Geographical patterns in Brazilian elections

Thiago Marzagão[1]

1. Empirical puzzle, background information, and existing literature

Why do people living in nearby counties tend to vote similarly? The existing literature provides abundant evidence that social interaction among relatives and friends is an important predictor of political behavior (Huckfeldt & Sprague 1991; Nickerson 2008). But relatives and friends usually live *within* the same county, so why are vote patterns spatially autocorrelated *across* counties? This paper tests three alternative hypotheses. The first one is that voting is influenced not only by within-county social interactions but also by cross-country social interactions. As gravitational models show, trade flows are inversely proportional to distance (Tinbergen 1962) so all else equal we should expect nearby counties to trade more than distant counties. Such economic interactions, in turn, produce social interactions: firms develop supply chains wherein people located in different counties interact on a regular basis. In these interactions people may come to discuss politics, which may influence their political behavior. The second hypothesis is that political campaigns focus on certain geographical clusters while neglecting others, thus inducing spatial autocorrelation. Candidates have finite resources so they have to allocate their "war chests" strategically. They do so by focusing their resources on those areas with the highest "yield" in terms of donations or votes (Cho & Gimpel 2007). To the extent that such "yield" is spatially autocorrelated (because of historical patterns), the allocation of campaign resources will also be spatially autocorrelated and, unless campaign resources have zero electoral effect, voting will also be spatially autocorrelated. Finally, our third hypothesis is that people in nearby counties vote similarly simply because they are socioeconomically similar and therefore have similar policy preferences. Socioeconomic indicators are not distributed randomly across space: poor counties tend to cluster together, as do prosperous ones. To the extent that objective material conditions influence electoral behavior, that should cause voting to be spatially autocorrelated.

In spatial econometrics terms, the first hypothesis (cross-county social interactions) implies a spatial lag model: the value of the dependent variable in a given county is determined by a weighted average of the values of the dependent variable in nearby counties (plus some set of exogenous

[1] PhD student in Political Science at the Ohio State University. This research was possible due to the funding provided by the Brazilian Ministry of Education (via The Capes Foundation), the Fulbright Commission, and the Brazilian Ministry of Planning.

covariates). Formally, this relationship may be expressed as $y_i = \lambda W_s y_s + \beta X + \varepsilon$, where $\boldsymbol{y_i}$ is the value of the dependent variable in observation $\boldsymbol{i}$, $\boldsymbol{W_s}$ is a matrix specifying the weight to be assigned to each surrounding observation, $\boldsymbol{y_s}$ is a matrix of containing the value of the dependent variable in each surrounding observation, $\boldsymbol{\lambda}$ is the effect of $\boldsymbol{W_s y_s}$ on $\boldsymbol{y_i}$, $\boldsymbol{X}$ is a matrix of exogenous covariates, $\boldsymbol{\beta}$ is a vector of coefficients, and $\boldsymbol{\varepsilon}$ is the random error. The $\boldsymbol{W_s y_s}$ term is called the spatial lag. Since $\boldsymbol{y}$ is being regressed on spatially lagged values of $\boldsymbol{y}$, this is called a spatial lag model or spatially autoregressive model (SAR). The second and third hypotheses, by contrast, imply a spatial error model: observations do not influence each other but are influenced together by common factors (campaign strategy or socioeconomic structure). When these common factors are explicitly included in the model the spatial autocorrelation that they cause disappears. When they cannot be explicitly included in the model they will be captured by the error term. Thus the error terms of nearby observations will be correlated: formally, $E[\varepsilon_i \varepsilon_j] \neq 0$. Such models are called spatial autoregressive error models (SARE). When a model has both a spatial lag (SAR) and spatial error (SARE) it is called a spatial autoregressive model with autoregressive disturbances (SARAR). (For an introduction to these models, see Ward & Gleditsch 2008.)

The three hypotheses outlined above will be tested with data from the Brazilian presidential election of 2010. (To the best of my knowledge this is the first time those data are used in spatial analysis.) The remainder of this section provides some background information on Brazilian politics and on the 2010 election. It also discusses the literature on geographical patterns in Brazilian elections. Section 2 shows that voting was spatially autocorrelated in the 2010 election and identifies in which areas of the country this autocorrelation was stronger. In addition, Section 2 assesses to what extent the geographical patterns observed in previous Brazilian elections have been reproduced or changed in the 2010 election. Section 3 tests the three hypotheses outlined above. Testing the first hypothesis (cross-county social interactions) requires testing whether the model has a spatial lag, which is done in two different ways: by applying Lagrange-multiplier (LM) tests to the residuals of (non-spatial) least squares estimations and by estimating a SARAR model and checking whether the spatial lag is significant. Testing the second (campaign clustering) and third (socioeconomic clustering) hypotheses requires testing whether the model has spatial error (i.e., whether the residuals are geographically correlated). As with the spatial lag hypothesis, this is done both by applying LM tests to the residuals of non-spatial least squares estimations and by estimating a SARAR model. But finding spatial error does not allow one to adjudicate between the second and third hypothesis. This is done by the explicit inclusion of socioeconomic indicators in the model. The logic is simple: if socioeconomic indicators turn out statistically significant and no spatial error remains in the model, then the second hypothesis is rejected and the third hypothesis is accepted; if socioeconomic indicators turn out statistically significant but spatial error remains, then both the second and third hypotheses are accepted; and if socioeconomic indicators are not statistically significant but spatial error remains, then the second hypothesis is accepted and the third hypothesis is rejected. (The scenario where socioeconomic indicators are not statistically significant and no spatial error remains is logically inconsistent – variables that have zero effect cannot account for spatial autocorrelation).

Thus the third hypothesis is tested directly while the first and second hypotheses are tested only indirectly. This is due simply to data availability: socioeconomic data are publicly available whereas data on social interactions and campaign strategies are not. (Ideally, if all data were available, the resulting model would show zero spatial autocorrelation.) Finally, Section 3 also assesses whether the spatial effects and the effect of socioeconomic structure are the same across different areas of Brazil. This is done by means of a geographically weighted regression and of subsampling.

Before proceeding to Sections 2 and 3 some background information is in order. In Brazil federal and state elections are held simultaneously every four years. On the same day (usually the first Sunday of October), all adults under 70 are required to vote for president, representative, senator, governor, and state representative (there are no state senators – state legislatives are unicameral). Voting is mandatory (which produces high turnout rates – usually above 80%) and direct (there are no electors or electoral colleges). The election of president and governor is based on majority rules: candidates who win an absolute majority are elected in the first round; those who win only a plurality face the second runner-up in a runoff ballot (normally in mid-November). For president and governors only two consecutive terms are allowed (there are no limits for non-consecutive terms). The election of representatives, senators, and state representatives is based on proportional vote. Every state has a fixed number of seats in the lower and upper houses but other than that there is no geographical representation: the constituency of each representative or senator is his or her entire state (i.e., there are no electoral districts). The same is true for state representatives (there are no county-specific seats in state assemblies). Thus there are no *institutional* reasons to expect vote clustering across Brazilian counties: within each state, the list of candidates is the same for all counties. Any geographical patterns must result from the factors discussed above: social interaction, campaign strategies, or socioeconomic similarities.

Brazil has a multiparty system with high fragmentation (there are 29 political parties, 23 of which currently have representatives in the lower house) and low party discipline. There are four major parties. The left-wing Workers' Party (*Partido dos Trabalhadores* – PT) has held the presidency since 2003, when Luiz Inácio "Lula" da Silva was elected (after being the second runner-up in 1989, 1994, and 1998). Lula was re-elected in 2006 and then succeeded by Dilma Rousseff, also from the Workers' Party, in 2010. Today the PT has the largest plurality in the lower house (85 out of 513 representatives) and five of the country's 27 governors. The center-left Social Democratic Party (*Partido da Social-Democracia Brasileira* – PSDB) held the presidency between 1995 and 2002 and currently has eight governors – including Sao Paulo's, which is the country's most industrialized state, and Minas Gerais', which is the most populous state. The Democratic Movement Party (*Partido do Movimento Democrático Brasileiro* – PMDB) is a catch-all party with no clear programmatic content (it has supported the PSDB presidency and it now supports the PT presidency) but it has the second largest share of the lower house (76 of 513 representatives) and five governors. Finally, the Socialist Party (*Partido Socialista Brasileiro* – PSB) has six governors and its coalition (which includes two other parties) has the third largest share of the lower house (63 of 513 representatives).

The geographical patterns of Brazilian electoral politics have been discussed in Carraro et al (2007), Nicolau & Peixoto (2007), Soares & Terron (2008), Hunter & Power (2008), and Terron & Soares (2010). Only Carraro et al (2007), Soares & Terron (2008) and Terron & Soares (2010) actually use spatial econometric methods; Nicolau & Peixoto (2007) and Hunter & Power (2008) merely discuss electoral outcome variation across states. Overall three main findings stand out from the literature. First, while most of the vote for Lula was concentrated in the urban, industrialized centers of the South and Southeast until 2002, once he won the presidency he managed to extend his electoral support to the rural, underdeveloped areas of the North and Northeast (Soares & Terron 2008). The shift is apparent in the election of 2006. Most authors point to *bolsa-família* – a cash transfer program expanded by Lula – as the primary cause of that shift. *Bolsa-família* provides poor families with a monthly stipend of R$ 68 (about US$ 37), plus an additional R$ 22 (about US$ 12) per children in school. The program costs about 0.5% of the country's gross domestic product (GDP) and benefits around 44 million people, mainly in the rural areas of the North and Northeast – precisely the areas where electoral support for Lula grew the most between 2002 and 2006. (Carraro et al 2007 are the only authors who claim that economic growth, not *bolsa-família*, is the main predictor of county-level variation in electoral support for Lula.) The second main finding is that spatial autocorrelation is usually high. The global Moran's I statistic for the county-level share of of votes for president varied between .60 and .81 between 1994 and 2006 (Terron & Soares 2010). Finally, the third main finding is that the electoral bases of Lula and of the PT have followed different paths since 2002. While Lula managed to expand his electoral support to the North and Northeast, the PT itself remained mainly a urban-based party. Soares & Terron (2010) show that the bivariate spatial correlation between votes for Lula and votes for PT representatives declined from 0.41 in 2002 to -0.04 in 2006. This is probably related to Brazil's typically personalistic electoral political environment: voters associate *bolsa-família* (or economic growth, in Carraro et al's 2007 interpretation) to Lula, not to his party.

In sum, the existing literature has uncovered interesting patterns. Its methodological shortcomings limit its usefulness though. None of the authors disclose which estimation method was used in their multivariate analysis: it may have been ordinary least square (OLS), maximum likelihood estimation (MLE), or the generalized method of moments (GMM). This omission is problematic because spatially lagged regressors always introduce endogeneity (since neighboring observations mutually influence each other) and ordinarily also heteroskedasticity (Anselin 2006) and these problems have different consequences for different estimation procedures (Ward & Gleditsch 2008; Kelejian & Prucha 2010). Moreover, most of the time the interpretation of the estimates is simply outright wrong. Consider Soares & Terron (2008), for instance. They use a spatial lag model for the difference in Lula's county-level vote shares between the first and second rounds of the 2006 election. They find a coefficient of .97 for the *bolsa-família*/average income ratio and conclude that, all else constant, "every 1% increase in the *bolsa-família*/average income ratio increases that difference by about 1%" (294). But because neighboring observations mutually influence each other, the coefficients of a regression with spatial terms cannot be interpreted as in a linear model (Ward & Gleditsch 2008). Every change in observation ***i*** produces a change in observation ***j***, which then feedbacks into further changes in observation ***i***, and son. While in the non-spatial case fitted values can be computed as $\hat{y} = X\hat{\beta}$, in the spatial lag case fitted values should be computed as $\hat{y} = (I - \hat{\lambda} W)^{-1} X\hat{\beta}$, where ***I*** is the identity matrix and the other terms are defined as before (the hats indicate estimates as opposed to parameters).[2] As Ward & Gleditsch (2008) put it, the $(I - \hat{\lambda} W)^{-1}$ term is a multiplier that measures what proportion of the change in ***X*** will "spill over" onto other

[2] That formula is a simple derivation of the spatial lag model: $y = \lambda Wy + X\beta + \varepsilon$. By moving all terms related to ***y*** to the left-hand side of the equation the result is $(I - \lambda W)y = X\beta + \varepsilon$, which can be rearranged as $y = (X\beta + \varepsilon)/(I - \lambda W) = (I - \lambda W)^{-1}(X\beta + \varepsilon)$. Since $E[\varepsilon] = 0$, that reduces to $E[y] = (I - \lambda W)^{-1} X\beta$.

nearby observations and then feedback into further changes in observation ***i***. Thus the final equilibrium is a product of the direct and indirect effects of every change in ***X***. Since each observation has different "neighbors", there are no unconditional effects here: every change in ***X*** will produce different changes in ***y*** for different observations. Soares & Terron's (2008) interpretation is thus doubly incorrect: it assumes that the .97 coefficient is the total effect (when in fact it is only the direct effect) and it assumes that the effect is the same across all counties. (Moreover, a coefficient of .97 would mean that a *one-percentage point* increase in the *bolsa-família*/income ratio produces a *one-percentage point* increase in the vote share difference. A log-log transformation would be required to permit the sort of interpretation Soares & Terron make.) Terron & Soares (2010) and Carraro et al (2007) make exactly the same mistakes. Besides, the choice of the weights matrix is sometimes questionable. Carraro et al (2007), for instance, use a distance band of 50 kilometers. As Section 2 will show, that excludes half of Brazil's territory from the estimation, due to the large size of the counties in the North region (we must keep in mind that distance-band weights matrices are based on distances between polygon *centroids*, not distances between polygon *borders*). Since the North is socioeconomically very different from the South and Southeast (it is much poorer and less populated), Carraro et al's (2007) inferences are based on severely biased estimates. Finally, none of the existing works discusses the possibility of heterogeneity. They all assume that the estimates are the same for the entire country, but as Section 3 will show, this is not true.

Thus besides testing the three hypotheses outlined before, this paper will engage the existing literature on Brazilian electoral geography. It will use data from the 2010 election to check the evolution of geographical patterns since 2006. Despite lacking charisma, having no experience with electoral politics, and being completely unknown to the public before the campaign, Dilma Rousseff won the presidency. What was the geographical distribution of her electoral support? Was it similar to Lula's in 2006? How did it compare to the PT's? Has the PT been able to "ride" on Lula's popularity and expand its own basis of support since 2006? This paper will correct some methodological shortcomings of the existing literature, which hopefully will help increase the quality of subsequent research on the subject.

2. Spatial patterns in Brazilian elections

This section assesses whether any spatial patterns exist (otherwise a simple aspatial model estimated by OLS may suffice to explain county-level vote shares in Brazilian elections), performs an exploratory analysis of the data, engages the existing literature, and discusses the choice of weights matrix.

Map 1 shows the vote share of Dilma Rousseff in each county in the second round of the 2010 presidential election.[3]

[Map 1 – click to view]

As Map 1 shows, Rousseff's votes were concentrated in the Northeast region and in the Northern states of Amazonas (almost all of which falls in the 76%-97% range) and Amapá. This pattern is remarkably similar to that of 2006: scaling differences apart, the map above looks very much like the one we find in Soares & Terron (2008), which is reproduced here as Map 2.

[Map 2 – click to view]

Map 1 also suggests spatial autocorrelation: voting patterns appear to be clustered. In order to formally test for spatial autocorrelation it is necessary to define what exactly a "neighbor" is in this case. As discussed above, using a distance band of 50 km, as Carraro et al 2007 do, is inappropriate. Because of early colonization patterns, Brazilian counties are very small in the South, Southeast, and Northeast but very large in the North and in part of the Center-West. (A somewhat similar pattern is observed in the United States, where counties become increasingly large as one moves from the East coast to the West coast). For instance, if you are in the centroid of Altamira (a Northern county of 159,695 km^2, larger than Portugal) you would need to travel 198.6 km to reach another county's centroid (incidentally, the state of Pará, where Altamira is located, is five times larger than the state of São Paulo, in the Southeast, but while Pará has only 143 counties São Paulo has 645). We created a weights matrix based on that 50 km band and it resulted in 334 neighborless counties. That may seem low compared to Brazil's total of 5566 counties, but combined those 334

[3] The source of geocoded data was the *Instituto Brasileiro de Geografia e Estatística* (IBGE). The source of electoral data was the *Tribunal Superior Eleitoral* (TSE). For details and summary statistics, see Appendix B.

counties represent some 50% of Brazil's total area (virtually 100% of the North region and more than 50% of the Center-West region), as shown in Map 3.

[Map 3 – click to view]

Excluding such large part of the country from the estimations would be a serious limitation: the entire analysis would only be applicable to the South and Southeast regions, and to part of the Northeast region. One might reply that, since counties in the North are so big, perhaps their populations are indeed isolated from each other and therefore it is appropriate that they be excluded from the estimation. But at the very least counties must trade (no county is an autarkic unit that produces everything it consumes and consumes everything it produces) and as gravitational models have long shown (Tinbergen 1962), all else equal nearby units trade more than distant units. Altamira and Gurupá (large neighboring counties in the Northern state of Pará) may interact less than, say, Florianópolis and São José (small neighboring counties in the Southern state of Santa Catarina), but they do interact and constraining that interaction to zero is a serious misspecification. Increasing the distance threshold to 80 km eliminates many "islands" in the Northeast but still leaves nearly all the North and most of the Center-West out of the estimation. Increasing the threshold to 110 km eliminates practically all "islands" in both the Northeast and the Center-West but still changes very little in the North. The minimum distance band necessary to produce no neighborless counties is 374 km. But the distance band cannot be increased arbitrarily. A distance band of 374 km does not adequately model the dynamics in the South and Southeast: e.g., that is the approximate distance between Ribeirão Pires and Ribeirão Preto (two counties in the state of São Paulo), but there are 16 other counties in between. Whatever influence Ribeirão Pires might have on the electoral outcomes of Ribeirão Preto (and vice-versa) is certainly dissipated along the way and becomes zero for all practical purposes. Since no single distance band could be adequate for the entire country, in this paper two alternatives are adopted. The first one – used in this section and in part of Section 3 – is a contiguity matrix. Specifically, queen contiguity is chosen over rook contiguity: with extremely irregular polygons (such as Brazilian counties) the difference is small and queen contiguity is less computationally demanding. The contiguity weights matrix produces only three neighborless units: the islands of Fernando de Noronha and Ilhabela and the nation's capital, Brasília (which is technically neither a state nor a county). The second one – used in most of Section 3 – is an inverse-distance weights matrix. In this alternative every county is related to all other counties but the strength of the relation is inversely proportional to the distance between them. The reason why the inverse-distance weights matrix is not used in this section is technical – see Appendix B for details. In any case, Section 3 will show that both the contiguity matrix and the inverse-distance matrix produce very similar results (although the magnitude of the spatial effect is larger under the latter).

The global Moran's I statistic of 0.783 ($p < 0.0001$) confirms that there is indeed spatial autocorrelation in Rousseff's vote. Map 4 shows which electoral clusters are "real" rather than mere graphical artifacts. It does that by highlighting the counties whose local indicator of spatial association (LISA) is statistically significant. These are categorized in four different groups. "High-high" counties are those with a high pro-Rousseff vote share that are surrounded by other counties with high pro-Rousseff vote shares. "Low-low" counties are those with a low pro-Rousseff vote share that are surrounded by other counties with low pro-Rousseff vote shares. "High-low" counties are those with a high pro-Rousseff vote that are surrounded by counties with low pro-Rousseff vote shares. And "low-high" counties are those with a low pro-Rousseff vote share that are surrounded by counties with high pro-Rousseff vote shares. These last two groups – high-low and low-high – are residual (only 50 counties out of a total of 2891 counties with a significant LISA) so we will focus on the high-high and low-low cases.

[Map 4 – click to view]

There are three big high-high clusters: the Northeast region; the state of Amazonas (located in the North); and the Northern portion of the state of Minas Gerais (located in the Southeast). On the other hand, there are five big low-low clusters: the state of Roraima (located in the North); the Northern portion of Mato Grosso (located in the Center-West) together with the central part of the state of Pará (located in the North); the state of Mato Grosso do Sul (located in the Center-West); and the entire state of São Paulo (located in the Southeast) together with the West portions of Paraná, Santa Catarina, and Rio Grande do Sul (all of them located in the South). Comparing Map 4 to its equivalent in Soares & Terron (2008), the only difference is the state of Amapá (located in the North). In 2006 Amapá was almost entirely a high-high cluster whereas in 2010 there is only a small high-high cluster in the middle of the state. Other than that there have been no major changes. Table 1 compares socioeconomic indicators across the four clusters.

[Table 1 – click to view]

Table 1 shows a clear socioeconomic divide between high-high clusters and low-low clusters. In high-high clusters (which comprehend 1,344 counties and 43.1 million people) the GDP per capita is R$ 8.865, 18% of the population is illiterate, 27.1% live in the rural area, 60.7% of the households have inadequate sanitation, and *bolsa-família* cash transfers represent 13.3% of the GDP. In contrast, in low-low clusters (which comprehend 1,464 counties and 61.3 million people) the GDP per capita is almost three times higher (R$ 22.088), the illiteracy rate is almost ten times lower (1.4%), the incidence of inadequate sanitation is only 26.1%, less than 10% live in the rural area, and the *bolsa-família*/GDP ratio is almost ten times lower (1.4%).[4] The divide is not only socioeconomic but also geographical: 99.7% of the high-high clusters are in the Northeast and North regions and in the Northern portion of Minas Gerais, whereas 93.2% of the low-low clusters are in the South, Southeast, and Center-West regions and in the Southern portion of Minas Gerais. In sum, the presidential race of 2010 pitted the poor, rural, "North" (loosely defined) against the rich, urban, "South". Since the electoral weight of the low-low clusters (61.3 million people) is much higher than that of the high-high clusters (43.1 million people), at first sight it is surprising that Rousseff won the race. Her victory is the result of two factors: the non-clustered counties (2,675 in total, with socioeconomic indicators between those of the low-low group and those of the high-high group) and her expressive vote share even in the low-low clusters. This last point deserves some elaboration. It looks as if Rousseff's vote share had a "lower bound": in no county she received less than 19% of the vote. For instance, consider the county of Capixaba, in the state of Acre. It has the largest LISA statistic of all low-low counties: 5.13. Yet Rousseff received 20.1% of the vote there, which is a substantial minority. Or, alternatively, consider the case of São Paulo city (the capital of the state of São Paulo), which is the most important PSDB stronghold in the country. Rousseff received 46.8% of the vote, narrowly losing the race there by only 6.25 percentage points. Thus even where Rousseff lost she still amassed an important share of the vote. The same was not true of her adversary, former governor of São Paulo and PSDB candidate José Serra. His vote share had no "lower bound": in high-high clusters he often received less than 10% of the vote share; in the county of Calumbi (in the state of Ceará), for instance, he received only 3.49% of the vote share.

---

[4] The source of all socioeconomic indicators except *bolsa-família* is IBGE. The source of bolsa-família expenditures is the *Controladoria-Geral da União* (CGU). For details and summary statistics, see Appendix B.

Maps 1 and 4 above showed that Rousseff's electoral support in 2010 followed the same geographical pattern of Lula's in 2006. What about the electoral support of the PT? As discussed above, while Lula managed to extend his geographical basis of support, the PT itself did not (Terron & Soares 2010). Did that change between 2006 and 2010? Maps 5 and 6 show the PT vote share for state assemblies and for the lower house in 2010.

[Map 5 – click to view]

[Map 6 – click to view]

Clearly, Brazilian politics remains highly personalistic: the geographical distribution of PT vote does not help us predict the geographical distribution of Rousseff's vote. This indicates that the decoupling found by Terron & Soares (2010) has not receded. The difference is especially salient in the state of Amazonas (located in the North) and in the Southeast and South regions. In Amazonas, Rousseff received 81.1% of the votes, but the PT received only 5.2% of the votes for the state assembly. On the other hand, in São Paulo (located in the Southeast), Rousseff received only 44.9% of the vote, whereas the PT received 17.1% of the vote for the state assembly. If anything, it seems that Rousseff was stronger where the PT was weaker. In fact, the bivariate Moran's I statistic is -0.0453. In other words, Lula managed to "transfer" his electoral basis of support to Rousseff, but not to the PT.[5]

In sum, this section has shown that presidential voting is strongly clustered geographically; that the electoral basis of Rousseff in 2010 was essentially the same as the electoral basis of Lula in 2006; and that the PT itself did not benefit from Lula's popularity. The next section tests alternative explanations for spatial autocorrelation in Rousseff's vote and assesses whether these explanations are equally valid across different areas of the country.

3. Spatial dependence in Brazilian elections

3.1 Hypothesis testing

---

[5] It is also worth noting that although the PT vote for state assemblies does show spatial autocorrelation the effect is much less pronounced than in the presidential race (the Moran's I is 0.3733, with $p < 0.0001$).

In this subsection the three hypothesis outlined in Section 1 are put to test. The first step is the estimation of a least squares model with all the relevant covariates but no spatial terms. Then Lagrange-multiplier (LM) tests are performed in order to choose a spatial model. If both the LM test for SAR and the LM test for SARE turn out statistically significant, robust LM tests are used. If both robust LM tests are also statistically significant, then (ordinarily) the researcher either picks the model with the highest LM statistic or chooses a SARAR model. The chosen model is then estimated either by GMM or by MLE. OLS is usually not viable with spatial models. The spatial lag introduces endogeneity into the model – county A influences county B but county B also influences county A –, which renders OLS estimators inconsistent, i.e., they do not converge to their true values as the sample size increases. As Ward & Gleditsch (2008) put it, "Instead, the feedback or dependency that is ignored by the OLS specification is likely to grow rather than be eliminated as the size of the data frame grows" (41). The spatial error, in turn, violates the assumption of independently distributed residuals; although the point estimates remain unbiased and consistent, OLS standard errors will be underestimated, leading to type I errors. MLE and GMM, on the other hand, have been shown to produce consistent estimators in large samples (Lee 2004; Kelejian & Prucha 2010).

Following the algorithm just presented, Table 2 reports least squares estimates of Rousseff's county-level vote share in the second round of the 2010 presidential election. The specification is straightforward. The goal is to filter out the effect of socioeconomic clustering so that any remaining spatial effects may be ascribed to social interactions (captured by the spatial lag) and campaign strategies (captured by the spatial error). Thus the included variables are the *bolsa-família*/GDP ratio, the natural logarithm of GDP per capita, the percentage of people living in the rural area, the illiteracy rate, and the percentage of households with inadequate sanitation. There is also a dummy for every state except Acre (which is the baseline category) and a dummy coded 1 if the mayor was from the PT in 2009 and 0 otherwise (the idea is that a PT mayor will probably campaign harder for Rousseff). (For data sources and details, see Appendix B.) Three different models are reported: OLS estimates, weighted least square (WLS) estimates using population as weight (WLS_P), and WLS estimates using state-level variance as weight (WLS_V). The plain OLS estimation is intended as a baseline. The WLS_P estimation addresses heteroskedasticity by taking into account that each observation is a proportion, not an actual individual observation. For instance, Rousseff's vote share was 46.8% in the county of São Paulo (capital of the homonymous state) and also 46.8% in the county of Água Limpa (located in the state of Goiás). But São Paulo has a population of 11.3 million people while Água Limpa has a population of 2.2 thousand people. According to the central limit theorem, the larger the number of data points, the closer the observed proportion is to the "true" proportion – i.e., the vote share we would expect given the relevant covariates and their respective parameters. In Água Limpa, an accidental fire in one of the county's polling stations might cause the observed vote share to deviate substantially from the true vote share. In São Paulo, by contrast, an accidental fire in one of the county's polling stations would probably only cause a marginal deviation from the true vote share. Therefore the model should provide a better fit for counties with large populations than for counties with small populations – in other words, the variance must be negatively correlated with population size. In order to correct for that, in the WLS_P estimation all variables are weighted by the natural logarithm of the county populations.[6] The WLS_V estimation, in turn, addresses heteroskedasticity by grouping the residuals from the plain OLS regression by state and using the respective state-level variances as weights in a subsequent regression. The idea is that the variance, although not constant across states, may be constant within each state. This approach is known as groupwise heteroskedasticity correction. The state dummies are omitted in all tables. The weights matrix used in the LM tests is based on queen contiguity (see discussion above).

[Table 2 – click to view]

Coefficients and standard errors are very stable across the three estimations. Interestingly, PT mayors actually reduced Rousseff's vote share. According to Strøm (1990), incumbency is usually an electoral liability. So were Brazilian voters using the presidential election to punish local incumbents? That sounds unlikely: that would require voters to be able to associate Rousseff with the PT. As discussed above, Brazilian politics is highly personalistic and the electoral bases of Rousseff (inherited from Lula) and of the PT are very different. The negative effect of PT mayors is not a fluke though: Nicolau & Peixoto (2007) and Zucco (2008) find that PT mayors also had a negative effect on Lula's vote share in the 2006 election. This is a curious result and probably deserves some further investigation (which is beyond the scope of this paper). Also unexpected is the lack of statistical significance of GDP per capita. This is probably due to wealth inequality in certain areas, which renders GDP per capita an imperfect measure of individual-level prosperity. All other variables have the expected sign though, and they are all statistically significant: the more a

[6] The log is chosen over sheer population in order to bring all variables to a similar scale, since scale differences may themselves cause heteroskedasticity.

county is rural, illiterate, dependent on *bolsa-família*, or lacking in good sanitation, the higher Rousseff's vote share. The R-squared is also very stable across estimations and shows that the model accounts for some 63% of the variation in Rousseff's vote share. One should not read too much into those estimations though. First, neither the WLS_P nor the WLS_V estimation eliminated heteroskedasticity: the Breusch-Pagan test rejects the null hypothesis of homoskedasticity in all cases. Second – and more importantly –, all three models show statistically significant spatial dependence. The global Moran's I statistic rejects the null hypothesis of spatial randomness in all three cases. As regards the specific spatial model to be estimated, the OLS estimation suggests a SARAR model whereas the WLS_P and WLS_V estimations suggest a SARE model. Since the tests in favor of the SARAR model are not robust to the simple weighting carried out in the WLS_P and WLS_V estimations, it seems that the SARE model should be preferred. Yet, to be on the safe side, both the SARAR and SARE models will be estimated and compared. If the SARE model is indeed the correct one, then the spatial lag of the SARAR model should not be statistically significant. In order to check whether the contiguity assumption is too restrictive, we also run the same models using an inverse-distance weights matrix (wherein every county is related to all other counties but the magnitude of the effect is inversely proportional to the distance between them). As for the estimation method, the main procedure is the GMM estimator developed by Kelejian & Prucha (2010). The big advantage of this estimator is that it is robust to heteroskedasticity. It is limited though in that it does not allow us to constraint the SAR term (or the SARE term) to zero. In other words, it only allows the estimation of the SARAR model. Thus the SARE model is estimated by MLE. Lee (2004) shows that the ML estimator is asymptotically consistent under homoskedasticity but Arraiz et al (2010) show that that is not the case under heteroskedasticity. Therefore the ML estimates here must be interpreted with care. Table 3 presents the results.

[Table 3 – click to view]

The spatial lag is either extremely close to zero or statistically insignificant in the unrestricted models where. The spatial error, on the other hand, is statistically significant in all models. Finally, the socioeconomic indicators are statistically significant and have the expect sign in all models (except for GDP per capita). In substantive terms, the results support the second and third hypotheses but not the first one: i.e., socioeconomic similarities and campaign strategies help explain spatial autocorrelation in voting across counties, but social interactions do not. Unfortunately there are no available data to test the campaign hypothesis directly. Only total campaign expenditures are publicly reported – there is no disaggregation by county. Moreover, in Brazil over 50% of all campaign donations are estimated to be "off the book", so any attempt to use publicly available figures would be doomed from the start. Any research design would have to be based on non-financial data: for instance, the number of times Rousseff visited the county during the campaign. But these non-financial data would capture only relatively unimportant aspects of the campaign – a candidate can only visit so many places so the campaign must rely on costly publicity to reach most voters. Thus in the end there would probably still be spatial error in the results. In sum, the above estimations are perhaps the best test of the campaign hypothesis that can be performed with the available data.

We also ran unrestricted ML estimations (using a queen contiguity weights matrix and an inverse-distance weights matrix) and both the exogenous regressors and the spatial terms are very similar to those of the GMM estimations reported in Table 3. All this suggests that heteroskedasticity is not causing too much damage in the SARE estimations. Another finding is that although the choice of weights matrix has little effect on the exogenous regressors, it has a substantial impact on the spatial error term. Using an inverse-distance weights matrix substantially boosts the spatial error term in the ML estimation: it increases from .113 to .702. Thus it seems that the simplifying assumption of contiguity effects was too restrictive: counties are related not only to their immediate neighbors but also to other counties as well. (It is important not to confuse weights matrices based on inverse distance with weights matrices based on distance bands. In inverse-distance matrices, every nearby observation is assigned a different weight, which is inversely proportional to the corresponding distance. Going back to the example discussed before, Ribeirão Preto may be affected by Ribeirão Pires – over 300 km away –, but it will be much more affected by nearby counties, such as Monte Alto and Batatais. In distance-band matrices – such as the one used in Carraro et al 2007 –, every nearby observation is assigned a weight of one, whereas those outside the specified range are assigned a weight of zero. It is possible to "combine" both types by computing an inverse-distance weights matrix with a threshold distance above which all weights are constrained to zero. That is not done here in order to avoid the imposition of arbitrary constraints on the model).

3.2 Heterogeneity

This subsection assesses whether the estimates are spatially stationary, i.e., whether they are stable across space. Two techniques are employed: geographically weighted regression (GWR) and subsampling. GWR is an estimation procedure that allows the effect of each independent variable to vary across observations. Thus rather than a unique set of estimates it produces as many sets as observations. The intuition is not entire unlike that of structural breaks. The difference is that GWR explicitly ascribes heterogeneity to geographical location and allows every observation (rather than groups of observations) to have its own set of estimates. At first sight the method may seem odd: the OLS estimator requires that there be more observations than variables, so how can a single observation produce coefficient estimates? The answer is that the estimates for each observation are actually based on other observations – specifically, those that are considered "neighbors". Since every observation has a different set of neighbors, the outcome is a different set of estimates for each observation. Formally, while the OLS estimator is given by $\hat{\beta} = (X'X)^{-1}X'Y$, the GWR estimator is given by $\hat{\beta}_i = (X'W_iX)^{-1}X'W_iY$, where $\boldsymbol{W_i}$ is a matrix that contains the weight of the other observations in relation to the *i*th observation. The weights must be inversely proportional to the distance, which is usually achieved by defining $w_{ij} = \exp(-d_{ij}/h)^2$, where $\boldsymbol{d}$ is the distance between observations $\boldsymbol{i}$ and $\boldsymbol{j}$ and $\boldsymbol{h}$ is a quantity called the bandwidth or kernel. The bandwidth is a distance threshold beyond which observations are no longer used in the estimation of $\hat{\beta}_i$. It can be a fixed number or it can be allowed to vary across $\boldsymbol{W}$ (in which case it is called an adaptive bandwidth).[7] Since Brazilian counties vary widely in size (see discussion in Section 2 above), here an adaptive bandwidth is used. Table 4 summarizes the results. The estimates of the GMM estimation using the inverse-distance weights matrix (from Table 3 above) are reproduced here to facilitate comparison. (For a comprehensive discussion of GWR see Fotheringham, Brudson & Charlton 2002).

[Table 4 – click to view]

Although the results differ markedly between the GMM estimation and the GWR estimation, the GWR estimates themselves do not vary much. The computation of standard errors for GWR estimates is involved since it raises issues joint tests (see Fotheringham, Brudson &

[7] A common choice, known as cross-validation, is to select the bandwidth that minimizes the mean square prediction error of the geographically weighted regression.

Charlton 2002 for a discussion of this point) but Table 4 suggests that most estimates would probably not be statistically different from each other. Urbanization aside, for practical purposes the GWR estimates summarized in Table 4 can be considered equivalent.

GWR estimates should be interpreted with care though. As the very proponents of the method acknowledge, GWR estimates are necessarily biased (Fotheringham, Brudson & Charlton 2002). Thus it is safer to complement the GWR procedure with some other check of spatial stationarity. Here this is done by partitioning the dataset and estimating separate models for selected states. We picked the state with the largest number of counties within each macro-region (North, Northeast, Center-West, South, and Southeast). Thus the states chosen are São Paulo (645 counties), Piauí (223), Pará (143), Goiás (246), and Rio Grande do Sul (498). The regressions are based on the GMM estimator and on the inverse-distance weights matrix discussed before. Table 5 summarizes the results.

[Table 5 – click to view]

Clearly the estimates are not spatially stationary. The *bolsa-família*/GDP ratio has no effect in Piauí, Goiás, and Rio Grande do Sul. This is a striking result: the existing literature considers *bolsa-família* the single most important predictor of PT vote. On top of that, Piauí is the poorest state in Brazil so one would expect *bolsa-família* to have had a particularly strong effect there. This lends credence to Carraro et al's (2007) claim that the electoral impact of *bolsa-família* has been overrated by the literature. (On the other hand, it is possible that this is just an artifact of the modifiable areal unit problem – MAUP: perhaps *within* every county *bolsa-família* has an effect. As Carraro et al 2007 warn, the proper level of analysis of *bolsa-família*'s electoral impact is the individual, not the county. We do not engage that debate here: the goal of this paper is not to explain county-level variation in Rousseff's vote, but county-level spatial autocorrelation in Rousseff's vote.) Urbanization, in turn, has a negative effect in Piauí and no effect in the other four states. Illiteracy has a positive effect on the highly developed states of São Paulo and Rio Grande do Sul and no effect on the much poorer states of Piauí, Goiás, and Pará. Sanitation has no effect in any of the five states. Finally, PT mayors have a negative effect in Piauí and zero effect in the other states. Given such high heterogeneity, any substantive interpretations of the magnitude of the estimates would be of little use. These results are impressive and they call into question the entire existing literature on Brazilian elections, all of which relies on aggregate, full-sample analysis. Moreover, these results call into question our third hypothesis, i.e., that socioeconomic similarities can explain spatial autocorrelation in voting patterns. It seems that this may be the case in some regions but not in others and that where socioeconomic similarities do have an effect, this effect is different in different areas. These results also show that GWR can be highly misleading as a measure of heterogeneity: if GWR estimates do not show heterogeneity the researcher must proceed to subsampling before drawing any further conclusions.

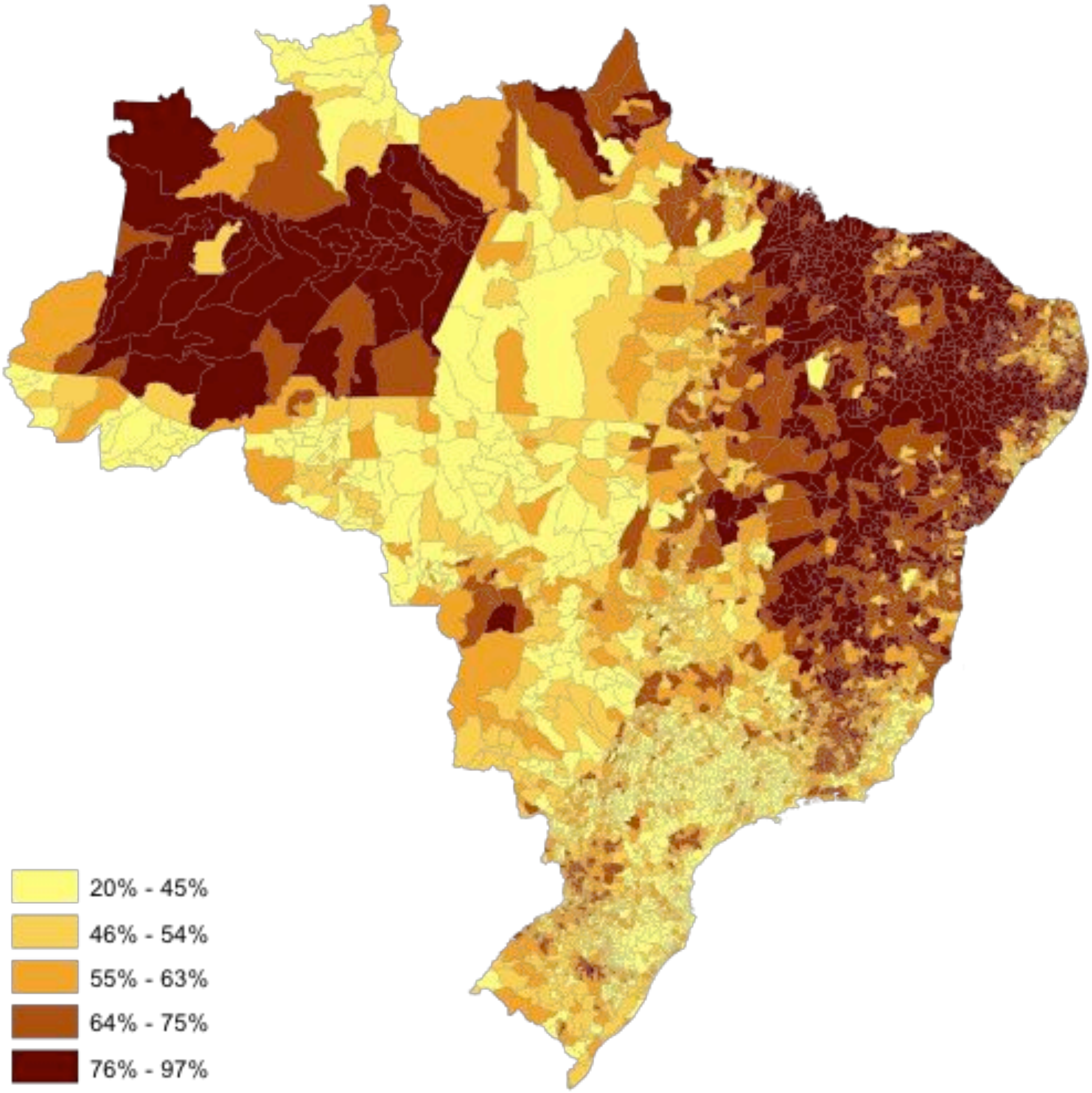
Reassuringly though, Table 5 shows that, regardless of heterogeneity, the spatial error model still prevails over the spatial lag model. The spatial error term is statistically significant in all five models while the spatial lag is statistically significant only in Goiás, and even then it is very close to zero.

4. Conclusion

This paper tested three alternative explanations for vote clustering across counties: social interactions, campaign strategies, and socioeconomic similarities. The empirical results showed no support for the first hypothesis, preliminary support for the second hypothesis (since the test was indirect, the evidence must be considered preliminary), and qualified support for the third hypothesis (global estimates support it, but state-specific estimates show high heterogeneity). This paper has also complemented the existing literature on Brazilian elections by showing that although Rousseff "inherited" Lula's expanded electoral basis of support (mainly the North and Northeast regions), the PT itself still remains a mostly urban-based party.

Appendix A – Maps and tables

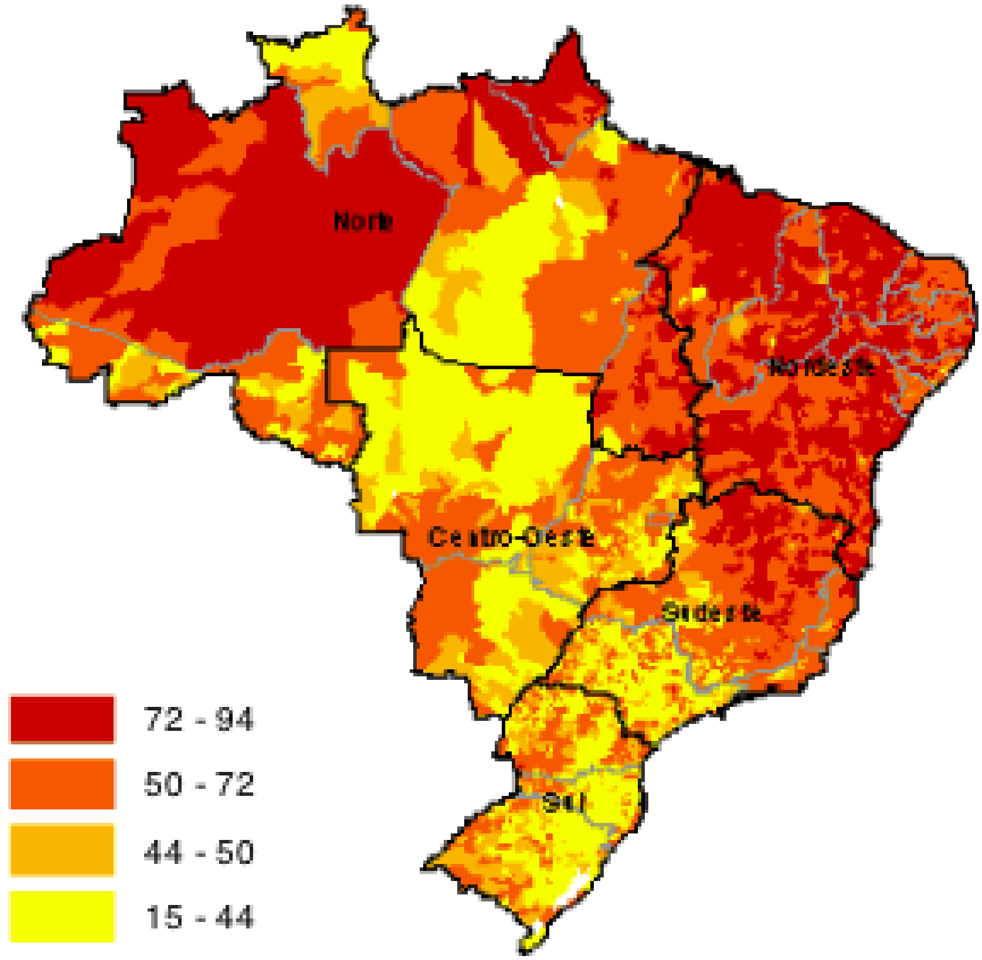
Map 1 – Dilma Rousseff's vote share in the second round (2010)



Sources: TSE. IBGE.

[Click to return to text.]

Map 2 – Lula's vote share in the second round (2006)



Source: Soares & Terron (2008).

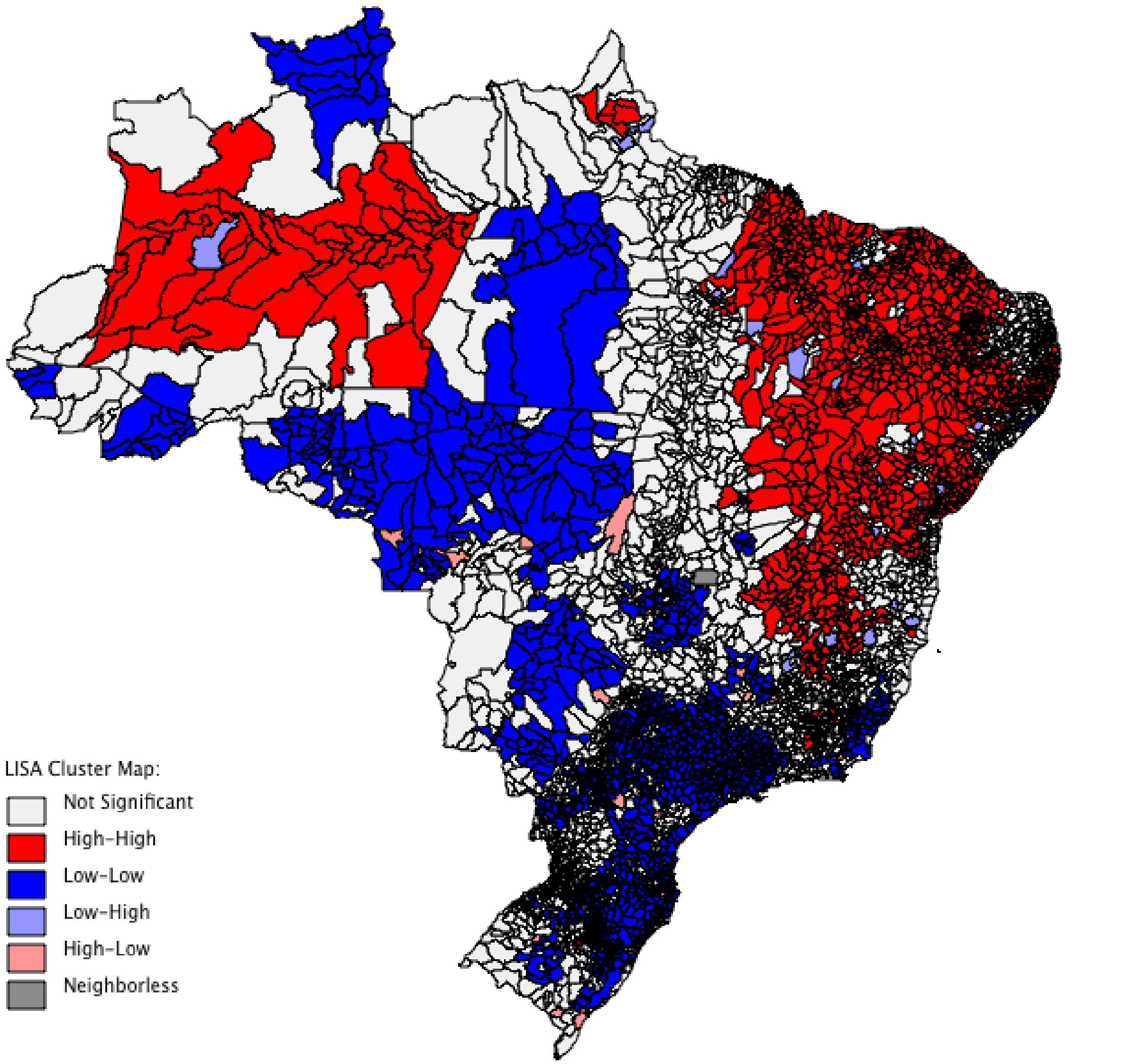
[Click to return to text.]

Map 3 – Neighborless counties resulting from a 50 km distance band (yellowed areas)

Source: IBGE.

[Click to return to text.]

Map 4 – LISA statistics for Rousseff's vote share in the second round (2010)



Sources: TSE. IBGE.

[Click to return to text.]

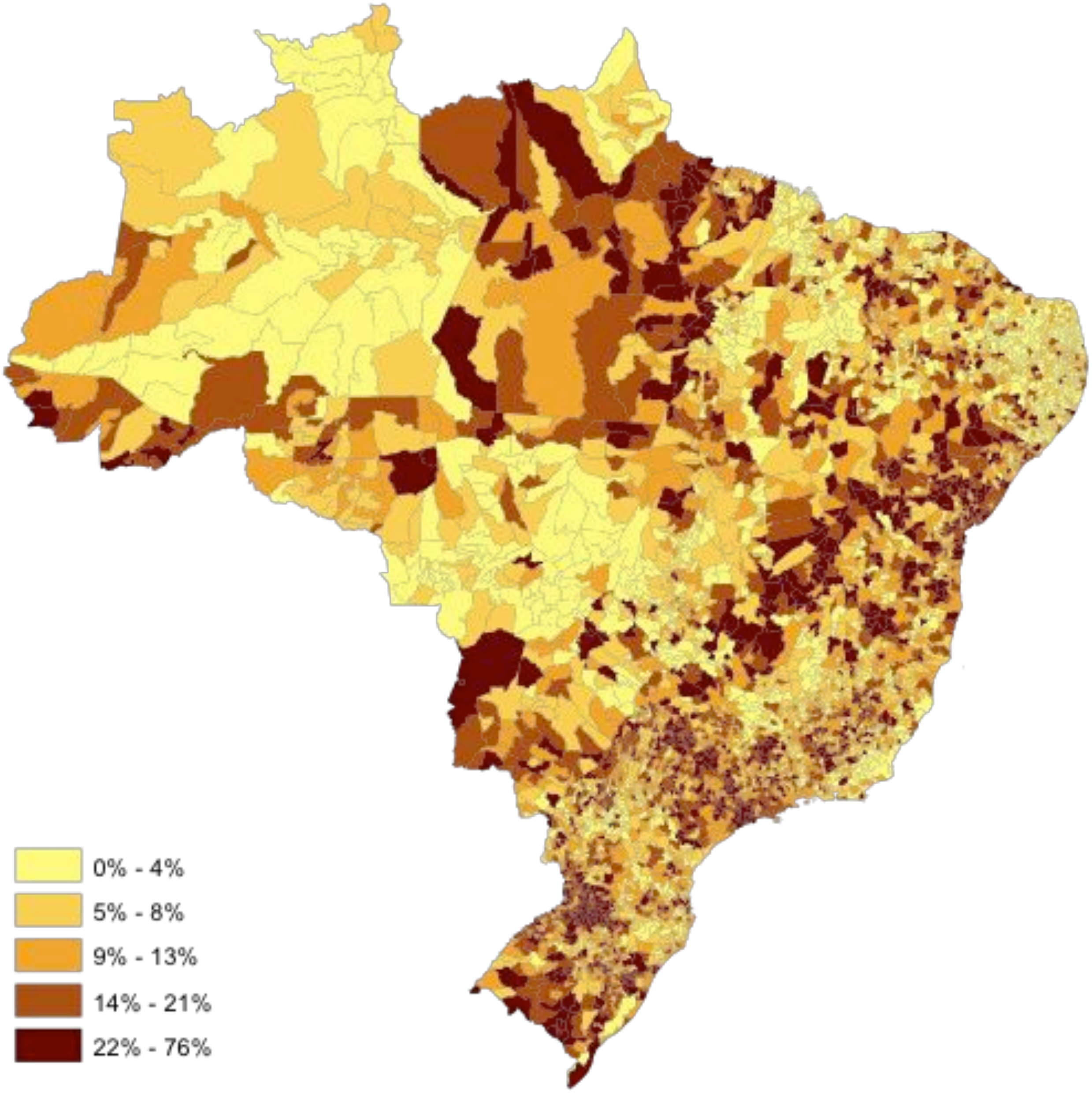
Table 1 – Socioeconomic indicators of LISA clusters

	High-High	Low-Low	Low-High	High-Low
Number of counties	1,344	1,464	30	50
Population	43.1 million	61.3 million	2.1 million	1.6 million
GDP per capita	R$ 8.865	R$ 22.088	R$ 9.185	R$ 14.799
Bolsa-família/GDP	13.3%	1.4%	10.5%	2.6%
Illiteracy rate	18.0%	5.4%	13.7%	6.6%
Rural population	27.1%	9.6%	1.4%	9.5%
Inadequate sanitation	60.7%	26.1%	61.7%	24.3%

Sources: IBGE (population, GDP per capita, illiteracy rate, rural population, inadequate sanitation); CGU (*bolsa-família* disbursements). Notes: All data refer to 2009. Illiteracy rate is the rate of people 15 or older that cannot read or write.

[Click to return to text.]

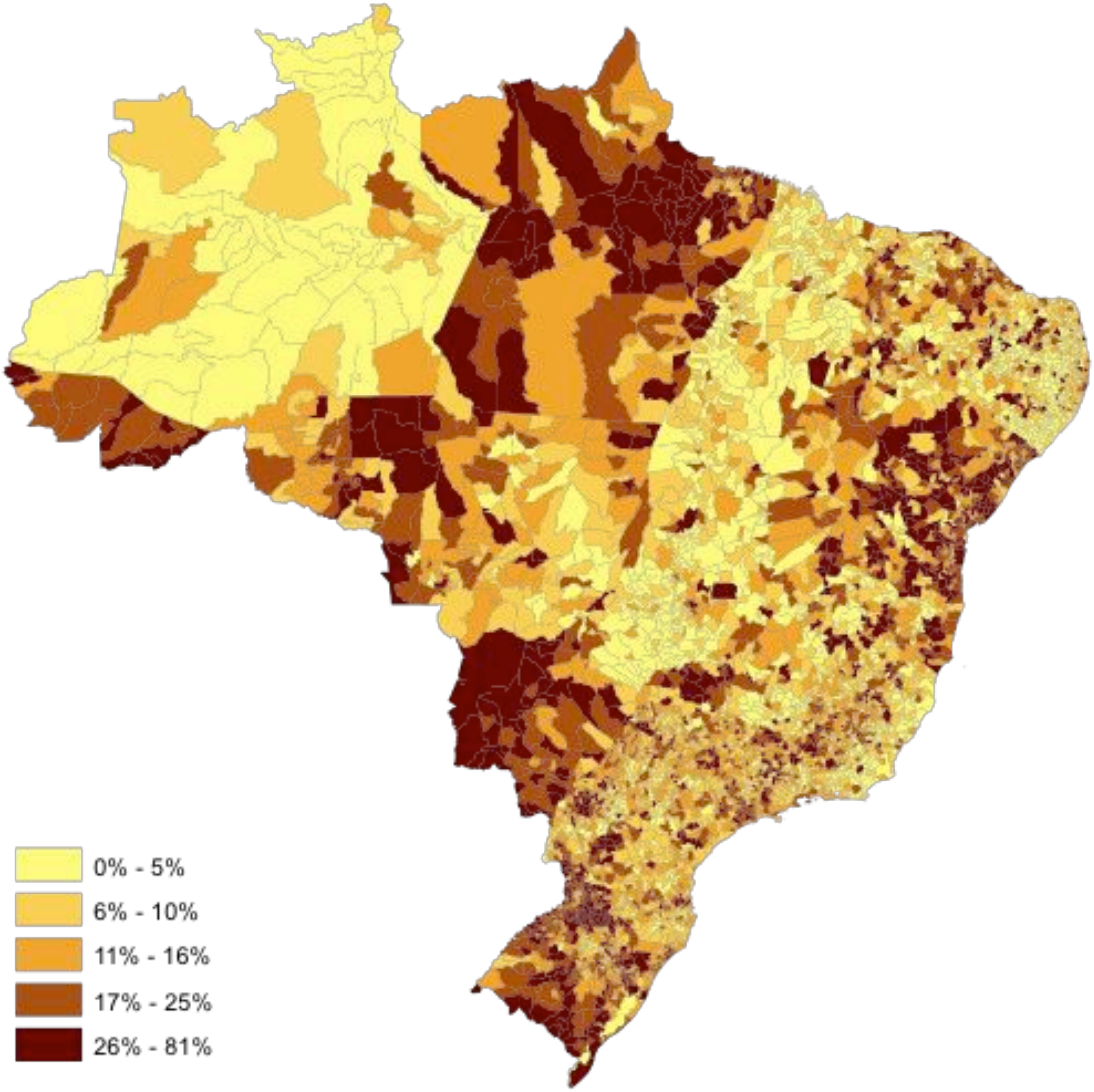
Map 5 – PT vote shares for state assemblies (2010)



Sources: TSE. IBGE.

[Click to return to text.]

Map 6 – PT vote shares for the lower house (2010)



Sources: TSE. IBGE.

[Click to return to text.]

Table 2 – WLS estimates of Rousseff's vote share in the second round (2010)

	OLS	WLS_P	WLS_V
intercept	18.518***	17.982***	19.233***
	(4.007)	(3.980)	(4.308)
PT mayor	-1.106**	-1.044*	-.935*
	(.422)	(.415)	(.428)
Bolsa-família/GDP	.198***	.204***	.186***
	(.018)	(.018)	(.018)
Ln(GDP per capita)	.270	.391	.179
	(.346)	(.344)	(.350)
% Rural population	.028***	.027***	.026**
	(.008)	(.008)	(.008)
% Illiterate	.169***	.142***	.171***
	(.031)	(.031)	(.031)
% Inadequate sanitation	.061***	.063***	.063***
	(.007)	(.007)	(.007)
N	5566	5566	5566
R^2	.638	.643	.636
F	305***	311***	303***
Breusch-Pagan	234.55***	234.55**	234.44***
Moran's I (residuals)	.438***	.443***	.438***
LM(error)	2922.6***	2911.3***	2925.3***
LM(lag)	2834.3***	32.653***	-.4012
Robust LM(error)	200.66***	2878.9***	2925.7***
Robust LM(lag)	112.4***	.1957	-.0059

Notes: *** p < .0001; ** p < .001; * p < .01. State dummies omitted.

[Click to return to text.]

Table 3 – GMM and MLE estimates of Rousseff's vote share in the second round (2010)

	SARAR(1,1) contiguity (GMM)	SARAR(1,1) inverse distance (GMM)	SAR(1) contiguity (MLE)	SAR(1) inverse distance (MLE)
intercept	26.353***	23.132***	28.213***	27.335***
	(4.382)	(4.035)	(3.825)	(2.126)
PT mayor	-1.831***	-1.494***	-1.853***	-1.761***
	(.321)	(.387)	(.326)	(.359)
Bolsa-família/GDP	.108***	.143***	.106***	.129***
	(.016)	(.017)	(.015)	(.016)
Ln(GDP per capita)	-.032	.100	-.174	-.342
	(.327)	(.347)	(.290)	(.307)
% Rural population	.038***	.035***	.039***	.040***
	(.007)	(.007)	(.007)	(.007)
% Illiterate	.242***	.170***	.242***	.240***
	(.036)	(.033)	(.032)	(.031)
% Inadequate sanitation	.018**	.028***	.020**	.027***
	(.007)	(.007)	(.006)	(.007)
Spatial lag	-.002*	-.0001	-	-
	(.001)	(.003)		
Spatial error	.117***	.381***	.113***	.702***
	(.001)	(.017)	(.003)	(.009)
N	5566	5566	5566	5566
Log-likelihood	-	-	-19262.6	-19461.5
Wald chi2	-	-	2810.66***	49686.7***

Notes: *** p < .0001; ** p < .001; * p < .01. State dummies omitted.

[Click to return to text.]

Table 4 – GWR estimates for Rousseff's vote share in the second round (2010)

	GMM (inverse distance)	GWR Min	GWR Median	GWR Max
intercept	26.353	58.400	59.500	60.800
PT mayor	-1.831	-1.680	-1.500	-1.260
Bolsa-família/GDP	.108	.369	.371	.375
Ln(GDP per capita)	-.032	-1.540	-1.440	-1.360
% Rural population	.038	.009	.016	.020
% Illiterate	.242	.278	.288	.294
% Inadequate sanitation	.018	.021	.022	.024

Notes: *** $p < .0001$; ** $p < .001$; * $p < .01$.

[Click to return to text.]

Table 5 – GMM estimates for Rousseff's vote share in the second round (2010), by selected states (São Paulo, Piauí, Pará, Goiás, Rio Grande do Sul)

	São Paulo	Piauí	Pará	Goiás	Rio Grande do Sul
intercept	21.110	102.456***	30.113	60.746***	34.773**
	(9.137)	(28.753)	(28.516)	(16.226)	(12.795)
PT mayor	1.739	-6.448*	-.271	-1.710	-.175
	(1.007)	(2.003)	(1.505)	(1.658)	(.863)
Bolsa-família/GDP	.751***	-.017	.414***	.191	.257
	(.164)	(.030)	(.091)	(.114)	(.182)
Ln(GDP per capita)	1.850	-4.955	3.089	-.999	.764
	(.749)	(2.990)	(2.472)	(1.541)	(1.238)
% Rural population	-.029	.135***	.082	.008	.031
	(.041)	(.029)	(.038)	(.034)	(.020)
% Illiterate	1.204***	-.171	-.596*	.398	.506*
	(.141)	(.132)	(.199)	(.171)	(.176)
% Inadequate sanitation	-.029	.123	-.101	.002	.009
	(.036)	(.051)	(.088)	(.021)	(.018)
Spatial lag	-.025	.023	.052	-.087***	.0171
	(.027)	(.051)	(.069)	(.020)	(.0121)
Spatial error	1.002***	2.174***	5.86**	4.778**	1.010*
	(.227)	(.496)	(1.85)	(1.789)	(.292)
N	645	223	143	246	498
Region	Southeast	Northeast	North	Center-West	South

Notes: *** p < .0001; ** p < .001; * p < .01. Weights matrix: inverse distance.

[Click to return to text.]

Appendix B – Data sources, software used, and summary statistics

Data sources:

• All geocoded data were retrieved from IBGE's website and are available at ftp://geoftp.ibge.gov.br/malhas_digitais/municipio_2007/escala_2500mil/proj_policonica_sirgas2000/ They reflect Brazil's counties as they were in 2007 and are based on conic projections (thus they are "safe" for Euclidean calculations). The coordinates are expressed in kilometers.
• Mayor's party, GDP per capita, urbanization, illiteracy and sanitation data were retrieved from several different IBGE surveys. *Bolsa-família* data were retrieved from CGU's website and are available at http://www.portaltransparencia.gov.br/ All these data are from 2009 (2010 data were not available for social indicators).
• A total of 37 counties had one or more variables missing. In non-spatial econometrics observations with missing values can simply be deleted, but in spatial analysis this is problematic since observations are dependent on each other. Thus I imputed state averages for continuous missing values (social indicators) and state modes for binary missing values (the PT mayor dummy).
• The compiled version of the dataset used in this paper can be downloaded from http://www.2shared.com/file/HxIOODWz/Brazil2010election.html

Software used:

• The graphical and statistical analyses were carried out with a variety of applications. Vote share maps were produced in ArcGIS 10. Map 3, which depicts the area excluded by the 50km distance band, and Map 4, which depicts LISA clusters, were produced in GeoDa 1.0.1. The statistical analysis was divided between Stata 12 and R 2.14.2. The statistical tests of Section 2, the preliminary least squares regressions (Table 2) and the GWR estimations were carried out mostly in R, using the 'spdep' and 'sphet' packages (the Stata tools for global Moran' I, LISA statistics, and LM tests require that the weights matrix be created via 'spatwmat' but this function does not work well with large datasets; the Stata tool for GWR does not require 'spatwmat' but its algorithm is less efficient than that of the R equivalent and it crashed after a few iterations). On the other hand, all GMM and ML estimations were carried out in Stata, using the 'sppack' package (here it was R that had less efficient algorithms that failed to converge, especially when it came to GMM estimation; moreover, R cannot handle inverse-distance weights matrices).

Summary statistics:

	Mean	Std. dev.	Min	Max
Rousseff's vote share (2nd round)	59.46	15.39	19.66	96.50
PT vote share (state assemblies)	12.60	10.53	.06	75.65
Ln(GDP per capita)	9.02	.69	7.56	12.79
Bolsa-família/GDP	17.48	18.24	.001	92.78
% Rural population	36.17	22.03	0	95.82
% Illiterate	16.15	9.83	.95	44.4
% Inadequate sanitation	64.22	29.95	1.15	100

Works cited

Anselin, L. (2006). Spatial econometrics. In Mills, T. & Patterson, K. (eds.), *Palgrave Handbook of econometrics*, vol. 1. Basingstoke: Palgrave Macmillan, 901–941.

Anselin, L., & Florax, R. (1995). Introduction. In Anselin, L. & Florax, R. (eds.), *New directions in spatial econometrics*. Berlin: Springer, 3-18.

Arraiz, I., Drukker, D., Kelejian, H. & Prucha, I. (2010). A spatial Cliff-Ord-type model with heteroskedastic innovations: small and large sample results. *Journal of Regional Science*, 50(2), 592–614.

Carraro, A., Araújo Junior, A., Damé, O., Monasterio, L. & Shikida, C. (2007). *'É a economia companheiro!': uma análise empírica da reeleiçnao de Lula com dados municipais*. IBMEC-MG Working Paper #41. Retrieved from http://ceaee.ibmecmg.br/wp/wp41.pdf

Cho, W. & Gimpel, J. (2007). Rough terrain: spatial variation in campaign contributing and volunteerism. *American Journal of Political Science*, 54(1), 74-89.

Fotheringham, A., Brudson, C. & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Jon Wiley & Sons.

Huckfeldt, R. & Sprague, J. (1991). Discussant effects on vote choice: intimacy, structure, and interdependence. *Journal of Politics*, 53(1), 122-158.

Hunter, W. & Power, T. (2008). Rewarding Lula: executive power, social policy, and the Brazilian elections of 2006. *Latin American Politics & Society*, 49(1), 1-30.

Kelejian, H. & Prucha, I. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157, 53-67.

Lee, L. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6), 1899-1925.

Nicolau, J. & Peixoto, V. (2007). *Uma disputa em três tempos: uma análise das bases municipais das eleições presidenciais de 2006*. Paper presented at the XXXI ANPOCS meeting. Retrieved from http://jaironicolau.iesp.uerj.br/artigos/NICOLAU%20&%20PEIXOTO%20ANPOCS2007%20ST%2024%20PARTIDOS%20E%20SISTEMAS%20PARTIDA%C2%A6%C3%BCRIOS.pdf

Nickerson, D. (2008). Is voting contagious? Evidence from two field experiments. *American Political Science Review*, 102(1), 49-57.

Soares, G. & Terron, S. (2008). Dois Lulas: a geografia eleitoral da reeleição (explorando conceitos, métodos e técnicas de análise geoespacial). *Opinião Pública*, 14(2), 269-301.

Strøm, K. (1990). *Minority government and majority rule*. Cambridge: Cambridge University Press.

Terron, S. & Soares, G. (2010). As bases eleitorais de Lula e do PT: do distanciamento ao divórcio. *Opinião Pública*, 16(2), 310-337.

Tinbergen, J. (1962). An analysis of world trade flows, the Linder hypothesis and exchange risk. In Tinbergen, J. (ed.), *Shaping the world economy*. New York: The Twentieth Century Fund.

Ward, M. & Gleditsch, K. (2008). *Spatial regression models*. Los Angeles: Sage.

Zucco, C. (2008). The president's 'new' constituency: Lula and the pragmatic vote in Brazil's 2006 presidential elections. *Journal of Latin American Studies*, 40(1), 29-49.