Going Local to Win the Nation:
Political Boosting in Mayoral Politics

Tiago Ventura

May 17, 2019

Abstract

In federal democracies, political parties often invest in local politics to increase their gains nationally. This strategy is crucial particularly in fragmented, nosy democracies in which parties need to find informational shortcuts to attract inattentive voters. I define this strategy as political boosting and investigate its dynamics for the Brazilian case, a textbook example of a contemporary fragmented polity. I use regression discontinuity design to investigate the electoral benefits political parties retrieve in upper-level elections after winning local races. I show that electing the mayor significantly boosts the party's national performance in upcoming elections. Horizontal and vertical party coordination mediate the positive effects *vis-a-vis* career ambition explanations. Applying a Bayesian LASSO algorithm to deal with data sparsity in RD designs, I estimate the treatment effects by parties and additionally find that large parties gain more on boosting. By disentangling the effects of winning local in between elections in federal democracies, the paper explains the general logic by which parties build electoral strength in fragmented polities.

Number of words: 9730

Introduction

Parties often invest in local politics to win national elections. I define this as political boosting, borrowing from the statistical process with the same name. In statistics, the term refers to a family of algorithms which converts weak learners on strong predictors using marginal gains through repetition. In statistics, the critical problem of boosting is to gain precision when noisy data prevents researchers from measuring a specific event. Similarly in politics, fragmented political environments provide equally weak information for politicians battling for a critical commodity: vote. Political boosting is the process of augmenting the political signal of parties by aligning local political capital to national gains; parties use repeated interactions between mayors to increase the strength and precision of their political signals in order to boost their national performance. I argue winning multiple local elections increases the strength of political signals in between elections because parties use the local offices as an information shortcut for inattentive voters who need to decide their candidates for upper-level positions. In this paper, I provide a theory and a statistical model to understand how parties go local to win nationally. I investigate this question systematically through a study of the electoral effects of electing mayors in Brazil on the performance of their parties for upper-level legislative elections. In doing so, the theory outlines a possible mechanism by which parties build electoral strength in fragmented polities. I analyze all Brazilian mayoral polls between 2000 and 2012 using a regression discontinuity design (RDD) to identify the effect of political boosting.

I divide the analysis into three major parts. At first, I ask whether winning a local election increases the party vote share for the coming House election in the respective municipality. I find that winning the local executive, even by a small margin, increases the vote share of the incumbent's party at the coming legislative election for the House. On average, the vote share of the incumbent's party on the coming House Election is 2% greater compared to the vote share of the party who barely lost the local dispute.

I then move a step forward to investigate the mechanisms behind political boosting in Brazil. I theorize about which set of conditions makes the signs emitted by mayors more effective on subsequent upper-level elections. I argue that party coordination capacity makes boosting more effective. Parties with strong representation in the national legislative branch and aligned with upper-level executive chiefs can more faithfully sign to inattentive voters about their partisan candidates running for congress. As the literature has consistently shown, partisan alignment means more access to a variety of federal grants in Brazil, with evidence showing voters reward such behavior (Brollo and Nannicini, 2012; Bueno, 2017; Firpo et al., 2015). I compared this mechanism with the conditional effect of career incentives on political boosting which has been largely used to explain local politics in Brazil (Titiunik et al., 2015; Novaes, 2017).

In the third part, I discuss and empirically assess the presence of heterogeneous subgroup treatment effects by parties. I define effect heterogeneity as the degree to which the treatment has differential causal effects conditional on each subgroup. In an environment with high levels of party fragmentation, it is unlikely that distinct parties will be able to extract the same electoral advantages from political boosting. I focus on identifying which parties matter more and through which channels parties succeed in enforcing their local politicians to comply with their national goals. The results indicate that political boosting works for large parties, while the small parties should rely on a different repertoire of strategies to increase their national presence (Calvo et al., 2015).

To estimate the effect of parties, I implement a novel methodological strategy by applying a Bayesian estimator, developed to deal with data sparsity across subgroups, to identify heterogeneous subgroup effects in the context of Regression Discontinuity Designs Anastasopoulos (2018). I use the LASSOplus method developed by Ratkovic and Tingley (2017) to identify the subgroup effects of the more than 30 political parties in Brazil who elected mayors from 2000 to 2012. The model consists of a Bayesian sparse extension of the LASSO model, a machine learning algorithm which simultaneously performs regulariza-

tion and variable selection for regression models with abundant parameters (James et al., 2013; Tibshirani, 1996). The use of Bayesian LASSO deals with two crucial dilemmas on identifying heterogeneous treatment effects using RDDs: a) sparsity on the data around the window threshold of the forcing variable where continuity holds, b) ad-hoc decisions about which subgroups, or political parties, should be tested. The first may lead to violations of asymptotic assumptions of local linear regressors while the latter may lead to fiding statistically significant results when no true relationship exists by using performing subgroup analysis in an ad-hoc manner (Imai and Strauss, 2011). This implementation is the methodological contribution of this paper to the emerging literature combining machine learning methods and causal inference (Grimmer et al., 2017; Athey and Imbens, 2015; Imai and Ratkovic, 2013; Ratkovic and Tingley, 2017; Green and Kern, 2012; Hainmueller and Hazlett, 2014). The results indicate that larger parties have a positive conditional treatment effect extracting more from the mayors than small parties. When investigating the partisan effects over the career incentives and party coordination hypothesis, again I find larger parties have higher positive effects when factors of the latter hypothesis hold.

In the next section, I expand on the idea of political boosting and the formulation of the hypothesis later assessed in the paper. Then, I introduce the reader to the institutional environment in Brazil. The next section presents the research design employed in each of the three parts of the paper in which I dedicate further attention to the use of Bayesian LASSO to estimate subgroup effect on Regression Discontinuity Designs. Then, I discuss the results for political boosting on federal democracies, the mechanisms, and the conditional partisan effects. I end by examining the broader contributions of the paper

Political Boosting in Multilevel Democracies

In this section, I outline the argument about Political Boosting. Does winning local midterm races make parties more competitive in national elections? Research on multilevel democracies has paid con-

siderable attention to how elections on different scales might affect each other. Using the concept of coattails, much of the literature supports that horizontal forces from executive races shape the performance and levels of fragmentation of legislative parties (Golder, 2006; Magar, 2012; MEREDITH, 2013; Samuels, 2000a), vertical effects between multilevel executive positions with presidents affect gubernatorial races (Borges and Lloyd, 2016; Rodden and Wibbels, 2010), and reverse effects of local politics can impacting - or not - national electoral competition (Ames, 1994; Broockman, 2009). A consistent pattern on this scholarship relates to the conditionality of time; coattails effects are notably mitigated when elections are non-concurrent (Golder, 2006; Borges and Lloyd, 2016; Samuels, 2000a).

While our understanding of the dynamic of coattails effects on concurrent elections has been consistently expanded by the scholarship above, we still do not understand how non-concurrent elections on federal democracies affect each other. Little research has discussed the paths through which winning local elections affect partisan perspectives on the national level when elections do not occur together. This is precisely the theoretical and empirical gap that our theory of political boosting aims to fill in the literature on partisan politics in federal democracies.

Economies of scale, information shortcuts in noisy federal democracies, and differential access to political resources explain why political boosting runs from the local to upper-level elections in the context of this work. In noisy democracies, upper-level candidates have incentives to structure their campaigns relying on local intermediaries. Having the support of players who control the domestic patronage machinery and the provision of public services become paramount to reduce costs and increase the efficiency of candidates running for upper-level positions in federal democracies. Voters also recognize the importance of mayors with behavioral evidence suggesting that voters incorporate information coming from local dynamics when deciding about upper-level offices (Feierherd, 2018; Bueno, 2017; Lucardi, 2016).As a consequence, voters typically pay attention to their mayor and use her party affiliation as an information

shortcuts to help them decide whom to vote for in House elections. These elements justify the idea of political boosting and investing in the local in between elections to improve parties' national performance.

Winning control over the local executive gives the local party leader access to clientelistic (jobs and pork) and non-clientelistic goods (build a local label, fulfill citizens demands, among others) that can be employed to help her co-partisan running for upper-level legislative positions. In federal democracies with high decentralization, as in Brazil, mayors usually control both types of goods and voters interact mostly with social services provided by the local offices. Therefore, controlling local public offices becomes a vital tool in this fragmented, noisy environment. In the first hypothesis of this paper, I argue that winning local elections increases the strength of parties' signs towards inattentive voters about their candidates running for upper-level non-concurrent elections.

However, political boosting depends on mediators; intermediaries in this economy of scale might be disloyal (Novaes, 2017),therefore, our theory needs to incorporate positive incentives on mayors to make them more likely to support their national co-partisans. My second hypothesis discusses which set of conditions makes the signs emitted by mayors more effective. The first mediator I investigate relates to party coordination capacity on the district level affecting political boosting. Evidence for the Brazilian case suggests mayors' affiliation with the President's party guarantee higher access to upper-level executive grants (Brollo and Nannicini, 2012; Bueno, 2017), and allocation of legislative budget amendments (Luz and Dantas, 2017). Related research also indicates voters indeed pay attention to local politics when making decisions about upcoming non-concurrent elections which in some cases reward or punish upper-level candidates according to politics that are essentially local (Feierherd, 2018; Bueno, 2017; Firpo et al., 2015).

Under the aforementioned structure of incentives, parties with strong representation in the national legislative branch and aligned with upper-level executives have access to more resources, and traditionally

use this as a campaign resource when interacting with voters. I argue that these features increase the precision of local parties' signs to inattentive voters, therefore, affecting the credibility of their promises about their co-partisans' candidates running for upper-level congressional seats. Then, in my second hypothesis, I test if political boosting from winning the local executive vary according to the affiliation of the mayor with the State Governor's Party and if the incumbents' party holds at least one seat in the House when the coming Federal election occurs.

I also assess whether a mayor's career ambition matters for political boosting. I analyze heterogeneous effects of political boosting conditional on the decision of the incumbent to run for reelection, to be in an open-seat, or be a lame duck. Employing the same logic outlined on previous studies of incumbency effects in Brazil (Klašnja and Titiunik, 2017), I expect that mayors facing a reelection race in the short run have higher incentives to help her co-partisans; having a co-partisan elected for the House can help the mayor to access federal resources which might be crucial to her reelection goals, for example. When a Lame-Duck or in an Open-Seat, the mayors theoretically have fewer incentives to send signs about her co-partisans running for the House elections. Therefore, I compare the groups and discuss which mediator works for political boosting.

Finally, I expect political boosting to be partisan dependent. Here, the effects for the political parties can occur directly and indirectly. More established parties are more likely to have House representatives across districts in Brazil as well as being more competitive for state governors' elections, therefore indirectly affecting political boosting through the mediator above discussed. However, it is also reasonable to expect a direct effect of partisan strength on political boosting. Larger parties, even in noisy democracies, exhibit advantages regarding the strength of their labels, their capacity to coordinate local brokers, and their access to upper-level clientelistic goods. All these factors, in my argument, increase the strength and precision of the signs from mayors affiliated with larger parties about their co-partisans running for

upper-level elections [1].

Institutional Background

Brazil is administratively divided into 27 states, further subdivided into 5,570 municipalities. The political structure at the local level mirrors the presidential arrangement of the central government, except that states and cities have a unicameral legislature while the national level has an Upper and Lower level chamber. The chief of the executive in all three federal levels are directly elected using a runoff majority rule, although, for mayors of municipalities with under 200,000 eligible voters who represent the absolute majority of the towns, mayors are chosen through plurality rule. Brazil has used the same electoral rules for electing its legislative representatives since the democratization of the country, including the period here discussed. Legislators are selected in an open-list proportional system in which coalition-building occurs at extraordinary levels (Calvo et al., 2015; Amorim Neto et al., 2003; Figueiredo and Limongi, 2000). Municipal governments answers for a relevant share of the provision of public goods and services in Brazil, mainly those related to education, health, and infrastructure projects. However, the municipal budget is heavily dependent on mandatory and discretionary grants sent by the Federal government. Elections for all national and state level offices take place at the same time every four years, whereas all municipal elections occur by two years and also take place every four years.

The Brazilian party system is famous for exhibiting high levels of fragmentation. After the most recent 2018 national election, the largest party in the House, the Worker's Party, holds only slightly more than 10% of the seats whereas thirty parties won at least one seat for the House. Even though the Legislative Branch has approved some minor changes over time to reduce the partisan fragmentation, the trend on the number of effective parties is still upward in the last two decades. Based on the Laakso-Taagepera index,

[1] See Feierherd (2018) for a discussion about conditions in which strong parties can hurt their up-ticket presidential candidates

Brazil had in 1998 a total of 8.16 effective parties in the Lower Chamber using the vote share as the unit of analysis. In 2014, the last year under scrutiny here, the total number of effective parties almost doubled achieving the impressive mark of 14.1 effective parties. Party fragmentation is not limited to the national level. Not surprisingly, the local level of party fragmentation resembles the findings depicted above for the House composition. In 2000, the number of effective parties winning local executive offices was 7.38, while in 2012, the value jumped to 11.

Figure 1 1 adds to these descriptive information about the parties' vote share for the House elections aggregated in the municipal level. This measure is what I use as the dependent variable of the statistical models for political boosting. The left figure plots the evolution of the number of effective parties using the vote share of the parties for the Lower Chamber. The right plot shows the average vote share of the parties in each municipality [2]. Both pictures indicate how partisan fragmentation has increased in Brazil over the years; the first clearly shows a movement in the distribution towards an increase in the number of effective parties, whereas the latter indicates that higher dispersion on the average vote share of parties for the Lower Chamber across the municipalities.

An important remark about the party system in Brazil relates to the uniqueness of the Workers Party (PT) The scholarship on political parties in Latin America points out to the PT of a leftist, mass based party in the continent, also viewing the emergence of the PT and their electoral expansion as singular fact in modern Brazilian politics. Previous research suggests various peculiarities about the PT compared to other large parties in Brazil. For example, the strength of the PT's label (Feierherd, 2018; Samuels and Zucco, 2018), the PT's capacity to mobilize and activate local grassroots organizations (Samuels and Zucco, 2015; Keck, 1992), and the party's capacity to discipline its members in the national and local level (Klašnja and Titiunik, 2017; Amaral, 2013; Hunter, 2010). I empirically assess this uniqueness

[2]The Superior Electoral Court does not provide information of the parties who had zero votes for the years of 2002, 2006 and 2010; therefore, I exclude these cases to calculate the average

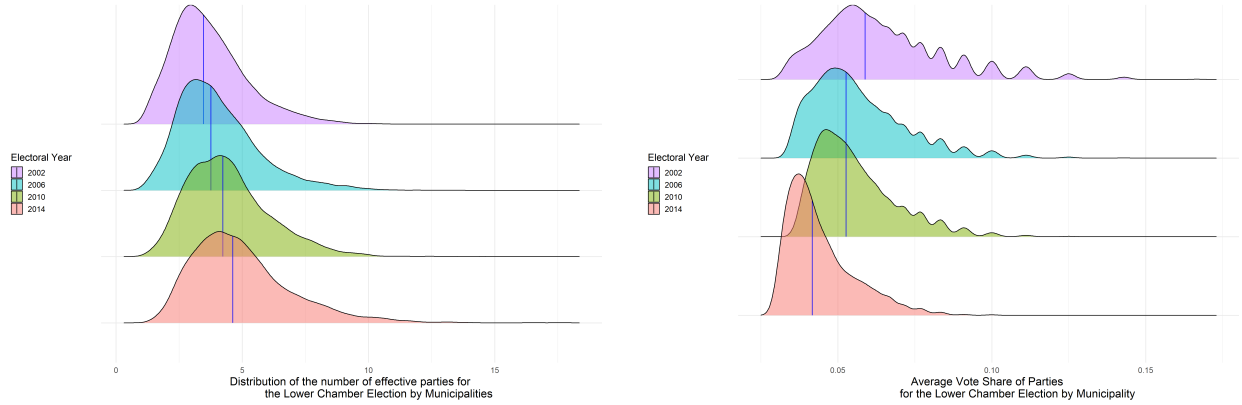


Figure 1: Distribution for the Effective Number of Parties and the average vote share for the House by municipalities. The vertical blue line indicates the median for each distribution. The mean over the years and the standard deviation for the number of Effective Parties are 4.29 and 1.74, respectively. Whereas, the mean and the standard deviation for the vote share are 5.5% and 1.1%Distribution of the Effective Number of Parties and the average vote share for the House by municipalities. The vertical blue line indicates the median for each distribution. The mean over the years and the standard deviation for the number of Effective Parties are 4.29 and 1.74, respectively. Whereas, the mean and the standard deviation for the vote share are 5.5% and 1.1%

about the PT in the analysis on subgroup partisan effects of political boosting in Brazil.

Brazil offers a unique opportunity for a thorough investigation of strategies of parties on fragmented polities. First, the Brazilian political system adopts institutions which produce few incentives for intra-partisan cooperation, as we have extensively discussed. Second, due to the non-coincidence of the local and national election, I can disentangle without further problems the action of local politicians on the national efforts of party construction, therefore filling a critical gap in the literature on coattails effects. Finally, Brazilian parties vary significantly in their strength; thus providing me enough variation to discuss the presence of heterogeneous effects across the political parties and allowing for a clear understanding of the factors that make the political boosting more effective. The next section discusses the research strategy developed in the paper.

Research Design

In this paper, I measure mayors' political boosting by assessing at which degree electing a local executive chief affects the party performance for the next Lower Chamber legislative election. I analyze all the elections between 2000 and 2014. The years of 2000, 2004, and 2008 and 2012 represent years with local elections (t), while 2002, 2006, 2010, and 2014 ($t + 2$) are years with state and national elections. Therefore, the analysis covers a total of four electoral cycles. The data is available from the Superior Electoral Court in Brazil. All the other variables related to the electoral dynamics are extracted from the same source.

The pressing empirical challenge for this paper relates to the chance of omitted or reverse bias affecting the party vote share for the House (y) and the election of the co-partisan mayor (T) simultaneously. In other words, in municipalities where the mayor for party j is elected, the level of support for the same party is likely to be higher; therefore, party j has from the beginning a distinct baseline in its likelihood of being better off on the upcoming legislative race in this particular district. Thus, positive results might be spurious due to this different baseline propensity instead of related to the effort of the mayor on delivering support for their co-partisans.

To overcome this empirical challenge, I use a regression discontinuity design. This design has become one of the most credible and accessible strategies in political science to deal with causal effects using observational data and have been extensively applied on close elections in Brazil (Avelino et al., 2012; Klašnja and Titiunik, 2017; Novaes, 2017; Boas and Hidalgo, 2011; Brollo and Nannicini, 2012), as well tested on many different countries (Lee, 2008; Eggers et al., 2015; Broockman, 2009).The designs make the assumption that the treatment effect is continuous at a particular cutoff allowing the researched to isolate possible cofounders affecting both the party's performance between the local election and the House election. Consequently the model provides a clear identification for political boosting measured as

the treatment effect of electing the mayor on the incumbents' party results for the upper-level legislative elections (De la Cuesta and Imai, 2016; Cattaneo et al., 2017).

In RD designs, the units have a score, usually called running variable in the literature, where groups below and above the score receive/not receive the treatment. The application of RDD on electoral studies for most of the cases uses the vote share of candidates as the running variable (Skovron and Titiunik, 2015).). In this paper, the running variable is the margin of victory between the winner and the runner up candidate for the local executive elections, and the unit of analysis is the municipality at a particular electoral cycle. I include the full sample of cities in the estimation for the four electoral cycles [3]. The paper's design obviously relies on a sharp discontinuity; in municipalities where the margin of victory is positive, parties receive the treatment (wins the election) while the opposite holds for parties who have a negative margin (lose the election).

A variety of methods have been proposed to estimate regression discontinuity designs. I follow the nowadays most recommended setting of using nonparametric local linear regression (LLR) to approximate the treatment effect at the cutoff point (Titiunik et al., 2015; Gelman and Imbens, 2018). I employ a local polynomial with degree one to fit two separate regression functions above and below the cutoff of zero in the margin of victory in the mayoral dispute, with the treatment effect being the difference in the limits of the cutoff, or in other words, the intercepts from each direction. To smooth the local regression function, I employ triangular kernel weights as a function to the distance between each observation's score and the cutoff. Therefore, observations far away from the cutoff are under weighted on the estimation strategy allowing for a better approximation of the treatment effect. I use a data-driven search to select an optimal bandwidth for the estimation which minimizes the Mean Square Error (MSE) of the model. To address bias on the treatment effects due to approximation errors, I report the robust

[3] For municipalities with more than 200 thousand voters, where a runoff between the top two candidates may occur, I use the data for the second round to build the discontinuity. I am grateful to Natalia Bueno, Leonardo Barone, and Lucas Novaes for making their code to estimate close races in Brazil publicly available.

confidence intervals developed by Calonico, Cattaneo, and Titiunik (2014). I report in appendix results using different bandwidths on the treatment effect for winning local elections.

For the basic setup of this paper, I estimate the effects of the party j barely wins (loses) the mayoral election at t on the vote share of the party j co-partisan candidates to the congressional race at t_{+2}. For this paper, as described above, I cover the last four electoral cycles in Brazil; hence t represents the local election, and t_{+2} represents the national elections that occur two years later. The dependent variable employed is the vote share of the two top candidates (the mayor and the runner up) from party j for the House election at t_{+2} in the municipality i. Therefore, the causal effect indicates the increase/decrease in the vote share of the incumbents' party at t_{+2} *vis-a-vis* the vote share of the runner up candidate's party.

The second part of the empirical strategy of this paper asks which factors increase the speed of political boosting. I argue two sets of explanations can condition political boosting: *incumbent party coordination* and *career ambition incentives*. To test these mechanisms, I measure whether the causal effect of controlling the local executive increases when: i) the state governor is from the same party of the barely elected mayor; ii) if the party of the incumbent has a House Representative running for reelection on t_{+2}. Table 1 describes these subgroups. I hypothesize partisan alignment and presence on the legislative district makes mayors signaling more effective to voters about their co-partisans.

The second mechanism argues career perspectives might push the mayors in the direction of supporting, or not, their co-partisan candidates for the House election. I expect that mayors elected at t, who run for reelection (t_{+4}) in the following local contest, have stronger incentives to help their co-partisans in their "midterm" legislative races (t_{+2}). In the future election, this candidate needs, even if in the case of a weak party, to have the support of her organization,therefore, the dominant strategy here for a candidate who faces reelection is to stay in good hands with her party leader. On the other side, lame-duck mayors and open-seat incumbents have weaker incentives to engage in this partisan effort. These factors

are theoretically derived from the model for the incumbency advantage proposed by Klasnja and Titiunik (2017). Table 2 describes the subsamples.

To estimate these models, I subset the data according to these variables. For the career centered incentives, I split the data into three subgroups: i) Incumbent, when mayors decided term limits could not run for reelection; iii) Open Seat, when mayor could but decided not to run for reelection. I follow the same strategy using the respective variables to test the party coordination hypothesis. In these subsamples, the sparsity of the data is not an issue since I have more than 40 thousand observations in the whole sample. This concern will only arise in the third section of this paper when I work with the effects by parties.

Table 1: Description of RDD for the Party Coordination Hypothesis

Governor Sample

- Treatment Group: Mayor who barely wins at t and is a member of the Governor's
- Control Group: Runner up candidate who barely loses at t for a mayor from the treatment group.
- Outcome: The party vote share for House at t_{+2} in municipality i for the treatment and control group.

Representative Elected Sample

- Treatment Group: Mayor who barely wins at t and her party has at least one elected representative for the House at the same district at t_{-2}
- Control Group: Runner-up candidate who barely loses at t for a mayor from the treatment group.
- Outcome: The party vote share for House at t_{+2} in municipality i for the treatment and control group.

Table 2: Description of RDD for the candidate centered Hypothesis

Reelection Sample

- Treatment Group: Mayor barely wins at t and runs for reelection at t_{+4}.
- Control Group: Runner up candidate who barely loses at t and the winner runs for reelection at t_{+4}
- Outcome: The party vote share for House at t_{+2} in municipality i for the treatment and control group.

Lame Duck Sample

- Treatment Group: Mayor barely wins the reelection at t and is forbidden to run again at t_{+4}.
- Control Group: Runner up candidate who barely loses at t for a Lame-Duck mayor
- Outcome: The party vote share for House t_{+2} in municipality i for the treatment and control group.

Open Seat Sample

- Treatment Group: Mayor barely wins the reelection at t and decides not to run for reelection at t_{+4}.
- Control Group: Runner up candidate who barely loses at t for a mayor who has decided not to run for reelection.
- Outcome: The party vote share for House t_{+2} in municipality i for the treatment and control group

Subgroup partisan treatment effects using Regression Discontinuity Designs.

The third part of the paper discusses how the treatment effects and its mechanisms vary across the family of political parties in Brazil. Previous analysis considering heterogeneous partisan effects using RDD on Brazilian electoral data have some noticeable shortcomings. Klašnja and Titiunik (2017)'s model argues that incumbency effects vary by party and to analyze this variation the authors handpick solely the five largest parties in the country. Boas et al. (2014)uses the same approach to identify the effects of winning a local election on accessing future government contracts. In this case, the authors are mainly motivated by the heterogeneous effects for the Workers Party (PT) which lead them to subsample the data for this party rendering at least one for their estimators with only 45 cases within the bandwidth.Novaes (2017) reports partisan effects only on the supplemental file, but not before stating in footnote 26: "calculating heterogeneous effects for individual parties is difficult because individual study groups are small."

While subgroup identification was paramount on these papers, particularly on the first two, their methodological solution is insufficient. First, the authors make an ad-hoc decision about which parties to include in their analysis which can generate false discoveries when no true relationship exists, notably under the condition of sparsity on the data (Pocock et al., 2002; Imai and Strauss, 2011). Second, every time the researcher discards part of the data, the estimate is inefficient; as researchers, the design is giving away variance for free. Finally, local linear models with such data sparsity as in the Boas et al. (2014) example are likely to renders instability in the asymptotic properties of the treatment effect, therefore, raising important suspicious about the findings.

A vivid range in statistics research has proposed ways to deal with such data limitation by applying

what has been called "sparse models" (Tibshirani et al. (2015) for an overview), yet these techniques remain out of the toolbox commonly employed in applied political science research. Sparse Modeling involves estimation and variable selection by zeroing out all but the most relevant of coefficients from hundreds or thousands of possible candidates. To identify the subgroup partisan effects, I combine sparse modeling with discontinuity design using a Bayesian variant of the LASSO estimator. This strategy allows me to include all possible subgroups in the estimation, it provides interpretable results, and it preserves the original variation of the data by using all the information in the sample. I next offer the readers with a general overview of the LASSO model and its Bayesian variant and discuss the modeling strategy in the context of a regression discontinuity.

Bayesian LASSO Model

Initially developed by Tibshirani (1996), the LASSO model focus on improving the predictive accuracy of OLS estimators and its interpretability when dealing with high-dimensional settings by zeroing out non relevant parameters. When the number of covariates k is large, researchers tend to rely on a "kitchen sink" approach adding all the variables to the model which will likely lead to over fitting, biased asymptotic distributional properties, and low predictive power. Park and Casella (2008) extended the LASSO estimator deriving a fully Bayesian version for implementation [4]. The Bayesian version of LASSO consists of estimating the parameters β_k with some form of a prior responsible for the regularization process. In other words, instead of shrinking the parameters toward zero by adding a penalty term decided by cross-validation, the Bayesian version adds a prior with high density distribution around zero.

The most employed candidate for this prior assumes β_k goes with a Double Exponential prior (Park and

[4]The Bayesian version of the estimator has been proved to exhibit better performance compared to its frequentist counterparts. First, the broad family of Bayesian methods has been shown to work better when groups have few observations (Stegmueller, 2013). Second, it deals appropriately with the estimation of standard errors in the LASSO setting; in particular, it may provide measures of uncertainty for the parameters which the penalty term shrinks to zero (Kyung et al., 2010). Third, simulations have shown the Bayesian version outperforms in terms of mean squared errors and prediction (Ratkovic and Tingley, 2017; Kyung et al., 2010)

Casella, 2008; Tibshirani, 1996) with the following form $pr(\beta_k|\lambda) = \frac{1}{2\lambda}exp(-\lambda|\beta_k|)$, in which λ works as a hyperparameter to be set by the researcher[5]. I use in this paper the recent LASSOplus proposed by Tingley and Ratkovic (2017) as the Bayesian LASSO version which slightly modify the general structure from Park and Casella (2008) [6] ; their model can be written as:

$$
\begin{aligned}
Y_i|X_i, \beta, \sigma^2 &\sim \mathcal{N}(X_i\beta, \sigma^2) \\
\beta_k|\lambda, w_k, \sigma &\sim DE(\lambda w_k/\sigma) \\
\lambda^2|N, K &\sim \Gamma(k(\sqrt{N-1}), \rho) \\
w_k|\gamma &\sim generalizedGamma(1, 1, \gamma) \\
\gamma &\sim exp(1)
\end{aligned}
\tag{1}
$$

Where $DE(a)$ denotes the double exponential density, $\Gamma(a, b)$ and a $generalizedgamma(a, b, c)$. Following the authors' suggestions, I set the hyperprior $\rho = 1$. For a further detailed explanation of the model, I direct the readers to the original piece. Some of the advantages of this estimation although must be mentioned here. First, LASSOplus produce sparse results derived from the posterior distribution of the parameters; in this sense, variable selection and estimation co-occur. Second, it returns confidence intervals calibrated to achieve nominal coverage. Third, in simulation studies, the model has been shown to outperform other LASSO algorithms showing a lower false positive rate and false discovery rate, while keeping competitive performance for a false negative rate.

In the original motivation, Ratkivic and Tingley (2017) motivates the use of LASSOplus as a strategy to identify subgroup effects in particular as experimental research designs grow more complex. Here, I expand this original logic arguing in defense of combining this Bayesian algorithm with regression discontinuity designs. To explain how I propose combining the LASSOplus model with RDD, I briefly

[5]The decision about the prior is justified mostly by conjugacy issues which makes the mathematical derivation more tractable and the estimation easier.

[6]For different Bayesian Lasso implementations see Park and Casella (2008), Kyung et al. (2010), Carvalho et al. (2010)

describe a simple RDD estimation for two subgroups $S = S_1, S_2$. Consider the following model, where R, T, τ_s represents the running variables, the treatment group, and the subgroup treatment effect.

$$\begin{aligned} Y_i = \beta_0 \cdot S_1 + \tau_{s1}(T \cdot S_1) + \beta_{1s_1}(X \cdot S_1) + \beta_{2s_1}(X \cdot TS_1) + \\ \beta_3 \cdot S_2 + \tau_{s2}(T \cdot S_1) + \beta_{4s_2}(X \cdot S_2) + \beta_{5s_1}(X \cdot TS_1) \end{aligned} \tag{2}$$

Assuming Ω as the linear combination between β and τ_S, we have the following Bayesian implementation of the model augmenting the parameter space by representing the double exponential distribution as a scale mixture of normals.

$$\begin{aligned} Y_i|X_i, \Omega, \sigma^2, \sim \mathcal{N}(X_i \cdot \Omega, \sigma^2) \\ \Omega|\lambda, w_k, \sigma \sim DE(\lambda w_k/\sigma) \implies \Omega|\lambda^2, w_k^2, \sigma^2 \sim \mathcal{N}(0, \lambda^2\sigma^2/w_k^2) \\ \tau_k^2 \sim exp(\tau^2/2) \end{aligned} \tag{3}$$

In the simple model of equation 2, I must estimate four parameters to recover the treatment effect for each group. In the data set of this paper, thirty three parties would generate 132 parameters, all with distinct sample size, and with some of them extremely sparse around the bandwidth for the RDD estimation. As the number of parameters explodes, the model needs some sort of stabilization. The use of sparse priors provides precisely this adjustment. More importantly, as political scientists, these subgroup effects are often in the center of our concern, with the use of machine learning techniques to recover heterogeneous treatment effects emerging as an innovative area of research in the last years (Grimmer et al., 2017; Imai and Strauss, 2011; Imai and Ratkovic, 2013; Athey and Imbens, 2015; Green and Kern, 2012). The strategy outlined above works as an applied contribution to this growing scholarship. More importantly, using simulation studies, Anastasopoulos (2018) shows that when data is sparse, Bayesian LASSO estimation of treatment effects has superior performance compared to LRR estimation. His work argues that Bayesian LASSO minimizes false negative and false positive rates under small sample sizes

and estimate more credible bounds.

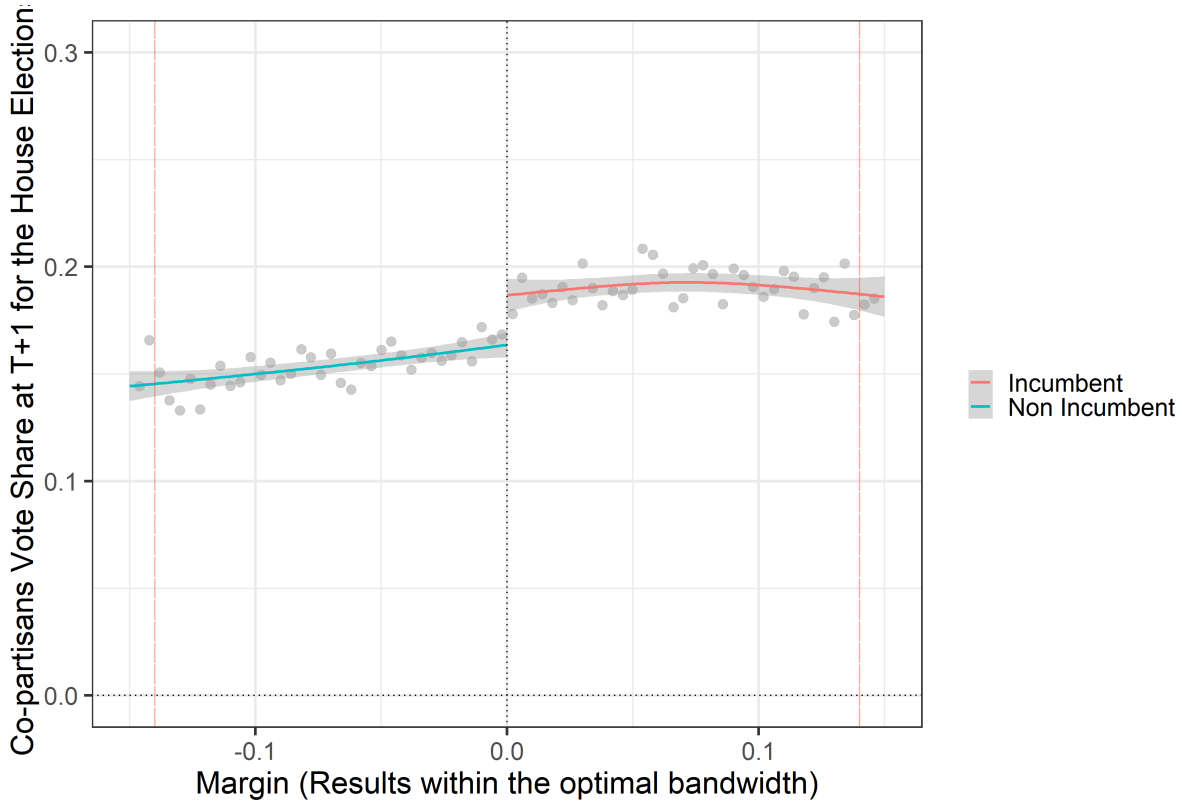
To estimate the Bayesian models, I start evaluating the treatment effect for each parties. Then, I interact each party parameter with the subsamples for the career incentives and party strength hypothesis. The former allows me to verify which parties matter whereas the latter model shows under which mechanism they matter. I use the same bandwidth calculated by the data-driven approach for each model (Anastasopoulos, 2018). Finally, I estimate the LASSOplus model using the Gibbs sampler with 1,000 burn-in iterations, 1,000 posterior samples, and thinning every 30 samples, which yielded 1,000 draws.

Results

I start providing evidence for the effects of winning the local executive on the party vote share for the House. Figure 2 presents the results for the RD estimation pooling all the years together. Analyzing all the electoral years, the average effect of controlling the local executive is an increase in two percentage points in the vote share of the incumbent's party for the House. The results are aggregated at the municipal level; therefore, it should be understood as an increase in one particular municipality. This finding confirms similar investigations on the importance of mayors in Brazil (Novaes, 2017; Avelino et al., 2012).

Figure 3 presents the treatment effect for each of the four electoral cycles, and the pooled results using the optimal bandwidth selection. Electing the mayor (t) increases on average 2.2 percentage points in the vote share for the mayors' co-partisans candidates for the House election (t_{+2})compared to the runner up party vote share. I provide the numerical results with different bandwidth choices in the supplemental files. All the results have positive values, although with some variation on statistical significance within very narrow windows. The effect increases over time which goes in the direction of the theoretical idea of learning in boosting dynamics. These findings corroborate with previous research on local politics and clientelism in Brazil. On a recent work, Novaes (2017) argues that institutional reforms in 2007 in

Figure 2: Smoothed Regression Discontinuity of the Treatment Effect (Red lines represent the optimal bandwidth decision)

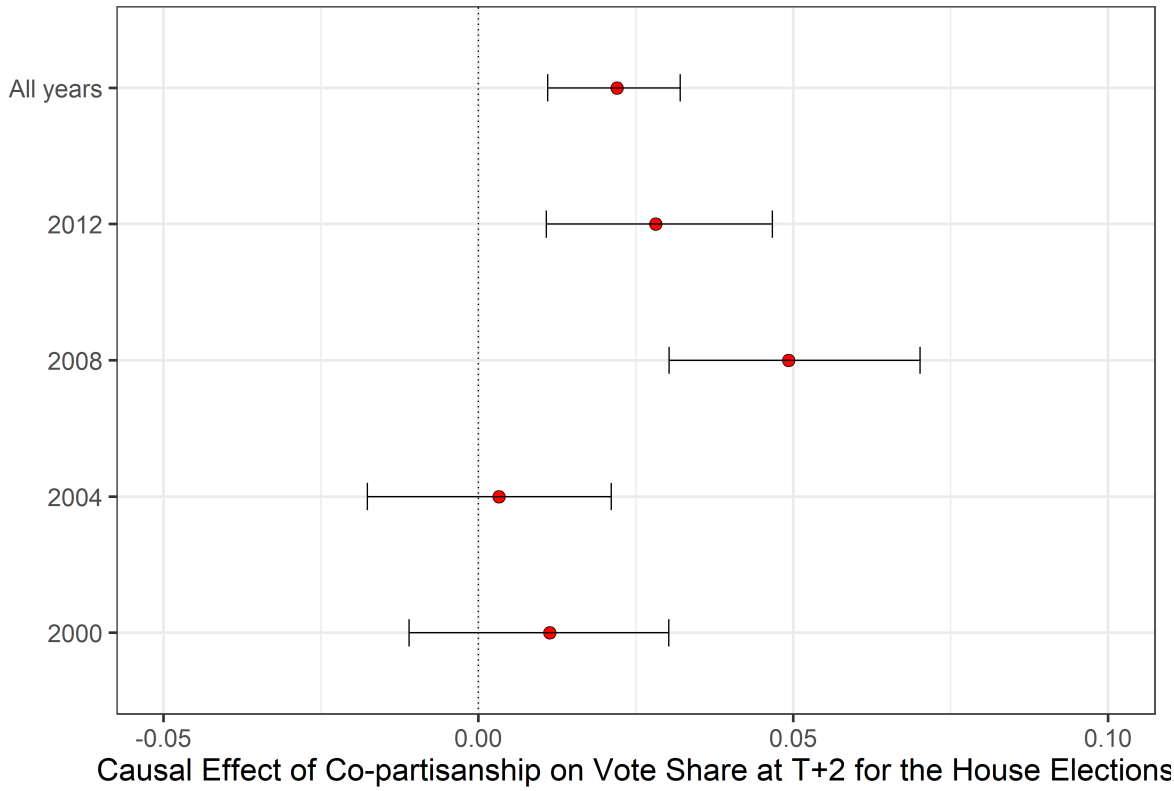


Brazil constrained the politicians' freedom to switch parties, therefore, increasing mayors commitment to delivery votes to their upper-level co-partisan politicians. In 2008, for example, the treatment effect was twice the size of the pooled average effect.

What is the substantive effect of a two percentage points advantage in the incumbents' party vote share for the House? The average vote-share per party for the House across all the electoral years is 5.5% with a standard deviation of 1.1% in each city (See figure 1).Therefore, electing the mayor increases on average two standard deviations in the vote share of the parties for the Brazilian House election in each municipality. Undoubtedly, an increase of two standard deviations is a substantial effect. Additionally, the dispute for the House Seats in Brazil are remarkably competitive and uncertain (Calvo et al., 2015);); therefore, a two standard deviation increase in any party vote share might be the difference between

winning or losing a seat.

Figure 3: Treatment effect over time (Four Electoral Cycles)



Mechanisms for Political Boosting

In this section, I present evidence for the importance of party coordination capacity on political boosting. The mechanisms related to party strength seem to play the role expected by our theories. Figure 4 presents the conditional treatment effect using the subsamples in table 1. The results show that conditional on being from the same party of the governor, winning the mayoral race increases the party vote share for the House election by 13.5% percentage points. This conditional effect on the governor subsample is six times larger than the average treatment effect indicating the importance of vertical alignment for political boosting.

Figure 4 shows a similar effect for the mayors who have at least one elected house representative at

the time of the election in t_{+2}. Political boosting results are substantively more significant when the incumbent's party also holds a seat in this district. The effect is twice the size of the average treatment effect [7].

These results make sense with our theory of political boosting. In an economy of scale, political elites have more loyal allies at the local level when party coordination capacity is higher. When the party controls the State level executive branch, avoiding free rider problems becomes more trivial since partisan elites have more instruments to punish and reward local partisans. Similar logic is present when the party already has a representative in this State. In the other side, the support of the governor also increases the strength of signs sent by the mayor about their co-partisans; voters pay attention and know the benefits their towns might receive by having politically aligned representatives with their mayors (Bueno, 2017). Finally, the results confirm the general finding of the specialized literature about the importance of governors for partisan and electoral dynamics in Brazil (Samuels, 2003; Abrucio, 1998).

Figure 4 plots the conditional treatment effect of winning the local election across the different subsamples in table 2. The results indicate a positive conditional treatment effect for all three subsamples which show that career paths do not matter for the mayor's willingness to help her co-partisans. The effects are similar in magnitude to the average treatment effect of a 2% increase in the vote share for the House elections, although in the Lame-Duck sample, the effect is not statistically different from zero using the robust 95% confidence intervals. To summarize, political boosting occurs mostly independently of the mayors' career ambition; these results go in a different direction from the previous research on incumbency effects in Brazil and mayoral effects on upper level elections (Klašnja and Titiunik, 2017; Novaes, 2017). Notably, partisan coordination capacity has stronger mediator effects on political boosting in federal democracies.

[7] See appendix where I provide results for the subsamples when the incumbent has no co-partisan running for reelection for the House and is from the same party of the State Governor

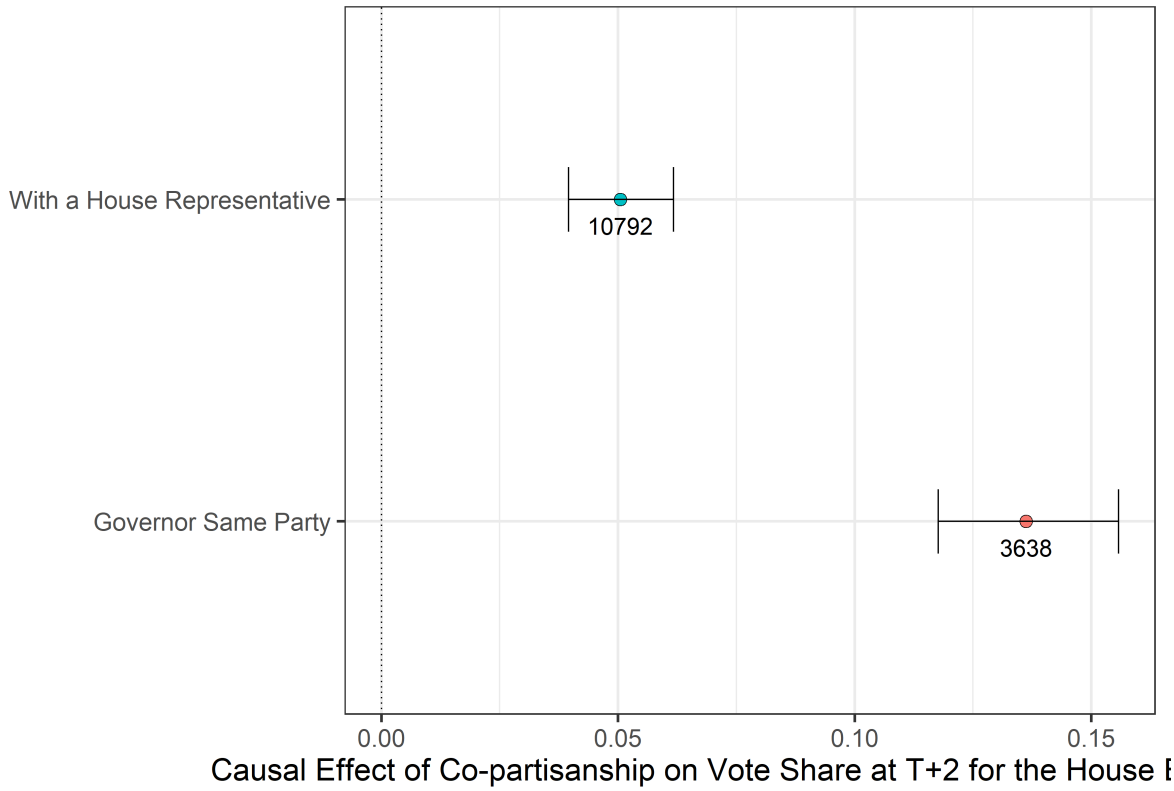


Figure 4: Conditional Treatment effect for the Party Coordination mediator on Political Boosting. The figures plot robust 95% confidence intervals. The number of cases is printed above each point estimate.

Do Parties matter for political boosting?

Figure 6 presents the Bayesian LASSO non-zero coefficients for the party effects. The results indicate five parties have treatment effects different from zero: the *Partido do Movimento Democrático Brasileiro* (PMDB), the *Democratas* (DEM, formerly Partido da Frente Liberal), the *Partido da Social Democracia Brasileira* (PSDB), and the *Partido Progressista* (PP), and the *Partido dos Trabalhadores* (PT). The effects are positive, statistically distinct from zero, and the PP has the largest conditional effect.

These five parties are the largest parties in the Brazilian system regarding seats in House and number of elected mayors over the last three decades [8] which indicates that party size matters notably for coordi-

[8]The exception here is the emergent PSL, the party of the president Jair Bolsonaro, which elected the President and the second largest number of representatives in the 2018 National election.

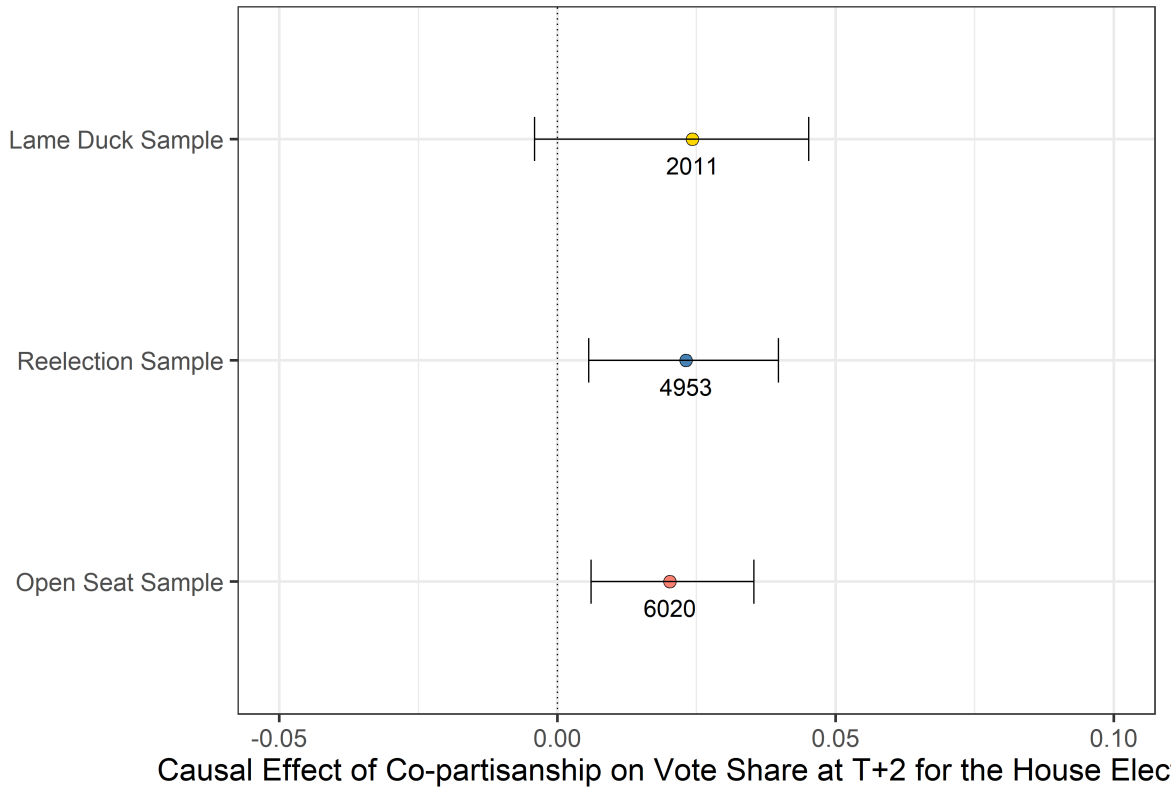


Figure 5: Conditional Treatment effect for Mayors' Career Ambition mediator on political boosting. The figures plot robust 95% confidence intervals. The number of cases is printed above each point estimate.

nation capacity. The PP, a non-programmatic, clientelistic party appears as the one with the larger effect suggesting that the treatment is hardly related to the strength of party labels, as the literature supposes when analyzing incumbency effects, for example, Titiunik et al. (2015). Second, the fact that I find positive outcomes for these parties corroborates with the positive findings of the subsamples for the party coordination hypothesis: larger parties in Congress have more representatives spread around the districts as well as governors elected, therefore exhibiting a higher ability to coordinate the mayors' behaviors. Finally, programmatic preferences and coattail presidential effects also seem to have no effects; the five parties represent a fair amount of variation in a programmatic aspect, and except for the PT and PSDB, none of the other three parties have had competitive presidential candidates over the recent election [9].

[9] See appendix for LLR models using only the parties with competitive presidential candidates. The results show no consistent effects.

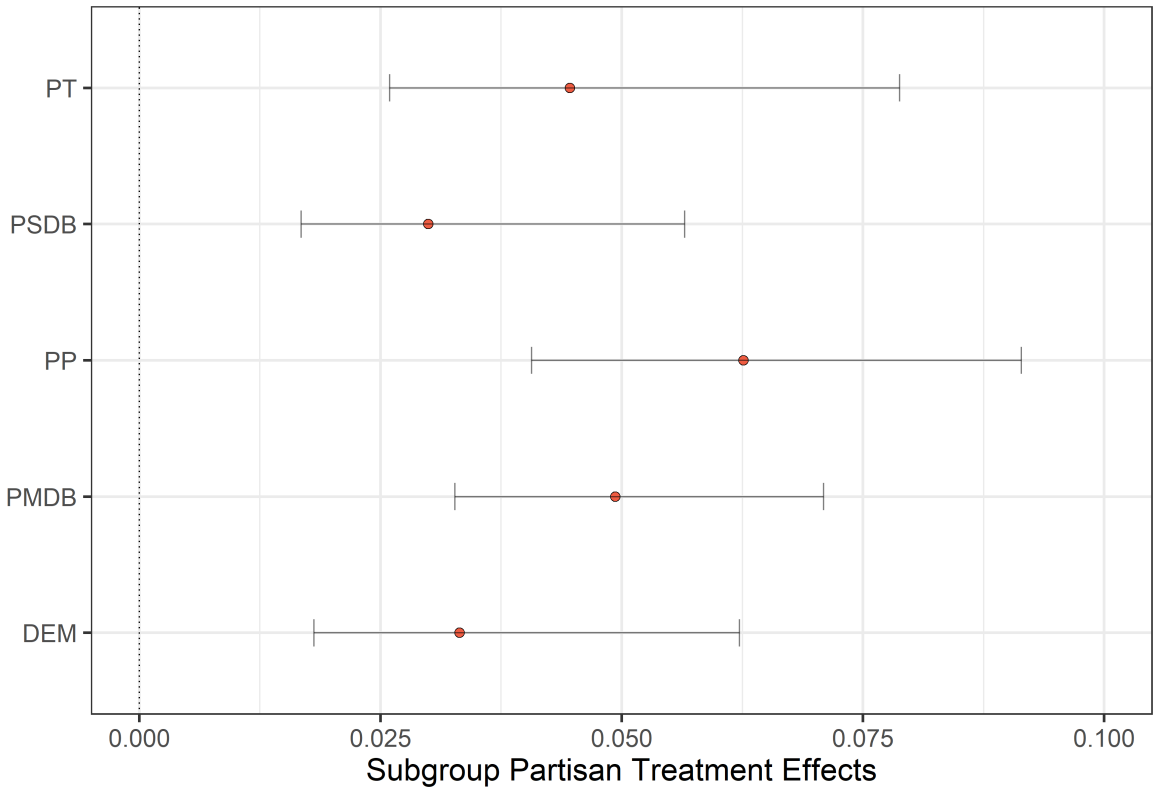


Figure 6: Estimated nonzero Conditional Treatment Effects by parties using Bayesian Sparse LASSO. Estimation with data-driven bandwidth selection. Median point estimates with 95% confidence intervals.

Figure 7 presents the partisan effects across the subsamples for career-ambition and party coordination mechanisms. I estimate seven distinct Bayesian LASSO Models each with 132 variables including the 33 subgroup party treatment effects. Across all the models, I find fourteen non-zero treatment effects for the subgroups. The results suggest that the PT exhibits positive effects across most of the subsamples, in particular, the most substantial effect appears when the Workers Party controls both the local executive and the State Governor. It is also noticeable that most of the partisan effects occur when party capacity moderates each of the partisan effects. In this sense, we find robust findings that political boosting increases with large parties and when these parties have a higher capacity to faithfully send signs to voters and coordinate their mayors' behaviors. I should also direct the reader to the fact that that the conditional

effects are all larger than the average treatment effect.

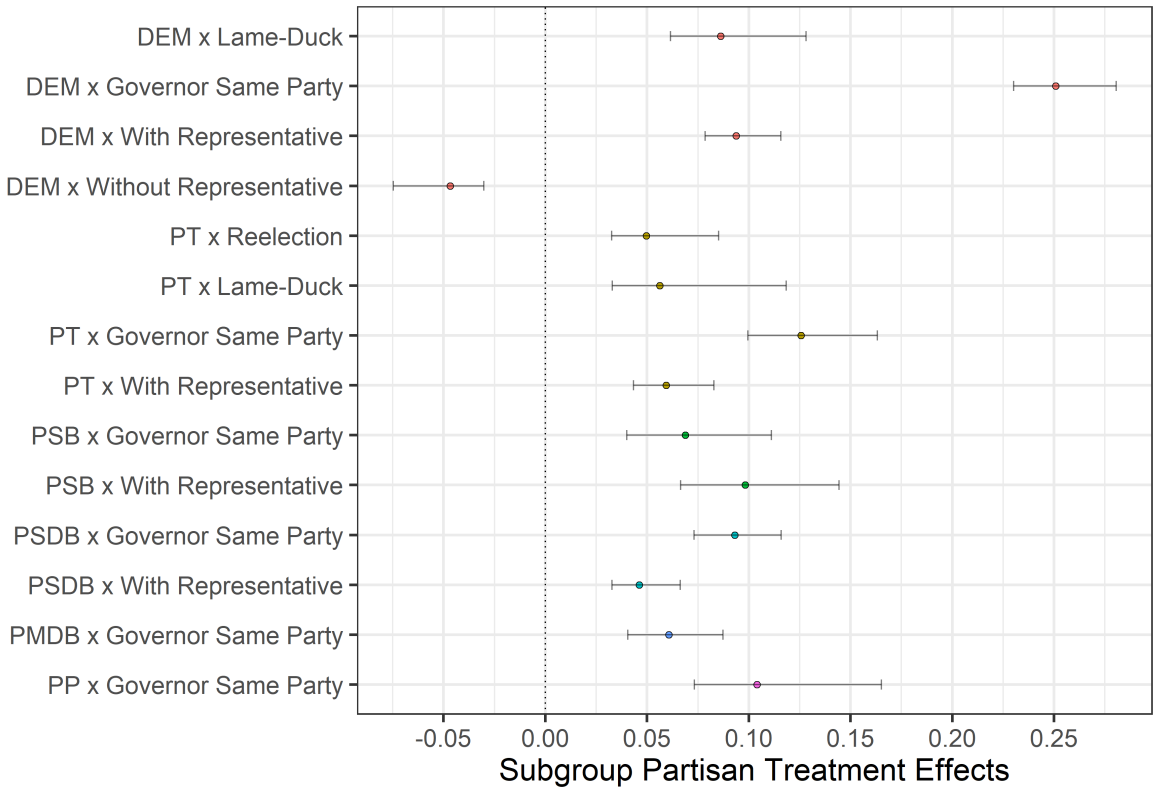


Figure 7: Estimated nonzero Conditional Treatment Effects by Parties using Bayesian Sparse LASSO within the subsamples. Estimation with data-driven bandwidth selection. Median point estimates with 95% confidence intervals.

Robustness Checks

For the discontinuity design to be internally valid, the continuity assumption must hold. Although identification assumptions are untestable, it is standard procedure to provide empirical evidence that no unobserved cofounders exist in the running variable and outcome. I perform two sorts of tests for the validity of my RD design. In the appendix, I first show no evidence of sorting on the margin of victory in the pooled data, the data by year, the subsamples, and the five parties with positive conditional effects; none of the cases have robust p-values for the null hypothesis of sorting smaller than .11 using the test

developed by Cattaneo, Jansson, and Ma (2018). I also run a "placebo" test regressing the treatment on lagged values for the vote share for House Election (t_{-2}) for the pooled data, for the year subsamples, and for the five parties. The result using the lagged vote share are particularly reassuring regarding the validity of the RD setup, as it implies that predetermined electoral outcomes are not related to close races. Except for the case of the PMDB, with a negative coefficient which is statistically distinct from zero, I find now evidence of spurious correlation with the pre-treatment vote share.

To further test the robustness of the results, I re-estimate the treatment effect in the supplemental files using the rate of change of party support for House elections before and after electing the mayor and the vote share of state legislators as the dependent variables. If winning locally matters to explain different electoral gains between the incumbent and the runner up in the upcoming elections, it is reasonable to expect the incumbents' vote share *vis-a-vis* the runner-up candidate also increases over time. Therefore, in the supplemental files, I re-estimate the models using the rate of change of the barely elected (loser) mayor's vote share from the previous to the next election as explanatory variable. The logic applies correspondingly for state legislators who are elected at the same day of House members in Brazil; if informational shortcuts run from below, I should expect that winning locally also matters to co-partisans state legislators. I direct the reader to the appendix of a more extended discussion, where I report in the details the results.

In brief, the findings perfectly converge for change of electoral support over time; indeed, the treatment effect is more substantial when the baseline of the vote share in previous elections comes to the model. However, the effects on state legislators do not seem as strong. In the latter case, only when the mediators of party coordination are present can mayors effectively help their co-partisans. In my view, these findings speak to the fact that state deputies usually are closer to the local than their co-partisans in the House, therefore, relying less upon the informational shortcuts provided by mayors. Finally, I also test in the

appendix if having nationally competitive candidates for the presidency works as a mediator for political boosting. The effects do not provide evidence for this argument.

Conclusion

In this paper, I propose a theory to understand if, how, and for which parties investing in local political capital renders substantial national gains. I focus on Brazil as a noisy, fragmented political environment because I argue it provides a hard test for the theory. I label this process as political boosting borrowing from the statistical term through which algorithms improve their performance on noisy environments. This paper finds consistent results for political boosting of the mayor's support on the party vote share for the House elections. The article also identifies the mechanisms behind this strategy: party-coordination capacity increases the effectiveness of controlling the local to win nationally. When the state governor belongs to the same party as the mayor or the organization holds at least one House Seat in the district, the effects of winning local elections are positive and substantively larger than the average treatment effect. Finally, I also show parties matter for boosting: larger parties have a stronger positive impact, while small parties exhibit no effect. I identify the subgroup treatment effect by applying a novel Bayesian model combined with regression discontinuity designs.

The findings of this paper make several contributions. I contribute to the literature on multilevel democracies by discussing the existence of political effects on non-concurrent elections. Our theory of political boosting and the empirical findings on this paper show how on non-concurrent elections, voters incorporate information from the local politics to decide upper-level elections. Therefore we propose a new explanation to the null finding in the literature on coattails when dealing with non-concurrent races (Golder, 2006; Samuels, 2000b; Borges and Lloyd, 2016).

I also contribute to the literature on information shortcuts in noisy democracies. In the opposite di-

rection of Rodden and Wibbels' (2010) argument, when elections are non-concurrent, our theory argues information runs from below; from the local to the upper level. In this sense, controlling the local becomes key for party building in fragmented polities. Inattentive voters pay attention to local politics, in our case, affiliation with the mayor, to reward her co-partisans running for House election. The importance of party coordination capacity identified in our empirical findings shows voters reward large parties, with higher capacity to access upper-level grants, instead of solely responding to mayors' decisions about their career.

Finally, the paper contributes to the literature on political parties in Brazil. Most of this scholarship converges to the conclusion about the uniqueness of PT in the Brazilian hostile environment (Klašnja and Titiunik, 2017; Feierherd, 2018; Samuels and Zucco, 2018). Even though the PT exhibits consistent positive effects in our models, other political parties also have some remarkable capacity to articulate local politicians to work towards their national interests. The results here seem to be more connected to national parties' strength and the occupation of political offices, than the importance of labels and party discipline features commonly argued to be exclusive of the PT in Brazil.

Additionally, the paper also adds a new piece to the puzzle of partisan-building in hostile environments. Previous research identified the presence of small parties bias for House elections in Brazil due to the rule for coalition building (Calvo et al., 2015). Putting together these findings with our theory, we can understand how parties use different strategies to pursue national goals; while small parties extract positive rewards from building heterogeneous coalitions, large parties have a different dominant strategy. To boost their political performance, large parties in Brazil go local to win the nation.

References

Abrucio, F. L. (1998). *Os barões da federação: os governadores e a redemocratização brasileira*. Editora Hucitec.

Amaral, O. (2013). *As transformações na organização interna do Partido dos Trabalhadores entre 1995 e 2009*. Alameda/Fapesp, São Paulo.

Ames, B. (1994). The Reverse Coattails Effect: Local Party Organization in the 1989 Brazilian Presidential Election. *American Political Science Review*, 88(1):95–111.

Amorim Neto, O., Cox, G. W., and McCubbins, M. D. (2003). Agenda Power in Brazil's Câmara Dos Deputados, 1989–98. *World Politics*, 55(4):550–578.

Anastasopoulos, J. (2018). Bayesian lasso estimation of treatment effects in regression discontinuity designs.

Athey, S. and Imbens, G. W. (2015). Machine learning methods for estimating heterogeneous causal effects. *stat*, 1050(5).

Avelino, G., Biderman, C., and Barone, L. S. (2012). Articulações intrapartidárias e desempenho eleitoral no Brasil. *Dados*, 55:987–1013.

Boas, T. C. and Hidalgo, F. D. (2011). Controlling the Airwaves: Incumbency Advantage and Community Radio in Brazil. *American Journal of Political Science*, 55(4):869–885.

Boas, T. C., Hidalgo, F. D., and Richardson, N. P. (2014). The spoils of victory: campaign donations and government contracts in Brazil. *The Journal of Politics*, 76(2):415–429.

Borges, A. and Lloyd, R. (2016). Presidential coattails and electoral coordination in multilevel elections: Comparative lessons from Brazil. *Electoral Studies*, 43:104–114.

Brollo, F. and Nannicini, T. (2012). Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil. *American Political Science Review*, 106(4):742–761.

Broockman, D. E. (2009). Do Congressional Candidates Have Reverse Coattails? Evidence from a Regression Discontinuity Design. *Political Analysis*, 17(4):418–434.

Bueno, N. S. (2017). Bypassing the Enemy: Distributive Politics, Credit Claiming, and Nonstate Organizations in Brazil. *Comparative Political Studies*, 51(3):304–340.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Calvo, E., Guarnieri, F., and Limongi, F. (2015). Why coalitions? Party system fragmentation, small party bias, and preferential vote in Brazil. *Electoral Studies*, 39(APRIL):219–229.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Simple local polynomial density estimators. *arXiv preprint arXiv:1811.11512*.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2017). Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality. *Journal of Policy Analysis and Management*, 36(3):643–681.

De la Cuesta, B. and Imai, K. (2016). Misunderstandings about the regression discontinuity design in the study of close elections. *Annual Review of Political Science*, 19:375–396.

Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and Snyder, J. M. (2015). On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races. *American Journal of Political Science*, 59(1):259–274.

Feierherd, G. (2018). How Mayors Hurt Their Presidential Ticket: Party Brands and Incumbency Spillovers in Brazil. *Journal of Politics*.

Figueiredo, A. C. and Limongi, F. (2000). Presidential Power, Legislative Organization, and Party Behavior in Brazil. *Comparative Politics*, 32(2):151–170.

Firpo, S., Ponczek, V., and Sanfelice, V. (2015). The relationship between federal budget amendments and local electoral power. *Journal of Development Economics*, 116:186–198.

Gelman, A. and Imbens, G. (2018). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, pages 1–10.

Golder, M. (2006). Presidential Coattails and Legislative Fragmentation. *American Journal of Political Science*, 50(1):34–48.

Green, D. P. and Kern, H. L. (2012). Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees. *Public Opinion Quarterly*, 76(3):491–511.

Grimmer, J., Messing, S., and Westwood, S. J. (2017). Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods. *Political Analysis*, 25(4):413–434.

Hainmueller, J. and Hazlett, C. (2014). Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach. *Political Analysis*, 22(2):143–168.

Hidalgo, F. D. (2010). Digital democratization: Suffrage expansion and the decline of political machines in Brazil.

Hunter, W. (2010). *The Transformation of the Workers' Party in Brazil, 1989-2009*. Cambridge University Press.

Imai, K. and Ratkovic, M. (2013). Estimating treatment effect heterogeneity in randomized program evaluation. *The Annals of Applied Statistics*, 7(1):443–470.

Imai, K. and Strauss, A. (2011). Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign. *Political Analysis*, 19(1):1–19.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Keck, M. E. (1992). *The Workers' party and democratization in Brazil*. Yale University Press, New Haven, CT.

Klašnja, M. and Titiunik, R. (2017). The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability. *American Political Science Review*, 111(1):129–148.

Kyung, M., Gill, J., Ghosh, M., and Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5(2):369–411.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, 142(2):675–697.

Lucardi, A. (2016). Building Support From Below? Subnational Elections, Diffusion Effects, and the Growth of the Opposition in Mexico, 1984-2000. *Comparative Political Studies*, 49(14):1855–1895.

Luz, J. and Dantas, H. (2017). Coligações majoritárias municipais e emendas parlamentares federais: em busca de conexões. *CADERNOS ADENAUER*, 4:79–101.

Magar, E. (2012). Gubernatorial Coattails in Mexican Congressional Elections. *The Journal of Politics*, 74(2):383–399.

MEREDITH, M. (2013). Exploiting Friends-and-Neighbors to Estimate Coattail Effects. *American Political Science Review*, 107(4):742–765.

Novaes, L. M. (2017). Disloyal Brokers and Weak Parties. *American Journal of Political Science*, 62(1):84–98.

Park, T. and Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686.

Pocock, S. J., Assmann, S. E., Enos, L. E., and Kasten, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practiceand problems. *Statistics in medicine*, 21(19):2917–2930.

Ratkovic, M. and Tingley, D. (2017). Sparse estimation and uncertainty with application to subgroup analysis. *Political Analysis*, 25(1):1–40.

Rodden, J. and Wibbels, E. (2010). Dual accountability and the nationalization of party competition: Evidence from four federations. *Party Politics*, 17(5):629–653.

Samuels, D. (2000a). Concurrent Elections, Discordant Results: Presidentialism, Federalism, and Governance in Brazil. *Comparative Politics*, 33(1):1–20.

Samuels, D. (2003). *Ambition, Federalism, and Legislative Politics in Brazil*. Cambridge University Press, Cambridge, New York.

Samuels, D. and Zucco, C. (2015). Crafting Mass Partisanship at the Grass Roots. *British Journal of Political Science*, (March):1–21.

Samuels, D. J. (2000b). The Gubernatorial Coattails Effect: Federalism and Congressional Elections in Brazil. *The Journal of Politics*, 62(1):240–253.

Samuels, D. J. and Zucco, C. (2018). *Partisans, Antipartisans, and Nonpartisans: Voting Behavior in Brazil*. Cambridge University Press, Cambridge.

Skovron, C. and Titiunik, R. (2015). A practical guide to regression discontinuity designs in political science. *American Journal of Political Science*.

Stegmueller, D. (2013). How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science*, 57(3):748–761.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Tibshirani, R., Wainwright, M., and Hastie, T. (2015). *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC.

Titiunik, R., Calonico, S., and Cattaneo, M. D. (2015). rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs. *R Journal*.

Appendices

Appendix 1: Numerical Results under different bandwidths

In this appendix, I provide the numerical results of the treatment effect using different bandwidths for the running variable. This strategy contributes to the robustness of the findings reported in the paper. Table 3 presents the results pooling all the years together.

Table 3: Regression Discontinuity Results: Impact of electing the Mayor

Outcome: Vote Share Co-partisans for the House Election				
Bandwidth (Margin of Victory)	Estimate	Lower Bound CI	Upper Bound CI	Number of Pairs
Optimal Bandwidth (14.4%)	0.022	0.011	0.032	12768
1%	0.019	−0.024	0.086	1080
5%	0.018	−0.006	0.036	5374
10%	0.019	0.001	0.031	9962
25%	0.024	0.012	0.032	17482
100%	0.025	0.017	0.030	21384

Appendix 2: Regression Tables for Figures 5 and 6

Table 4: Regression Discontinuity Results across the subsamples

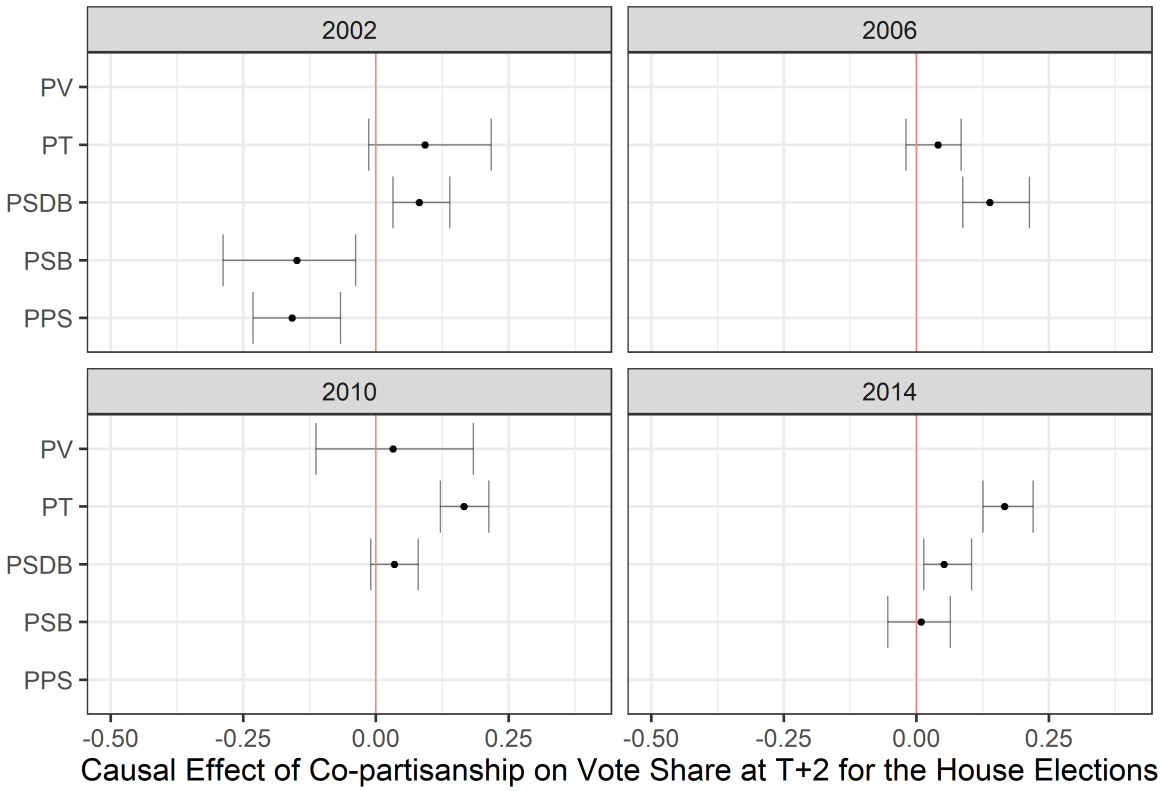
Outcome: Vote Share Co-partisans for the House Election					
Subsample	Optimal Bandwidth	Estimate	Lower Bound 95% CI	Upper Bound 95% CI	Number of Cases
Career Ambition Hypothesis					
Open Seat Sample	16.390	0.020	0.006	0.035	6020
Reelection Samples	12.686	0.023	0.006	0.040	4953
Lame Duck Sample	13.518	0.024	−0.004	0.045	2011
Party Strength Hypothesis					
Governor Different Part	13.459	−0.011	−0.024	−0.001	9661
Governor Same Party	20.339	0.136	0.118	0.156	3638
Without a House Representative	16.337	−0.122	−0.144	−0.102	2209
With a House Representative	13.948	0.050	0.040	0.062	10792

Appendix 3. Parties with competitive Presidential Candidates.

Being competitive nationally might be crucial to political boosting. This idea has remarkable semblance with the arguments made by the scholarship on presidential coattails horizontal effects; therefore, it offers us an interestign robustness check about the argument of political boosting in between elections. In Brazil, Presidential elections occur at the same time of the House elections; hence, it is a plausible argument that national competitiveness might work as the causal mechanism explaining heterogeneous effects of the parties. To discuss this alternative, I estimate local polynomial models splitting the data by year and by each competitive presidential candidates. The PT and the PSDB appear in all the estimations since these parties have been the main contenders of all the races in my dataset. I then add the third and fourth runner up parties changes for each election up to all sum more than 95% of the valid vote share in the first round for each presidential election. Figure 8 presents the results.

Having a competitive national candidate have no consistent effect on partisan coordination. As figure 8 indicates, none of the other parties running against the PT and PSDB, exhibit consistent positive co-partisan effects of winning local elections. Even among the PT and the PSDB some heterogeneity occurs. In 2002 and 2006, only the PSDB has positive effects, while the PT exhibit positive effects alone in 2010m and together with the PSDB in 2014. The results also rule out the effect of presidential incumbency affecting local boosting.

Figure 8: Treatment effect for nationally competitive party using the local linear estimation by electoral year



Appendix 4. Effects for Non-Governor and No Representative Sample

.

In the subsection *Mechanisms for Political Boosting*, I assess whether alignment with the Governor Party and Party incumbency for the House election explain the degree to which parties gain from political boosting. Here, I provide an extra piece of information for this section by analyzing the treatment effect for the opposite cases of these two subsamples. I discuss the impact of winning the local election when the mayor does not belong to the governors' party, and it has no representative running for the reelection in the House race. The table below presents the treatment and control group for each subsample.

Table 5: Description of RDD for the Party Coordination Hypothesis

No Governor Sample

- Treatment Group: Mayor who barely wins at t and is not a member of the Governor's.
- Control Group: Runner up candidate who barely loses at t for a mayor from the treatment group.
- Outcome: The party vote share for House at t_{+2} in municipality i for the treatment and control group
- Outcome: The party vote share for House t_{+2} in municipality i for the treatment and control group.

No Representative Elected Sample

- Treatment Group: Mayor barely wins the reelection at t and is forbidden to run again at t_{+4}.
- Control Group: Runner up candidate barely loses at t for a mayor from the treatment group.
- Outcome: The party vote share for House t_{+2} in municipality i for the treatment and control group.

Figure 9 presents the results. For both subsamples, the treatment effects are negative. Mayors who are barely elected reduce the vote share of their party in subsequent elections in cases of non-alignment with the governor. The same effects occurs for incumbents affiliated with a party with no representative elected in the district. Considering the size of the Brazilian districts, these parties are likely to be small parties, which gain no benefit from political boosting according to our findings on subsection *Do Parties Matter for Political Boosting?*. In this sense, as the paper more broadly show, political boosting seems to work effectively for larger parties, while smaller parties probably rely on different strategies for gaining national political offices.

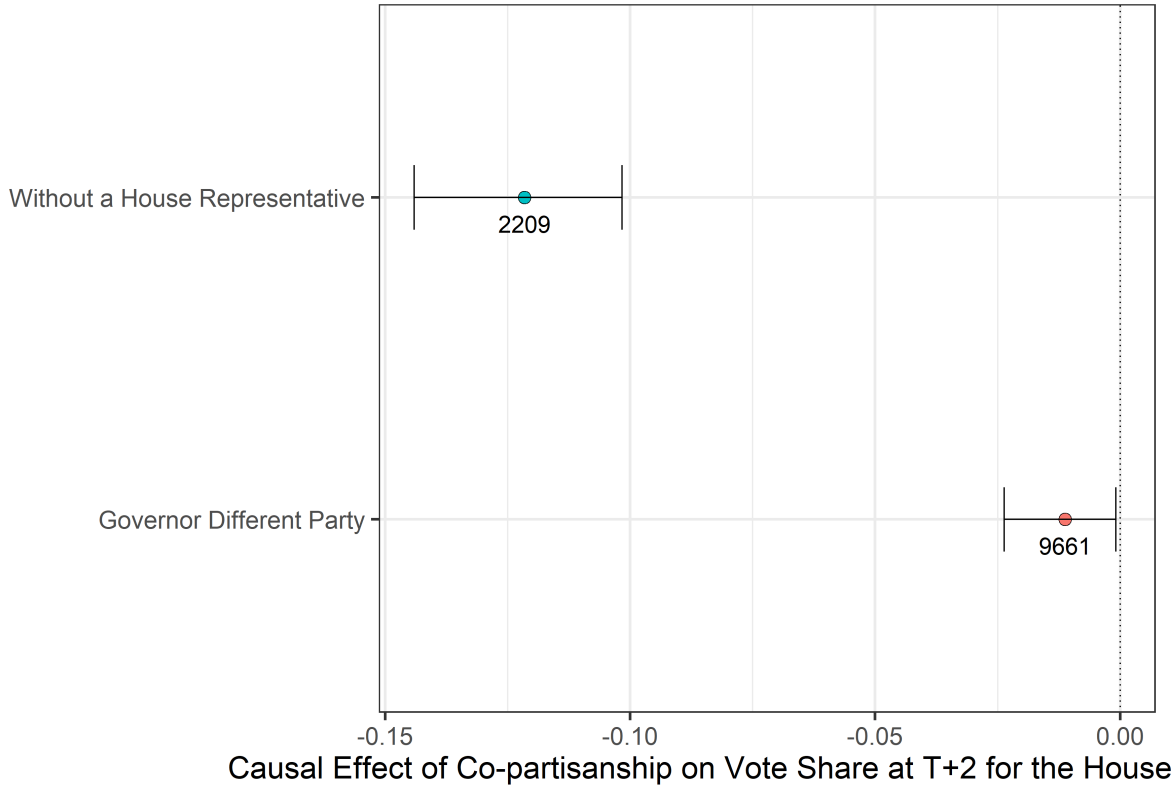


Figure 9: Conditional Treatment effect for the No Governor Alignement and No Representative Elected subsamples. The figures plot robust 95% confidence intervals. The number of cases is printed above each point estimate.

Appendix 5. Local Linear Estimators for Partisan Effects Effects for Non-Governor and No Representative Sample

.

In this appendix, I provide the estimate for party subgroup treatment effects using the local estimators proposed by Calonico, Cattaneo, and Titiunik (2014). Table 6 presents the full results and figure 10 plots them graphically. The results provide a crucial intuition about the advantages of using the Bayesian Lasso estimation for subgroup effects. First, data is sparse for most of the parties what might add result in bias in the local estimation of the parameters. Second, the results somewhat converge to those reported in the paper. Except for the PCdoB, a strong programmatic communist party in Brazil, which have positive and statistically significant treatment effects, all the other 5 five parties identified by the Bayesian Lasso model also appear in the local estimation with positive and significant results. The main issue here is on the identification of false negatives; the local linear model finds a negative and statistically significant effect for some other parties for which the samples are ridiculously small. For example, I direct the reader to the cases of the PRTB and the PSL with 21 and 53 cases, respectively.

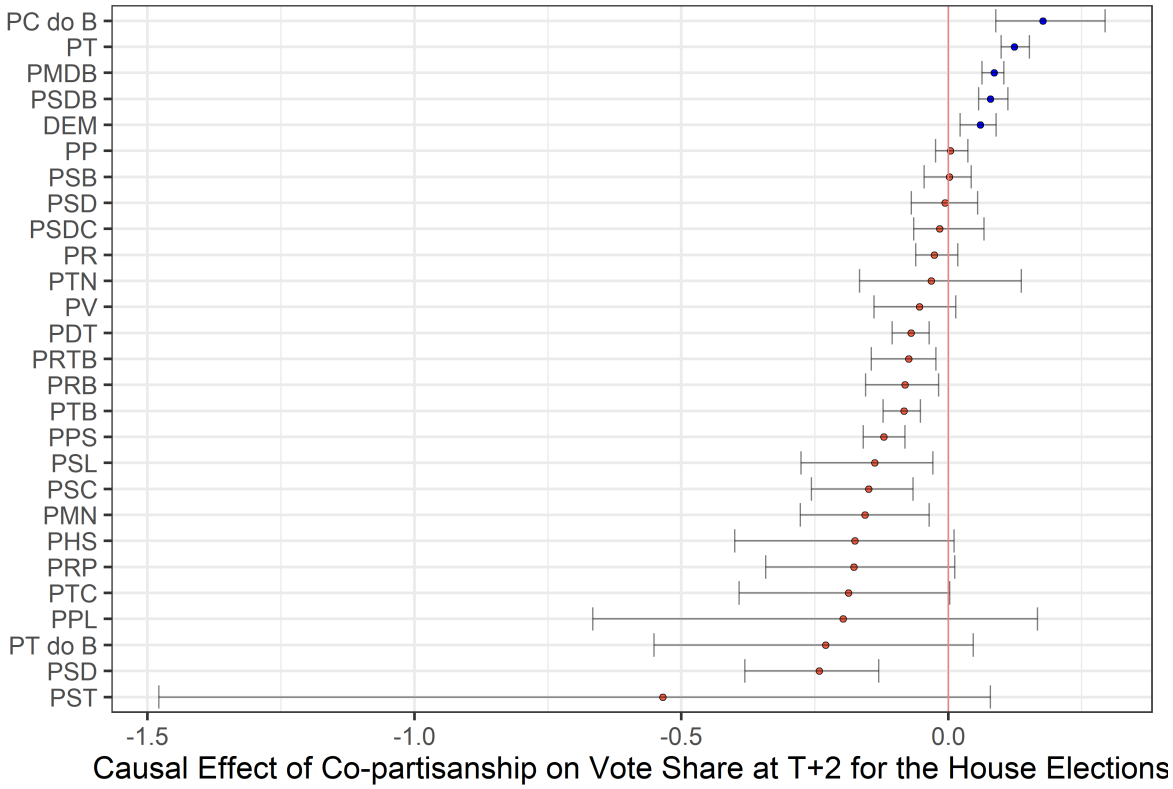


Figure 10: Treatment subgroup Party Effects using Local Estimators.

Table 6: Regression Discontinuity for Party Effects using Local Estimators

Outcome: Vote Share Co-partisans for the House Election					
Party	Number of Cases	Estimate	Lower 95% CI	Lower 95% CI	Bandwidth
PC do B	64	0.177	0.089	0.294	16.089
PT	1292	0.124	0.099	0.152	19.563
PMDB	2829	0.086	0.063	0.104	15.391
PSDB	1504	0.079	0.057	0.112	11.076
DEM	1423	0.060	0.022	0.090	13.518
PP	1295	0.004	−0.024	0.037	13.308
PSB	616	0.002	−0.045	0.043	14.366
PSD	288	−0.006	−0.069	0.055	12.764
PSDC	14	−0.016	−0.065	0.067	7.526
PR	833	−0.026	−0.061	0.018	17.493
PTN	18	−0.032	−0.166	0.137	16.824
PV	184	−0.054	−0.139	0.014	17.450
PDT	881	−0.070	−0.105	−0.036	18.926
PRTB	21	−0.074	−0.144	−0.023	10.874
PRB	82	−0.081	−0.155	−0.018	13.321
PTB	853	−0.083	−0.122	−0.052	13.949
PPS	530	−0.121	−0.159	−0.081	20.445
PSL	53	−0.138	−0.276	−0.029	17.574
PSC	115	−0.149	−0.256	−0.066	12.840
PMN	100	−0.156	−0.277	−0.036	19.996
PHS	27	−0.175	−0.400	0.011	11.235
PRP	52	−0.177	−0.342	0.012	11.976
PTC	37	−0.187	−0.392	0.003	18.221
PPL	3	−0.197	−0.666	0.167	7.602
PT do B	25	−0.230	−0.551	0.047	8.754
PSD	68	−0.242	−0.381	−0.130	12.578
PST	8	−0.535	−1.479	0.079	6.782

Appendix 6. Validity of RD Designs

.

In this section, I discuss the validity of the RD design employed in the paper. To analyze the change in the outcome occurring at the cutoff as being the causal effect of winning local elections, I must show that the continuity assumption for the RD design holds in this case. I show now some empirical evidence supporting the validity of the assumption. The main empirical challenge here is to show no other factors related to the outcome and the running variable are discontinuous at the cutoff.

I first plot the density of the running variable for the pooled data and by year, for each subsample of the mediators, and for the five parties with positive conditional effect identified using the Bayesian LASSO model. If parties had the ability to influence whether they lose or win, one would be likely to observe very few parties that barely lose, and many more parties that barely win. At least since the adoption of electronic ballots, Brazil has tremendously reduced cases of electoral manipulation (Hidalgo, 2010). However, it is paramount to verify this empirical pattern to validate the RD design. Figures show histograms of the margin of victory also reporting p-value of the null hypothesis that the density of the running variable is continuous at the cutoff using the local polynomial density estimator developed by Cattaneo, Jansson, and Ma (2018). The graphs 11, 12, and 13 indicate that there is no evidence of sorting in any of the cases (p-values range from 0.11 to 0.99).

Furthermore, I estimate "placebo" RD effects on some pre-treatment covariates. The treatment here being assigned after these covariates are measured, then I can assume winning local elections have effects indistinguishable from zero. Significant effects would be an indication of unobserved confounders affecting the outcome. I run tests estimating the treatment effects of winning at t on the party vote share at t_{-2} in the House elections for the pooled data, for each year and for the five parties I find significant subgroup effects. Table 7 presents the results. The assumption here is the same; if winning at t has any

relationship with the dependent variable measured before the treatment occurs, we have an indication of unobservable covariates. I should note I do not run this test on the election of 2000 because in the lagged case - house election for 1998 - the country was not yet using electronic ballots, and then electoral fraud was still a possibility. Except for the PMDB, actually, with negative results, none of the models are statistically indistinguishable from zero at 95% level of confidence using the robust standard errors and data-driven bandwidth s

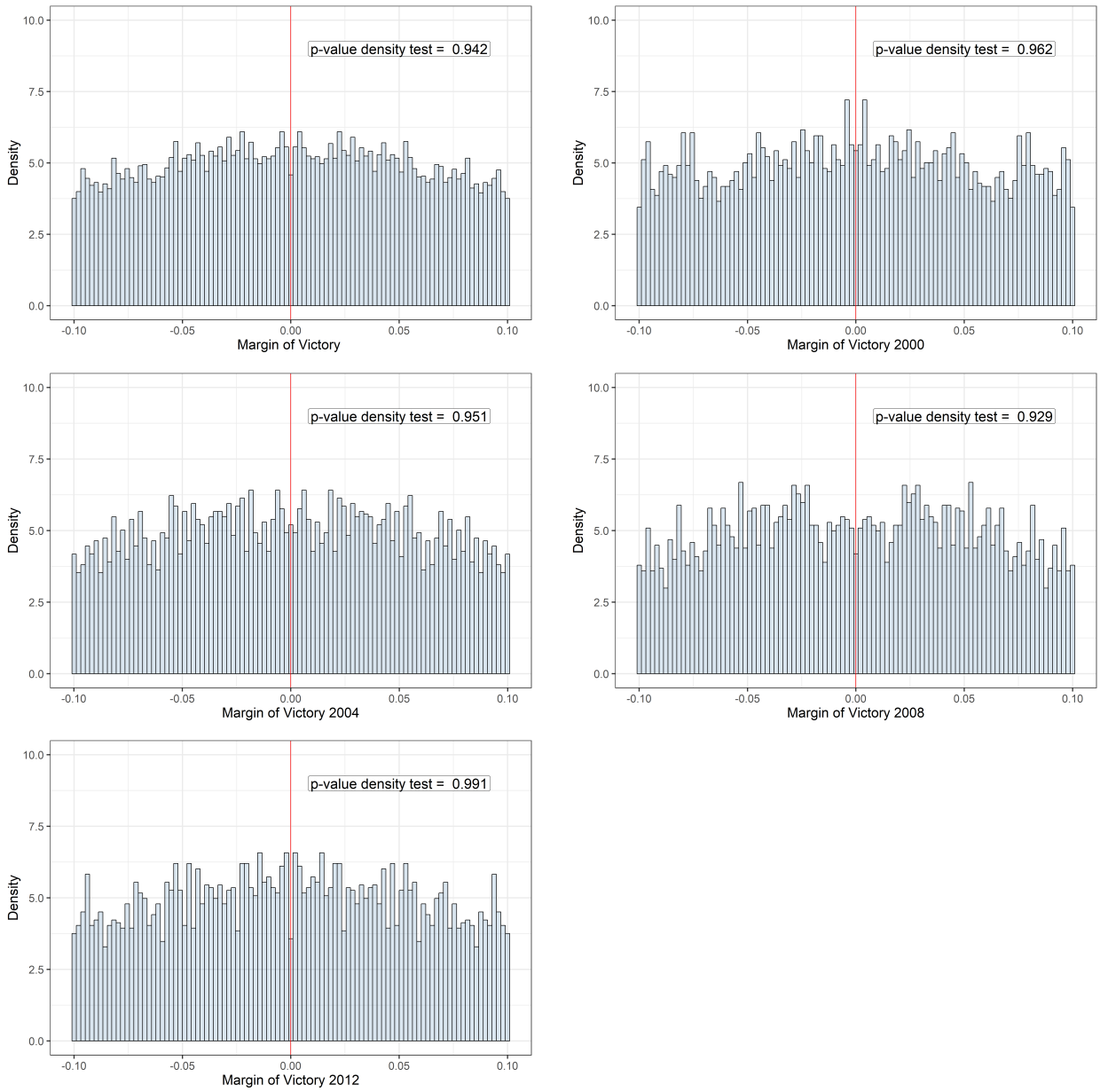


Figure 11: Histogram of margin of victory for the pooled data and divided by electoral cycle. P-value robust density tests developed by Cattaneo, Jansson, and Ma (2018) in each figure.

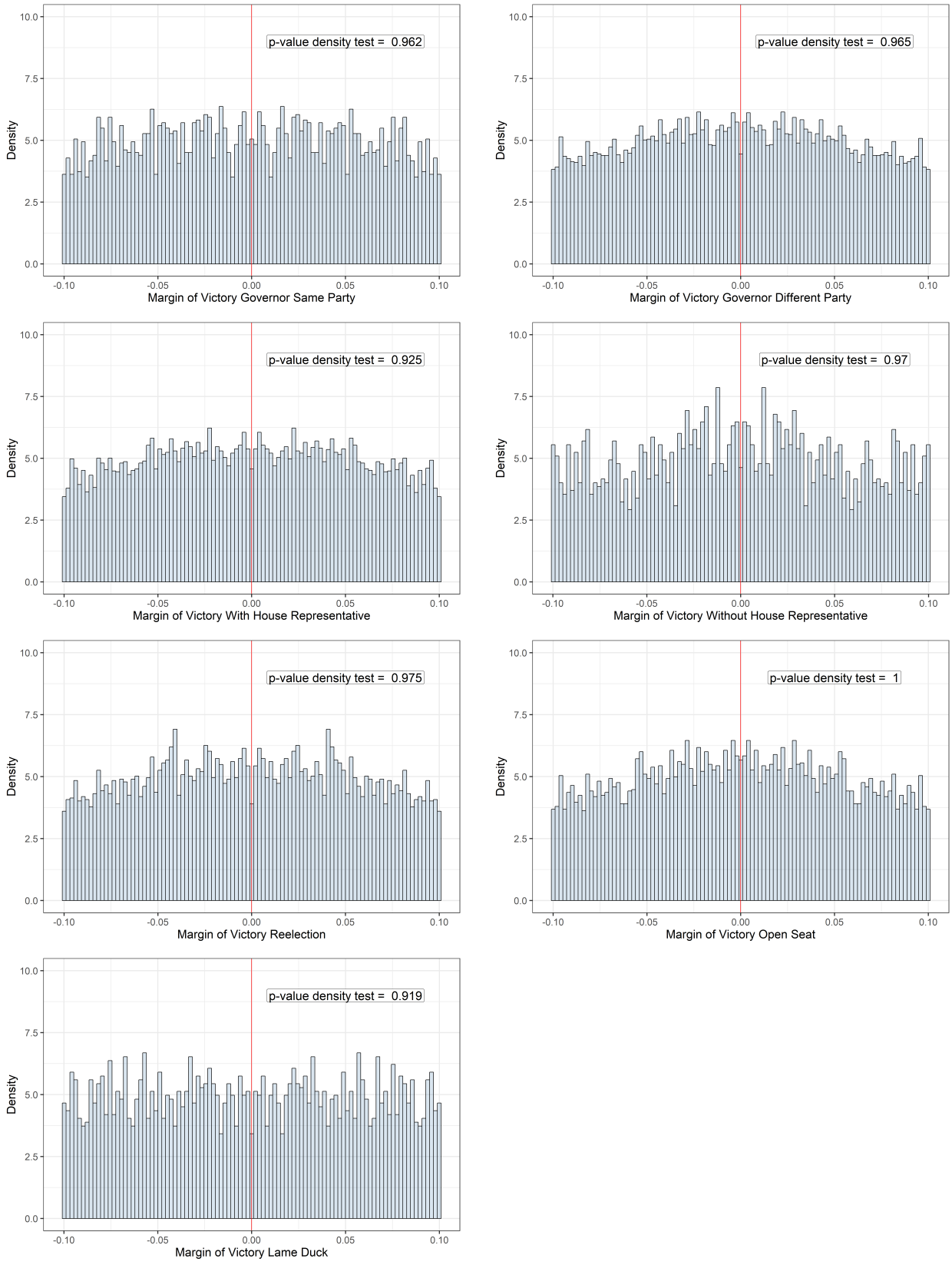


Figure 12: Histogram of margin of victory for the mediators subsample. P-value robust density tests developed by Cattaneo, Jansson, and Ma (2018) in each figure.

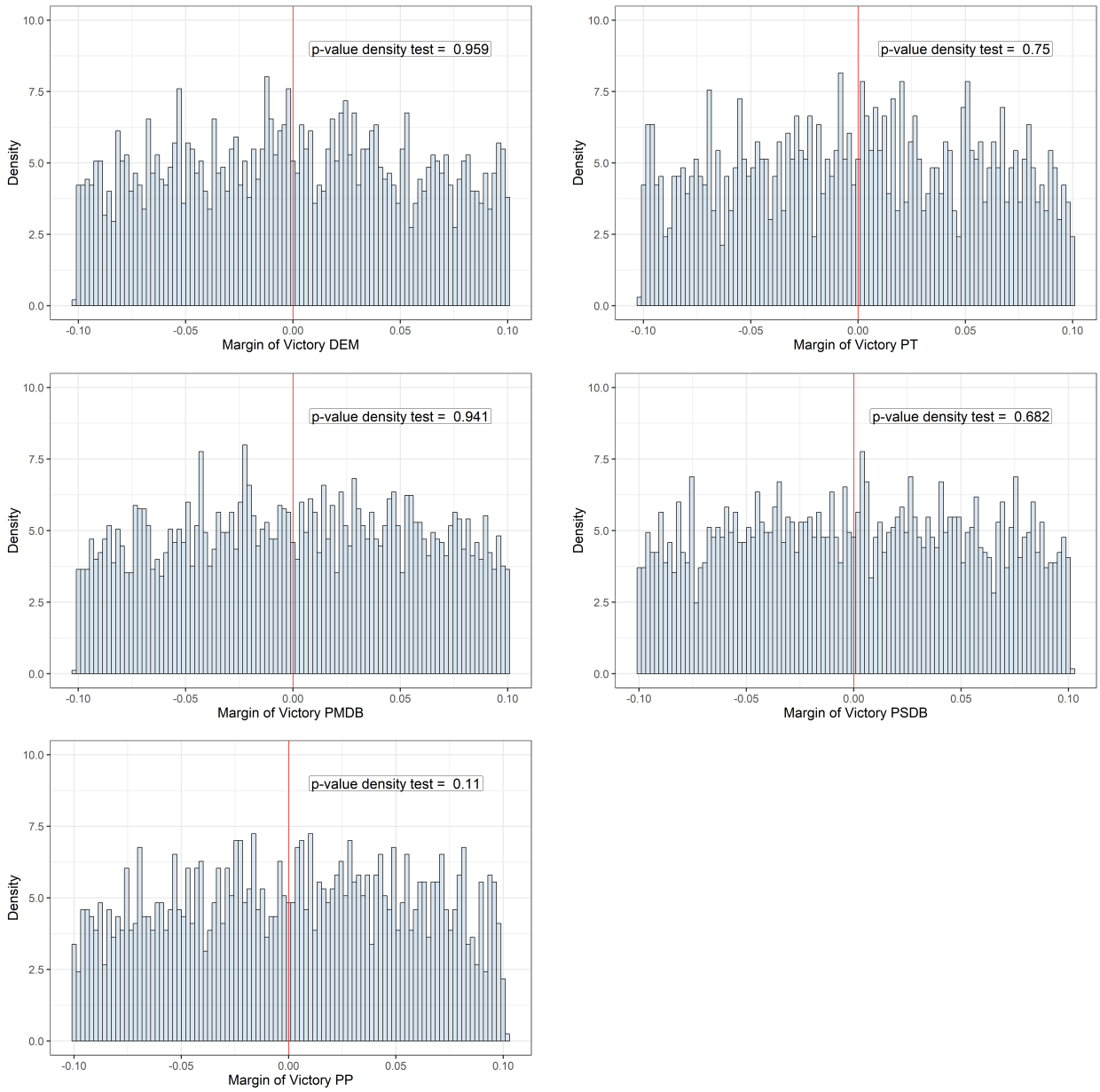


Figure 13: Histogram of margin of victory for each of the five parties with positive conditional treatment effects. P-value robust density tests developed by Cattaneo, Jansson, and Ma (2018) in each figure.

Table 7: Regression Discontinuity on Lagged Values Validity

Outcome: Vote Share Co-partisans for the House Election in the previous election						
Outcome	p-value	Estimate	Lower 95% CI	Lower 95% CI	Number of cases	Bandwidth
Lagged Vote Share: Incumbent	0.325	−0.005	−0.018	0.006	9894	14.264
Lagged Vote Share: 2004	0.071	−0.019	−0.039	0.002	3941	18.389
Lagged Vote Share: 2008	0.085	0.018	−0.002	0.038	3431	16.109
Lagged Vote Share: 2012	0.306	−0.009	−0.027	0.009	3661	16.453
Lagged Vote Share: PT	0.752	0.002	−0.027	0.038	1380	15.668
Lagged Vote Share: PSDB	0.226	0.017	−0.011	0.047	937	15.146
Lagged Vote Share: PMDB	0.040	−0.026	−0.056	−0.001	2001	14.725
Lagged Vote Share: PP	0.280	−0.018	−0.067	0.019	796	11.679
Lagged Vote Share: DEM	0.899	−0.003	−0.050	0.044	931	13.787

Appendix 7. Robustness Checks: Rate of Change of Political Boosting for House Elections and support for State legislators

.

In this appendix, to assure robustness of the findings, I re-estimate the models for the main effects and mediators using two different dependent variables in which my theory also predicts the existence of political boosting. I first estimate the treatment effect on the rate of change of the partisan vote share for House elections before and after electing the mayor; second, I use vote share of state legislators between the incumbent and the runner up's party. If winning the local matters to explain different electoral gains between the incumbent and the runner up in the upcoming election, it is reasonable to expect the incumbents' vote share *vis-a-vis* the runner up candidate also increases comparing the moment before and after winning the control over the local executive. I replicate precisely the same models of the paper and present the results below. Then, I do the same tests for State legislators election which also occur on the same day of House elections for all the States in Brazil.

The effects of political boosting for the rate of change are higher than those reported on the paper for my original dependent variable. In other words, winning the local executive generates a substantial impact on the growth of the incumbent party vote share compared to the previous House election. All the results are aggregated at the municipal level. For the mediators, the results are somewhat similar in the case of the career ambition variables, while party coordination capacity seems to explain most of the variation. In conclusion, the results confirm our theory that controlling the local matter for partisan national gains.

Regarding the political boosting for state legislators, the effects are weaker than we expected. Here, only when the mediators for party coordination exist, political boosting is detected. As I discuss in the paper, I read this result as an indication that the economy of scale I suggest shape the effects of boosting

are weaker for state legislators. In other words, these candidates are closer to the local, therefore, the mayors matter less for their electoral mobilization.

Figure 14: Smoothed Regression Discontinuity of the Treatment Effect using the variation in vote share for the incumbent's party before and after winning the local election (Red lines represent the optimal bandwidth decision).

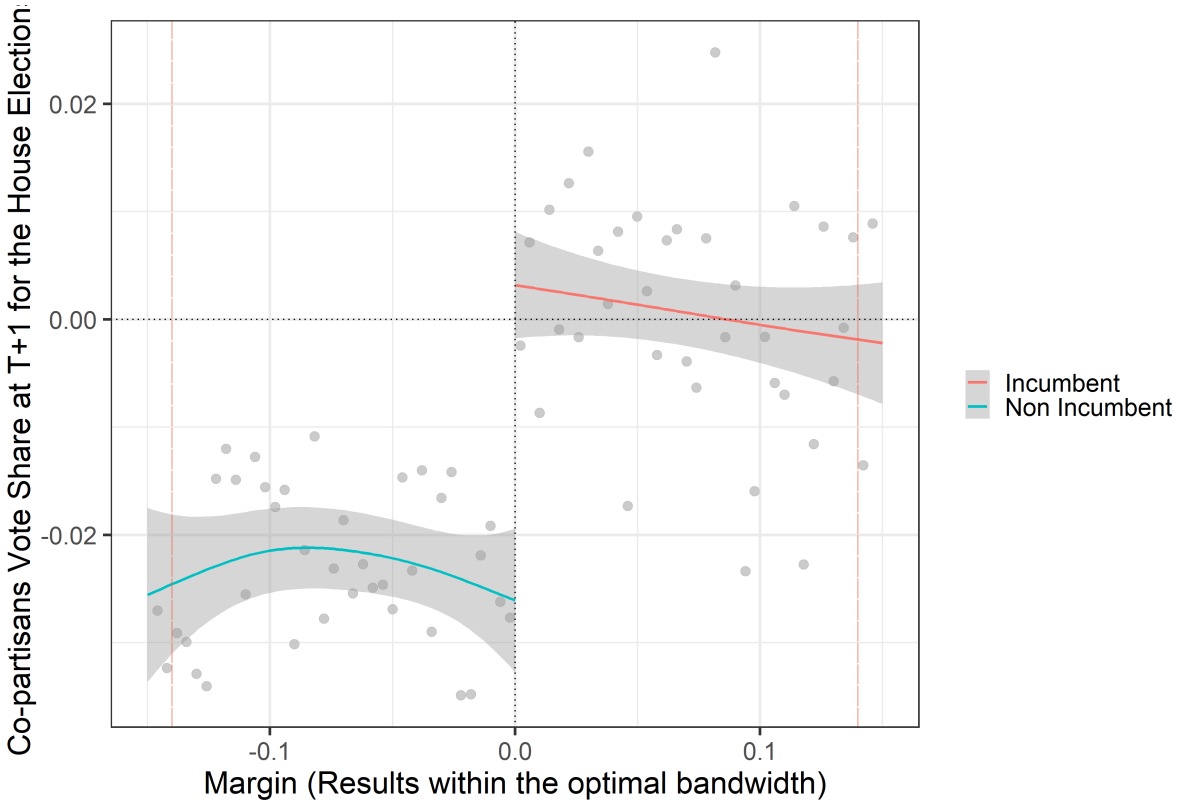


Table 8: Regression Discontinuity Results

Outcome: Rate of Change of the Vote Share of co-partisans for the House Election					
Subsample	Optimal Bandwidth)	Estimate	Lower 95% CI	Lower 95% CI	Number of cases
Incumbent	14.7	0.031	0.022	0.041	13239
Year: 2000	18.561	0.030	0.012	0.052	3637
Year: 2004	15.617	0.022	0.001	0.041	3576
Year: 2008	15.563	0.033	0.016	0.052	3352
Year: 2012	12.088	0.034	0.019	0.053	2966
Open Seat Sample	17.711	0.016	0.002	0.028	6311
Reelection Sample	11.456	0.061	0.046	0.080	4569
Lame Duck Sample	18.441	−0.014	−0.037	0.010	2552
Governor Different Party	16.817	0.033	0.023	0.043	11178
Governor Same Party	20.889	0.015	−0.003	0.036	3686
Without a House Representative	16.847	0.018	−0.003	0.040	2251
With a House Representative	16.594	0.032	0.023	0.043	12033

Figure 15: Smoothed Regression Discontinuity of the Treatment Effect using the state legislators vote share (Red lines represent the optimal bandwidth decision).

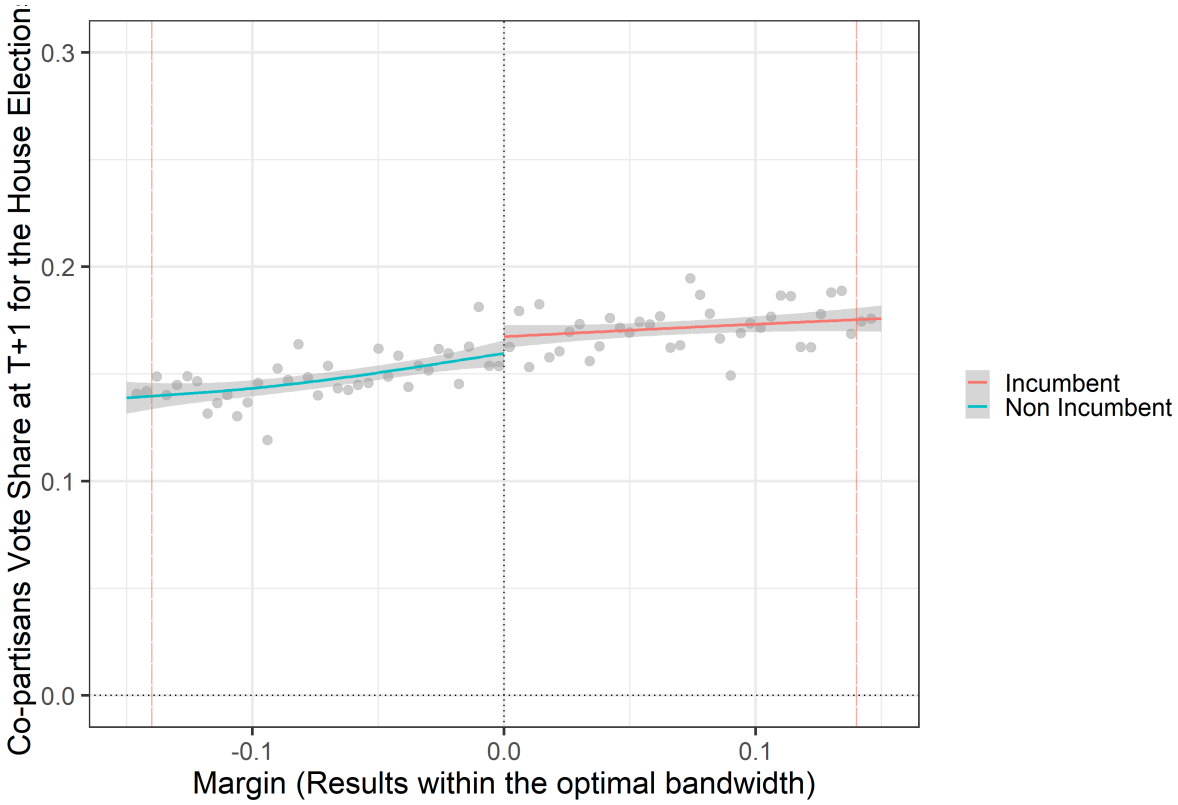


Table 9: Regression Discontinuity Results

Outcome: Vote Share Co-partisans for the State Legislators					
Subsample	Optimal Bandwidth (Margin of Victory)	Estimate	Lower 95% CI	Lower 95% CI	Number of cases
Incumbent	0.131	0.004	−0.007	0.013	12201
Year: 2000	15.826	0.001	−0.020	0.020	3265
Year: 2004	13.756	−0.006	−0.029	0.011	3305
Year: 2008	14.390	0.022	0.004	0.043	3203
Year: 2012	18.008	0.004	−0.012	0.019	3874
Open Seat Sample	13.912	0.009	−0.005	0.024	5438
Reelection Sample	13.633	−0.005	−0.022	0.009	5178
Lame Duck Sample	16.906	0.026	−0.001	0.047	2391
Governor Different Party	12.674	−0.032	−0.045	−0.023	9299
Governor Same Party	18.893	0.132	0.114	0.153	3490
Without a House Representative	20.906	−0.071	−0.091	−0.053	2539
With a House Representative	13.563	0.019	0.007	0.029	10569