

Web Supplement

July 29, 2016

These materials constitute an online supplement to the research reported in:

Ingram, Matthew C., and Marcelo Marchesini da Costa, “Spatial Analysis of Homicide Across Brazil’s Municipalities” *Homicide Studies* (forthcoming; accepted Aug 2016)

The journal does not support web appendices or supplements, so this supplement is offered by the corresponding author (Ingram) in an effort to provide readers with additional materials to better understand and replicate the analysis reported in the journal.

The supplement covers two main areas:

- (1) Estimation strategies, and
- (2) Alternative visualizations of results.

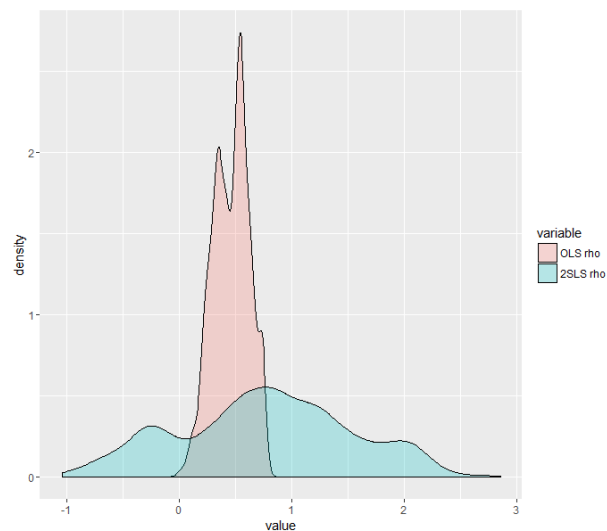
We occasionally update this supplement if we find additional relevant or useful information to post, so please visit this document again.

Estimation Strategies

An early, original analysis used only OLS with a spatially lagged dependent variable. We re-analyzed the data using a two-stage least squares model (2SLS) and all results of the GWR-SL analysis now report only these results because they are technically more appropriate, the analytic strategy matches that of a previously published study using a GWR-SL model (Shoff et al. 2014), and for economy of presentation. See revised discussion on page 14 (in pre-publication version), including footnote 12.

The 2SLS model first estimates y using spatially lagged versions of all explanatory variables as predictors (following work by Anselin and Rey 2014, as well as Dow 2007; 2008; Ingram 2016). We then generated a spatially lagged version of predicted values (\hat{y} -hats), which entered right-side of second stage equation.

In footnote 12 of the article, we offered to provide a comparison of the OLS and 2SLS estimates in a web appendix. The graph below reports this comparison, showing density plots of our estimates of ρ using both OLS and 2SLS (actual data).



Some readers may be interested in this comparison since – though the 2SLS approach is methodologically more appropriate in order to capture the endogeneity underlying the spatial lag process – two points are worth noting: (1) the OLS and 2SLS estimates are highly correlated ($r = 0.67$), and (2) none of the OLS estimates fall outside the range of 2SLS estimates. Regarding the second point, while the 2SLS estimates are technically more appropriate, the OLS estimates can be interpreted as more conservative. That is, the OLS estimates do not identify any magnitude in the relationship that is not identified by the 2SLS estimates. Also, while the OLS results reveal a small number of negative rho values, only the 2SLS estimates reveal a substantial portion of observations with these negative rho values, which could be important in order to identify the full nature and range of rho's effects across a geographic space.

This discussion suggests the possibility of Type “M” and Type “S” errors. Type M error occurs when the magnitude of a relationship is misidentified. Type S error occurs when the direction (sign) of the relationship is mis-identified.¹ Notably, any non-local or non-GWR technique is vulnerable to both Type M and Type S error, since the coefficient estimated by non-local techniques would necessarily be global, i.e., any non-local approach assumes a single coefficient appropriately captures the nature of the relationship between explanatory and outcome variables.

In the future, we may add graphs showing distribution of significance values (e.g., p-values) to illustrate whether significance distribution varies meaningfully across estimation strategies.

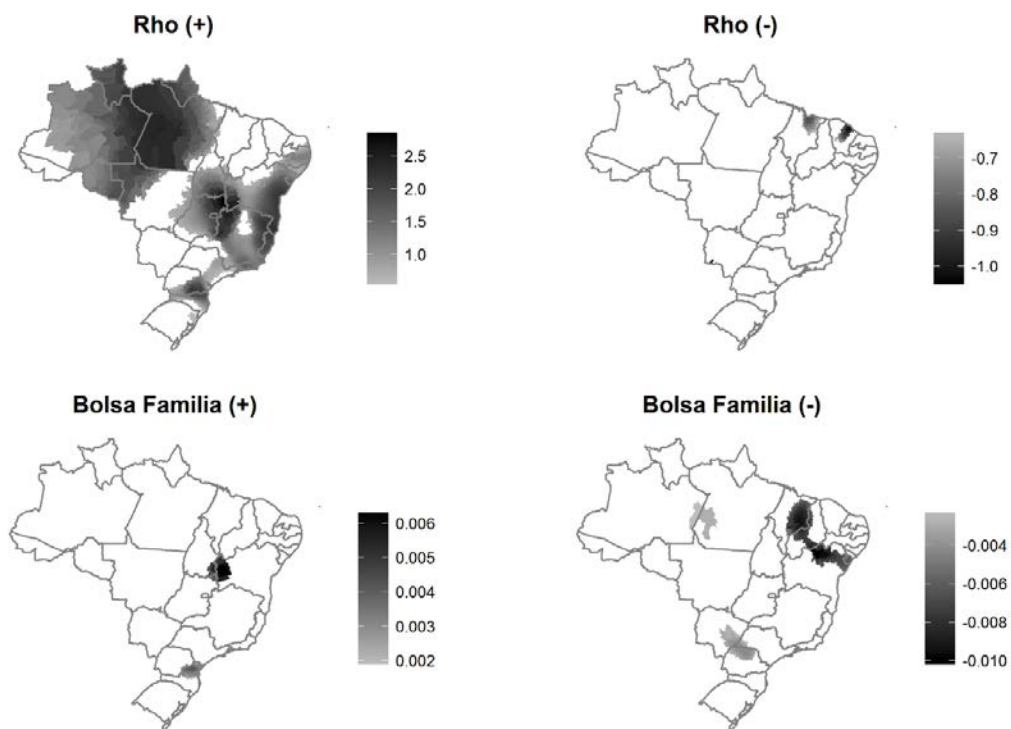
Alternative Visualizations of Results

The results reported in the Homicide Studies article were presented in black and white at the request of the editor (journal does not support color images). We had originally generated color maps, but these were removed and replaced with greyscale images to facilitate interpretation and readability in the print version of the journal.

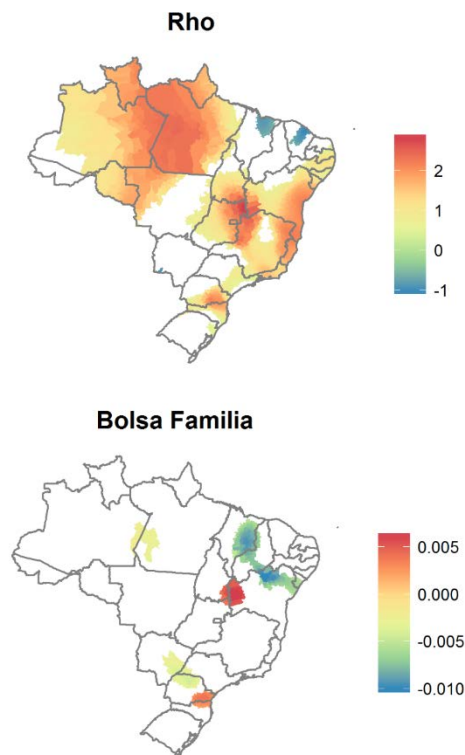
For interested readers, we report the original, published figures (greyscale) below followed by alternative visualizations in color. We occasionally update this document if we come across an interesting and informative way of visualizing the data, so please visit this site again. Also, please let us know if you have suggestions for visualizing data in different way.

¹ See Andrew Gelman on Type M and S errors). URL: http://andrewgelman.com/2004/12/29/type_1_type_2_t/

Figure 1. GWR-SL Results for Lag Effect (ρ) and BF Coverage



Alternatively, the results could be reported in color, simplifying the figures to just two maps by including both positive and negative coefficients for each variable in same map. We prefer these color images, and use them in our own presentations of the research.



We did the same for the control variables that exhibited non-stationary behavior.

Figure 2. GWR-SL Results for Non-Stationary Controls

