A Variable Selection Approach to Spillovers

Gustavo Diaz[1]

@gusvalo

diazdia2@illinois.edu

[1] *Department of Political Science, University of Illinois at Urbana-Champaign*

Overview

Most methods to analyze spillovers in experiments assume the researcher observes all relevant networks directly (see Aronow et al 2020 for a review). However, most of the time researchers only have a guess of how units are connected in a network. In some cases, researchers can overcome this problem by designing experiments that manipulate the network directly (e.g. Baird et al 2018 and Sinclair et al 2012), yet this solution is not feasible in experiments already conducted or in observational studies.

I propose a variable selection protocol to model spillovers in such settings. This approach assumes the researcher has identified the theoretically relevant network(s) and needs to make an informed choice about how units are connected to each other. For example, if the relevant network is geography, then the task is to identify a reasonable upper bound for how far spillovers travel.

This approach can accommodate different data generating processes and both binary and continuous treatments. The key assumption is that spillovers can be expressed as a linear combination of the outcome variable of interest against a set of predictors that represent the distribution of treated units in the network.

Protocol

1. Identify relevant network
2. Express network as a similarity metric that indicates units' placement on it
3. Calculate pairwise distances between exposed and treated units
4. Create predictors counting treated units within distance ranges
5. Use supervised learning algorithm of choice to perform variable selection (I use adaptive lasso, see Zou 2006 for details)
6. Estimate or test for spillovers

Application

Electoral Irregularities in Ghana

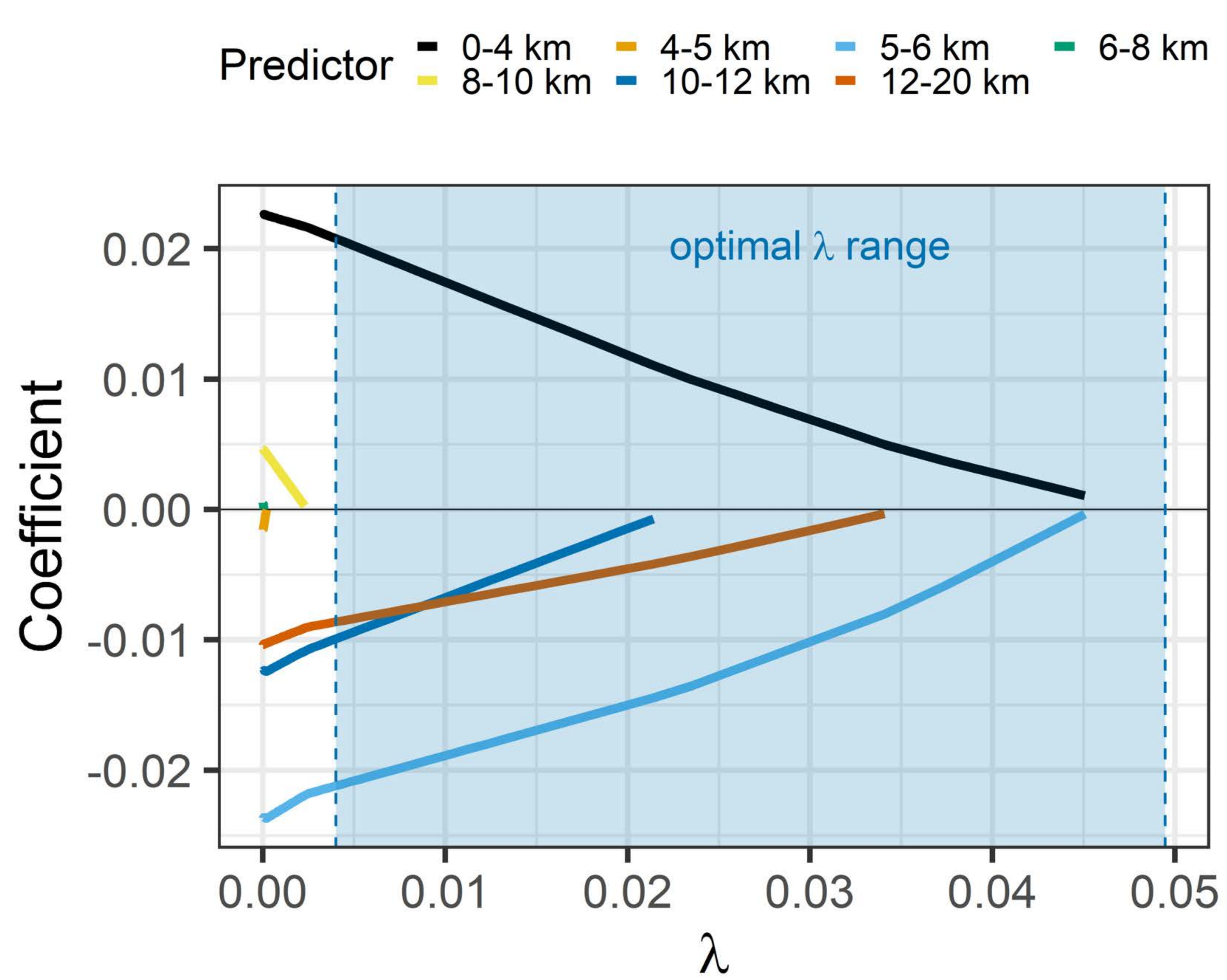
Ichino and Schündeln (2012) conduct an experiment to evaluate the spillover effect of election observers on voter registration irregularities. They use the following two-level design:

Constituencies (clusters)	Electoral areas (ELAs)	Observations
Control	Control	592
Treatment	Control	199
Treatment	Treatment	77

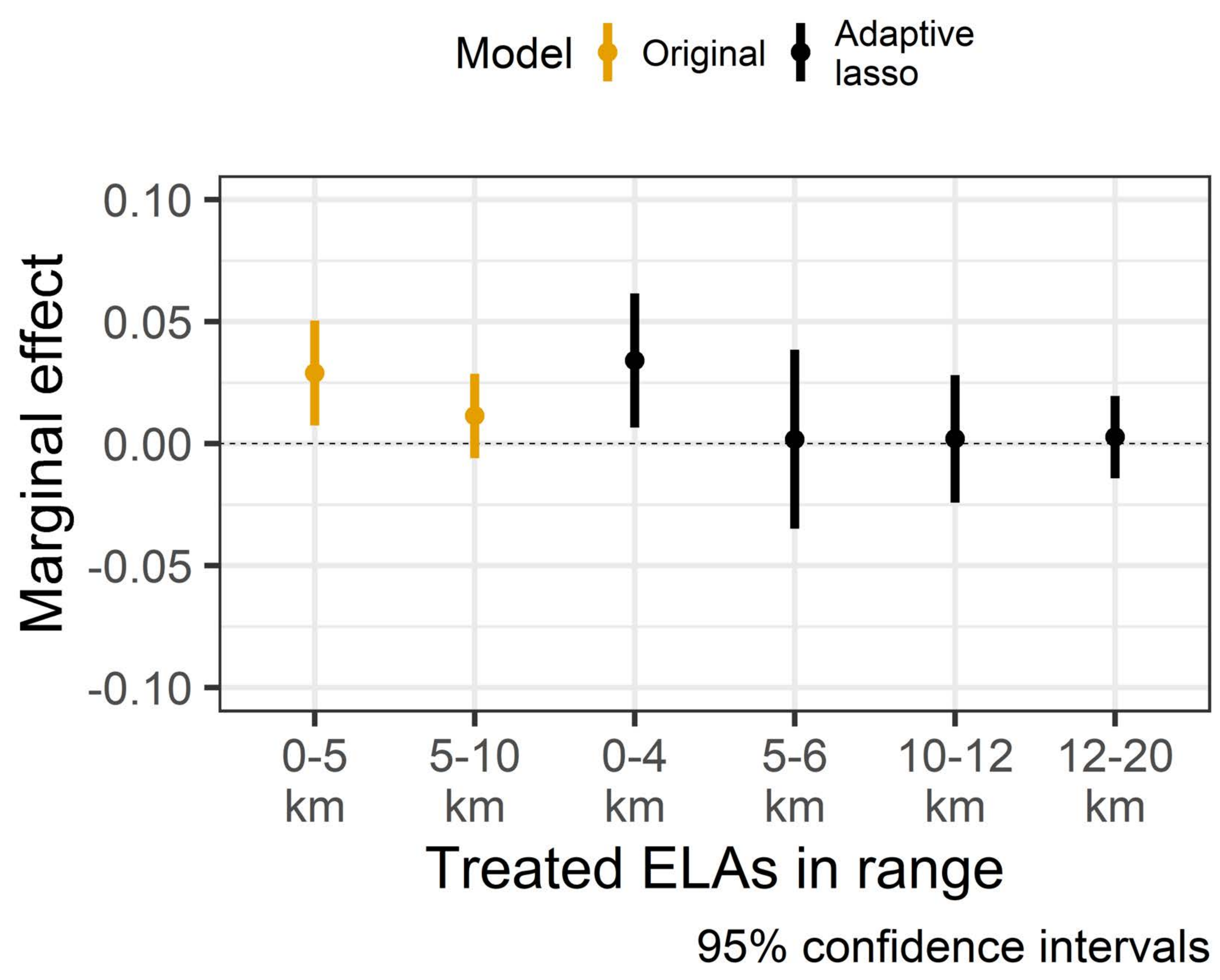
Their main argument is that election observers displace (rather than deter) irregularities, which they measure as the percent change in voter registration from the 2004 to the 2008 election. In their analysis, one unit increase in the number of treated ELAs within five kilometers leads to a 3% increase in irregularities among control units. They find no evidence for spillovers beyond the 5 kilometer range.

Results

The figure below fits an adaptive lasso with voter registration irregularities as the outcome, and the number of treated units within the distance ranges available in the replication materials as predictors. The adaptive lasso shrinks OLS coefficients towards zero based on a shrinkage penalty governed by the tuning parameter $\lambda \geq 0$. I select the appropriate value of λ via 10-fold cross-validation.



As lines approach zero, the corresponding predictor should be omitted from the model. The shaded area denotes the optimal values of λ based on root mean squared error. I estimate an OLS model including all non-zero predictors within the optimal λ range, along with block fixed effects and clustered standard errors for comparability. The next figure compares my results with the original findings.



Conclusion

Results based on the adaptive lasso correspond to the findings of an experiment designed to capture spillovers. This highlights the protocol's usefulness in future applications.

Future Work

Visit my website for updates on this project. I am currently working on:

- Including research design features in the variable selection step
- Including shrinkage penalty in the estimation-testing step
- Developing standards to choose predictor ranges
- Assessing the contribution of multiple networks

Acknowledgements

A previous version of this project received the best poster award in the 2019 Latin American Political Methodology Meeting in Montevideo, Uruguay.