

Interpretation: The Final *Spatial* Frontier

Supplementary Files

Guy D. Whitten

Laron K. Williams

Cameron Wimpy

October 11, 2018

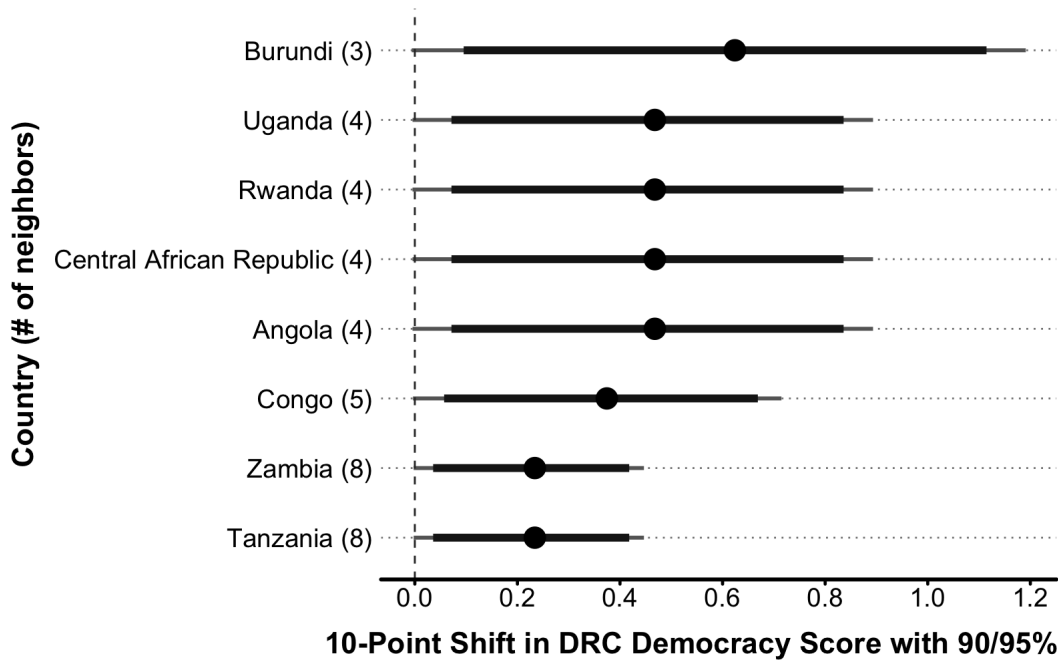
Overview

In this document, we provide further documentation and elaborate on examples discussed in—though not included in—the manuscript.

SAR quantity of interest: effects of an unmodeled shock in the outcome

Another quantity of interest from the SAR model explores how an unmodeled shock to the outcome in observation i influences its neighbors through Equation 15 in the manuscript. In [Figure A.1](#) we demonstrate how a 10-point increase in the Freedom House score for the Democratic Republic of the Congo (DRC) in 1994 influences its first-order neighbors' democracy levels. All eight of its contiguous neighbors experience a positive spillover (statistically significant at the 90% confidence level), and the effects range from +0.23 for Tanzania and Zambia to +0.62 for Burundi. Though each country in [Figure A.1](#) is contiguous, the size of the effects vary as a result of the number of neighbors in the row-standardized weights matrix. For example, Burundi experiences over twice the spillover that Zambia experiences, because the DRC is one of three neighbors of Burundi, but one of eight neighbors for Zambia. Since spatial spillovers influence each of the eight neighbors equally (due to the row-standardization), the positive spillover as a result of the DRC's democratic improvement will be smaller for those observations with more neighbors.

Figure A.1: Spatial Diffusion Patterns Following a 10-point increase in the Freedom House score for the Democratic Republic of the Congo



Model comparison and the coefficient interpretation approach

We argue that there are many perils associated with the *coefficient interpretation approach*. For example, consider the analysis of welfare state size during the Cold War by [Obinger and Schmitt \(2011\)](#). The authors theorize that the reason for the drastic expansion of the welfare state in the Cold War was the regime competition between those in the capitalist West and those in the communist Soviet bloc (247). The authors theorize that three sets of connections between capitalist and communist states determine the degree of competition: all connections weighted equally, geographic proximity (inverse distance between capitals), and economic size. They then estimate three SAR models and three SLX models (with the spatial lags temporally lagged one year). Their interpretation is limited to pointing out that the spatial coefficients are positive, as expected, and that the coefficients for the control variables—such as economic growth and economic wealth—are also signed in the expected direction (263). Unfortunately, this approach provides no sense of the average effects due to positive spatial dependence, the average total effects, or *which* countries’ connections are most meaningful. In short, the *coefficient interpretation approach* neglects a litany of theoretically-interesting inferences.

Moreover, the authors compare the coefficients across the SAR and SLX models to conclude that the SLX models have a “more substantial effect for the regime competition variable” (262). The authors then caution in a footnote immediately following that comparison that “due to different estimation techniques the results cannot be strictly compared. They can only indicate tendencies about the relative importance of particular effects” (262). The authors are correct

that in the *coefficient interpretation approach* comparing the coefficients is a fool's errand. On the other hand, the *general approach* allows scholars to directly compare the effect sizes across different spatial model specifications with the use of the partial derivative matrix.

Row-standardization illustration

A common strategy is to row-standardize the weights matrix. However, the row-standardization process forces the average total effects to be the same; each observation is equally influenced by other influences. Recall the example in the manuscript of four observations that are located at points 0.5, 4.5, 5.5 and 7 on a single dimension. In the un-row-standardized version, the elements of \mathbf{W} capture the inverse absolute distance (\mathbf{W}), or $\frac{1}{\text{abs}(p_a - p_b)}$. Row-standardizing the matrix involves dividing each element by the row total. The resulting partial derivatives matrix and total effects are shown in Equation 1.

$$\mathbf{w} = \begin{pmatrix} 0 & .41 & .33 & .25 \\ .15 & 0 & .61 & .24 \\ .11 & .54 & 0 & .36 \\ .13 & .33 & .55 & 0 \end{pmatrix}, \quad (\mathbf{I}_N - \rho \mathbf{W})^{-1} \beta = \begin{pmatrix} .501 & .022 & .019 & .014 \\ .008 & .502 & .031 & .014 \\ .006 & .028 & .503 & .019 \\ .007 & .018 & .029 & .502 \end{pmatrix}, \quad TE = \begin{pmatrix} .556 \\ .556 \\ .556 \\ .556 \end{pmatrix} \quad (1)$$

As demonstrated by the last column, each observation has the same total effects.

Quantities of interest and estimates of uncertainty

As we discuss in the manuscript, there are often challenges and shortcomings when gleaning proper point estimates and associated estimates of uncertainty from spatial models. For the examples in the manuscript, we employ simulation methods to derive quantities of interest from the explanatory variables and spatial parameters. We then derive confidence intervals based on the percentile method described in (Carsey and Harden 2014, 232-234). Examples of this are included in the associated reproduction files. It is worth noting that several statistical software packages now offer straightforward routines by which different types of spatial impacts (direct, indirect, and total) can be derived along with both simulated and/or analytical solutions. In our case, we employed parametric simulation in order to avoid working out analytical solutions using different formula across the specifications. Examples of the analytical approach are presented in Franzese and Hays (2007); Franzese Jr and Hays (2008). In either case, we recommend taking the additional step to derive estimates uncertainty associated with any spatial parameters and post-spatial explanatory effects.

Extensions

There are important extensions of the approach we outline here. The partial derivatives approach sheds light on how the effects of covariates change as a function of spatial location, or the degree of spatial heterogeneity. Other approaches can deal with spatial heterogeneity, either in a discrete (geographic regression discontinuity design, see [Keele and Titiunik \(2014\)](#) and [Keele, Titiunik, and Zubizarreta \(2015\)](#) for examples) or continuous manner (geographically weighted regression, see [Brunsdon, Fotheringham, and Charlton \(1996\)](#); [Darmofal \(2008, 2015\)](#) for examples). All of these approaches reveal any underlying spatial heterogeneity in effects, whether it is through observation-specific parameters (GWR) or a global parameter that is weighted by spatial location (partial derivatives). The latter approach is especially useful in the presence of continuous spatial heterogeneity and has applications primarily in geographic-based connectivity, although it need not be limited in that way.

References

- Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical analysis* 28 (4): 281–298.
- Carsey, Thomas M., and Jeffrey J. Harden. 2014. *Monte Carlo Simulations and Resampling Methods for Social Science*. Sage Publications.
- Darmofal, David. 2008. "The Political Geography of the New Deal Realignment." *American Politics Research* 36 (6): 934–961.
- Darmofal, David. 2015. *Spatial Analysis for the Social Sciences*. New York: Cambridge University Press.
- Franzese Jr, Robert J, and Jude C Hays. 2008. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance." *Comparative Political Studies* 41 (4-5): 742–780.
- Franzese, Robert J., and Jude C. Hays. 2007. "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15 (2): 140–164.
- Keele, Luke J., and Rocio Titiunik. 2014. "Geographic Boundaries as Regression Discontinuities." *Political Analysis* 23 (1): 127–155.
- Keele, Luke, Rocio Titiunik, and José R. Zubizarreta. 2015. "Enhancing a Geographic Regression Discontinuity Design through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178 (1): 223–239.
- Obinger, Herbert, and Carina Schmitt. 2011. "Guns and Butter?: Regime Competition and the Welfare State during the Cold War." *World Politics* 63 (2): 246–270.