Final Replication

Katie Cao

4/29/2020

Contents

0.1 Introduction

Matthew C. Ingram and Marcelo Marchesini da Costa's (2019) paper offers a statistical analysis of the uneven, geographically-varying effect of municipal political predictors of violence across Brazil's 5562 municipalities while controlling for dominant socio-structural accounts. More specifically, the authors investigate the local varying effect of four political variables –party identification of mayors, partisan alignment of mayors and governors, electoral competition, and voter participation—on homicide rates at the municipal level.

The outcome of interest—change in homicide rate—is measured as the difference in the two-year average of the homicide rate between 2011-2012 and 2007-2008. The change in homicide captures the change in violence over the four-year term of mayors who entered office in January 2009, and left office in December 2012, indicating if the mayor may have implemented actions that made their municipality more or less peaceful. Looking at the change in homicide variable across municipalities in Brazil, there is noticeable variation spanning positive and negative values, indicating that the location of the data observed may be important in predicting violence outcomes—that is, the political environment may have a differing effect on homicide rates in different municipalities. To understand this variation, determine if there may be an uneven effect of predictors of violence, and identify potential sources of unevenness, the authors use a geographically weighted regression, or GWR.

The GWR method produces local coefficients for predictors, allowing different relationships to exist at different points in space and facilitating the analysis of spatial heterogeneity. Spatial autocorrelation can be important in statistics when the locations of data are related to the data itself because many global tests assume all data is randomly distributed, or that the values at one location do not depend on values at other neighboring locations so that the location of the data points wouldn't affect the information content of the data. Randomly distributed data are rare, and in many cases, this unmodeled heterogeneity is a serious form of model misspecification.

The authors conduct two main tests using data on homicide rates from Brazilian ministry of health, data on political predictors comes from high electoral tribunal, and data for socio-structural controls from a variety of sources: (1) estimate four global ols models and test the stationarity of coefficients across space: and (2) visualize GWR models in a series of maps. Tabular reports of GWR results do not lend themselves to intuitive interpretation, requiring a comparison between the central tendency of the local coefficients and the OLS estimate, as well as a closer examination of which local coefficients are statistically significant. Thus, we do not present any tables of GWR results. Fortunately, maps are an effective and efficient vehicle for communicating the variation across local coefficients estimated by GWR analyses (Matthews & Yang, 2012; Shoff et al., 2014), especially across multiple models and robustness checks (Ingram & Marchesini da Costa, 2017), so the discussion of GWR findings focuses on the results visualized in the following maps.

The authors find that the effect of the party identification of mayors and voter participation are statistically significant and display non-stationarity, meaning the estimators are geographically-varying. Neither partisan alignment of mayors and governors nor electoral competition are statistically significant; the authors focus on the party identification of mayors, voter participation, and socio-structural controls for the GWR analysis. GWR map analysis shows that the largest left party in Brazil, Workers' Party (PT), had a beneficial effect, reducing violence in large parts of Brazil, the center party that held most local governments (PMDB) had a harmful effect in certain areas of Brazil, and the largest center-right party (PSDB) had mixed effects – helpful in some parts of Brazil and harmful in others. Lower voter participation was associated with higher homicide in large parts of the northeast and south. The mixed results of PSDB indicate PSDB governments are either behaviorally different in certain parts of Brazil, or (b) there is a local, contextual factor present. The authors synthesize these findings with a small case study to suggest that the municipal's government's partisan alignment with the federal government may explain this heterogeneity, because in the period studies, the federal government was affiliated with the Workers' Party (PT).

To replicate this analysis and conduct my own analysis, I use R, an open source programming language. Ingram and Costa's original code and data are made publically available on the Harvard Dataverse. Running the original code, I produced and exported the tables and figures detailing the results of the paper, namely the global ols results, and GWR map visualizations for explanatory variables of interest. Please see the Appendix for the replication. I failed to reproduce the Monte Carlo tests because of computing restraints. Diving deeper, I examine the authors' claims that the non-stationarity of the relationship between mayor party identification and homicide rates is a result of municipality-federal party alignment. I run an OLS regression with interaction effects to predict the effect of a federal social program on homicide rates, given municipality-federal party alignment. I do not find support for the author's proposed explanation. For a more detailed explanation and the results of this regression, see Extension (Section 6). All analysis for this paper is available in my github repo.[1]

0.2 Literature Review

Past research examining predictors of lethal violence has focused on socio-structural accounts emphasizing demographic, economic, and environmental factors. The authors contribute to the literature by examining the role of electoral and partisan politics, which has received little attention previously. In addition, the authors approach this analysis with special attention to spatial interdependence, which is a rising trend in social sciences.

To review, the authors focus on four political variables –party identification of mayors, partisan alignment of mayors and governors, electoral competition, and voter participation. The authors expect that more programmatic parties on the left or right, will reduce violence relative to centrist parties because existing research shows programmatic parties emphasize crime and public safety in their platforms. Prior research regarding partisan alignment shows that policy coordination and funding are better with municipality-state party alignment, so alignment is expected to reduce violence. Electoral competition has been found to improve government performance across many areas and thus is expected to reduce violence. Voter participation is expected to exert downward pressure on violence because of improvements to social organization.

[1]Katie Cao Github Repo

0.3 Replication

The authors generously made their data and code available on Dataverse and is written in R. I was able to replicate almost all the main findings in the Results and Discssion section of the paper. These results, viewable in the Appendix (Section 6), include a global OLS regression with four models of varying specifications to test robustness, and GWR map visualizations for the explanatory variables and control variables that demonstrated non-stationarity. I exported the results as images to include in this paper.

I failed to replicated the Monte-Carlo tests for stationarity preformed on the global OLS regression that was used to identify predictors that vary by magnitude, statistical significance, and even direction across municipalities. As detailed by the authors in their ReadMe file, the code includes 5 Monte Carlo tests that take approximately 25 hours to run. My several attempts to run these tests all resulted in either R crashing or the tests running for over 25 hours without completing. I could not track down the cause of this issue and was ultimately not able to include the results.

0.4 Extension

```
ols_ext <- formula(
  DifHRElec ~ CoverBF + PT + CoverBF*PT +
    lpopdensity + lpctpopym +
    GINI + IDHM +
    HHsinpar + Ocup18male
)
ols_ext <- lm(ols_ext, data=shpbra@data)
stargazer(ols_ext)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, May 14, 2020 - 18:14:59

0.5 Conclusion

0.6 Appendix

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, May 14, 2020 - 18:14:59

Table 1:

	Dependent variable:
	DifHRElec
CoverBF	0.003
	(0.010)
PT	−2.401
	(2.899)
lpopdensity	0.266^{*}
	(0.158)
lpctpopym	8.965^{***}
	(1.936)
GINI	−3.918
	(3.339)
IDHM	-17.169^{***}
	(4.004)
HHsinpar	0.303^{***}
	(0.068)
Ocup18male	0.057^{*}
	(0.030)
CoverBF:PT	0.014
	(0.029)
Constant	24.201^{***}
	(4.998)
Observations	5,562
R^2	0.017
Adjusted R^2	0.016
Residual Std. Error	14.331 (df = 5552)
F Statistic	10.900^{***} (df = 9; 5552)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2:

	OLS Models DV: change in homicide rate (HR Change)			
	DifHRElec			
	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
Margin of Victory	-0.730	-0.702	-0.707	-0.744
	(0.807)	(0.807)	(0.807)	(0.808)
Alignment	-0.765	-0.812^{*}	-0.509	-0.613
	(0.511)	(0.470)	(0.495)	(0.466)
Abstension	1.078^{***}	1.077^{***}	1.051^{***}	1.052^{***}
	(0.380)	(0.380)	(0.380)	(0.380)
PMDB	1.241^{**}	1.360^{***}		
	(0.507)	(0.473)		
PSDB	-0.082		-0.406	
	(0.625)		(0.588)	
PT	-0.751			-1.042
	(0.668)			(0.643)
PopDensity	0.317^{**}	0.316^{**}	0.309^{*}	0.314^{**}
	(0.159)	(0.159)	(0.159)	(0.159)
YoungMalePct	7.825^{***}	7.801^{***}	7.689^{***}	7.723^{***}
	(1.996)	(1.996)	(1.997)	(1.997)
GINI	-6.234^{*}	-6.416^{*}	-6.497^{*}	-6.137^{*}
	(3.423)	(3.417)	(3.422)	(3.423)
HDI	-17.682^{***}	-17.748^{***}	-17.820^{***}	-17.970^{***}
	(4.045)	(4.033)	(4.047)	(4.034)
SingleMotherHH	0.287^{***}	0.285^{***}	0.280^{***}	0.283^{***}
	(0.068)	(0.068)	(0.068)	(0.068)
Employment	0.047	0.046	0.050^{*}	0.052^{*}
	(0.030)	(0.030)	(0.030)	(0.030)
BolsaFamilia	0.004	0.004	0.004	0.003
	(0.010)	(0.010)	(0.010)	(0.010)
Constant	23.145^{***}	23.191^{***}	23.264^{***}	23.194^{***}
	(5.018)	(5.018)	(5.021)	(5.020)
Observations	5,562	5,562	5,562	5,562
R^2	0.020	0.020	0.019	0.019
Adjusted R^2	0.018	0.018	0.017	0.017
Note:		*p<0.1; **p<0.05; ***p<0.01		

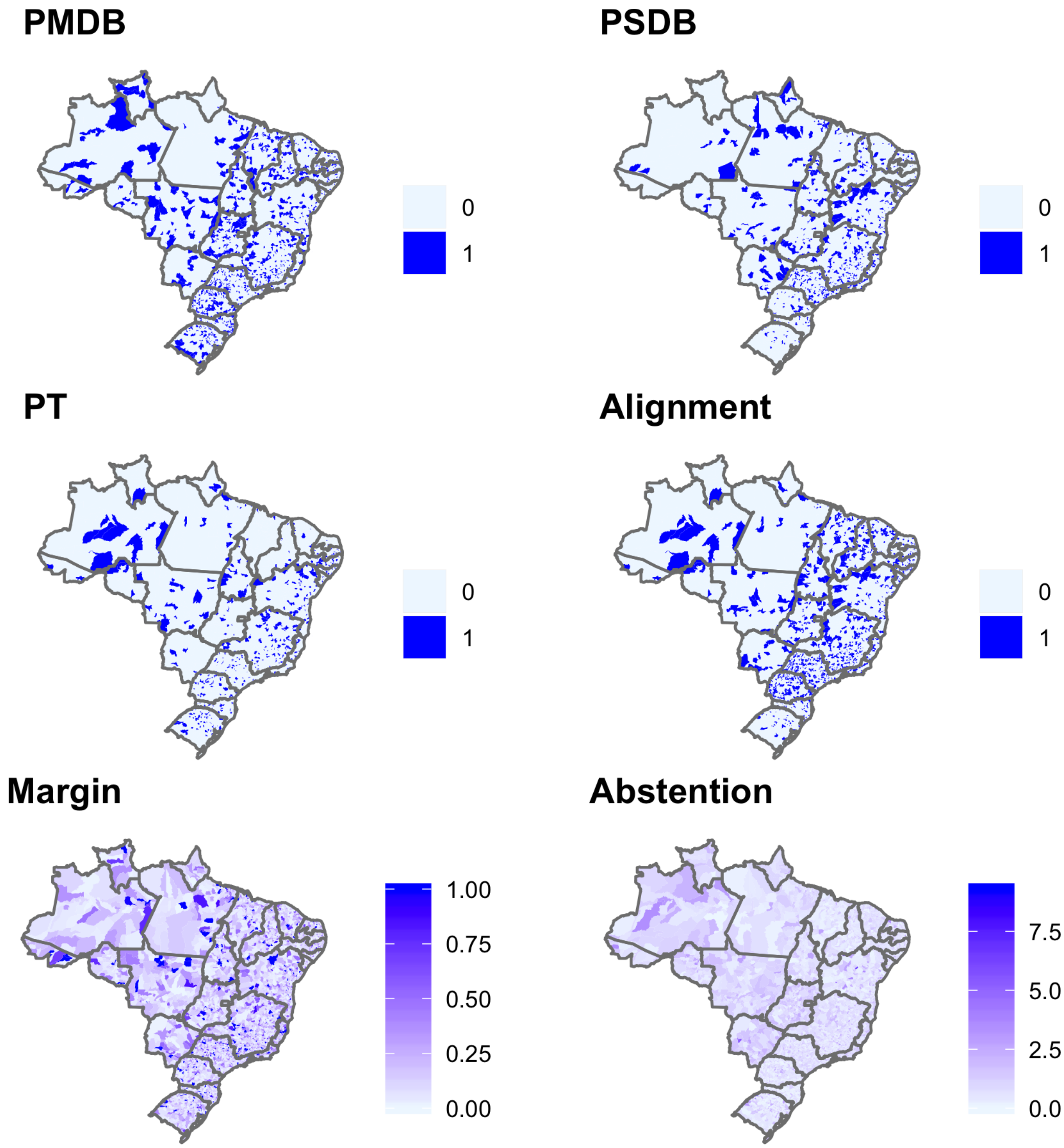

Figure 1: Figure 1: Geographic Distribution of Key Explanatory Variables.

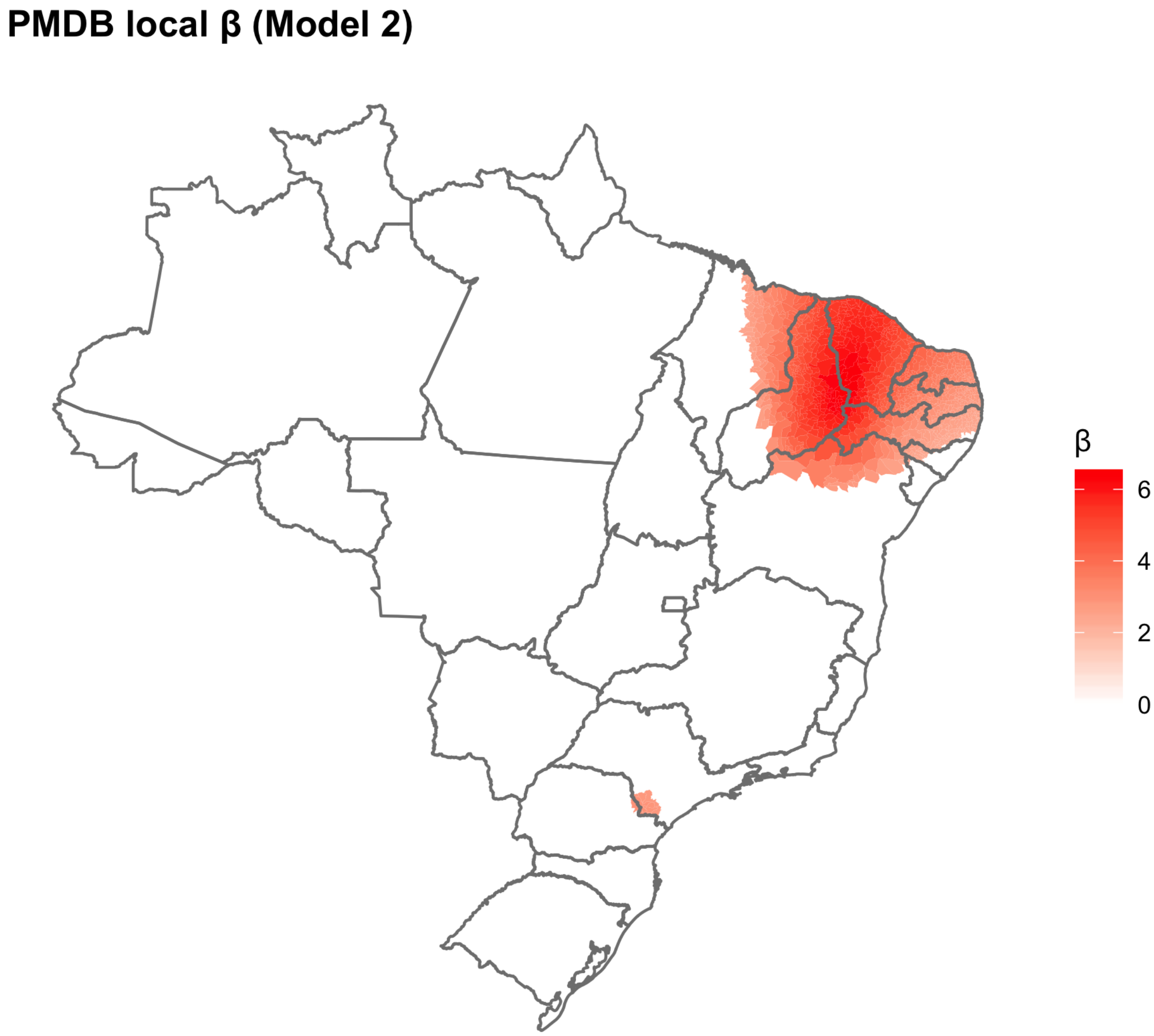

Figure 2: Figure 2: Local coefficients for PMDB (GWR Model 2).

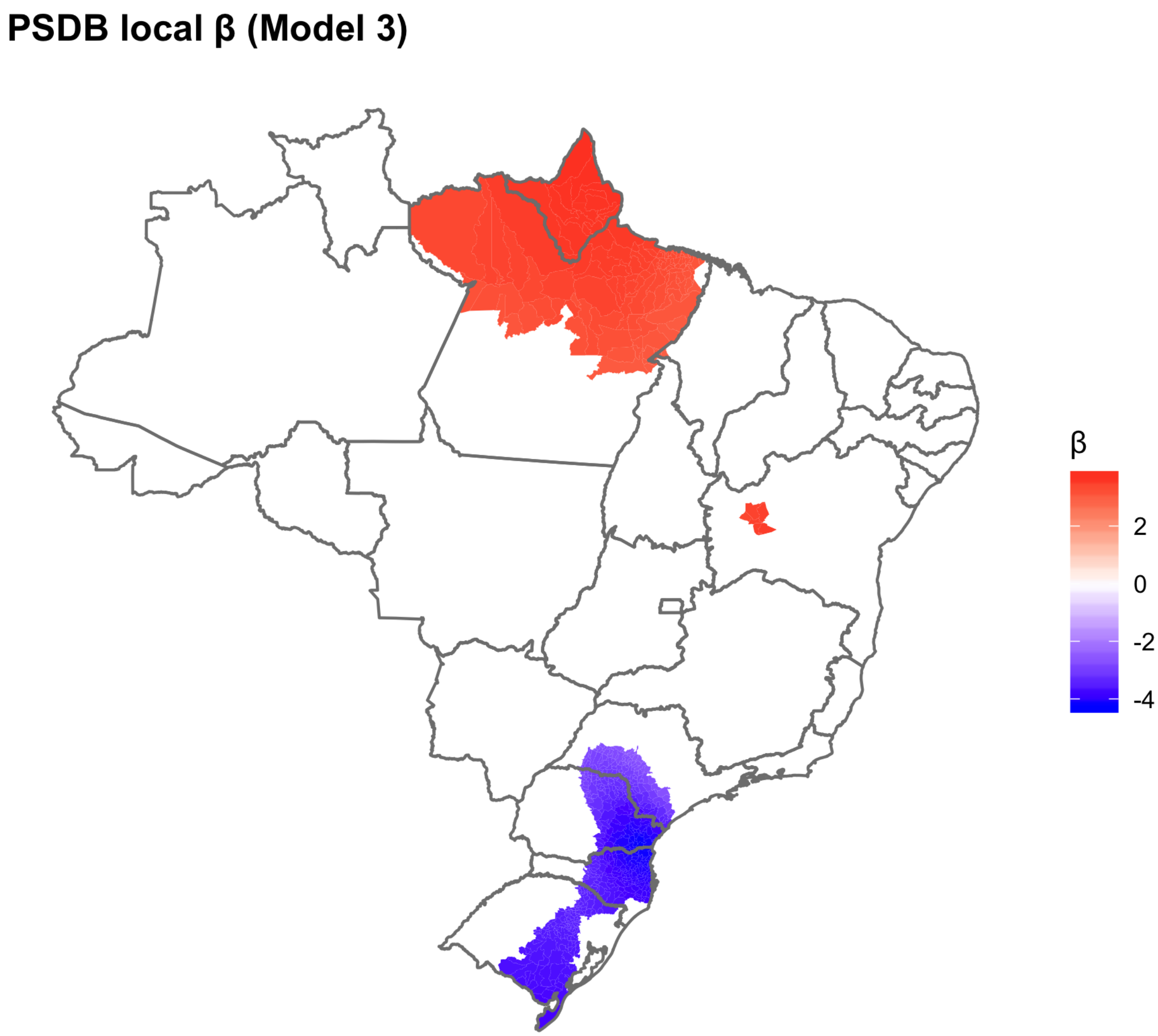

Figure 3: Figure 3: Local coefficients for PSDB (GWR Model 3).

PT local β (Model 4)

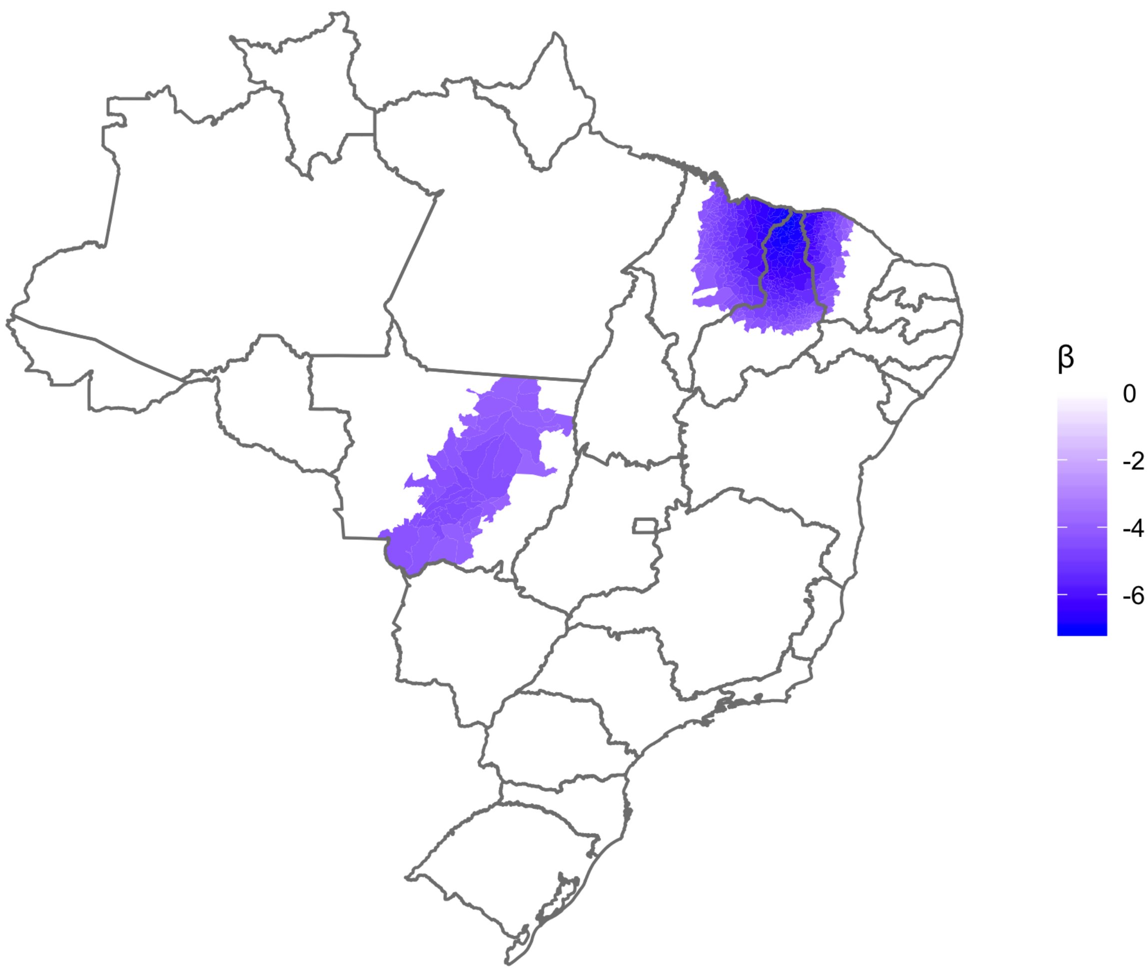

Figure 4: Figure 4: Local coefficients for PT (GWR Model 4).

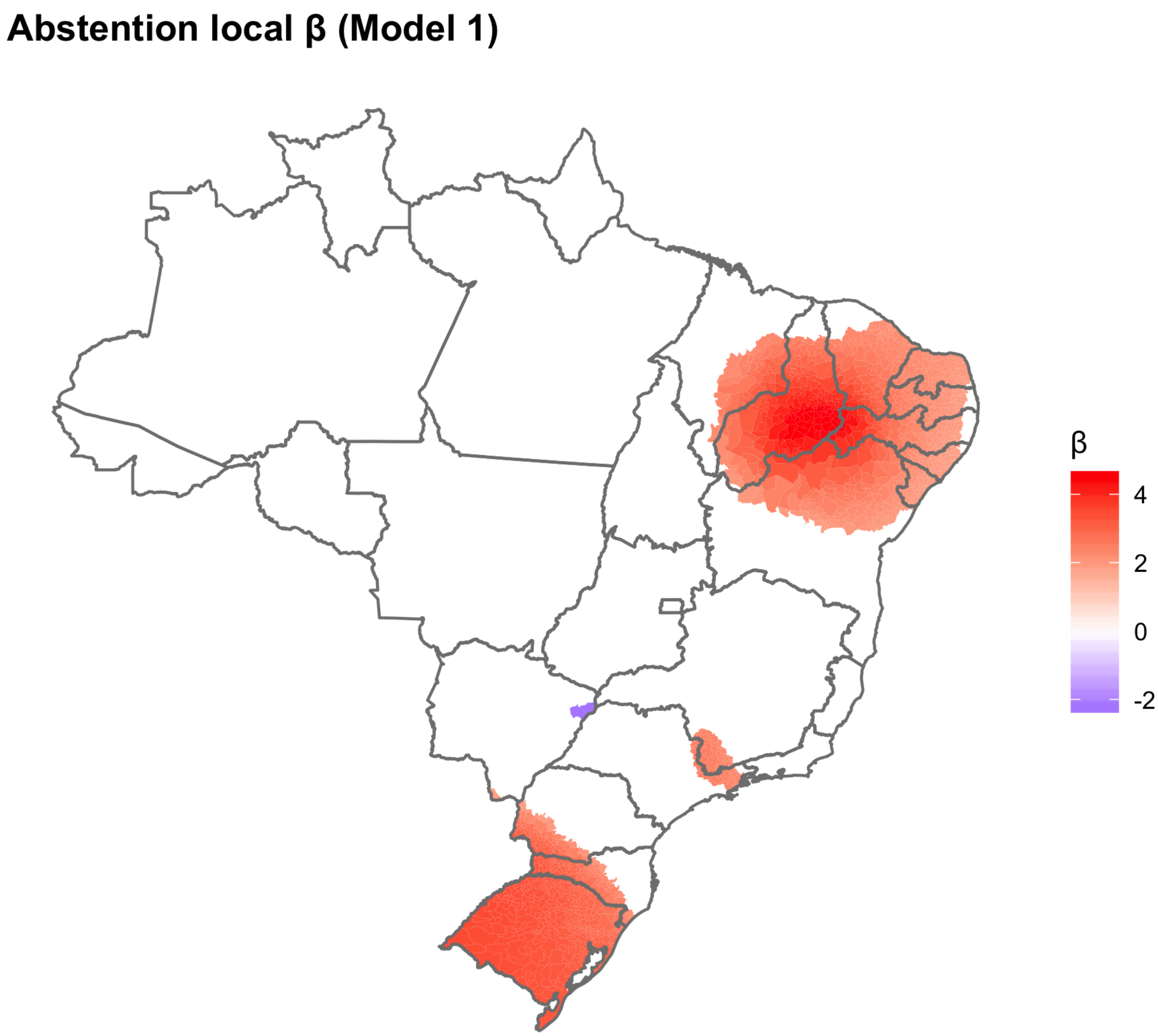

Figure 5: Figure 5: Local coefficients for Abstention (GWR Model 1).

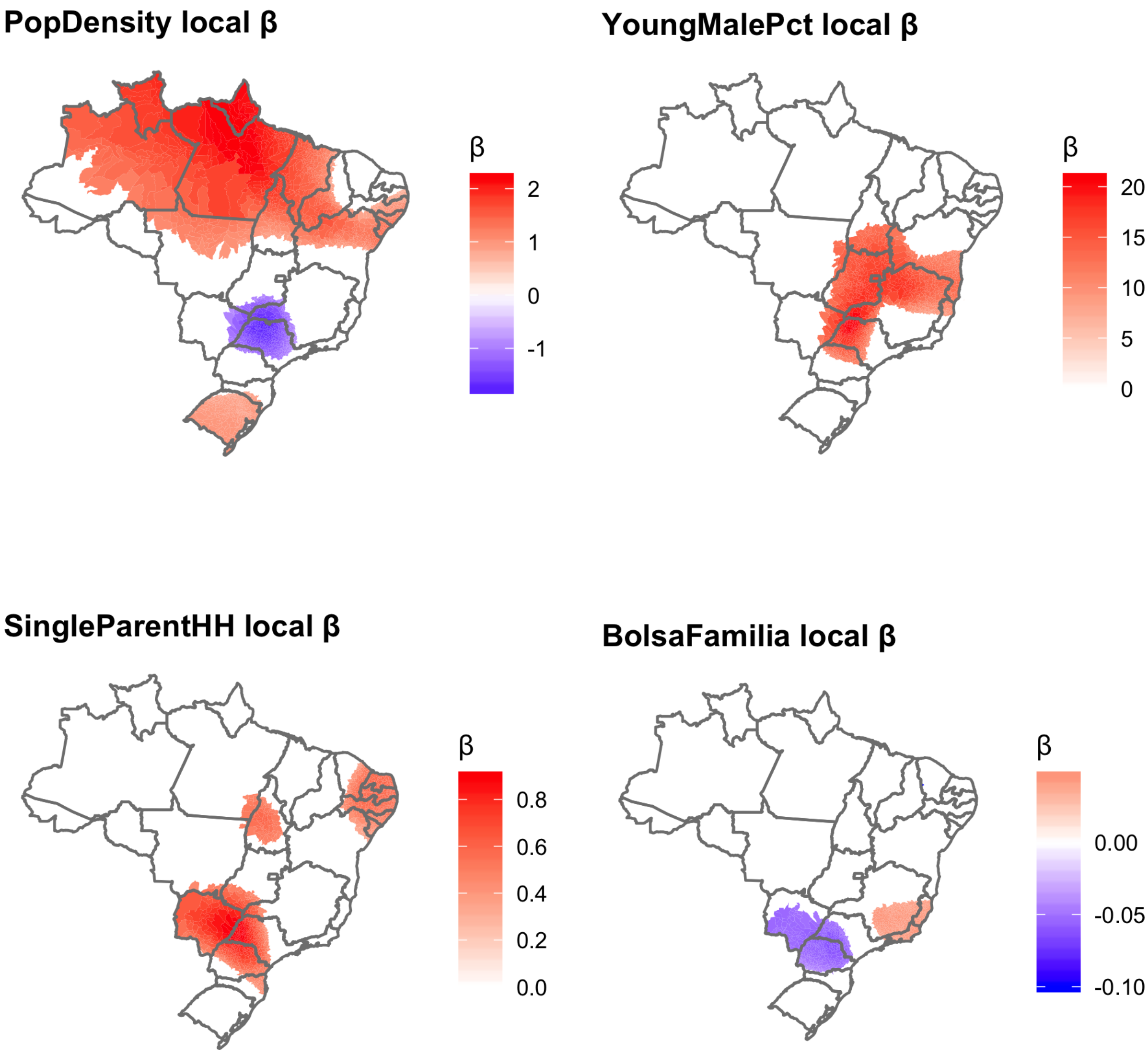

Figure 6: Figure 6: Local coefficients for non-political control variables (GWR Model 1).

0.7 References