DEPARTMENT OF STATISTICAL SCIENCES, UNIVERSITY OF TORONTO

# Spatial Statistical Analysis for Influencing Factors on Donald Trump in 2016 Elections in Wisconsin Based on Bayesian Inference

Jiaqi Bi
December 2, 2021

ABSTRACT This report analyzes important factors that Trump won the election in 2016 specifically in Wisconsin. The results show statistical significance in population density has small indexes in affecting the voters, a strong significance that White populations are supporting Trump, and Indigenous people have higher probability voting against Trump. The analysis proves the spatial effects as well.

## 1 INTRODUCTION

With the American election results in 2016 showed a predominate trend for Donald Trump, we are interested in if there are demographic factors or spatial patterns for the election result including urban and rural phenomenon, racial factors that Trump is more attractive to White voters. Besides, there are probably other factors that influence the result that need to use spatial data analysis to see the pattern, such as distribution of Trump supporters throughout Wisconsin.

## 2 METHODS

### 2.1 STATISTICAL MODELLING

Since we are interested in if an individual region $i$ is voting for Trump with finite independent $N_i$ times of observation, we are going to use Binomial distribution for the outcome $Y_i$:

$$Y_i \sim \text{Binomial}(N_i, \rho_i)$$

where $\rho_i$ is the probability of county $i$ voting for Trump. We are interested in the odds of voting, and some possible factors influencing the vote:

$$\log[\rho_i/(1-\rho_i)] = \mu + X_i\beta + U_i$$

that $\mu$ is the intercept of the regression, and $X_i$ is the vector of possible influencing factors such as proportion of each region which is White and Indigenous, Log Ratio of total population and surface area in $km^2$. We have $\beta$ priors of $\beta_0 \sim N(0,\infty)$, and $\beta_1, \beta_2, \beta_3 \sim N(0,1000)$ as the default parameter. $U_i$

is spatial random effects with spatial structured variance $\sigma^2$ and spatial independent variance $\tau^2$.

$$U_i \sim \text{BYM}(\sigma^2, \tau^2)$$

Using BYM model in Bayesian inference incorporates spatial random effects into the log-linear model for the relative risk. The inclusion of these random effects allows smoothing relative risks at certain level. We also have the marginal standard deviation $\theta_1$ and spatial proportion $\theta_2$ where

$$\theta_1 = \sqrt{\sigma^2 + \tau^2}$$

$$\theta_2 = \sigma / \sqrt{\sigma^2 + \tau^2}$$

with the prior distribution of $P(\theta_1 > \log(2.5) = 0.5$ and $P(\theta_2 > 0.5) = 0.5$. Based on the data of Wisconsin, we can get the prior distributions have the prior median of $\log(2.5)$ for $\theta_1$ and 0.5 for the spatial proportion $\theta_2$.

## 2.2 Spatial Data Visualization

In Wisconsin, the majority of electoral districts tends to have more reds than blues as shown in Figure 2.1. The redder means the county is in favor of Donald Trump, and the bluer indicates the county is voting against Donald Trump. The white areas illustrate a more neutrally distributed ratio. Trump supporters are highly located in the center to the northwest, and on the southeast except a small area near the Michigan Lake. Figure 2.2 shows a population density in Wisconsin, and the redder areas are highly populated and the bluer areas are lesser populated. We can see that more people are living in the southeast territory of Wisconsin by the Lake Michigan.

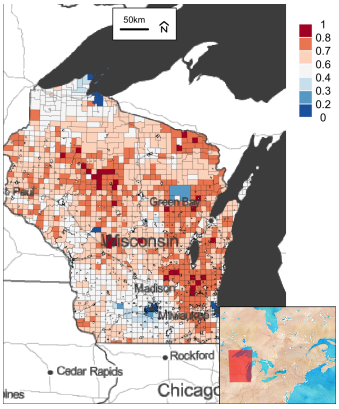


Figure 2.1: Trump Supporter Spatial Map

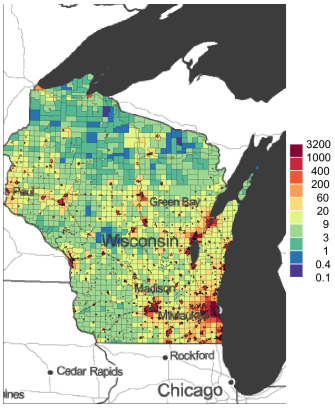


Figure 2.2: Population Density Spatial Map

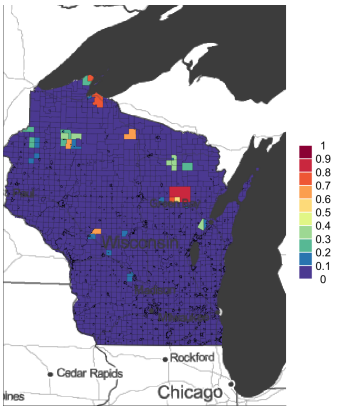


Figure 2.3: Indigenous Population Spatial Map

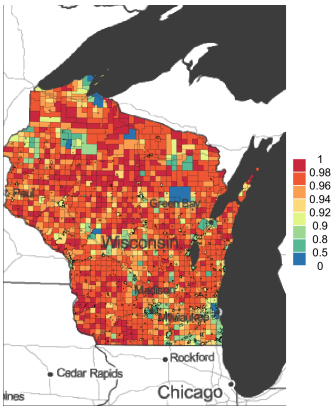


Figure 2.4: White Population Spatial Map

Based on Figure 2.3, Indigenous populations are identified where the redder grids indicate more percentage of Indigenous people are living in relatively, and more blue indicates less percentage of Indigenous people are residing. The dark red is the whole population for the area is considered as indigenous, and the purple area conversely has no indigenous populations. Moreover, in Figure 2.4, White people are residing in redder areas, and more blue means less White people are living in.

## 3 Results

According to spatial maps of relative risks about population density, Indigenous population and White population, we can observe a small patterns of blue area in the center and more northeast in Figure 2.3 and 2.4. Indigenous people and non-White populations are active in that area. However, we need a deeper analysis of random effects and fitted model as shown in Figure 3.1 and Figure 3.2.

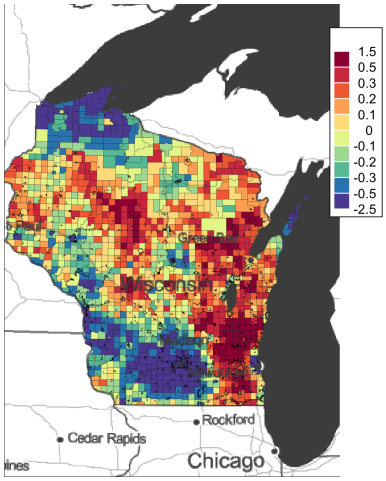


Figure 3.1: Random Effect Spatial Map

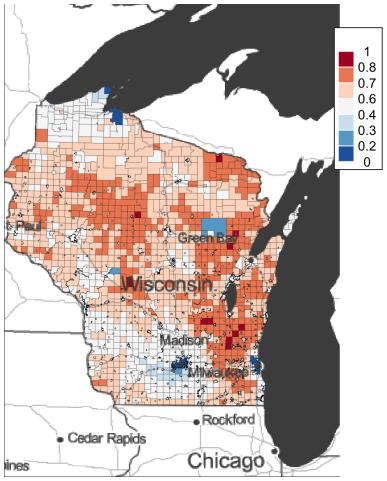


Figure 3.2: Fitted Model Spatial Map

Comparing the random effects $E(U|Y)$ and the fitted model $E(\lambda|Y)$ for two spatial maps in Figure 3.1 and 3.2, we can observe the spatial patterns that areas with larger than average rates (redder) have the corresponding higher probability of voting in favor for Trump.

Table 3.1: Posteriors Log Odds Ratios & SD of Spatial Effect and Dependence Parameter

| | Mean | 2.5% Quantile | 97.5% Quantile |
|---|---|---|---|
| Intercept | -0.56 | -0.83 | -0.30 |
| Population Density | -0.08 | -0.09 | -0.07 |
| White Population | 1.42 | 1.15 | 1.68 |
| Indigenous Population | -0.79 | -1.13 | -0.45 |
| SD of Spatial Effect | 0.32 | 0.30 | 0.33 |
| Dependence Parameter | 0.96 | 0.92 | 0.99 |

Table 3.2: Posteriors Odds Ratios & SD of Spatial Effect and Dependence Parameter

| | Mean | 2.5% Quantile | 97.5% Quantile |
|---|---|---|---|
| Intercept | 0.57 | 0.44 | 0.74 |
| Population Density | 0.92 | 0.91 | 0.93 |
| White Population | 4.13 | 3.17 | 5.38 |
| Indigenous Population | 0.45 | 0.32 | 0.64 |
| SD of Spatial Effect | 1.38 | 1.36 | 1.40 |
| Dependence Parameter | 2.61 | 2.50 | 2.68 |

Table 3.1 shows a log odds of parameters of our initial question of interests. We take the exponential for all of the terms and obtained Table 3.2, which shows a more direct odds ratio on each parameter if it plays a significant role in voting for Trump and we obtained the odds for standard deviation of random effects $sd(U_i)$ as well. The average odds of Population Density significantly concludes that less dense in population for 1 unit will gain a 8% higher voting probability for Trump. White Population in Wisconsin shows a strong significance with the odds of 4.13 in voting for Trump, and Indigenous people are strongly voting against Trump as it has the odds of 0.45. The log odds of dependence parameter is very close to 1 and standard deviation of spatial effect have shown a significant spatial effect. All 95% credible intervals do not contain 0, so we say the result is sufficient to make above conclusions.

# 4 Conclusion

Thus, based on our findings in the spatial maps and tables, we can conclude that less density will have a slightly higher probability of voters voting for Trump. White populations tend to have strong willing to vote for Trump across Wisconsin. Conversely, Indigenous people are more likely voting against Trump. Spatial maps of posterior means and predicted probability are significant where we make the conclusion that the predicted probability is accurate and there are spatial effects.