R Module 10

Harrison Brown

November 07, 2021

Spatial Regression

Queen's Rule Neighbors

Using `poly2nb` from the `spdep` package gives the following results

```
neighbors <- poly2nb(election_sf, queen = T)
```
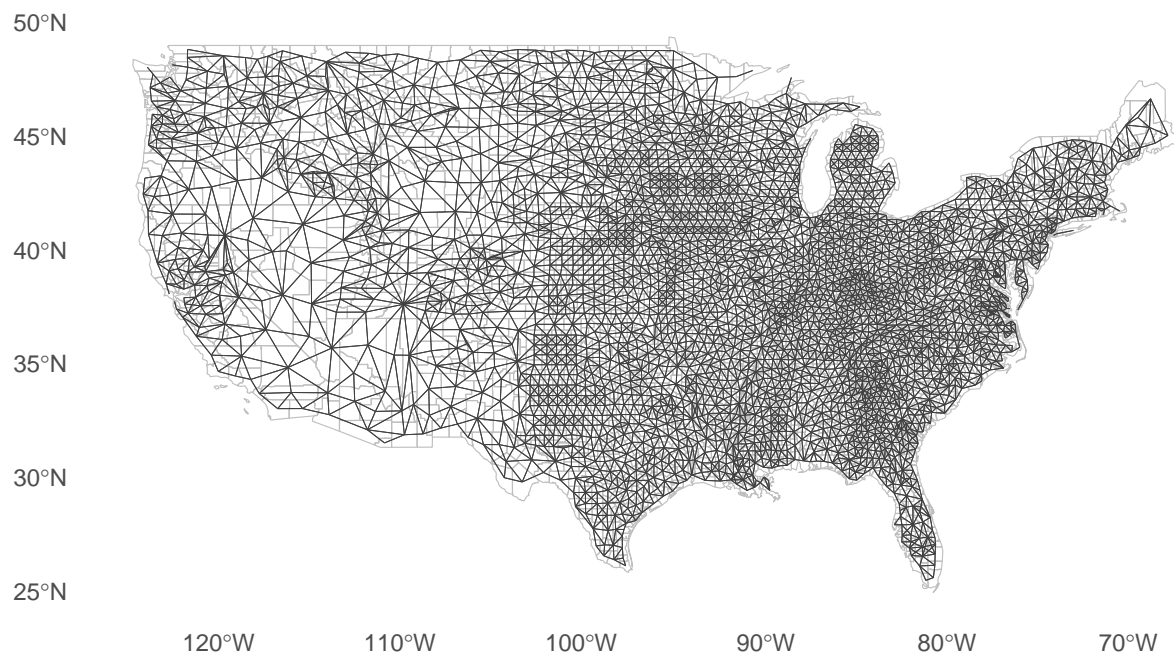

Figure 1: Queen's Method neighbors

1. Residuals

Plotting the residuals of `Bush_pct ~ pcincome` from the results of the election gives the following results:

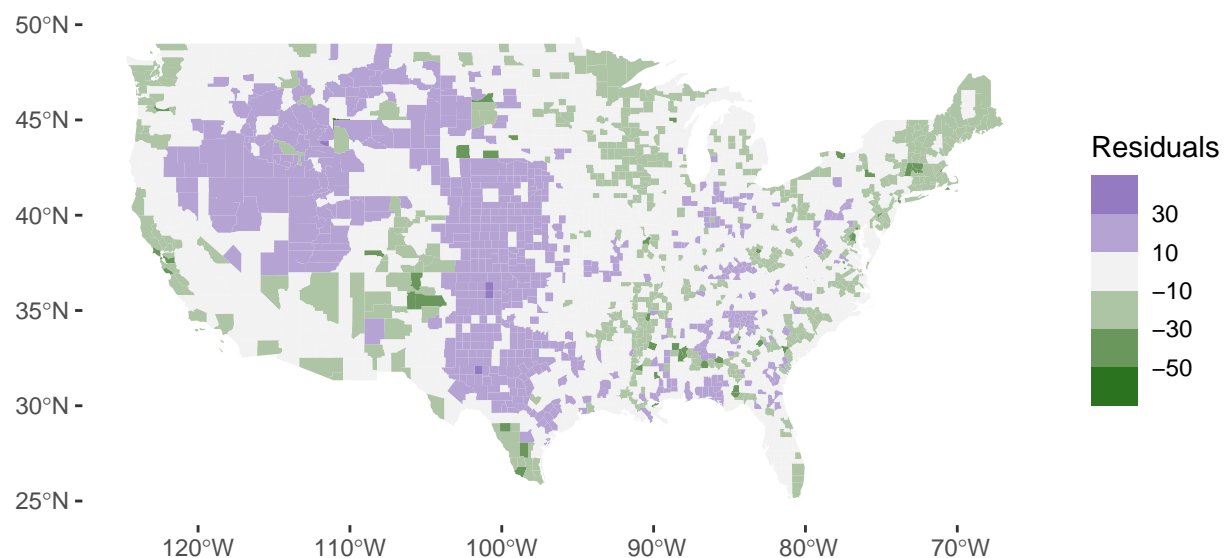
Figure 2: Residuals of Bush_pct ~ pcincome

The results of this plot of residuals appear to correlate with an urban/rural split, with rural, traditionally Republican counties (e.g., Midwest) showing higher residuals, and urban, traditionally Democrat areas (e.g., New England) showing lower residuals.

2. Spatial Lag Model

Spatial lag is a spatial regression model that accounts for autocorrelation among neighbors, defined by the function $\gamma = \rho W \gamma + \beta x + \epsilon$, with the coefficient rho (ρ) representing spatial autocorrelation between neighbors.

Using the formula for spatial lag from `spatialreg::lagsarlm()`, the residuals for the same analysis are as follows:
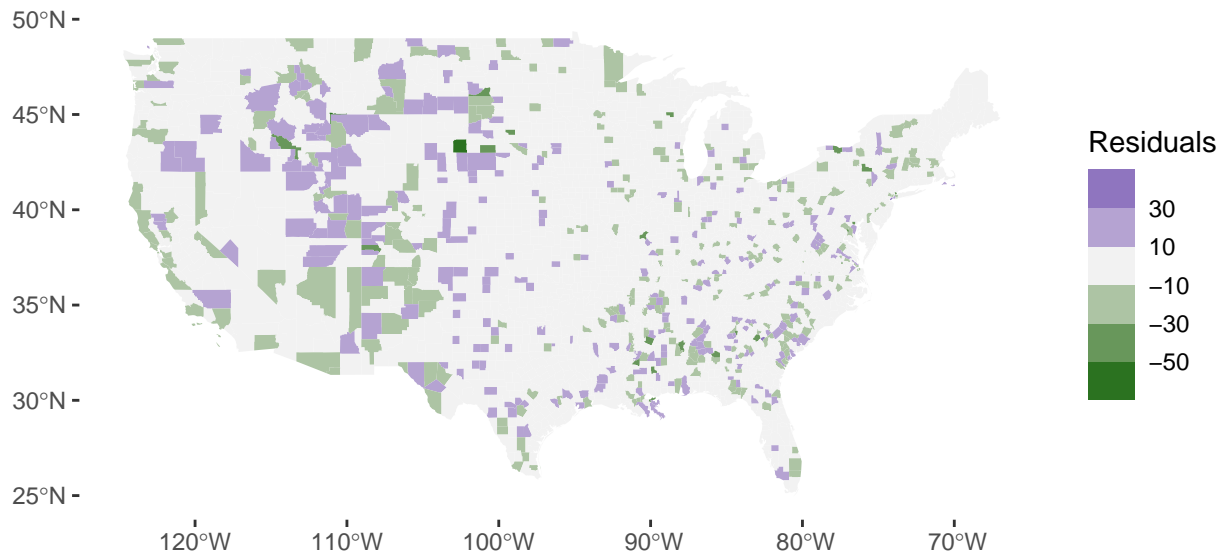

Figure 3: Residuals of Spatial Lag model

3. Moran's I

```
moran.spatiallag <-
  moran.test(x = election_sf$rsdls_l,
             listw = W_cont_el_mat,
             zero.policy = TRUE,
             alternative = "two.sided")
```

Additionally, calculating Moran's I on the residuals of the Spatial Lag model gives a p-value of 1.5756285×10^{-4} and a Moran's I statistic of **-0.0409755**, indicating that when adjusting for spatial lag, the remaining factors caused by spatial autocorrelation are largely removed. The value of the coefficient for spatial autocorrelation among neighbors, ρ, for the spatial lag model is **0.7510418**, indicating a strong positive spatial autocorrelation.

4. Spatial Error

The function for spatial error, $\gamma = \beta x + \epsilon; \epsilon = \lambda W \epsilon + \zeta$, represents spatial autocorrelation separated into two terms, with λ representing the correlation coefficient, and ζ modelling random error.

Estimating the same function when modeling spatial error (with the `spatialreg::errorsarlm()` function), the results are as follows:
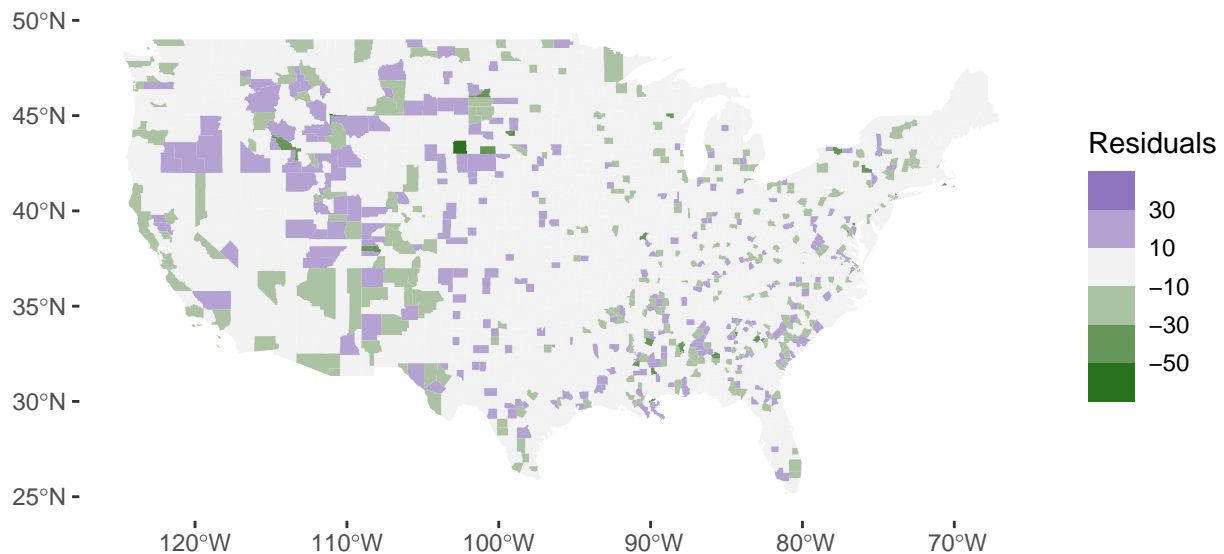

Figure 4: Residuals of Spatial Error model

Moran's I

```
moran.spatialerror <-
  moran.test(x = election_sf$rsdls_e,
             listw = W_cont_el_mat,
             zero.policy = TRUE,
             alternative = "two.sided")
```

The residuals from the spatial error model are quite similar to the spatial lag model, with a Moran's I statistic of **-0.0510921** and a p-value of 2.3674777×10^{-6}. With the Moran's I statistic being close to 0, there is little spatial autocorrelation when accounting for spatial error. As with the spatial lag model, this indicates that the original factors of the data are highly spatially autocorrelated, with a value of $\lambda =$**0.7619652**.

Summary

The summary from the analysis of spatial regression indicates that there is a strong spatial effect on the relationship between per-capita income and percentage of population who voted for Bush within U.S. counties. In addition to per capita income, other variables to consider could include demographic makeup, rural/urban percent, and unemployment rates.