

Causal Evidence on State Effects from a Geographic Regression Discontinuity Design

Markus Neumann
Department of Political Science - Penn State University

Overview

- ▶ There is a well-established, but unproven notion in state politics that states affect the opinions of their citizens.
- ▶ Some scholars have theorized this effect to be the result of 'culture' (Erikson et al. 1993)
- ▶ There is little doubt that there is a correlation, but a *causal effect* requires a greater burden of proof.

Geographic Regression Discontinuity Design

- ▶ Assignment to treatment by the state border
- ▶ Differences between the citizens on the two sides of the border assumed to be as-if random

Distance to the State Border

Given the radius r of the earth (assumed to be 6,378,137 m), the latitude ϕ_1, ϕ_2 and longitude λ_1, λ_2 of two points and the Haversine function:

$$\text{hav}(\theta) = \sin^2\left(\frac{\theta}{2}\right) = \frac{1 - \cos(\theta)}{2}$$

the Haversine distance d between those points is given by:

$$\gamma = \text{hav}(\phi_2 - \phi_1) + \cos(\phi_1) * \cos(\phi_2) * \text{hav}(\lambda_2 - \lambda_1)$$

$$d = 2r * \arcsin(\sqrt{\gamma})$$

This function is then applied to geo-coded data point and every point on the shared state border, determining the shortest distance.

Public Opinion in Geo-Coded Twitter Data

- ▶ Geo-coded (latitude/longitude) tweets from the 'firehose'
- ▶ Dictionary-based sentiment analysis on tweets mentioning Donald Trump

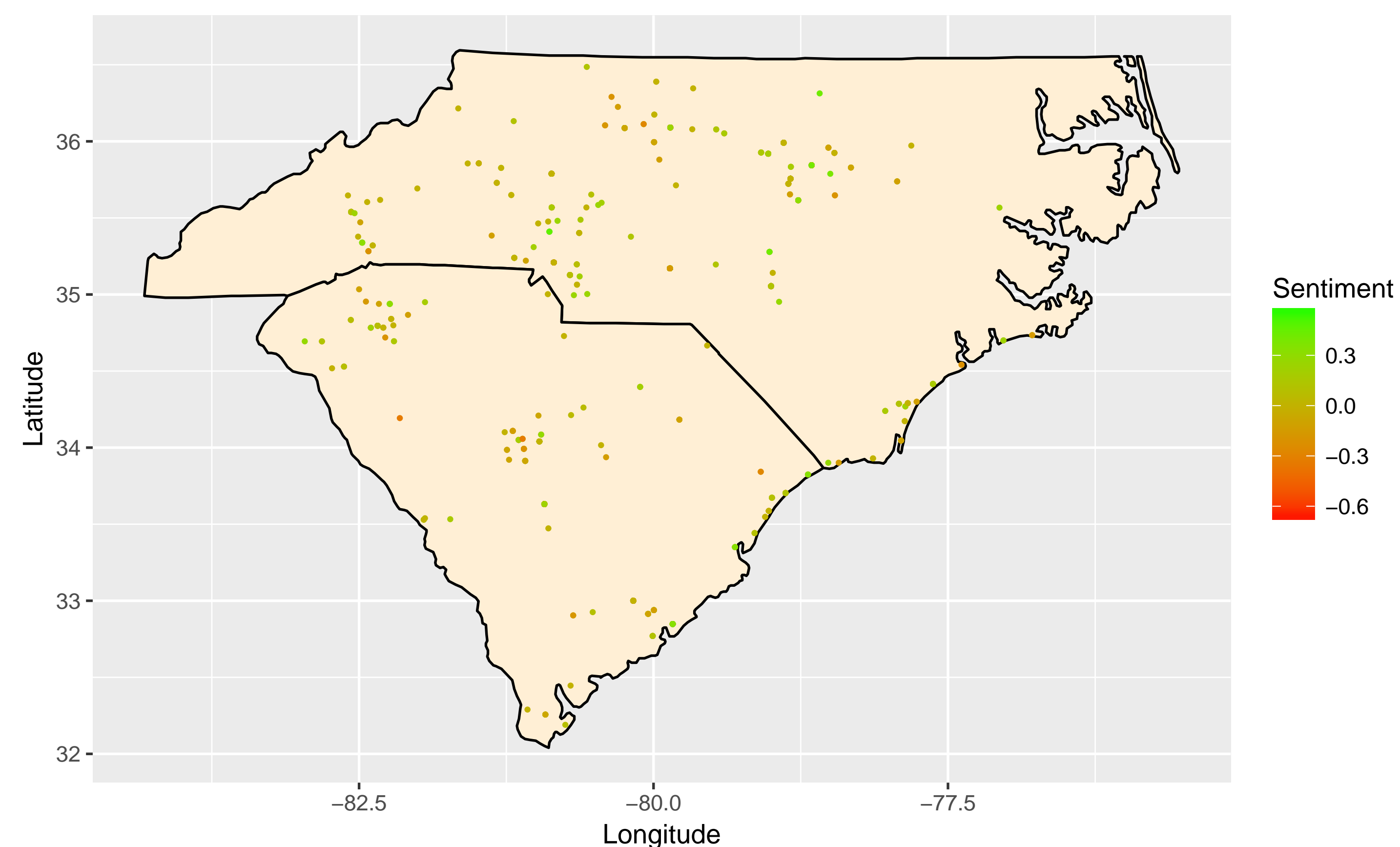


Figure: Map of North Carolina and South Carolina, with tweets superimposed, colored by their sentiment with respect to Donald Trump.

Results

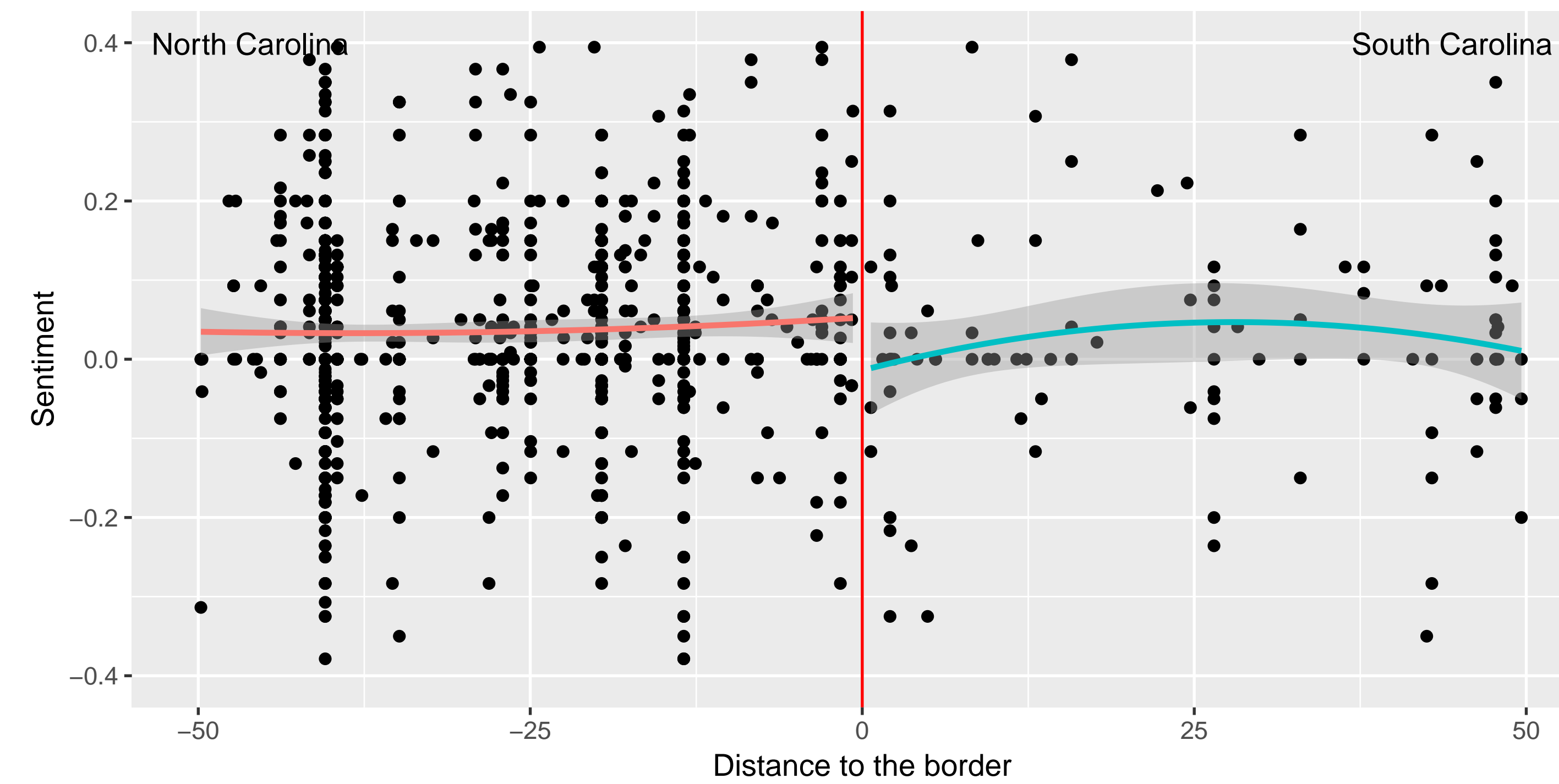


Figure: Sentiment of tweets in North Carolina and South Carolina, relative to their distance to the border. In a regression discontinuity design, the expected result is a sharp change around the cutoff (not present here).

Churches in the States

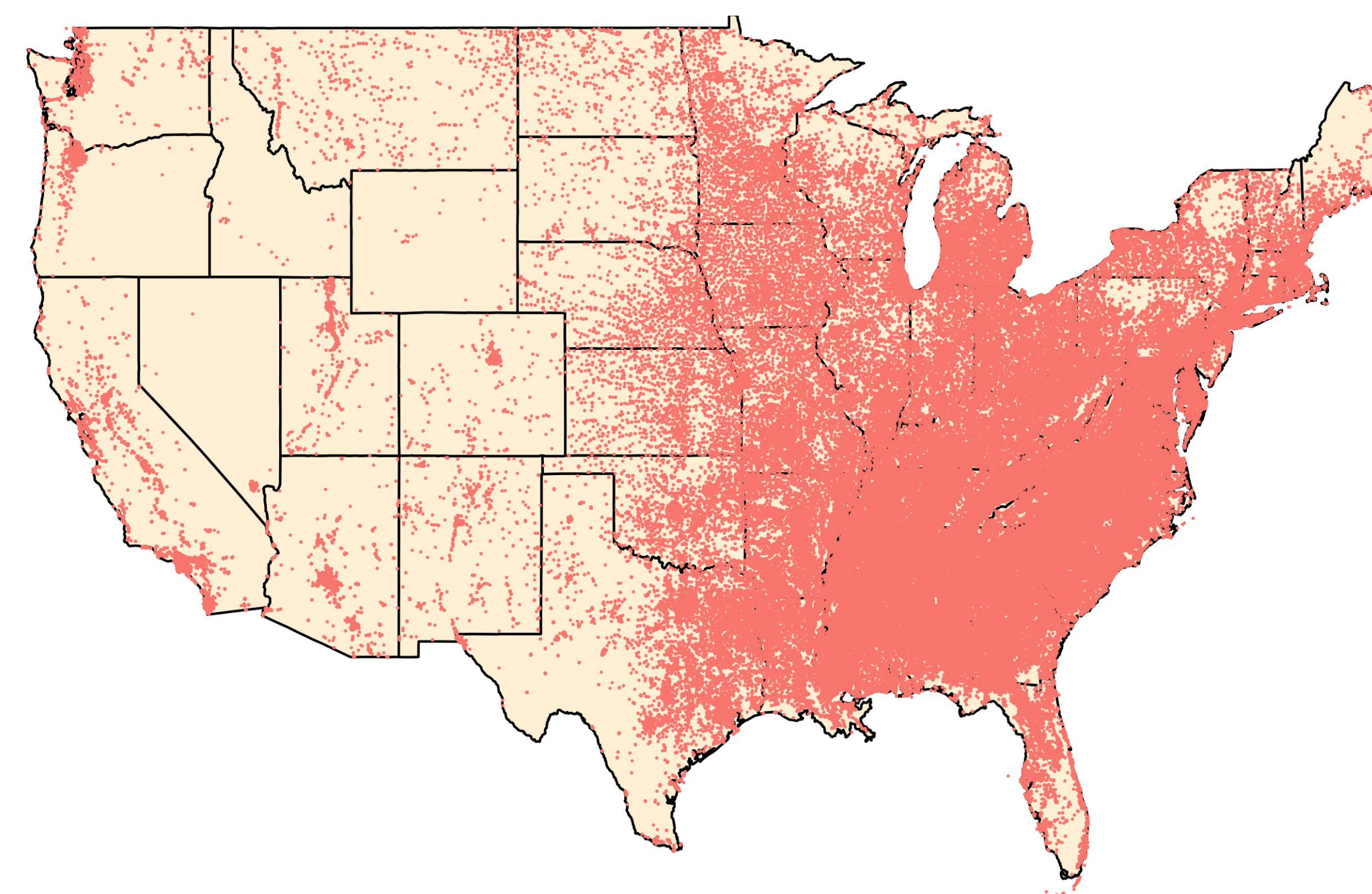


Figure: Geographic locations (latitude/longitude) of churches in the United States.

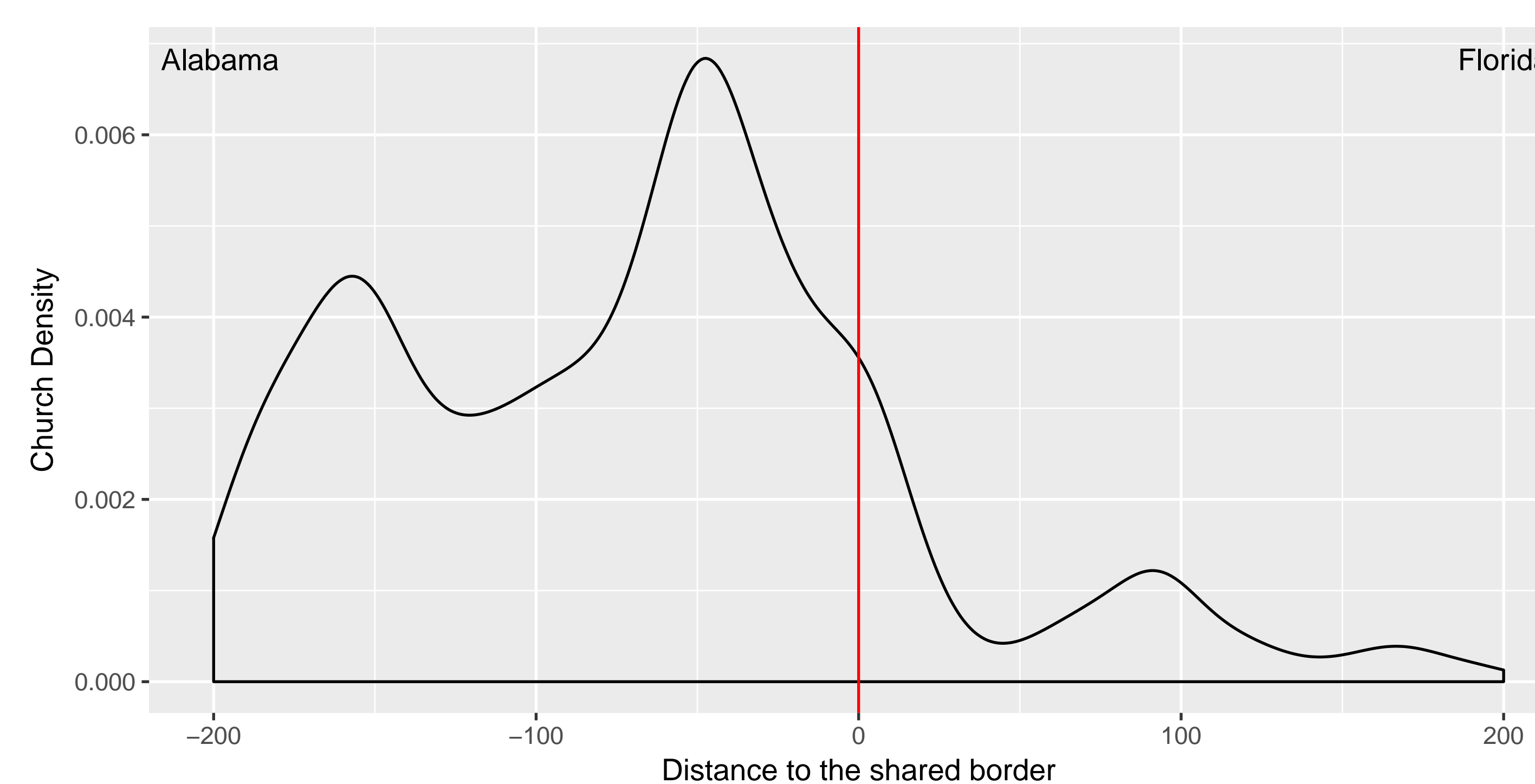


Figure: Church density in Alabama and Florida, relative to the border. In a regression discontinuity design, the expected result is a sharp change around the cutoff, which does occur in this case.

Gun Stores in the States

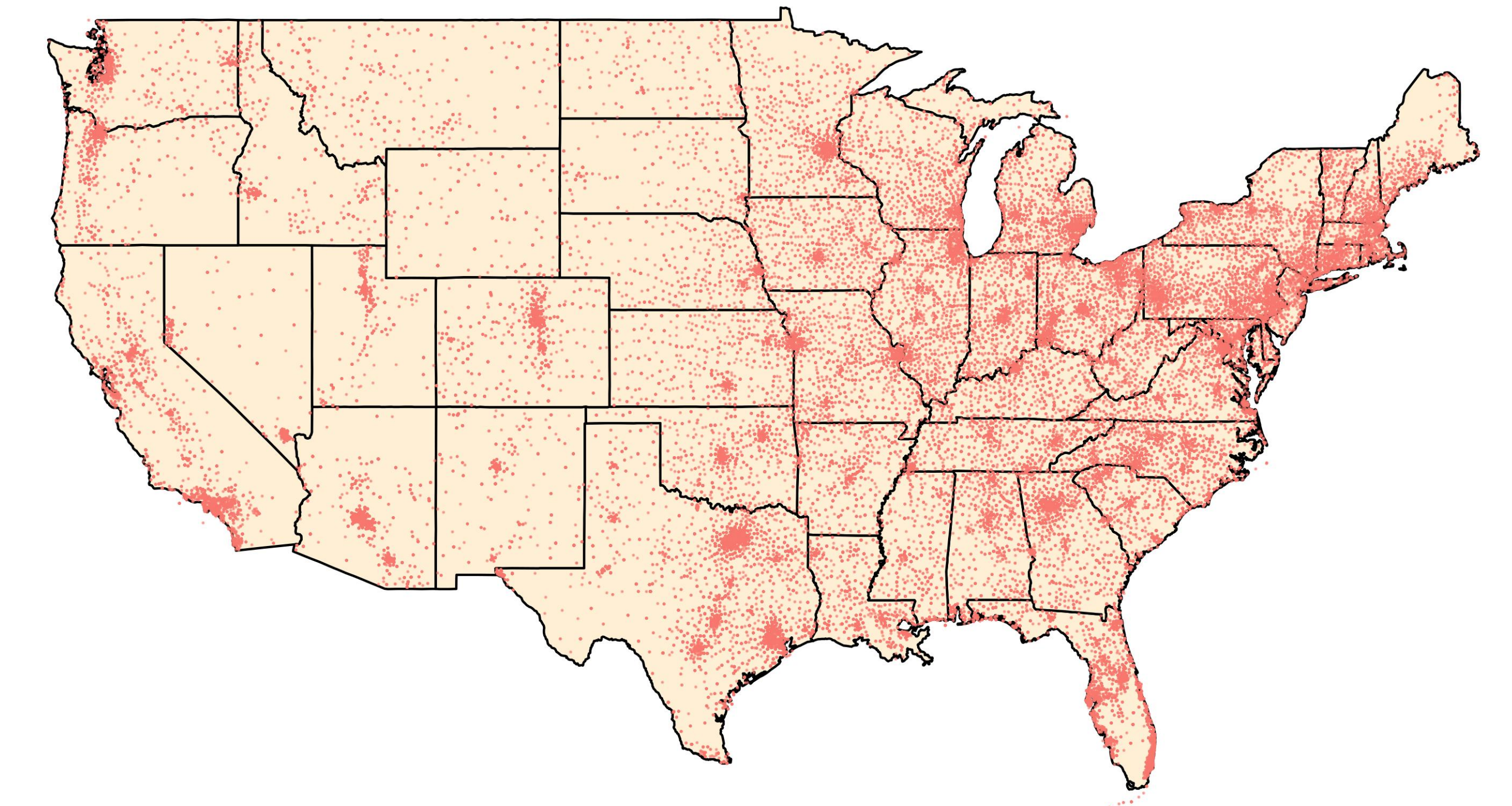


Figure: Geographic locations (latitude/longitude of the geographic midpoint of the respective zip code) of gun stores in the United States.

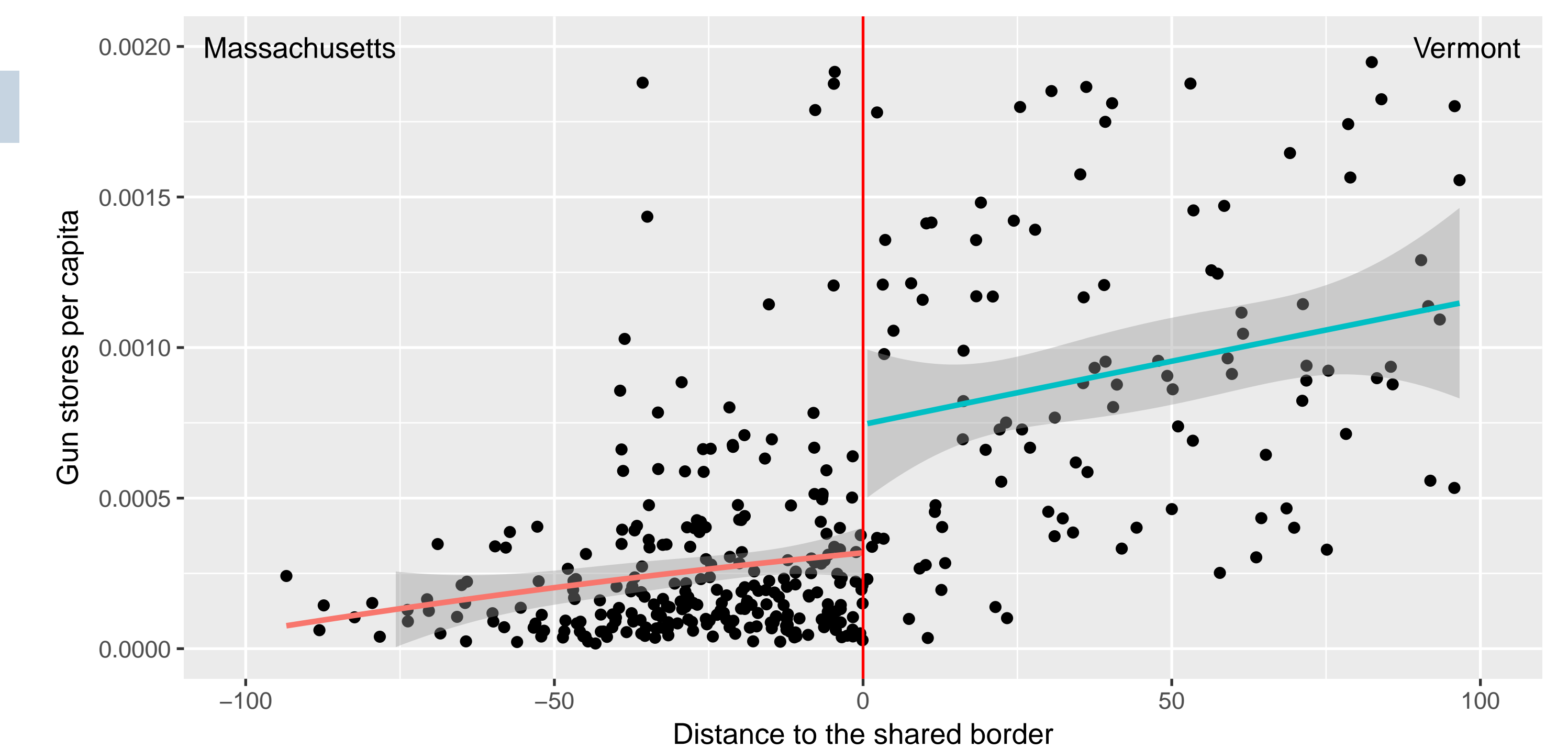


Figure: Gun stores in Massachusetts and Vermont per person in each zip code, relative to its distance to the border. In a regression discontinuity design, the expected result is a sharp change around the cutoff, which does occur in this case.

Conclusion

- ▶ For **public opinion**, no real regression discontinuity is visible at the border - both in the comparison shown here, as well as in all other states
- ▶ The sentiment analysis could be improved by doing stance detection with a neural network
- ▶ There is a clear causal effect visible for **churches** - both in the example here, as well as in many other states
- ▶ There is currently no control for population. This could be improved by either using zip codes instead of exact locations or, alternatively, spatial densities
- ▶ For **gun stores**, the example shown here demonstrates a causal state effect. In many other cases, there is still a discontinuity between the point estimates but not always the confidence intervals.
- ▶ The use of zip codes makes controlling for population easier. The downside is a lack of precision. Street addresses of gun stores could be geo-coded.
- ▶ **Conclusion: causal state effect on culture, but not opinion**