PMKT
BRAZILIAN JOURNAL OF MARKETING, OPINION, AND MEDIA RESEARCH

Chief Editor: Fauze Najib Mattar | Associate Editor: Claudio Alberto de Moraes
Evaluation System: Triple Blind Review
Published by: ABEP - Brazilian Association of Research Companies
Languages: Portuguese and English

Party organization and votes in the legislature: Study of the municipal organizations of the PT and PSDB in Sao Paulo state from the composition and geographical influence

Organização partidária e votos no Legislativo: Estudo das organizações municipais do PT e PSDB no estado de São Paulo a partir da composição e influência geográfica

Artur Junqueira Lascala*, Bruno Mitio Assano Silva, Eduardo de Rezende Francisco
Sao Paulo School of Business Administration, Getulio Vargas Foundation (FGV-EAESP), SP, Brazil

ABSTRACT

Does the type of party organization in a municipality (directories or provisional commissions) affect the number of votes attributed to a given party? Socio-economic-demographic factors are related to the electoral performance of the parties? Seeking to answer these questions, this study used data from the 2010 municipal elections for the State of São Paulo. It was analyzed the spatial distribution of the amount of votes attributed to Congressman, less susceptible to influences of party coalitions than elections for Mayor, from the configuration of party organization and IPVS - Paulista Social Vulnerability Index of each municipality. The leading parties PT and PSDB were analyzed. This is a georeferenced analysis of the potential influence of locally installed partisan organization and voter social profile in the municipality on the result of the election. Exploratory geographic techniques and Spatial Statistics models, such as Spatial Auto-Regressive model (SAR) and GWR (Geographically Weighted Regression) were used. The results show the great importance of the geographic aspects as an explanation of the local political phenomena, and point to the importance of the variable space on the explanation of the performance of the parties in the electoral process.

KEYWORDS: Party Organization; Elections; Geographical Intelligence; Spatial Statistics; IPVS.

RESUMO

A forma de organização partidária em um município (diretórios ou comissões provisórias) pode afetar a quantidade de votos atribuídos a determinado partido? Fatores socioeconômicos e demográficos se relacionam com o desempenho eleitoral dos partidos? Para responder a essas questões, este estudo utilizou dados das eleições municipais de 2010 para o estado de São Paulo. Analisou-se a distribuição espacial da quantidade de votos atribuída a Deputado Federal, menos suscetível a influências de coligações partidárias do que eleições para Prefeito, a partir da configuração da organização partidária e do Índice Paulista de Vulnerabilidade Social (IPVS) de cada município. Os partidos "polo" PT e PSDB foram analisados. Trata-se de análise georreferenciada acerca da potencial influência da organização partidária localmente instalada e do perfil social do eleitor, no município, sobre o resultado da eleição. Técnicas geográficas exploratórias e modelos de Estatística Espacial, como o Spatial Auto-Regressive model (SAR) e Geographically Weighted Regression (GWR) foram utilizados. Os resultados encontrados demonstram a grande importância dos aspectos geográficos como explicação dos fenômenos políticos locais, e apontam para a importância da variável espaço sobre a explicação do desempenho dos partidos no processo eleitoral.

PALAVRAS-CHAVE: Organização Partidária; Eleições; Inteligência Geográfica; Estatística Espacial; IPVS.

Submission: December 28, 2017
Approval: May 30, 2018

***Artur Junqueira Lascala**
Master in Public Policy and Management from the Getulio Vargas Foundation. Public Policy and Management Analyst at São Paulo City Hall.
Address: Viaduto do Chá, 15, 01020-900, Centro, São Paulo, SP, Brazil.
E-mail: arturlascala@gmail.com

Bruno Mitio Assano Silva
Master in Public Administration and Government from the Getulio Vargas Foundation. Assistant Professor at FECAP Economics Department.
E-mail: brunomitio@gmail.com

Eduardo de Rezende Francisco
PhD in Business Administration from the Getulio Vargas Foundation. Associate Professor at Public and Business Administration graduate courses at FGV-EAESP and ESPM. Founding partner at *Meia Bandeirada*, a car sharing management and urban mobility system and founding partner at *GisBI*, a Big Data and Geointelligence think tank.
E-mail: eduardo.francisco@fgv.br

1 INTRODUCTION

Among some authors of the contemporary literature such as Guarnieri (2009) and Sousa Braga, Rodrigues-Silveira and Borges (2012), the election for the Legislative in Brazil is, in general, a game at the local level. The way in which the institutional bases are articulated locally is fundamental to understand how the "power game" is distributed among the actors and how they strategically organize to compete for territory in the electoral space. And, in this scenario, how does the figure of the political party fit in?

Guarnieri (2009) observes, from the analysis of the statutes, that the political parties organize themselves through their affiliated bases. That is, zonal and municipal conventions choose the regional conventions, which choose the national convention. Each convention determines a directory that defines the regular direction of the party, when it does not meet the convention, and arbitrates on lists of candidacies and coalitions in their respective spheres (municipal, state and national). In this sense, the party organization would be given "from the bottom up."

However, as Guarnieri (2009) observes, Sousa Braga et al. (2012), Sousa Braga and Pimentel (2013) and Ribeiro (2013), current Brazilian legislation and party statutes have allowed a mechanism to dissolve local commissions and directories for the establishment of a provisional committee, with an average duration of around 90 days (depending on the status of each party), named from a national directory or commission. That is, in this second scenario, the control is done "top-down". The leaders, who make up the national commission, determine the provisional commissions that, for example, select the candidates in each sphere.

Notwithstanding, as the studies mentioned above point out, especially in Guarnieri (2009), provisional commissions are not being used exclusively or homogeneously between the parties. This, in turn, suggests different strategies of party organization, which assign to the directorates or provisional commissions, relevant decisions, such as the formation of the list of candidates or the use of funds from the party fund.

Some results pointed out by Sousa Braga and Pimentel (2013) show higher incidence of directories in parties with greater representation in the Chamber of Deputies. Why? According to the authors, there is a positive correlation between the concentration of directories and electoral success for each party, since these types of organizations would be locally more institutionalized and with greater capacity to mobilize the local electorate. However, such correlation has been stronger in parties with greater representation. This relationship can help to understand the result found by Sousa Braga et al. (2012), which point to a difference in the organizational evolution of the "larger" and "minor" parties - in relation to the number of affiliates - in the post-Constitution period of 1988. "Larger" parties tend to disperse their affiliated base (territorial stretch process) to soon after locally densifying its base ("massification" process). While the "smaller" parties have accomplished the two stages almost simultaneously.

Another relevant variable seems to be the size of the municipality. Sousa Braga and Pimentel (2013) point to a positive correlation between municipalities with higher numbers of inhabitants and those with higher incidence of directories. In fact, Sousa Braga et al. (2012) identify three regions with the highest concentration of directories in Brazil, in the following regions: Rio Grande do Sul, São Paulo, Minas Gerais and Espirito Santo, and in the Paraíba and Alagoas axis. A preliminary hypothesis that Sousa Braga and Pimentel (2013) indicate is that the size of the party (in terms of representation in the Chamber of Deputies and affiliated basis) and the size of the municipality (in number of inhabitants) are relevant variables and that determine different types of strategies. In other words, the party organization in terms of directories seems to have much more effect on electoral success for large parties in larger cities.

2 OBJECTIVES

The recent literature points to different forms of organization as an indication of different electoral strategies and power structure between parties. This view seems to weaken the classical view of purely "weak" parties and sets the precedent for some unanswered questions. Among them, how does geographic space, in its idiosyncrasies, influence the behavior of political parties? If the local component has its importance, why has the geographical evaluation not yet been effectively analyzed in the literature?

In the analysis described in this article, we include a variable that has not been verified in depth in the recent literature on party dynamics: geographic space. That is, not only to include the variable in the debate, but also to evaluate it critically in statistical models. In fact, as Francisco (2010) points out, this bias is not an exclusivity for party analysis. A "more traditional approach" to the evaluation of social, economic and environmental data in general uses useful and valid analytical techniques such as hypothesis testing and variance analysis. However, they precisely exclude the local aspects of these phenomena. One of the premises of this article is that geographical aspects - the "where" political facts occur - matter and are relevant to their understanding.

Thus, this work looks for a pattern of distribution of municipal party organization in space. The hypothesis is that, in municipalities where the party organization is in the form that the literature considers more "institutionalized", that is, with a directory installed, it is able to mobilize a higher level of local electorate. Thus, the objective is to analyze the explanatory power of electoral performance from different models compared: (i) classical regression, which does not consider the geographic variable, and models that include the geographic variable, (ii) Global Approach Spatial Auto-Regressive model (SAR), and (iii) in a local approach, that differentiates the parameters by means of the characteristics of each region - Geographically Weighted Regression (GWR) or Geographically Weighted Regression clusters.

With great potential of adoption for studies of phenomena linked to public policies, the effective uses of Spatial Statistics techniques in this context are still incipient. More recently, Morandi et al. (2016), we used spatial autocorrelation and spatial regression techniques to evaluate the relationship between public transport on rails, income and average travel time.

If the recent debate is correct, it is expected that the municipal party organization instituted through directories will perform better when compared to municipalities and parties of similar trajectory or structure. Moreover, it is expected that this relationship (organization x electoral performance) has different effects between the different types of party and municipalities. Finally, the aim is to analyze the electoral outcome from the perspective of a possible association with another social index, the Paulista Social Vulnerability Index (IPVS), in order to assess the potential relationship of this variable to the electoral result.

3 LITERATURE REVIEW

As a way of including the geographic aspect in social studies, Mittal, Kamakura and Govind (2004) evaluate which reasons explain the satisfaction (dependent variable) of the customers of an automotive factory, based on the seller's experience variables, vehicle quality and service. However, the great differential was to consider the geographical aspect in these models. The authors considered that such aspects as seller, vehicle and dealership did not have the same effect on customer satisfaction across all US regions. This conclusion can be made by constructing the Moran's I index, which verifies the autocorrelation of the variables in space. The presence of autocorrelation found by the authors means that the variables are not random in space. In other words, space is an important variable in the representation and understanding of customer satisfaction with the automotive sector. From this information, they used the GWR model, appropriate for heterogeneous and non-stationary spaces, and performed local regressions to estimate the satisfaction of each North American district, considering their respective local samples and the independent variables of the neighbors. The authors obtained a very efficient predictive model, obtained a good solution to support the hypothesis of error

exogeneity, taking geographic information as relevant information. The result was a fairly efficient predictive model.

Francisco, Fagundes, Ponchio and Zambaldi (2010) conceived a model with some predictive power on the propensity of frauds on the use of electric power by AES Eletropaulo, considering socioeconomic variables (such as the Paulista Index of Social Vulnerability) and variables related to the process of operation and inspection of fraud by the company. The spatial model used in this case was the SAR that includes the dependent variable itself as an autoregressive term (spatially lagged) in the explanation of the phenomenon, in addition to other independent variables.

Unlike the least squares (OLS) linear regression model, where only the variables within the geographic selection to which the analysis unit belongs are considered, the GWR and SAR models include the effect of the independent (or dependent) variables of their neighbors on the dependent variable analyzed, through the geographic space.

The mere "running the model", which has already aroused interests as well as excesses in the past, is only a small detail in the evaluation of the approaches that are proposed empirically. In order for the calculated estimator to represent the isolated effect, that is, in seeking to simulate the ceteris paribus effect, so common in theory, it must be ensured that certain assumptions are guaranteed. Otherwise, there is only one indicator of descriptive statistics and not a scientific explanation of cause, nor a statistically reliable predictor.

4 METHODOLOGY

In order to achieve the objectives, the relationship between the existence of the party organization in the municipality and the electoral performance was evaluated, that is, how many votes the party obtained in the same municipality when compared to the total votes obtained in the whole federation (the "share" of party votes in the municipality). For this, the year 2010 was used as the period of analysis, due to the advent of the elections to the Chamber of Deputies.

The bases of this analysis are:

- Scope: municipalities of the state of São Paulo;
- Elections of the year 2010 for the position of federal deputy;
- Party base: parties PT and PSDB;
- Set of organization models: municipal directory, municipal provisional committee, no organization;
- IPVS of 2010.

The selected territorial scope was the state of São Paulo, with the municipalities being the units of analysis. Only this state was chosen as a preliminary form of inquiry about the hypotheses present in the literature. The project included data on party organization available on the website of the Superior Electoral Court (TSE) (2010). For the result of the elections, the aggregated electoral results by municipality were used on the website of the Public Sector Policy and Economy Center of the Getulio Vargas Foundation (Cepesp / FGV) (2010), initially for the 2010 elections.

The choice of the position of Federal Deputy was due to the fact that, for this specific situation, party coalitions have no real influence, as it happens in positions of majority (executive and senate), and also because, by being national level, some unmeasurable regional influences are avoided. The adoption of metrics only in relation to PT and PSDB stems from the fact that, in these particular elections, and even in general, they are considered leading parties, able to promote a polarization of interests around them. Thus, it is believed to be better capturing a heterogeneity of preferences among voters. Thus, it is understood that results of georeferenced analysis involving only these two parties have the ability to present results capable of mirroring a good part of the electoral reality.

In order to carry out such evaluations, a classical linear model (OLS – Ordinary Least Squares) was used. Two new models were added with the intention of including the spatial aspect in the explanation of the phenomenon. The models are the SAR, which has global explanatory power, from

a neighborhood matrix, and the GWR, which performs a local regression model for each municipality. The two spatial methods can be seen in detail in Francisco (2010). In addition to the spatial aspect, IPVS was included, a highly reliable index available in public databases, able to configure the average profile of the inhabitants - and voters - in each city of São Paulo, making possible this important analysis of the potential influence of the social stratum in that the voter finds himself, in the electoral result. The IPVS of each municipality was the result of the IPVS average of each district, weighted by the number of households.

In summary, the analysis consists of four steps, as detailed below:

1) Classical Model (OLS)

$$y = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \varepsilon$$

Where:

y: is the share of votes;
d_1: is the dummy variable that represents the presence or absence of directories;
d_2: is the dummy variable that represents the presence or not of provisional commissions.

2) Spatial Model (SAR)

$$y = \beta_0 + \beta_1 d1 + \beta_2 d2 + \rho W y + \varepsilon$$

Where:
y: is the share of votes;
d_1: is the dummy variable that represents the presence or absence of directories;
d_2: is the dummy variable that represents the presence or not of provisional commissions.
W: is the neighborhood matrix (contiguity: queen case), which considers the influence (weight) of y of each of the other (neighbor) districts on the model.

3) Spatial Model (SAR) with the inclusion of the IPVS variable

$$y = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \rho W y + \beta_3 IPVS + \varepsilon$$

Where:
y: is the share of votes;
d_1: is the dummy variable that represents the presence or absence of directories;
d_2: is the dummy variable that represents the presence or not of provisional commissions.
W: is the neighborhood matrix (contiguity: queen case), which considers the influence (weight) of y of each of the other (neighbor) districts on the model.
$IPVS$: is the IPVS average of all districts in the municipality.

4) Spatial Model (GWR) with the inclusion of IPVS

$$y_i = \beta_{0i} + X(g_i)\beta(g_i) + \varepsilon$$

Where:
y_1: is the share of votes calculated for all municipalities;

$X(g_i)$: is the vector of independent variables (dummy of the presence of directories, dummy of the presence of provisional commissions, and IPVS) of the neighboring municipalities, for each point g_i (centroid of the district) of the state of São Paulo.

The technique used considered the creation of points based on the centroids of the municipalities to perform the GWR. At the end of the research, it was intended to evaluate: (i) the impact of the party organization (which may be a directory, provisional committee or no organization) on the number of votes that the PT and PSDB parties received in each municipality of São Paulo; and (ii) the influence of the average social condition of the electorate of each city of São Paulo, as measured by the IPVS, in the total votes that the aforementioned parties received in 2010.

5 EXPLORATORY DATA ANALYSIS

Through the use of statistical and geographic software (ArcView GIS, ArcGIS, QGIS, GeoDA and R), some maps have been produced to suggest that there is a correlation, in geographic terms, between the type of party organization and the electoral performance. First, a mapping of the types of party political organization was carried out for the two main parties in the state of São Paulo, the PT and the PSDB. Also, the mapping of each party's share in the result of the municipal election to federal deputy. The results are shown in Figures 1 to 4.

Figures 1 and 2 show the party organization and PSDB electoral performance (in this case, the percentage of valid votes). There are some areas where the concentration of the share of votes is greater than in others. That is, apparently the concentration of votes does not take place homogeneously or randomly in space. Figures 3 and 4, on the other hand, present PT's party organization and electoral performance (in this case, the percentage of valid votes).

There is a clear predominance of directories in the two parties observed, the PT party with 462 and the PSDB party with 442.Both are therefore well-organized parties in the state of São Paulo, with a small number of municipalities with commissions (the PT party with 154 and the PSDB party with 146) or no municipal organization (PT, 29 and PSDB, 57). A preliminary analysis of the data seems to indicate higher levels in relation to the share of votes for municipalities with constituted directories, and not with mere commissions.

It is noted that the areas where the PT party dominates, in general, are not areas where the PSDB party is more important in terms of proportion of the distribution of their votes. These are the expected results.

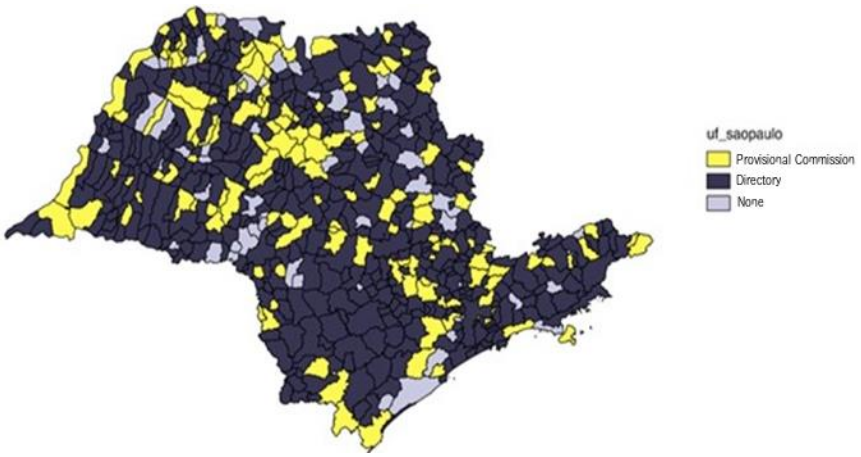


Figure 1 – PSDB party organization in 2010
Source: The authors, using QGis software.

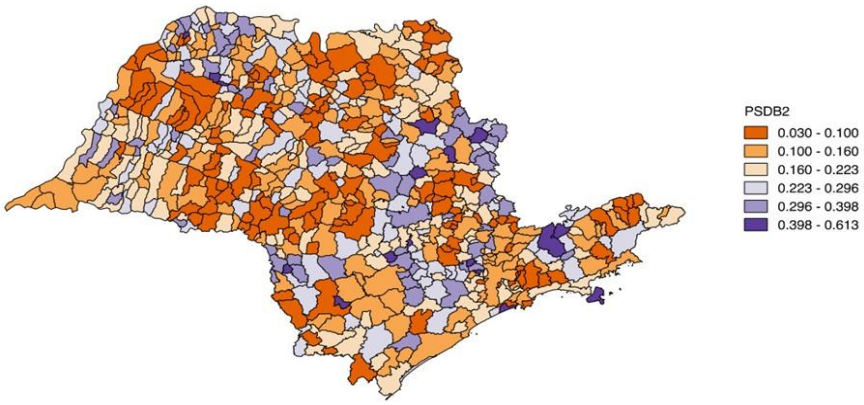


Figure 2 – Electoral Performance (percentage of valid votes) from PSDB to Federal Deputy in 2010
Source: The authors, using QGis software.

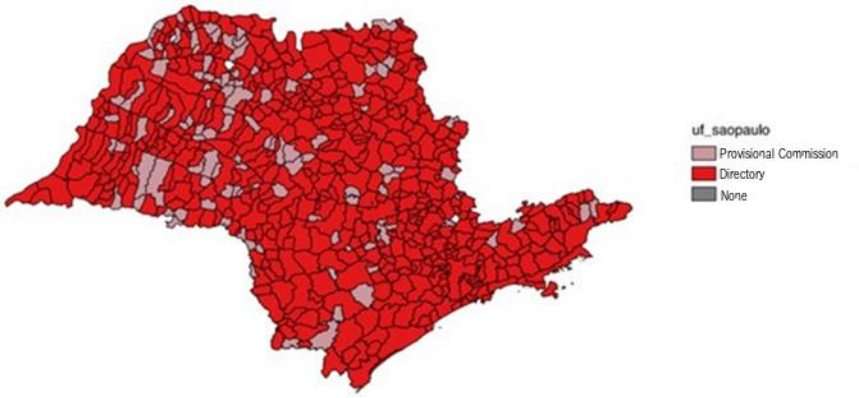


Figure 3 – PT party organization in 2010
Source: The authors, using QGis software.

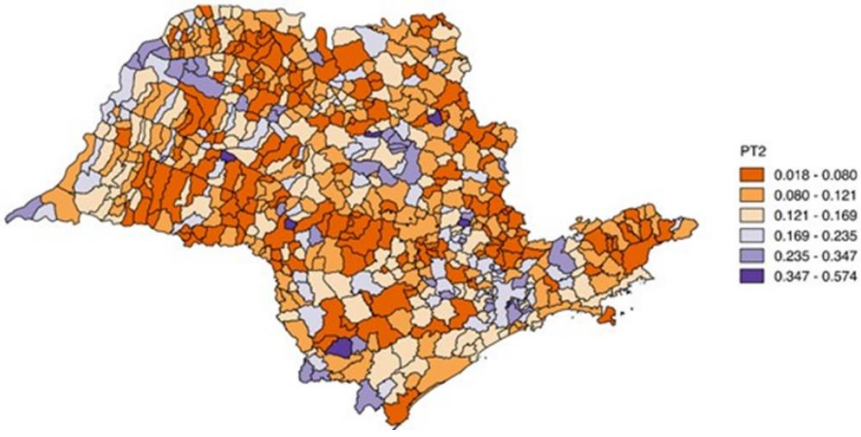


Figure 4 – Electoral Performance (percentage of valid votes) from PT to Federal Deputy in 2010
Source: The authors, using QGis software.

6 SPATIAL AUTOCORRELATION OF THE VOTES AND INCLUSION OF THE SPATIAL VARIABLE IN THE MODELS

It is not possible, however, to proceed with the analysis of the models without first making reference to the spatial distribution of the dependent variable. Francisco (2010) points out that, as a rule, spatial analysis addresses two classes of data: environmental and socioeconomic. Environmental data, basically resulting from surveys of resources and phenomena of nature, due to their

characteristic of stationarity and perenniality are treated in the scope of Geostatistics. This is not the case.

Socioeconomic data, which may include party organization, elections and their results - are much more complex to interpret: they come from censuses and samples of heterogeneous individuals and collectivities that are constantly changing. For this reason, they must have distinct, more accurate treatment under the techniques of Spatial Statistics. In this line, Francisco (2010) states that socioeconomic data can have two approaches: (i) as a set of homogeneous, disjoint and adjacent polygons, covering the entire study area, with descriptive attributes for each region, and (ii) as a set of samples (where each sample is associated with a [centroid, seat or arbitrary] point of the survey unit), revealing the complexity of its treatment and analysis.

In this sense, the fundamental difference between Geostatistics and Spatial Statistics concerns the concept of spatial stationarity. A stationary random process is similar to the analyzes that involve the study of the distribution of varable in time (time series), identified by having statistical properties (mean, variance, distribution, correlations, etc.) that do not vary in space. Thus, the Geostatistics approach represents a set of techniques to construct a surface based on a characterization of the spatial similarity between the samples, since it studies continuous and stationary space phenomena in space. Spatial Statistics considers the heterogeneity of the distribution of the variable in space, and is therefore more appropriate for most social and socioeconomic phenomena (Francisco, 2010).

In order to understand how the votes are influenced in each place, we must observe how they are spatially distributed, in order to detect the possible correlations of this variable with itself, the correlation between a value of an observation and the value of the observations close to this same variable and also the degree to which the characteristics of a location are similar - or distinct - to those in the vicinity. Figures 5 and 6 present some elements of the spatial autocorrelation of the share of votes of each party and its statistical significance.

It is possible to verify a Moran's I index of low to moderate electoral performance for both parties (25.1% for PT and 23.8% for PSDB). However, there are some clusters where the voting pattern seems to be more common. This differentiated behavior for each cluster found may explain the low overall Moran's I. This also seems to be an indication that local spatial regression models will perform better.

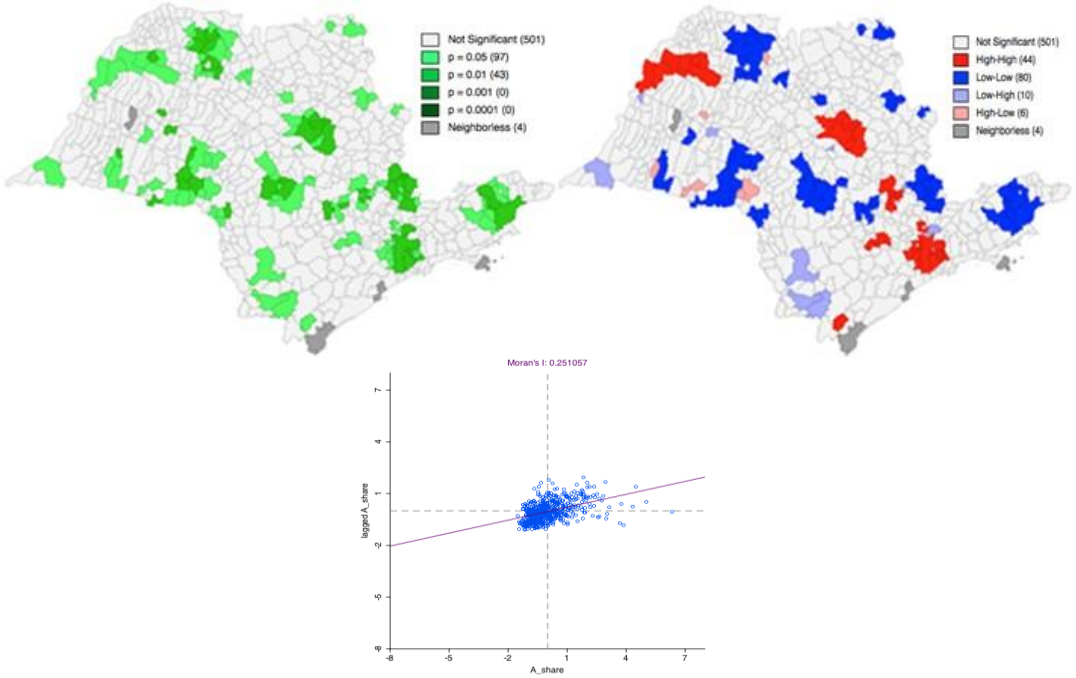


Figure 5 – Autocorrelation of Share of Votes to PT's Federal Deputy in 2010: Significance (left), Category (right) and I of Moran (center)
Source: The authors, using the GeoDA software.

Some considerations, however, merit reference on this occasion. Why it is necessary to adopt autocorrelation in these analyzes? We should return a little and refer to the classical linear least squares (OLS) method, without doubt, the statistical method of analysis most used in social sciences. This method constructs a relation (initially theoretical) between the dependent variable (which is meant to explain) and the explanatory variable (exogenous to the model and, as the presupposition of rationality, is not initially intended to explain).

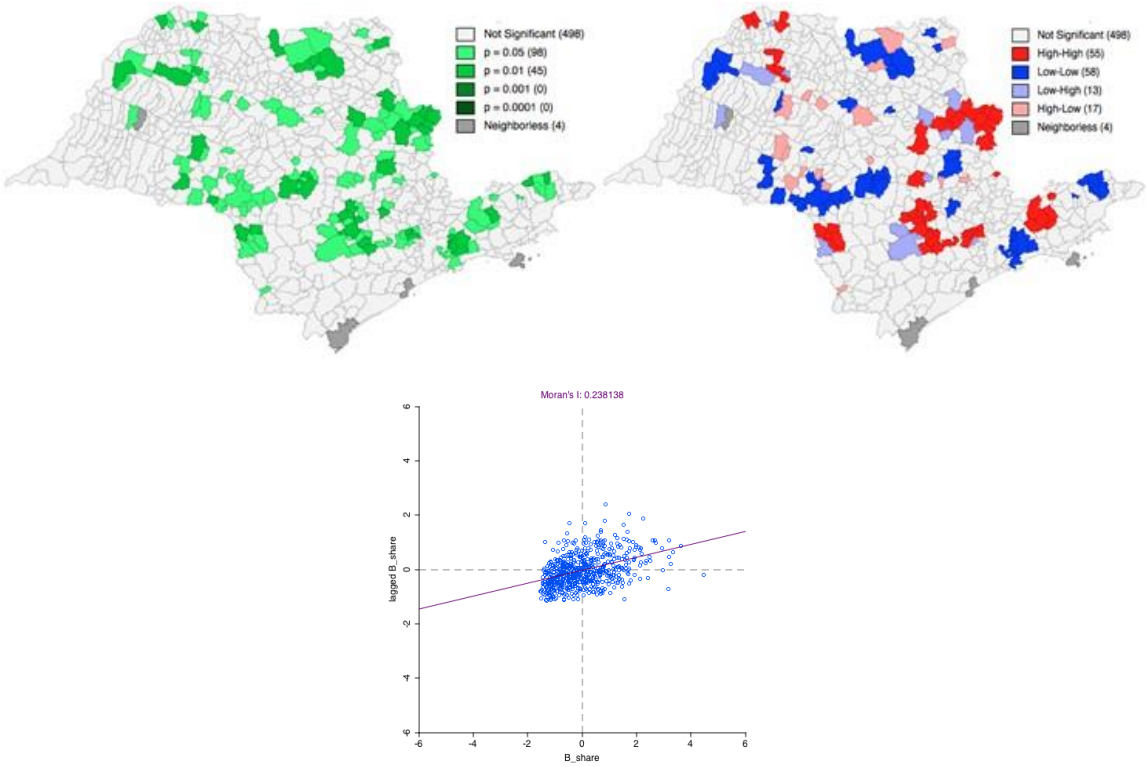


Figure 6 – Autocorrelation of Share of Votes to PSDB Federal Deputy in 2010: Significance (left), Category (right) and Moran's I (center)
Source: Authors, using the GeoDA software.

It should be noted here that the OLS method seeks to define a curve between all observed points of the independent variable that relates to all observed points of the dependent variable. This curve is constructed from the "residuals minimization", that is, the minimum distance from it with all observed points of the dependent variable. However, the simple calculation of the curve does not necessarily lead to an explanatory and predictive model of the dependent variable. The mere calculation leads to a correlation between variables, information that may have considerable relevance in this analysis, but does not necessarily lead to the main objectives referred to at the beginning of this text. It is at this point to cite the most valuable teaching of econometrics in the form of jargon: "correlation does not imply causality", since isolating causality has been one of the great challenges of the Social Sciences.

Returning to the OLS model, what guarantees the causal relationship between variables is not mathematics but rather the guarantee of certain theoretical assumptions that guarantee the statistical validity of the model. The most famous and important theoretical basis are the assumptions of Gauss-Markov. Some of these assumptions guarantee only the structuring of the model, such as linearity in the parameters (and not in the variables, which gives the freedom to construct more precise models) and randomness in the sampling. When more than one independent variable is observed, it is necessary to ensure that the variables are also not perfectly related (multicollinear), so that the dependent variable understands only the effect of a variable, but it is also necessary that the explanatory variable "vary" that is, has no variance equal to zero.

The assumptions mentioned are not entirely problematic nor have they been the major challenges in social studies. However, other assumptions are more problematic. Perhaps the most problematic is that it requires that the dependent variable is not correlated with the model errors. This means that if there is a correlation between error and explanatory, it means that the model has some kind of bias, that is, some other variable affects the variable, but it was not considered in the model. The way to try to "fit" the model, is to consider this variable. That is, to transform it into an independent variable to remove its effect from the error term. This is what is called the "control variable". The problem with this is that, in general, it is very difficult to control all effects on the dependent variable, that is, to have all the necessary control variables.

When working with sampling, that is, only part of the data is collected in a random way, in such a way that this set represents the entire population of data, one has to assume two more assumptions about the inference of the estimators (that is, if the estimated parameters even represent the population). The first one is that the sample set has a normal distribution. As for the Central Limit Theorem, it is generally assumed that, as the data set increases, the distribution of data tends to the normal distribution, this hypothesis is usually solved by the quantity of the sample. Another, more complex, assumption is that the data have a constant dispersion (homoskedasticity). Some more complex tests verify if the model has constant variance; and some (also complex) models adapt models with non-constant variance (heteroskedastic).

Consider now a rather likely hypothesis that space will affect the dependent variable. What would happen to the OLS model? Even assuring all other assumptions, there would be the bias of the uncontrolled variable (the space). It is, therefore, the spatial autocorrelation that evaluates and analyzes the degree of dependence between observations in space. In general, spatial dependence relations and their effects are estimated from neighborhood matrices of the areas studied on a map. Several criteria are used to define neighborhood matrices. One way is to determine the matrix based on the boundaries, by means of a dichotomous dummy that would represent the presence or absence of neighborhood between the localities. The Moran I index is one of the most used indicators to approximate the spatial autocorrelation measure calculated from the neighborhood matrix. In this case, it indicates the correlation between the value of the share of votes of a variable observed in a region and the values of this same variable observed in its neighboring regions (Francisco, 2010).

Given the presence of spatial autocorrelation, as Figures 5 and 6 seem to indicate, it is possible to include two models. The first form is the Spatial Auto-Regressive Model (SAR), which uses spatial neighborhood matrix incorporation as part of the independent variables of the regression model. Finally, another technique adopted is the Geographically Weighted Regression (GWR), which describes "a family of regression models in which the coefficients, β parameters, can vary spatially." The model estimates a regression for each observed point, pondering the other observations as a function of the distance of this point. In this study, the centroid of each municipality was considered. The result of the model is the estimation of different beta parameters for each municipality, with different values (Francisco, 2010).

7 RESULTS OF REGRESSION MODELS

The results of the models (i) OLS, (ii) SAR and (iii) GWR, specified in the methodology of this article, are listed in the tables below. Tables 1 and 2 show some significance of the presence of directories in higher votes gain in the municipality. However, the R-Squared of these models demonstrates little power of explanation between the two variations. In the models referring to Tables 3 and 4, the neighborhood matrix was included in the previous models. The gain in explanatory power is quite expressive. The R-Squared increases from 2.67% to 19.14% in the case of the PT and from 3.05% to 15.30% in the case of the PSDB. In the latter two cases, the presence of a directory installed in the municipality remains statistically significant in relation to the presence of a higher level of the share of votes.

Table 1 – Classic Model Estimators (i) – PT

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,1081699	0,005159	20,96	0,00	***
β_1	0,02633798	0,006096	4,32	0,00	***
Adjusted R-Squared:	0,028210				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

Table 2 – Classic Model Estimators (i) – PSDB

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,1559693	0,006710	23,24	0,00	***
β_1	0,0373700	0,008106	4,60	0,00	***
Adjusted R-Squared:	0,030486				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

An important observation is that, by performing the Breusch-Pagan test, there is a P-value (observed significance) of 0.19 for the PT party (Table 3) and a P-Value of 0.12 for the case of the PSDB party (Table 4), it is not possible to reject the hypothesis of homoscedasticity for both models. Therefore, for the two cases, the estimators are robust as to the variance of the error. The same, however, cannot be said for the results of Table 5, when the IPVS variable is included in the SAR models.

Table 3 – SAR Model Estimators (ii) – PT

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,0516503	0,007513	6,87	0,00	***
β_1	0,0373700	0,005547	3,87	0,00	***
ρ	0,4829759	0,046009	10,49	0,00	***
R-Squared:	0,191395				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

Table 4 – SAR Model Estimators (ii) – PSDB

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,0516503	0,007513	6,87	0,00	***
β_1	0,0215348	0,005547	3,87	0,00	***
ρ	0,4829759	0,046009	10,49	0,00	***
R-Squared:	0,153016				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

Table 5 – SAR Model Estimators with IPVS (iii) – PT

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,0185795	0,01051	7,95	0,00	***
β_1	0,0192943	0,005525	4,36	0,00	***
ρ	0,4637989	0,046162	8,31	0,00	***
β_3	0,0135470	0,003297	4,10	0,00	***
R-Squared:	0,209177				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

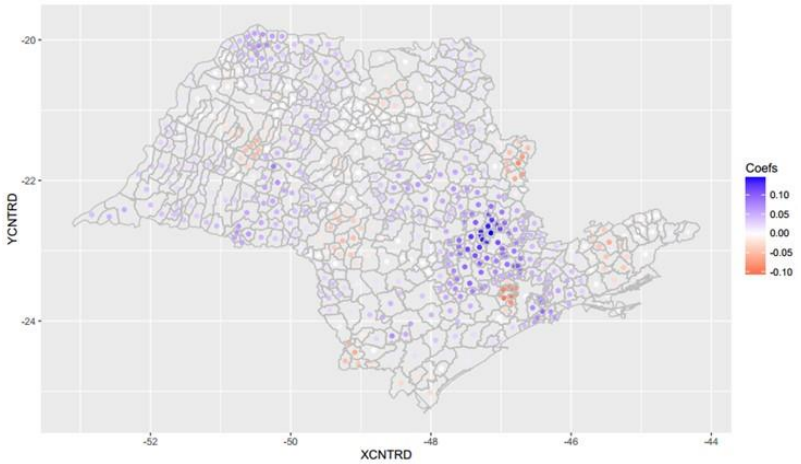
Table 6 – SAR Model Estimators with IPVS (iii) – PSDB

Variables	Estimator	Standard Error	Stat. Z	P-Value	
β_0	0,0993961	0,016456	6,03	0,00	***
β_1	0,0335725	0,007583	4,42	0,00	***
ρ	0,4071884	0,0492275	8,27	0,00	***
R-Squared:	0,154182				

*** Significant at 1% error

(No estimator is statistically significant at 10% error)

An interesting aspect in Tables 5 and 6 is that the variable is not statistically significant to explain the behavior of the PSDB party (for this reason it is expunged from Table 6), but is significant for the PT party (Table 5). However, it is not possible to assure robustness for this result, since the hypothesis of homoscedasticity for the PT case cannot be maintained. Therefore, it is not possible to conclude a difference in the importance of the variable between the two parties. Figures 7 and 8 highlight some results of the application of GWR models.



R^2 = 40,36%

Figure 7 – Spatial distribution of the "dummy presence of directories" parameter of the GWR regression of electoral performance for Federal Deputy in 2010 for PT
Source: Authors, using the R software.

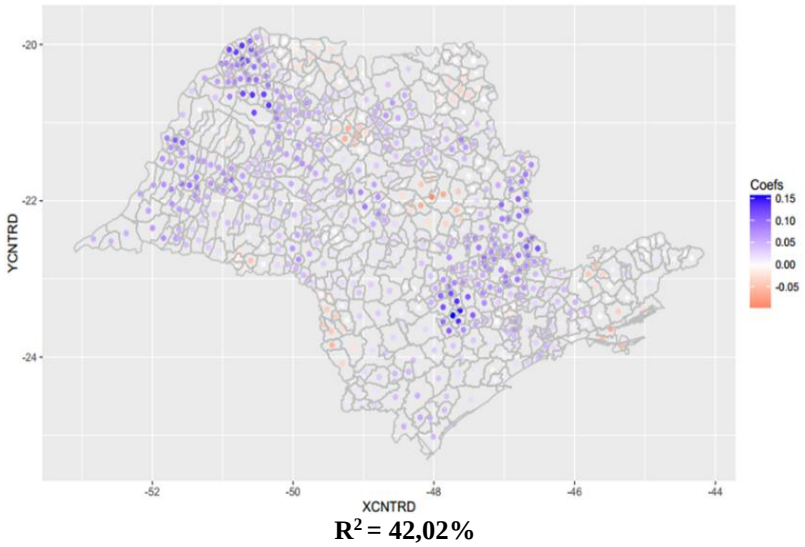


Figure 8 – Spatial distribution of the "dummy presence of directories" parameter of the GWR regression of electoral performance for Federal Deputy in 2010 for the PSDB
Source: Authors, using the R software.

The results were particularly relevant for the regression using the GWR technique, using software R *3.1.3* and the *spgwr* package. In the two maps that were created using this methodology, one can perceive the formation of clusters in different regions for each of the parties. The GWR models elevate the R-Squares for each model to a level of 0.40 (PT) and 0.42 (PSDB), more than double when compared to the SAR model in the first two cases. The result seems to reflect what has already been observed in the autocorrelation analyzes, according to Figures 5 and 6: there is no global voting standard where more directories necessarily answer the most votes. What exists, in fact, are clusters where the area of influence of the directories is larger, and some areas where, surprisingly, in some regions, the presence of directories leads to lower electoral performance.

8 FINAL CONSIDERATIONS

In order for the calculated estimator to really represent the isolated effect of party performance, that is, if it is even necessary to simulate the party organization's ceteris paribus effect on party performance in elections, it must be ensured that certain assumptions are guaranteed. Otherwise, there will be only one indicator of descriptive statistics and not a statistically reliable predictor. Among the assumptions required, perhaps the most complicated is to ensure that the model has no bias over its variables. That is, all the effect captured by the parameter is given only by its independent variables. This problem has had repercussion on recent studies that, instead of trying to control all the independents possibly related to the model, which also brings an undesired effect to increase the variance of the estimator, try to work the design to which the research is related

One way of attempting to "control" the omitted variable bias and demonstrating quite efficiency, as has been attempted here, has been to consider the spatial aspect in the models. Such inclusion is only possible with the technological development observed, since the end of the last century, in informational instruments. The relevance of spatial information goes beyond the spatial information itself, which in most analyzes is omitted because it can bring almost intangible information, such as local "culture" or "consumer preferences". If such information is relevant in the explanations of social relations and correlated to the geographic locus, bringing the geographic variable minimizes the bias contained in the error term of these models, reducing uncertainty and thus improving the causal explanation and prediction of the models, objectives of the applied models. The results demonstrate a valid and still widely used alternative to improve the ability to interpret a complex reality.

When we reflect about the party organization, many aspects tend to be left out in this model. In other words, many are the biases of omission in the models. How do the parties adapt to get the votes

of their constituents? This "conquest of votes" or formation of constituents occurs, in general, at the local level. From the moment the spatial variable is included, part of the local aspects of each electorate is also included. This is the reason for improvement in these models.

This study aimed to include the geographic aspect in explaining the electoral performance of PT and PSDB parties, in conjunction with their distribution of party structures. It was found that the party organization organized as a directory can be a good predictor for electoral performance, but especially in some specific regions of the state of São Paulo. The party organizations known as Provisional Commissions were not observed as significant, even though there were not enough degrees of freedom for a more robust analysis for this variable. As a fundamental lesson, the potential that the spatial dimension confers to statistical analysis is glimpsed. It was based on a model with very low explanatory power, towards the GWR model, in which it was concluded that it is not a random phenomenon in space, but a fact that is provable, given the significant increase of the R-Squared and the configuration in clusters of the spatial distribution of the electoral performances of the two parties analyzed. It is concluded that the introduction of the spatial dimension greatly enhances the quality of models for explaining electoral performance.

REFERENCES

Centro de Política e Economia do Setor Público da Fundação Getulio Vargas. (2010). *Resultados eleitorais agregados por município*. Recuperado de http://cepesp.fgv.br/pt-br/nossos-bancos

Francisco, E. de R. (2010). *Indicadores de renda baseados em consumo de energia elétrica: Abordagens domiciliar e regional na perspectiva da estatística espacial* (Tese de Doutorado em Administração de Empresas). Escola de Administração de Empresas de São Paulo da Fundação Getulio Vargas, São Paulo, SP, Brasil.

Francisco, E. de R., Fagundes, E. B., Ponchio, M. C., & Zambaldi, F. (2010). Development of an indicator of propensity to energy commercial losses using geospatial statistical techniques and socio-economic data: the case of AES Eletropaulo. *RAM. Revista de Administração Mackenzie*, *11*(4), 178-197.

Guarnieri, F. H. E. (2009). *A força dos "partidos fracos"- Um estudo sobre a organização dos partidos brasileiros e seu impacto na coordenação eleitoral* (Tese de Doutorado em Ciência Política). Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, SP, Brasil.

Mittal, V., Kamakura, W. A., & Govind, R. (2004). Geographic patterns in customer service and satisfaction: An empirical investigation. *Journal of Marketing*, *68*(3), 48-62.

Morandi, E., Ribeiro, R., Hernandes, E., Camara, B., Spinola, L., & Francisco, E. de R. (2016). *Análise Geoespacial da relação entre transporte público sobre trilhos, renda e tempo médio de deslocamento.* XL Encontro da EnANPAD.

Ribeiro, P. F. (2013). Organização e poder nos partidos brasileiros: Uma análise dos estatutos. *Revista Brasileira de Ciência Política*, n. 10, 225–265.

Sousa Braga, M. do S., & Pimentel, J., Jr. (2013). Estrutura e organização partidária municipal nas eleições de 2012. *Cadernos ADENAUER, 2,* 13–36.

Sousa Braga, M. do S., Rodrigues-Silveira, R., & Borges, T. (2012). Organización, territorio y sistema partidario: Difusión territorial de la organización de los partidos y sus potenciales impactos sobre la estructura del sistema partidario en Brasil. *América Latina Hoy*, *62*(18).

Tribunal Superior Eleitoral. (2010). *Dados sobre organização partidária*. Recuperado de http://www.tse.jus.br/