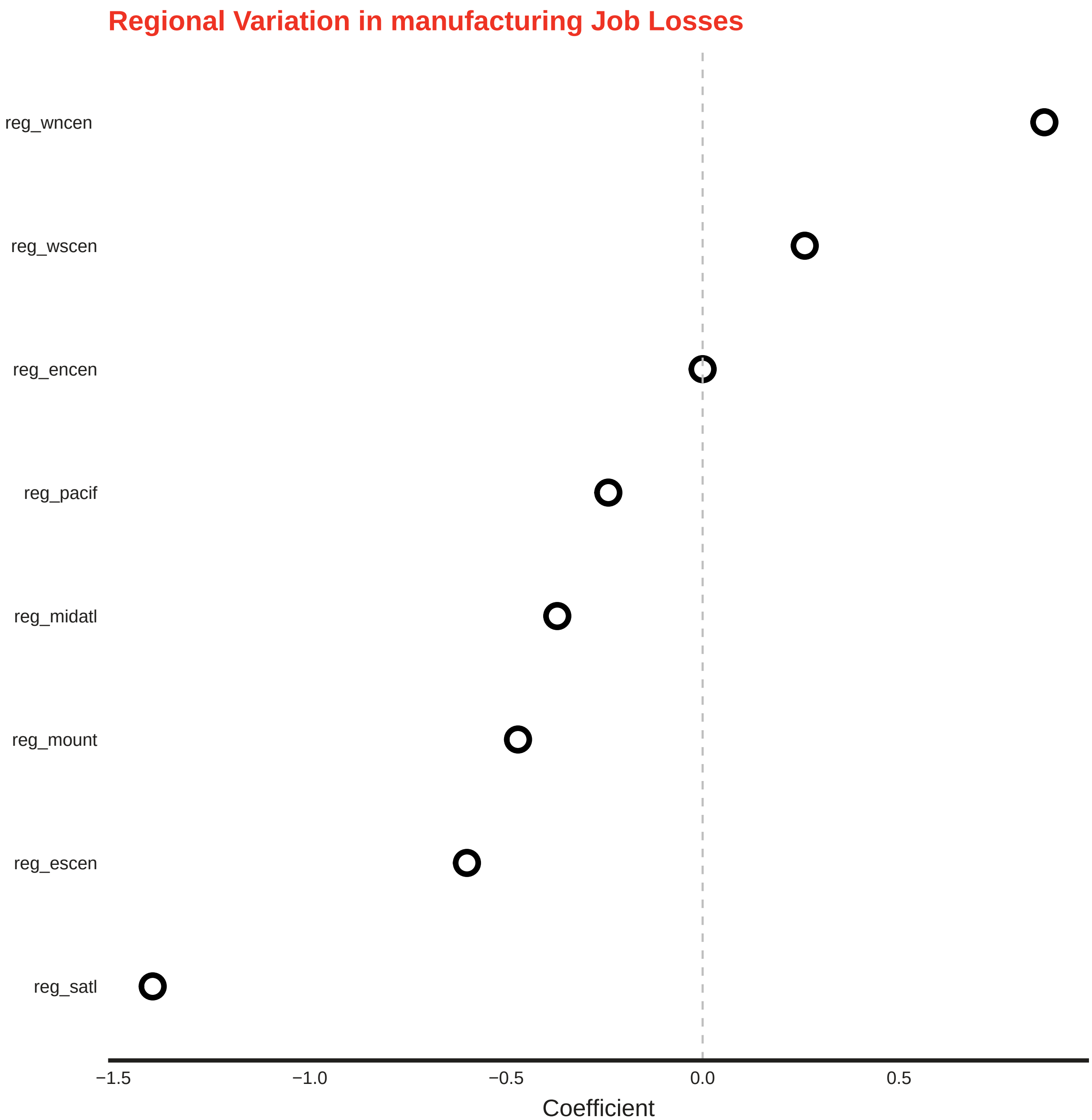


Figure 7: Appendix 2: Regression Coeffients on Regional Disparity in Loss of manufacturing Jobs.

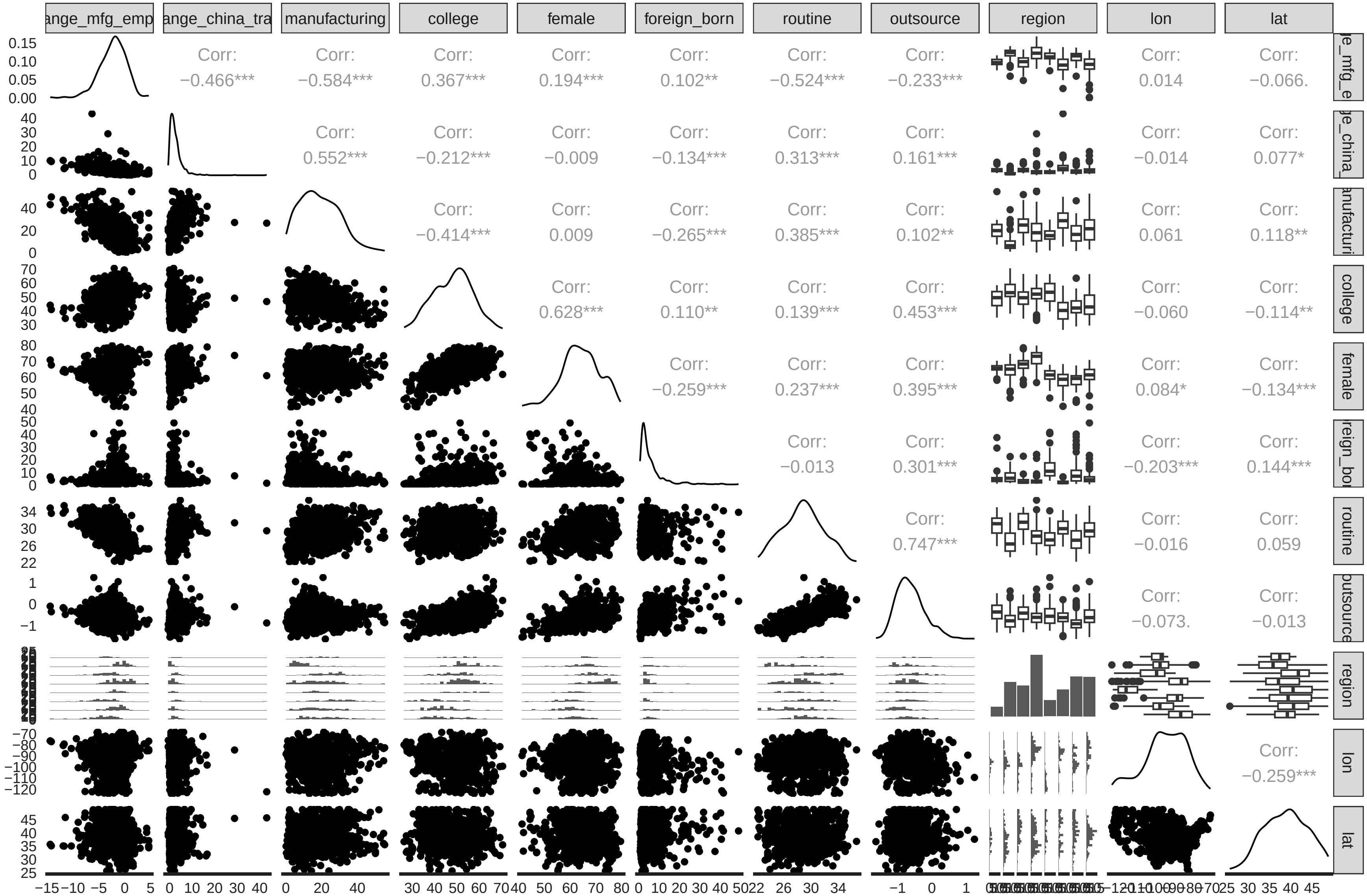


Figure 8: Pairs Plots

Table 4: Summary of Numeric Variables

| Variable | Mean | SD | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|---|
| change_mfg_employ | -2.27 | 2.58 | -14.38 | -3.83 | -2.03 | -0.49 | 4.6 |
| change_china_trade | 2.63 | 3.04 | -0.63 | 0.86 | 1.90 | 3.43 | 43.1 |
| manufacturing | 19.08 | 11.15 | 0.11 | 10.50 | 17.87 | 26.79 | 55.2 |
| college | 48.17 | 8.54 | 26.32 | 41.96 | 49.02 | 54.03 | 70.6 |
| female | 64.21 | 7.12 | 41.34 | 59.58 | 64.08 | 68.97 | 79.6 |
| foreign_born | 5.98 | 6.51 | 0.62 | 2.16 | 3.61 | 6.86 | 48.9 |
| routine | 28.79 | 2.88 | 22.23 | 26.78 | 28.83 | 30.75 | 36.7 |
| outsource | -0.62 | 0.42 | -1.64 | -0.92 | -0.67 | -0.38 | 1.2 |
| lon | -93.94 | 12.10 | -124.06 | -100.78 | -92.93 | -84.61 | -68.6 |
| lat | 38.92 | 5.05 | 25.67 | 35.11 | 39.09 | 42.79 | 48.8 |

Table 5: Summary of Character Variables

| skim_variable | character.min | character.max | character.empty | character.n_unique | character.whitespace |
|---|---|---|---|---|---|
| region | 7 | 18 | 0 | 8 | 0 |
| state | 4 | 14 | 0 | 46 | 0 |