Divided Government versus Incumbency Externality Effect

Quasi-experimental Evidence on Multiple Voting Decisions

Florian Ade
Ronny Freier

Divided Government versus Incumbency Externality Effect
Quasi-experimental Evidence on Multiple Voting Decisions

Abstract

This paper explores the interdependency of political institutions from the voter's perspective. Specifically, we are interested in: (1) Does the partisan identity of the mayor influence the voter's decision in the subsequent town council election?; (2) Does this partisan identity influence the vote in ensuing higher level elections?; and (3) Do voters condition their vote for the mayor on the result of the last council election? We rely on a regression discontinuity design focusing on close election outcomes based on municipal level data for Germany. We find (1) that the party of the mayor can receive a bonus of 4-6 percentage points in vote share in the next town council election (depending on the timing of the local elections). (2) The mayor partisan identity does not affect federal or European election outcomes within the same municipality. And (3), we show that voters punish mayor candidates of parties that performed strongly in earlier council elections. Throughout the paper, we explore how the findings can be related to an incumbency externality effect and to the theory of voter preferences for divided government.

Florian Ade	*Ronny Freier*
DIW Berlin	*Stockholm School of Economics*
Department of Public Economics	*Box 6501*
Mohrenstraße 58	*Sweden – 11383 Stockholm*
Germany – 10117 Berlin	*ronny.freier@hhs.se*
fade@diw.de	

This Version: August 30, 2011

1 Introduction

In modern democracies political power is divided in two different ways: between tiers of government (supranational, federal, state, and municipal levels) as well as between several institutions within a given tier (e.g. president and parliament, mayor and town council). As the policy outcomes depend on the complex interactions between these tiers of government and their institutions, we expect voters to be concerned with making the best decision within this system of political integration.

This paper studies voting behavior in the presence of interdependencies between political institutions. We focus on how voters react to the realization of the outcome in the election of one institution when they decide on another political body. Specifically, we investigate three questions: (1) Does the partisanship of the mayor influence the vote outcome of her party in the subsequent town council election?; (2) Is the party identity of the mayor relevant to the election results in subsequent higher level elections?; and (3) Do voters condition their votes in the mayoral election on the result in the town council election that was just previously held?

For elections at the local level, we find (1) that the partisanship of the mayor can matter for the election outcome of her party in the next town council election. Crucially, the results depend on the timing of the two elections. If the elections are held jointly (meaning that the next town council election is held at the same day as the next mayoral election), the party of the incumbent mayor receives a significant and sizable bonus of around 4-6 percentage points in vote share. If elections do *not* run simultaneously, however, the party of the mayor does *not* profit from holding the office. (2) From the analysis for German federal and European elections, we conclude that the partisan identity of the mayor exerts *no* effect on elections at those levels.

Finally, we show (3) that a party's outcome in a council election also affects the next mayoral election. For an additional 10 percentage points in the vote share for the town council election, the party's candidate a the run-off mayor election will see her vote share be reduced by 2.5 percentage points.

We consider two theories that apply to our analysis: an incumbency externality effect from the mayor's office and the theory of voter preferences for divided government. For the case of council elections, the first is expected to increase votes for the mayor's party while the second would decrease the voters' support for her party. The empirical evidence suggests that under specific circumstances both the incumbency externality effect as well as an effect of the preference for divided government are of importance. Under some simplification, we

argue that the incumbency externality effect is evident in the analysis of joint local elections, while the divided government effect can be observed in the mayoral run-off elections, and that both effects are present when the mayoral election precedes the council election.

The main conjecture of the incumbency externality effect is that the position of the mayor provides access to resources that determine the election outcomes for other institutions. Such resources can be both direct financial resources[1] as well as non-monetary aspects, such as media presence for the party and time spent with the electorate during the election period.[2] For example Hainmueller and Kern (2008) find that in mixed electoral systems a party can increase its vote share in the proportional vote when it provides the direct representative of the electoral district, thus identifying an incumbency externality effect of about 1.5 percent. Folke (2010) presents evidence for a spillover effect from government personnel nominated by the patronage system to their party. In spirit, the incumbency externality effect is similar to the well-documented electoral advantage that incumbent office holders receive (see Lee (2008) for US house representatives, Ferreira and Gyourko (2009) for US mayors, Freier (2011) for German mayors).

It is important to note, that the size of the incumbency externality effect can be different depending on the precise timing of the interacted elections. When elections are held at the same time, the election campaign effort of a party can, for example, capitalize on synergy effects or on a positive personality of a candidate running for the other election (e.g. Mondak (1990)).[3] Moreover, psychological aspects (sometimes referred to as coattail effects) may also apply to jointly held elections. Mondak and McCurley (1994) find that to increase their cognitive efficiency voters employ their evaluations of presidential candidates in the US to decide which House candidate to support. They also argue, that an individual might be inclined to cast her vote to the same party in both elections simply because splitting the vote on the same election day might increase psychological costs.

The theory of preferences for a divided government prescribes a strategic rational to vote

---

[1] In local German politics it is, for example, very common that elected officials contribute part of their wages/compensation directly to the account of their party. As funding of political campaigns is not very lavish in German politics, particularly for local politics, the funds offered by elected officials often make for a major share of the campaign budgets in the local elections.

[2] Apart from managing the administration of the community and framing public policy, the schedule of German mayors is packed with many social events in the community. At openings of kindergartens, elderly homes, and bowling alleys, the mayors not only promote themselves but is also expected to advertise their party.

[3] It is a reasonably well-documented fact, for example, that electoral campaigns for positions in the US Senate and the House of Representatives have better prospects for individual candidates, if their campaigns run simultaneously with the election for a US president and this race has a charismatic, winning presidential candidate from their party (e.g. Campbell and Sumners (1990), Ferejohn and Calvert (1984)). The same is shown to apply to gubernatorial elections, e.g. Hogan (2005).

for different parties in elections for distinct institutions to establish a political balance. This effect works in the opposite direction as the incumbency externality effect (see Alesina and Rosenthal (1996), Kern and Hainmueller (2006)). The voter is presented with two distinct opportunities (mayor and council election) to make a decision on the political actors governing the community. A rational median voter might prefer to hedge against extreme policy positions in local government by splitting the vote in the two elections and dividing the governmental power between different parties. By electing, for example, a conservative mayor yet a social democratic council the median voter can assure that policy outcomes will be balanced within the ideological spectrum.

At the Federal level in the United States, it is well established that there is a preference for divided government (or electoral balancing). The president's party typically loses seats during Congressional midterm elections (see e.g. Erikson (1988)). The same pattern can also be observed for other countries and at other levels of government (e.g. Norris and Feigert (1989), Folke and Snyder (2010)). While it is straightforward that divided government can occur in a presidential democracy (when the president and the majority in parliament are not from the same party), divided government can also exist in parliamentary systems. Here, divided government can occur via a second chamber of parliament. Kern and Hainmueller (2006) present evidence for electoral balancing in the German parliamentary system at the federal level.

The timing of the elections is decisive in to what extent the voter can engage in such strategic behavior. When elections are held sequentially, the voter can actively condition her vote on the outcome of the first election. If, for instance, a conservative candidate wins the mayoral election, the voter can take a deliberate decision to elect social democrats in the next council election. When the elections are held simultaneously, however, the incentives to behave strategically in this way are blurred as the outcomes of the respective elections are uncertain (see Alesina and Rosenthal (1995)).

Our analysis is based on a unique and detailed data set of election results at the municipal level in Germany. Overall, we use more than 9,500 elections in the analysis for interdependencies at the local level. To evaluate the effect on higher levels of elections we rely on 18,000 observations for the European elections and 35,000 German federal election observations.

Methodologically, the main analysis relies on a regression discontinuity design to draw causal inference. The crucial argument is that we can identify exogenous variation in the party identity of the mayor by focusing on close election outcomes. We assume that the specific

election outcomes are subject to some random component (e.g. shocks in the popularity of the party due to recent media coverage, shocks in turnout due to the weather, see Knack (1994.)). Given that a mayoral election was a very close race, the outcome of that election (the party identity of the mayor) can be seen as a quasi-random event. Lee (2008) was the first to use this methodology to investigate the incumbency advantage for individual seats held by a party in the US House of Representatives.[4] Following Lee's analysis, Ferreira and Gyourko (2009) study the effects of incumbency on a large number of mayoral elections in US cities. They find an effect of incumbency of about 32 percentage points for the probability of reelection. For German mayors, Freier (2011) also documents a substantial party incumbency effect in the order of 38-40 percent in the probability of winning the mayor office.[5]

Following the growing body of literature using the regression discontinuity design in election analysis, Caughey and Sekhon (2010) reassess the study of Lee (2008) and show that the identifying assumptions of the design do not necessarily hold, thus, casting a shadow of doubt on the identification strategy.[6] However, our detailed tests on the validity of the RDD show no sign that the credibility of the identifying assumptions are of concern in our analysis.

Apart from the main RDD analysis, we consider an additional design that makes use of the precise timing of local elections in the German electoral system. This design relies on second ballots (run-off elections) in cases where the first round of the mayoral race

---

[4]Lee analyzes the effect of obtaining the incumbency status for a party on the probability of winning the race for the district in the present election. He estimates the incumbency advantage to constitute a 45 percentage points higher probability of winning the race for a district. With his quasi-experimental approach, he is the first to produce reliable estimates of the magnitude of the effects of incumbency. Note that, Pettersson-Lidbom (2008) was the first to use the idea of close elections to identify party effects in spending of Swedish municipalities.

[5]There are several other studies applying the ideas of the regression discontinuity design to elections. E.g. Chamon, de Mello, and Firpo (2009), Titiunik (2009) for Brazilian municipalities, Eggers and Hainmueller (2009) for Members of Parliament in the UK, Linden (2004), Uppal (2005) for Indian parliamentary elections, Miguel and Zahidi (2004) for national elections in Ghana, Meyersson (2009) for Turkish municipalities, Gagliarducci and Nannicini (2009) for Italian politicians, see Caughey and Sekhon (2010) for an overview.

[6]They show that there are significant predetermined differences between candidates who just won and just lost their races. Among those differences are that marginal winners have more campaign money, are more likely to belong to the party that won the last election, and are more often the predicted winner in the journal "Congressional Quarterly". The authors conclude that tight elections might differ from other elections, as their tightness is often known to both campaigners and voters, hence, e.g. more money is directed at those tight races since its expected payoff is highest. While these arguments are of great importance for US-House elections which are characterized by detailed media coverage, a multitude of polls and high campaign spending, these arguments do not apply to the elections we investigate: as they are at the local level that is characterized by very small districts (communities and cities) there are hardly ever polls available, relatively low campaign spendings, and media coverage is usually limited to the local section of the regional newspaper.

is held simultaneously with the council election and no candidate received the necessary majority. This provides an unique opportunity to evaluate whether the voter subsequently conditions her vote in the run-off election on the outcome of the town council election. Identification can be obtained by conditioning on the results of the first round of the mayor election. Thereby, we can exclude unobservable characteristics (individual mayor candidate characteristics, popularity shocks for individual parties, etc.) and consistently estimate the impact of the town council election result. Given that there is no incumbency externality effects at work, the result can be attributed only to a preference consideration of the voter.

The remaining part of the paper is structured as follows: Section 2 introduces the basic features of local government in Germany that are relevant for the empirical analysis. Section 3 highlights the characteristics of the underlying data. In section 4, we describe the empirical methodology and derive the empirical model, before section 5 presents the results. Conclusions are drawn in section 6.

2 Institutional background: voting rules and local government

To understand the implications of the local institutional features for our analysis we present information on the voting systems and rules as well as the key properties of local governments in Germany. The results of this paper are based on elections in eight of the sixteen German states for which data is available and institutions are comparable (see table 1 for an overview of the states and elections included).

2.1 Voting system and rules

There are two important elections at the community level in Germany: elections for the position of the mayor and elections for the members of the town council. In all states under consideration the voters in a community elect the mayor through a majority vote and the council in a proportional election.

In the majoritarian election for the mayor, a party gains control over the mayor's office if the candidate obtains a simple majority of the votes. Mayoral elections are held every five to eight years[7] and are held on fixed state-wide dates in Bavaria, Thuringia, and Brandenburg (part-time mayors) and on community-individual dates in the other states.[8] Another feature that is crucial for the correct implementation of our design is the presence of a second (or

---

[7]Five years in Brandenburg (part-time mayors) and Thuringia, six years in Hesse and Bavaria, seven years in Saxony and Saxony-Anhalt, eight years in Saarland, Rhineland-Palatinate, and in Brandenburg (full-time mayors).

[8]In Brandenburg the elections of the full-time mayors are also community-individual.

Table 1: Data set - number of observations and elections

	Design 1	Design 2	Center -Left	Center -Right	# communities in 2001	# of election dates[a]
Bavaria	6595		2421	4174	2100	5
Thuringia	495		120	375	992	2
Brandenburg[b]	156	52	109	99	420	1
Rhineland-Palatinate		93	45	48	212	2
Hesse		1399	768	631	430	3
Saarland		108	54	54	52	2
Saxony-Anhalt		280	103	177	1118	1
Saxony		398	115	283	520	1
Total	7246	2330	3735	5841	5844	17

a The council election years are: Bavaria 1984, 1990, 1996, 2002, 2008; Thuringia 1999, 2004; Brandenburg 2008; Rhineland-Palatinate 1999, 2004; Hesse 1997, 2001, 2006; Saarland 1999, 2004; Saxony-Anhalt 2004; Saxony 2008.

b: Brandenburg is the only state in which we have observations in both designs. This is, because larger cities have different term lengths for the mayor and the council election whereas in small communities the term lengths of mayor and council coincide and elections are held simultaneously. *Source*: Own calculation based on the data provided by the state offices for statistical services.

run-off) ballot. If no candidate reaches the majority of 50% in the first ballot, a second ballot is held between the two leading candidates.[9] As the second ballot determines the victor in the mayoral race, we use the results of this second ballot if applicable. It is also interesting to note that it can occur that only one candidate runs for office and that, in some cases, parties can join forces to nominate a common candidate (see detailed description in the data appendix).

Council elections are held every four years in Rhineland-Palatinate, five years in Hesse, Saarland, Saxony, Saxony-Anhalt, Brandenburg, and Thuringia[10], and six years in Bavaria. Hence, in some states the term length of the mayor and the council are identical. In these states (Bavaria, Thuringia, and in communities with non-salaried mayors in Brandenburg) the elections for mayor and council are held on the same day. We denote these states as "Design 1" and the states with asynchronous elections as "Design 2". The choice of the two designs is exogenous to our analysis: First, the term length of local institutions and the alignment of elections is subject to state legislation. Hence, the communities have no direct influence on these rules. Second, the choice of a asynchronous or harmonized election scheme typically dates back to the time when the states were founded or results from legal obligations. Rhineland-Palatinate, Hesse, and Saarland introduced the direct election of

---

[9]In the state of Saxony all candidates can join in the second round of a mayor elections and the candidate with the largest vote share wins the office.

[10]In Hesse, prior to 2001, every four years.

the mayor only in the 1990s. When doing this the communities had to honor the term of the previous, purely administrative, heads of the local administrations. Their terms were of different length and ended at different dates hence causing mayoral and council elections to be asynchronous until today.

In the analysis we also consider the effect of the mayoral position on elections in other tiers of the state structure. Specifically, we look at elections for the European parliament and the German federal parliament (*Bundestag*). It is important to note that those higher level elections are hardly ever simultaneous with the local elections in any of the data that we collected.

2.2 Local government

The political and administrative structure of Germany is organized into four tiers. The four tiers consist of the federal level, sixteen states, about 450 counties and just over 12500 municipalities (as of 2006). While the federal government is involved in nearly all branches of state activity, the remaining tiers each have specific responsibilities.[11] The municipal level takes direct responsibilities for the provision of public goods in the areas of child care, cultural expenditures, sport and recreational facilities, local infrastructure investments as well as a number of minor tasks. Communities also often oversee public firms that deliver local services (e.g. energy and water supply) and administer mandated spending allocated by higher tiers (like social services, investment in schools, and certain infrastructural investments). In total, the local governments administrates and oversees about two thirds of all state investment. About 40 percent of all state employees work for municipalities.[12]

The mayor and the council jointly govern the community. The responsibilities of the mayor include the daily administration of the community, leading personnel, representing the community and making urgent decisions. The council is in charge of all general decisions of both legislative and administrative quality. It follows that, on the one hand, the mayor needs the support of a majority in the council to realize her policy goals and, on the other hand, any council majority will have a hard time pushing through projects against the will of the mayor. This is true in all states considered in this analysis. The exact distribution of power between the two political institutions at the local level is regulated by state law that specifies the precise nature of the interaction (e.g. does the mayor have an active voting right in council meetings).

---

[11]Education, for example, is in the responsibility of the state, whereas counties are generally concerned with the administration of public order issues (police, fire rescue) and health (hospitals, ambulances).

[12]See Bundesbank (2007).

While these regulations differ in each state, we stress that these differences in legislation would only matter to our analysis if they were recognized by the voters. A clear indication that this is not the case is that the electoral advantage of mayors in different states is generally very similar (around 38-41 percentage points). Freier (2011) studies the mayor incumbency advantage for German mayors in more detail. He finds no significant difference for the incumbency effects in different German states. However, we further test this assumption implicitly in the results section.

3 Data and descriptive statistics

We use data from the German states of Hesse, Rhineland-Palatinate, Bavaria, Saarland, Brandenburg, Saxony-Anhalt, Saxony, and Thuringia. We obtained information on both mayor and town council elections for several years from the statistical office of each respective state. Data for further years and the remaining German states was either not available or cannot be used due to differing institutional design (e.g. in the city-states). We also collected data on the results of European and German federal elections at the municipal level. On the basis of this data set, we are able to link the specific election outcomes in a municipality with information about the party identity of the mayor at the time when the election was held.

For all mayoral elections in the data we have the number of valid votes as well as the result for each candidate in the race for the mayor's office. From this data we can extract the exact vote share for each candidate (party) and identify how close each election was. For all other elections (town council, European and federal elections), we collected the number of valid votes as well as the votes for each participating party. For the analysis, we are interested in linking the results from the mayoral races with the outcomes from the other elections. We limit the analysis to results for the center-left (SPD - Sozialdemokratische Partei Deutschlands) and the center-right (CDU/CSU - Christlich Demokratische Union/Christlich Soziale Union) party because they are the only parties that participate regularly at all levels of the political structure (for further details see data appendix).

Table 1 provides an overview of the main data set in which we link each town council election with the preceding mayoral election in the respective community. Overall we have 9,576 such election pairs in 5,844 communities. The number of observations is derived as follows: In each state, we observe a number of town council election dates (see column 6) when each community (see column 5) in the state elects its town council. In Hesse, for example, we observe community councils elected three times (1997, 2001, 2006) in all 430 communities. For each election we match the preceding mayoral election to the data. One

observation for our analysis now is given if we can identify either the center-left or the center-right party to have participated in the town council election *and* there is a result for a candidate of the respective party also in the preceding mayor election (see data appendix for detailed description of the data limitations). It is important to note, that a particular community may thus generate two observations in one election (one for each of the two parties).[13]

As discussed in the section on the institutional background, some states hold council and mayor elections simultaneously (see observations in column 1 denoted 'Design 1') while other have differing term lengths and, hence, different election days and years (column 2 denoted 'Design 2'). We have a total of 7246 election pairs in Design 1 and 2330 in Design 2. Overall the state of Bavaria plays a central role in our data set with 6595 of the 9576 observations. This is due to good data availability combined with a large number of communities.

Table 2: Data set - defining different samples

	Design 1			Design 2		
	Total	Center -Right	Center -Left	Total	Center -Right	Center -Left
# of observations						
total	7246	4626	2620	2330	1215	1115
within 60% margin	5677	3371	2306	1968	1011	957
# of narrow observations						
within 5% margin	514	335	179	201	111	90
within 2% margin	219	140	79	83	45	38
within 1% margin	95	59	36	43	24	19

Notes: 'Margin' in this table means the difference in percentage points between the winner and the best opponent. For example, in the case of only two candidates a margin of 5 percent means that the winner may have gotten at most 52.5 percent while the other got 47.5 percent. *Source*: Own calculations based on the data provided by the state offices for statistical services.

Table 2 presents the number of observations in different subsamples relevant for our analysis. We estimate our models within the 60 percent margin of victory in the mayor election to make sure that extreme cases do not drive our estimates at the threshold via the control function. That means that we exclude all cases where the difference between the candidate

---

[13]Note that those two parties are seldom the only parties running in mayoral or council elections. To respect the fact that the results are nevertheless linked within the same council election, we cluster all standard errors in the analysis at the level of each election within the municipality.

with the most votes and her best opponent is more than 60 percent.[14] Applying this limitation we lose 1569 observations, of which most are cases where only one candidate ran for office. Table 2 also gives information on the observations that represent tight races in the mayor election. For example, in the case of only two candidates a margin of 5 percent means that the winner received less then 52.5 percent while the challenger got more than 47.5 percent. For Design 1 we have 514/219/95 tight races in the 5/2/1 percent margin respectively. For Design 2 there are 201/83/43 tight races in the three margins, respectively.

To test the effect of the incumbent mayor's party on higher level elections, we collected data on European and German federal elections in the municipalities. The numbers of observations and the corresponding elections years are presented in table 6 in the appendix. As in the case of council elections, the total number of observations is different from the product of the communities times the election date for the same reasons.[15] In general, data for higher government level elections is almost always available (in contrast to local election results). We have a total of 18,000 observations for the European elections and 35,194 observations for the federal elections. Note that the number of actual observations in the estimation tables will be lower as we exclude elections pairs with extreme mayor elections (more than 60 percent margin of victory). European elections are held every five years and federal elections every four years. Both elections are almost never held simultaneously with local elections.

4 Empirical model and methodology

The main empirical strategy is based on a Regression Discontinuity Design (RDD). The treatment under consideration is the mayor incumbency of a party. We denote treatment with the dummy variable $d_{i,t}^p$, where i refers to the community or town for which we observe elections, t refers to the election year[16] and p refers to the party under consideration. The variable that unambiguously determines treatment in our application is the margin of victory, $v_{i,t}^p$, which is defined as the distance in the vote share of party p in the mayoral election in t to the best opponent. The decisive threshold that the so-called score variable

---

[14]For clarity, consider the following example. In the case of only two candidates we exclude cases where the winner has more than 80 percent of the vote (hence, the second best opponent has 20 percent and the margin of victory is 60 percent). It would be hard to argue that these cases could contribute anything to (even via a flexible control function) the tight races in which we are interested in.

[15]However, we did not have to exclude any observation for non participation in the European/federal election as both parties participated in these elections in all communities.

[16]The timing with the t index is specific. The incumbency status of a certain mayor is determined in $t-1$. Henceforth, the party's candidate holds the relevant mayor position, which is indicated by $d_{i,t}^p$. We use the index t (instead of $t-1$) to illustrate that the mayor holds this position also when the next election is held in t.

$v_{i,t}^p$ has to cross is $v_0 = 0$. The relationship between $d_{i,t}^p$ and $v_{i,t}^p$ is as follows:

$$d_{i,t}^p = 1\left[v_{i,t-1}^p > 0\right] \tag{1}$$

This implies that the incumbency is earned if the party won first place in the preceding election. The outcome variables that we are interested in are denoted $y_{i,t}^p$. In particular, we consider the election outcomes of party p in the subsequent elections. Firstly, we measure the vote share of the party in the following town council election. The object of interest here is the evaluation of the effect of an incumbent mayor on the election results in the following town council election for her party. Secondly, we investigate the effect for the party in subsequent elections for both the European parliament and the *Bundestag* (German Parliament).

To estimate the causal effect of incumbency of a party on subsequent election outcomes we follow Pettersson-Lidbom (2008), Lee (2008) and Hainmueller and Kern (2008) (see also Hahn, Todd, and van der Klaauw (2001)). The basic model is given by:

$$y_{i,t}^p = \delta_0 + \delta_1 d_{i,t}^p + \epsilon_{i,t} \quad for \quad |v_{i,t-1}^p| < \Delta \tag{2}$$

where δ_1 is the parameter of interest. This approach is often denoted as a limited sample approach. The regression relies solely on observations within some small margin Δ close to the threshold. As an alternative to the limited sample approach, we also implement a RDD using a control function design:

$$y_{i,t}^p = \delta_0 + \delta_1 d_{i,t}^p + h(v_{i,t-1}^p, \theta) + \epsilon_{i,t} \tag{3}$$

where $h(\cdot)$ is some function that represents the influence of the margin of victory in the preceding mayor election on the voting outcome $y_{i,t}^p$. In this alternative specification, we run the estimation on the entire sample.[17]

The RDD makes use of quasi-experimental variation by focusing on the observations just around the threshold. For observations close enough to the threshold v_0 the reasoning is that they are comparable and differ only in the treatment status. Essential for the validity of the RDD are two arguments. First, the final vote count leading to the margin of victory $v_{i,t-1}^p$ must contain a random component. If it were true that a party had absolute certainty about the final vote outcome, the design would be invalid. Instead of

---

[17]In practice, we also limit the sample somewhat by excluding observations where the margin of victory is greater than 60 percentage points. Outside this range, observations are often based on elections where only one candidate contested the election. When we include those non-meaningful observations in the analysis, the control function approach shows greater dependence on the specific function that is chosen.

producing good counterfactual observations on the left and right of the threshold, perfect foresight of the election outcome would enable parties to sort around the cut-off point. The estimates might be affected by a selection bias. We believe, however, that close elections are reasonably subject to some random component (e.g. weather conditions, shocks in the party popularity on an aggregate level and other factors that determine outcomes on the election day, see Knack (1994.)). Secondly, for the RDD to hold, the margin of victory must have a continuous density. The example that Lee (2008) uses to illustrate the need for this assumption is electoral fraud. If one party had the ability to administrate electoral fraud and influence the final vote count in their favor, then in repeated draws the density of the margin of victory would be discontinuous around the threshold. We have no suspicion that electoral fraud is of any concern in local elections in Germany.[18]

By implementing the limited sample approach, one makes direct use of the counterfactual observations argument. By focusing solely on observations that are within this margin, one isolates the causal effect of the treatment. If it was a random event determining the election outcome, there should be no characteristic (be it observable or unobservable) that differs for the observations just to the right and left of the threshold.

The inclusion of the control function $h(\cdot)$ allows for estimation to proceed using the entire sample while maintaining the same identifying assumptions. A correctly specified control function will capture any correlation of $d_{i,t}^{p}$ with the error term that might be of concern. The issue is that a misspecified control function leads to an inconsistent estimate of the treatment effect. For the control function in our application, we used different parametric polynomial specifications. We offer more detail on how the results depend on the specification of the control function in the results section. Note that all control functions that we use allow for different parameters on either side of the threshold.

In the analysis of the effect of the mayor position on subsequent town council elections we distinguish between two designs relating to the timing of the local elections. Design 1 uses data only on simultaneous mayor and council elections while Design 2 is based on asynchronous elections. The empirical RDD strategy is, therefore, implemented using two distinct subsamples. The distinction is not relevant when analyzing the effect on higher level elections as these elections are never held simultaneously with local elections.

To complement the analysis, we also implement a completely different design in the final part of the empirical investigation. This last part does not rely on a RDD but makes

---

[18]Figure 5 in the appendix highlights the frequencies observed in the data. Focusing on just the region around the threshold, there is no indication that the frequencies are abnormally high or low just right or left to the deterministic cut-off-point.

use of another distinct institutional feature in the German electoral system. When mayor elections and town council elections are held simultaneously, there is the chance that the mayoral election is undecided and that a run-off election is needed, which takes place one or two weeks after the first round. By that time the result of the town council election is publicly known. Therefore, the voter can condition her vote on the outcome of the town council election in the second ballot election. We examine the impact of the town council election results on the mayoral run-off election. This can be done consistently because we can use the results from the first round of the mayor election to control for all unobservable characteristics (individual mayor candidate characteristics, popularity shocks for individual parties, etc.) that will affect the election results.

5 Results

In this section we highlight the empirical findings. First, we present our estimates from the RDD to identify whether voters are influences by the mayor's party identity when they elect the local town council. Here, we focus on the implications of the timing of local elections. Secondly, we turn to the estimation of the partisan effect of mayors on elections in higher tiers of the federal structure (German federal elections and European elections). Thirdly, we investigate the additional design in which we analyze the impact of town council election results on run-off mayor elections.

5.1 Effect of the mayor's political identity on subsequent town council elections

Table 3 highlights the results from the RDD analysis of the causal effect of the mayor's office on the subsequent town council elections. The object of interest, thus, is the election outcome for a party in the town council election after a candidate from that party has either narrowly won or lost the preceding mayoral election. As indicated above the analysis is conducted for observations in Design 1 and Design 2 respectively.

Panel 1 in table 3 shows the effects when mayor and town council elections are held simultaneously (denoted Design 1). In columns 1-6 we present different implementations of the RDD using both the limited sample approach (columns 1-3) as in eq. 2 and the control function approach (columns 4-6) as in eq. 3. We observe a sizable and significant effect of around 4.3-5.9 percentage points throughout almost all specifications. This implies, that (on average) a party receives a 4-6 percentage point bonus in vote share in the town council elections when its candidate just won the preceding mayoral election.

The highest effect, with 5.9 percentage points in vote share, is indicated in the limited sample with a 5% margin (column 1), which is arguably still a relatively large margin

Table 3: Interdependency between mayor's office and town council elections

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Vote share in TCE in t					
	Panel 1 : Design 1					
d	0.059***	0.056***	0.025	0.052***	0.043***	0.047***
	(0.009)	(0.015)	(0.024)	(0.019)	(0.006)	(0.011)
N	514	219	95	514	5677	5677
R2	0.08	0.07	0.01	0.08	0.26	0.26
	Panel 2: Design 2					
d	0.029*	0.009	0.010	0.015	0.021**	0.010
	(0.015)	(0.022)	(0.033)	(0.029)	(0.010)	(0.017)
N	201	83	43	201	1968	1968
R2	0.02	0.00	0.00	0.03	0.27	0.27
Sample	5 %	2 %	1%	5 %	60%	60%
Control function	none	none	none	linear	linear	3rd order

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of a party (which had a candidate in the preceding mayoral race) in a town council election. The regressions in columns 1-3 are based on a limited sample within the respective margins and include only a constant and the treatment dummy. The estimations in columns 4-6 include a polynomial control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

around the threshold. The result of using a 2% margin (column 2) shows a similar effect of 5.6 percentage points. The only insignificant estimate is displayed in column 3 for the 1% margin, based on very few observations. With the control function approach, the estimates are similar in size. We find an effect of 4.3 (4.7) percentage points using the full sample and a linear (cubic) control function (see columns 5 and 6). Column 4 combines the limit sample and control function approach and estimates the effect using a linear control function within the 5% margin sample. Again the effect is of a similar size and statistically significant. The effect is visualized in Figure 1, which illustrates a clear discontinuous jump at the threshold. Table 8 in the appendix shows that the effects are present irrespective of the individual parties.[19]

The magnitude of this effect is sizable. Depending on the size of the council, an effect of around 5% in vote share implies that the party holding the mayoral position earns 1 or even 2 additional seats. To further illustrate the importance of the effect, we present estimates

---

[19]Point estimates illustrate that the effect might be slightly higher for the center-left party (social democrats), but such differences are not statistically significant. Note that most of the data for Design 1 are from the state of Bavaria, which is traditionally a conservative party stronghold. This might explain why the effect is slightly higher when a social democrat narrowly makes it into office.

Figure 1: Design 1
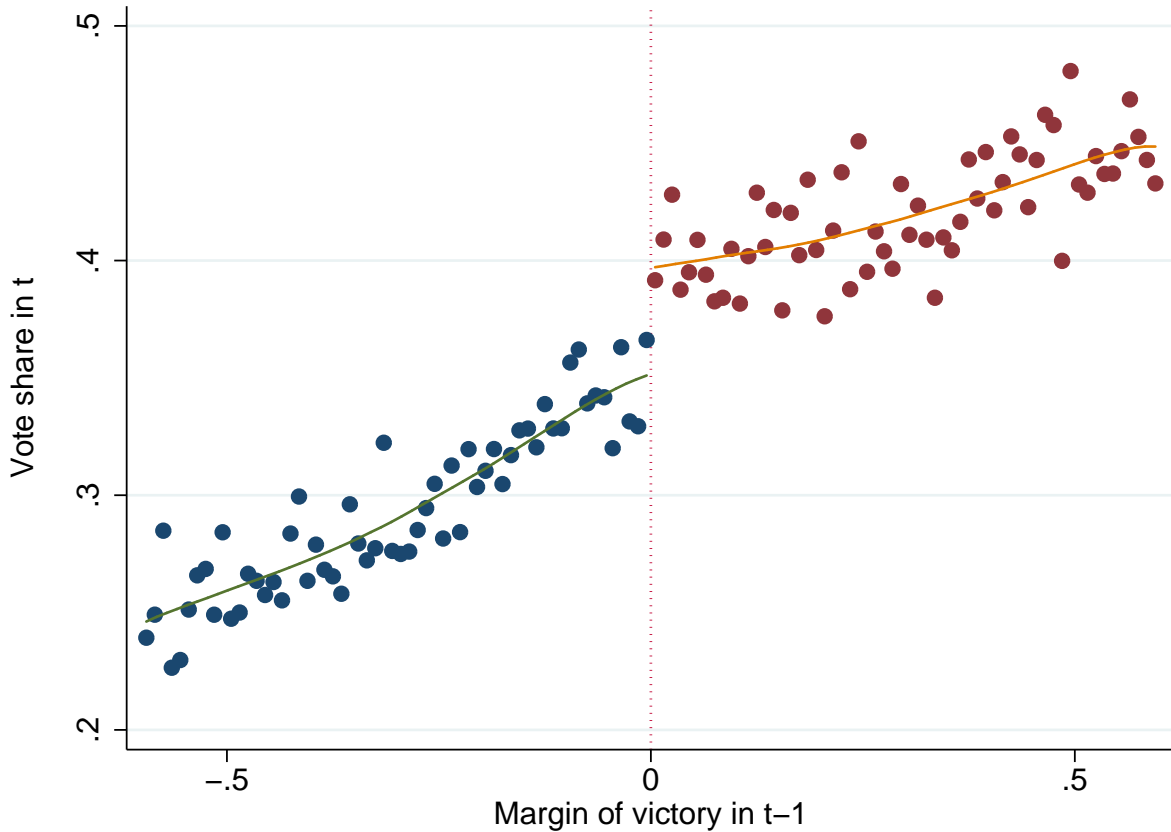
Notes: This figure graphically illustrates the discontinuous jump of the vote share in the town council election for party x when party x narrowly won the mayor election in t-1. The data used are those from Design 1. For clarity the data have been grouped in bins, each bin representing an interval of 1 percent in the margin of victory. The line fitted onto the scattered data is based on a local kernel regression using endogenous Epanechnikov weights. *Source:* Own calculations.

using the probability of becoming the strongest party in the council (instead of vote share) as an alternative outcome variable. This is particularly interesting as the party holding the most seats in the council is traditionally the first mover when it comes to coalition formation and policy making. Table 7 in the appendix illustrates the estimates from this specification. The evidence points to an effect as large as 20% percentage points in the probability of becoming the strongest party in the council.[20]

The pattern of the mayor influence is completely different for observations in Design 2, in which mayor and town council elections are *not* held simultaneously. The estimation results for this subset of the data are presented in table 3 (Panel 2) and figure 2. We find that when

---

[20]Note, that those results are not quite as stable to different specifications of the RDD. In our three preferred specifications we observe an insignificant 18.3 percentage point effect (column 4), one effect of 17.6 percentage points (column 2), and a clearly significant effect of 20.4 percentage points in column 6.

mayor and town council are asynchronously elected we do not observe a consistent effect of the mayor's partisanship on the town council election results. Point estimates are small and positive in the range of 0.9 to 2.1 percentage points but generally insignificant throughout most specifications.[21] Figure 2 does not indicate a significant break of the overall trend at the discontinuity.

Figure 2: Design 2
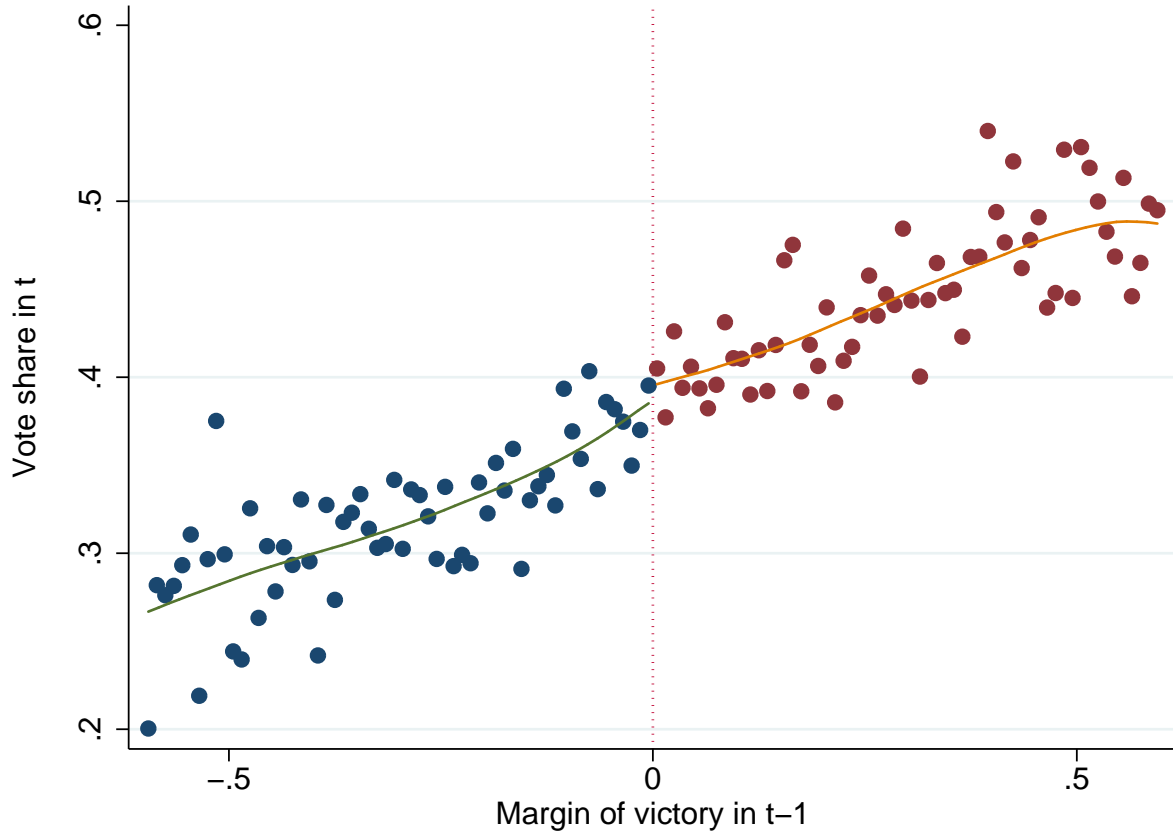
Notes: This figure graphically illustrates the evolution of the vote share in the town council election for party x when party x narrowly won the mayor election in t-1. The data used are those from Design 2. For clarity the data have been grouped in bins, each bin representing an interval of 1 percent in the margin of victory. The line fitted onto the scattered data is based on a local kernel regression using endogenous Epanechnikov weights. *Source:* Own calculations.

How are these results to be interpreted with regard to the two mechanisms previously discussed? We expect the incumbency externality effect to increase the vote share of the

_____

[21]The results also do not change if we include state and year dummies to increase efficiency. Moreover, specifications of the RDD, which we do not present here, do not show any indication that could point to a conclusion that there is an effect of the mayor position on the town council elections in Design 2. The difference to the estimation results in Design 1 are sizable and statistically significant in the preferred specifications. In column 2, the difference is 0.0478 points with a standard error of 0.025. It is thus statistically significant at least at the 10 percent level, and when including year fixed effects even at 5 percent level. The same is true for the estimated differences in columns 5 and 6.

mayor's party whereas a divided government effect would decrease the support for the party. We find, that even if there might be a preference effect for divided government, it does not prevail against the externality effect. In both designs the effect of the mayor on the subsequent town council is non-negative.

The substantial disparity between the results from Design 1 and Design 2 indicates that the precise timing of the elections for the two local political institutions (mayor and town council) is crucially important. The observed differences are consistent with the notion that either one or both of the mechanisms exert a differential effect with respect to timing.

It is reasonable to assume that the incumbency externality effect could be larger (more positive) in the case of joint elections. In both designs the party that has access to the mayor's office can profit from the use of this resource (media presence, increased financial resources, time spent with the electorate etc.). However, it is only when elections are held simultaneously that the party can benefit also from synergy effects of joint campaign efforts as well as a potential coattail effect. The differences in the two designs might therefore be consistent with a heterogeneous incumbency externality effect.

However, we also consider the theory that voters prefer divided government as a driving factor for the difference in the effects. According to this theory voters try to balance the government by voting for different parties in different elections, thus the mayor's party would lose support. As pointed out by Alesina and Rosenthal (1995), timing might be of the essence for the opportunity to act on this strategic incentive. When elections are not held simultaneously voters can condition their choice directly on the state in the second institution (here the mayor position). A "pro divided" voter can cast her vote accordingly as one part of the government is already known. When the elections are held at the same time, voters have to form an expectation about the outcome of the second election (here the mayor election). A voter that might otherwise prefer a divided government might decide to cast both votes for the conservatives to prevent an all social democrat leadership. The incentives to behave strategically are weaker under the uncertainty of the other election outcome. Hence, also heterogeneity in the divided government effect might explain the observed differences in Design 1 and 2.

The empirical observations from the two RDD designs are not conclusive as to distinguish between the two competing theories. As both effects might be heterogeneous with respect to timing, we have in essence four parameters, but only two estimates. At this juncture, we can only conclude that (1) the joint effect of the two mechanisms is non-negative under both designs, hence, the divided government effect never prevails; (2) As we observe a positive

effect in Design 1, it must be that an incumbency externality effect exists and is positive (at least in this case); and (3) The observed differences in the two designs are consistent with either a larger externality effect or a smaller divided government effect when elections are held simultaneously.

Apart from not being able to pin down the magnitude of the effects, it is very unsatisfying that we cannot yet derive a statement whether a divided government rationale is of importance at all. In the following subsections, (see subsections 5.2, 5.3), we examine additional evidence on the nature of the interaction between the mayoral office, the town council, federal, and European elections. We find proof of a divided government effect from the additional design studying the interaction between council and mayor elections. Although the institutional setting is slightly different, we use this result to further conclude that: (4) Given a negative divided government effect, it must be that an incumbency externality effect exists when the timing of elections is asynchronous. A positive incumbency externality effect of similar size as the divided government effect is needed to explain a zero joint effect. This is interesting, because it shows that the incumbency externality effect is of importance also when we exclude arguments such as joint campaign efforts or joint ballot coattail effects.

5.2 Effect of the mayor's political identity on elections for higher levels of government

In table 4 we highlight the effects of the mayor position on elections in higher tiers. In particular, we evaluate whether having a mayor of a certain party exerts an effect for her party in subsequent European and German federal elections.[22]

The panels 1 and 2 in table 4 refer to elections for the European parliament and the German *Bundestag* respectively. We find no significant effect of the mayor's office on either of the two levels. The mayoral position appears to be irrelevant for the electoral outcome of the party in higher tiers throughout all specifications.

Compared to the results of the town council, these outcomes can be clearly interpreted as evidence that there is no incumbency externality of the mayor's office at these levels. We assume that voters do not strategically balance their votes between the local elections and higher level elections. The policy making interaction between communities and the

---

[22]Note that these higher level election are held simultaneously with local elections only in exceptional cases. There are a few communities in Hesse, Rhineland-Palatinate, Saarland and Bavaria where mayoral elections were, in fact held, simultaneously. The number of elections however, is too few to run a meaningful, separate analysis.

Table 4: Interdependency between mayor's office and higher level elections

| | Dependent variable: Vote share in the respective election | | | | | |
| | Obs from Design 1 | | | Obs from Design 2 | | |
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel 1: European Elections					
d	0.008	-0.004	0.009	-0.013	0.021	-0.006
	(0.022)	(0.023)	(0.016)	(0.027)	(0.031)	(0.022)
N	469	1110	12057	86	194	1826
R2	0.00	0.01	0.12	0.00	0.01	0.05
	Panel 2: Federal Elections					
d	0.005	-0.002	-0.002	-0.001	0.008	-0.009
	(0.016)	(0.015)	(0.012)	(0.016)	(0.015)	(0.011)
N	790	2053	22849	160	357	3558
R2	0.00	0.00	0.09	0.00	0.00	0.08
Sample	2 %	5 %	full	2 %	5%	full
Control function	none	linear	3rd order	none	linear	3rd order

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual election. The dependent variable is the vote share of a party (which had a candidate in the preceding mayoral race) in a higher level election (European, federal election respectively). Columns 1-3 estimate the three preferred specifications for the observations from design 1, while columns 4-6 refer to estimates for observations from design 2. The regressions in columns 1 and 4 are based on a limited sample within a margin of victory of 2 percentage points and include only a constant and the treatment dummy. The estimations in columns 2,3 and 5,6 include a polynomial control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

federal government, let alone the European Union, is hardly relevant for an individual voters decision. We would therefore argue that there is no preference effect in this analysis. Thus, any effect would have been a pure incumbency externality effect.

Note that we also subdivided the results for the higher level election outcomes again into the observations from Design 1 and Design 2. We did this to examine whether the results are different at higher tiers along this margin. We observe no differences in those estimates. These findings are reassuring, as we can exclude that specific state differences are driving our results on the mayor's effect on town council elections.

5.3 Effect of town council elections on run-off mayoral elections

In this last part of the analysis, we turn away from the RDD and use another institutional feature of the German electoral system. Whenever mayor and town council elections are held simultaneously, there is the chance that the first ballot in the mayor election does not establish an absolute majority winner of the election. In these cases, a second ballot

between the two leading candidates is held. This feature in the electoral design presents an interesting opportunity to study the effects of the town council election on the second ballot (run-off) election for the mayor.

Specifically, we investigate whether the vote shares in the town council election affect the vote outcome in the second-ballot election for the mayor. This analysis is interesting as an incumbency externality effect should not play a role in this setting and any remaining influence can hence be attributed to a preference for divided government. If voters prefer divided government, we expect a high vote share in the town council election to decrease the vote share of a candidate of this party in the run-off election. Empirically, we can identify this effect because we also have the information on the first-ballot round results. While it is true that in general, town council election results and second-ballot outcomes are determined by common non-observed variables, any such factors should already be featured in the election result of the first round.

Table 5 highlights the results from this analysis. In column 1, we only regress the vote outcome for a party in the run-off mayor election on the result of the same party in the town council election, which was held one or two weeks previously. In this regression one variable simply proxies for the other and we simply pick up that a community generally prefers a certain party.[23] In column 2 and 3, we also include the vote share of the party (and its square) in the first round of the mayoral election. This not only controls for any unobservable characteristic that determines the attitude of the voters towards the party, it also controls for the personal characteristics of the candidates in the mayor elections.[24] We find that in this specification the effect of the town council election turns negative. For an additional 10 % in vote share in the council elections the voters in the community will cast 2.5% fewer votes to a candidate of this party in the run-off election. This effect remains perfectly stable if we include further information on the first-round mayor election such as the distance to the best opponent, a dummy for being second in the first round election and a set of dummies for the number of candidates that participated in round one (see columns 3-5).

We interpret this outcome as clear evidence for balancing behavior of the voters. At least some of the voters prefer not to cast their vote for this party's candidate when the vote

---

[23]Say, for example, that there was a shock in the community before the rounds of local elections that made the voters dislike the conservative party. This will turn up both in the town council election and in both rounds of the mayoral elections.

[24]The R^2 clearly indicates that the vote share from the first round is the ideal predictor for the outcome of round two. Note, that the effect on the vote share in round one is larger than 1 as there are, by definition, fewer candidates on the second-ballot.

Table 5: Run-off mayor elections

| | Dependent variable: vote share in run-off mayor election | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Vote share CE	0.947***	-0.211***	-0.253***	-0.246***	-0.246***	-0.240***
	(0.034)	(0.027)	(0.026)	(0.025)	(0.025)	(0.025)
Vote share main ME		1.365***	2.299***	2.090***	2.117***	2.217***
		(0.020)	(0.075)	(0.077)	(0.090)	(0.091)
Vote share main ME squared			-1.505***	-1.688***	-1.729***	-1.706***
			(0.121)	(0.116)	(0.133)	(0.129)
Dist. to next main ME				0.250***	0.239***	0.162***
				(0.026)	(0.035)	(0.037)
Behind in main ME					-0.004	-0.005
					(0.008)	(0.008)
4 candidates in main ME						0.014***
						(0.004)
5 candidates in main ME						0.036***
						(0.006)
6 candidates in main ME						0.039***
						(0.010)
7 candidates in main ME						0.005
						(0.022)
Constant	0.104***	0.015***	-0.097***	0.008	0.006	-0.044***
	(0.012)	(0.005)	(0.009)	(0.014)	(0.015)	(0.017)
N	1454	1453	1453	1453	1453	1453
R2	0.337	0.818	0.835	0.844	0.845	0.849

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. ME: Mayor election, CE: Council election. The dependent variable is the vote share of a party in the second ballot (run-off) election for town mayor. The main variable of interest is vote share of a party in the town council election held one or two weeks before the run-off election of the mayor's office. In further columns we control for the vote share in the first-round mayor election (linear as well as in squares), the distance to the next best opponent, a dummy whether the candidate was behind in the first round and a set of dummies that specifies the number of candidates that participated in the initial mayor election. *Source*: Own calculations.

share in the town council was high. We conclude that those results are evidence for a negative divided government effect.

5.4 RDD validity and robustness tests

In this section we present evidence for the validity of the RDD approach used in the estimations. Additionally to the results reported above, we also implement a number of robustness tests and further specifications in different subgroups.

There are two tests to examine the crucial assumption of random assignment of the treatment around the threshold. The first check is that the group of observations on either side are "perfect" counterfactuals. That implies that any predetermined characteristic of those communities just right and just left of the threshold must be similar on average. A second test to examine the underlying identifying assumption is to look at the frequencies of obser-

vations just around the threshold. If treatment is random at the threshold, there should be no systematic difference between the number of observations around the discontinuity. Any significant difference would speak for the ability to sort around the threshold that would invalidate the research design.

Figures 3 and 4 in the appendix examine this property in the data for five predetermined measures. We investigate whether the communities differ with respect to the voter turnout in the mayor election in t-1, whether the party ran a candidate in the mayor election in t-2, the level of spending in t-1, the unemployment rate in t-1 and the outcome of the town council election in t-1. As the treatment was not yet determined at the moment when we observe these characteristics, there should be no effect on those variables. The figures show that there is no discontinuous jump in any of these measures around the threshold. Figure 5 in the appendix presents the frequencies in our data. There is no indication that those frequencies are systematically different around the threshold. Hence, we conclude that the quasi-experimental design is valid according to both of those tests.

Freier (2011) estimates a RDD to evaluate the electoral advantage for German mayors in subsequent mayor elections. The argument is based on a similar design using close elections. He confirms the validity of the RDD with a number of additional tests using several other predetermined variables and frequency tests in different margins from the threshold.[25] He reports no significant differences in either the predetermined variables or frequency plots.

To ensure that the effects measured in the Design 1 analysis are indeed valid and not simply an artifact of random noise we run placebo tests simulating different thresholds determining our treatment. Table 9 in the appendix shows the results of two of those tests. In the first test (columns 1-3) we simulate that a party gained access to the mayor's office even when it lost the election with a 5 percentage point margin. In the second placebo, we test the results assuming that the mayor's office could only be obtained if the party won by more than 5 percentage points. Both tests yield the expected result that those devised thresholds -that should not matter for the interaction between mayor and town elections - indeed do not show any significant effects. This provides further evidence for the reliability of the effects estimated above.

---

[25]Freier (2011) reports no differences in the following predetermined variables: Incumbency status of the mayor in the previous two election periods (also 2nd moments), votes share in the last two mayor elections (also 2nd moments), participation in the mayor race in $t-2$ and $t-3$, status of the mayor in $t-2$ and $t-3$ (full-time or part-time employed), number of candidate in the race for mayor in $t-1$ and $t-2$, number of voters in $t-1$ and $t-2$, turnout rate in $t-1$ and $t-2$. As well as the fiscal information on total expenditures (per capita), total debt, revenue from trade tax, as well as tax power (measured by the ratio of tax revenue to total revenue) each in the year preceding the elections in t-1 and t-2. For the frequency plots, he looks at 1 percent bins in a large margin around the threshold as well as 0.25 percent bins just around the threshold and shows that there is no significant difference.

To further investigate the robustness of our results, we test a number of additional specifications of the RDD in which we varied the degree of the polynomial control function and experiment with other margins for the limited sample approach. The results always compare well to the estimates reported in the two designs. Further, we investigated whether the type of mayor position (full-time or part-time) made a significant difference. This might be especially interesting since communities in the sample of Design 1 are on average smaller and less likely to have a full-time mayor.[26] To check whether differences in the two designs are simply driven by differences in the community size/contract type of the mayor we run estimations for only those communities in Design 1 that have full-time mayors. The estimation results indicate that there is no significant difference between the estimates for full-time mayors only and the estimates in the overall Design 1 sample. If anything, point estimates are slightly higher on average for the full-time mayors, indicating that a differences between the designs might be even larger.

Moreover, for the analysis in Design 2 of the effect on the town council, we split the sample with regard to the time that the mayor and town council will have to govern together. For some observations this time period is only one or two years before a new mayor is elected, for others it is three or more years. This division is of potential interest as the voter might find it even more worthwhile to balance the government if this decision determines the political structure in the community for a long lasting period. However, estimation in these subgroups does not show any significant differences between the individual subgroups and the overall result for Design 2.

Last, but not least, we examine whether the interaction between mayor and town council election is different (in Designs 1 and 2) when we control for the past balance between mayor and council. We asked whether voters behave differently when the past mayor and council majority were of the same party compared to a case were the local government were already divided in the two political institutions. Again, estimation results do not exhibit any significant differences to the effects presented above.[27]

6 Conclusions

In this paper we examine interdependencies in voting decisions for different political institutions. In particular, we analyze how voters consider the election result of one political

---

[26]Most of the data for Design 1 is from the state of Bavaria, which has a large number of very small communities, many of which only have part-time mayors.

[27]Estimation results for all robustness tests and further specifications are available from the authors upon request.

body when they vote for a different institution. We compile a new data set containing community level results for mayoral, town council, federal, and European elections from eight German states to conduct the analysis. Methodologically, we mainly rely on a regression discontinuity design, which uses close election outcomes to identify quasi-random variation in the partisanship of the local mayor.

For the local elections, we find that the party of the incumbent mayor receives a bonus of around 4 to 6 percentage points in the subsequent council election if these elections are held jointly. If the elections for the mayor and the council are on different dates the mayor's party does not benefit from holding the office. In the other direction, we investigate the effect of the council election result on run-off mayor election outcomes. We find that a ten percent increase in vote share in the council election for a party is associated with a decrease of 2.5 percentage points in vote share for the party's candidate in the run-off mayor election.

Throughout the analysis, we link the empirical evidence to two underlying mechanisms: the incumbency externality effect from the mayor's office and a preference effect for divided government. Our findings support the following conclusions. We find that both a positive incumbency externality effect and a negative divided government effect are present in the observed voting behavior. In evaluating the impact of the mayoral position on the next council election, the incumbency externality effect (weakly) dominates the divided government effect, irrespective of timing. The intriguing timing differences are consistent with the either a larger externality effect or a smaller divided government effect when elections are held simultaneously.

For elections at the federal and European level, we show that there is no effect of the local mayor's partisanship on these election outcomes. Assuming that there is no incentives to also balance votes between institutions at different tiers, we conclude that there is no significant incumbency externality effect. Voting decisions for the higher tiers are hence independent from the local institution of the mayor.

For the analysis above, we show that interaction in voting behavior is not of significance between different tiers of government, however, within one tier we find substantial interdependencies between political institutions. We presented a number of reliable empirical observations that prove that the mayor's office can be of importance in the next town council election and vice versa. These findings are consistent with two mechanisms that we identify from the theoretical literature. The described effects have a significant influence on government composition and hence policy outcomes. We conclude that one should carefully consider the strategic and psychological aspects in the voters decision making when

designing electoral rules, in particular when it comes to timing of elections and run-off procedures.

Acknowledgements

References

ALESINA, A., AND H. ROSENTHAL (1995): *Partisan politics, divided government, and the economy.* Cambridge University Press.

——— (1996): "A Theory of Divided Government," *Econometrica*, 64(6), 1311–1341.

BUNDESBANK (2007): "Zur Entwicklung der Gemeindefinanzen seit dem Jahr 2000," Monatsbericht.

CAMPBELL, J. E., AND J. A. SUMNERS (1990): "Presidential Coattails in Senate Elections," *The American Political Science Review*, 84(2), 513–524.

CAUGHEY, D. M., AND J. S. SEKHON (2010): "Regression-Discontinuity Designs and Popular Elections: Implications of Pro-Incumbent Bias in Close US House Races," Travers Department of Political Science, UC Berkeley.

CHAMON, M., J. M. P. DE MELLO, AND S. FIRPO (2009): "Electoral Rules, Political Competition and Fiscal Spending: Regression Discontinuity Evidence from Brazilian Municipalities," *Escola de Economia de Sao Paulo Discussion Paper*, 208.

EGGERS, A. C., AND J. HAINMUELLER (2009): "MPs for Sale? Returns to Office in Postwar British Politics," *American Political Science Review*, 103 (4), 513–533.

ERIKSON, R. S. (1988): "The Puzzle of Midterm Loss," *The Journal of Politics*, 50(04), 1011–1029.

FEREJOHN, J. A., AND R. L. CALVERT (1984): "Presidential Coattails in Historical Perspective," *American Journal of Political Science*, 28(1), 127–146.

FERREIRA, F., AND J. GYOURKO (2009): "Do Political Parties Matter? Evidence from U.S. Cities," *The Quarterly Journal of Economics*, 124(1), 399–422.

FOLKE, O. (2010): "Parties, Power and Patronage," Ph.D. thesis, Stockholm University.

FOLKE, O., AND J. M. SNYDER (2010): "Gubernatorial Midterm Slumps," *Columbia University Working Paper*.

FREIER, R. (2011): "Incumbency as the major advantage - The scopes and determinants of the electoral advantage for incumbent mayors," *DIW Discussion Paper*, 1147.

GAGLIARDUCCI, S., AND T. NANNICINI (2009): "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *IZA Discussion Paper*, 4400.

HAHN, J., P. TODD, AND W. VAN DER KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

HAINMUELLER, J., AND H. L. KERN (2008): "Incumbency as a source of spillover effects in mixed electoral systems: Evidence from a regression-discontinuity design," *Electoral Studies*, 27, 213–227.

HOGAN, R. E. (2005): "Gubernatorial Coattail Effects in State Legislative Elections," *Political Research Quarterly*, 58(4), 587–597.

KERN, H. L., AND J. HAINMUELLER (2006): "Electoral balancing, divided government and midterm loss in German elections," *The Journal of Legislative Studies*, 12, 127–149.

KNACK, S. (1994.): "Does rain help the Republicans? Theory and evidence on turnout and the vote," *Public Choice*, 79 (1-2), 187–209.

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

LINDEN, L. L. (2004): "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections," *MIT working paper*.

MEYERSSON, E. (2009): "Islamic Parties and the Emancipation of the Poor and Pious. Evidence from Turkey," IIES, Stockholm University.

MIGUEL, E., AND F. ZAHIDI (2004): "Do Politicians Reward their Supporters? Regression Discontinuity Evidence from Ghana," *UCB working paper.*

MONDAK, J. J. (1990): "Determinants of coattail voting," *Political Behavior*, 12, 265–288.

MONDAK, J. J., AND C. MCCURLEY (1994): "Cognitive Efficiency and the Congressional Vote: The Psychology of Coattail Voting," *Political Research Quarterly*, 47(1), 151–175.

NORRIS, P., AND F. FEIGERT (1989): "Government and third-party performance in mid-term by-elections: The Canadian, British and Australian experience," *Electoral Studies*, 8(2), 117 – 130.

PETTERSSON-LIDBOM, P. (2008): "Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach," *Journal of the European Economic Association*, 6(5), 1037–1056.

TITIUNIK, R. (2009): "Incumbency Advantage in Brazil: Evidence from Municipal Mayor Elections," *University of Michigan Working Paper.*

UPPAL, Y. (2005): "The (Dis)advantaged Incumbents: Estimating Incumbency Effects in Indian State Elections," in *Symposiom: Democracy and Its Development*, no. Paper G05-06. Center for the Study of Democracy.

Appendix

In the appendix we present more detail on the properties and processing of the original data and additional detail on the econometric implementation. The appendix furthermore contains tables and graphs with further descriptives, additional analysis, and validity checks.

We restrict our analysis to the two major parties in Germany (the center-right CDU/CSU and the center-left SPD). We do this for three reasons. First, these two parties participate in most council elections and are the two key sources of mayoral candidates. The remaining mayoral candidates are most frequently from voter lists or even independent candidates. If we wanted to include those candidates we would also need that a voter list that nominates a candidate for the mayor election also runs a list for the following council election to obtain an observation. For independent candidates this is obviously not possible and for voter lists it does not occur very frequently. Furthermore, there is a lot of entry and exit from the political sphere when it comes to voter lists. Second, the votes for both voter lists and independent candidates are often poorly documented in the original data. They are often just coded as 'independent party 1', 'independent party 2' instead of providing the actual party name. This makes it impossible to track them over time. Sometimes the votes for all voter list are even only provided as total sum. Third, the smaller parties FDP, the Greens, and the left party do not frequently run mayoral candidates and are rarely among the top two candidates in terms of votes received. Including these parties would therefore not add much valuable information at the threshold and those cases would be exceptional situations.

Often several parties and voter lists jointly nominate a candidate. In these cases we only count the candidate as affiliated with the center-left or the center-right party when the respective party was mentioned first. E.g. when the party affiliation of a candidate was given by 'CDU/Independent voter group' we would count her as member of the center right party (CDU) while we would not do so if the affiliation was given by 'Independent voter group/CDU'.

The mayoral election in some communities may be counted as two observations. This occurs when the term of the mayor is longer than the term of the council. For example, a mayor is elected in year x, then there could be a council election in year x+1 and another council election in year x+5 before the next mayor election is held in year x+8. We include both cases in the analysis.

As discussed in the section on the data, we may use one election pair twice to generate two observations as we look at both big parties. To prevent that e.g. common shocks in

a community at a certain council election bias the standard errors we cluster the standard errors by community and council election year.

There are three different types of ballot structures for the council elections in the investigated states. In the first setting, e.g. in Hesse, each voter has as many votes as there are seats in the council. In the second setting, e.g. in Thuringia, each voter has three votes. In both cases the voters can freely attribute their votes between different candidates, which they can choose from the candidate lists of one or different parties. The voters can allocate up to three votes to one candidate. Alternatively, they can simply assign all votes to a specific party list. To calculate the number of seats of a party all votes received by candidates of this party are added and then compared to the total number of votes of the other parties. For the properties of these two settings there may be more votes than voters in the data. In the third setting voters have only one vote, which they can cast for a list. This is the case e.g. in Saarland.

It can occur that only one candidate runs for the mayoral office. In these elections voters can either accept or refuse the candidate. If, e.g., 90 percent accept the candidate, a margin of victory of 80 percent is recorded in the data. As those elections cannot be compared in any reasonable way to the elections with more than one candidate we exclude them from the analysis. About 15 percent of all mayor elections in our data are single candidate elections.

Table 6: Data set - higher level elections

	European elections		Federal elections	
	# Obs	Years	# Obs	Years
Bavaria	15059	$'79 - '09$	29656	$'49 - '09$
Thuringia	556	$'04$	1059	$'02 - '05$
Brandenburg	201	$'04$	207	$'05$
Rhineland-Palatinate	142	$'04 - '09$	147	$'98 - '05$
Hesse	1122	$'04 - '09$	2341	$'98 - '09$
Saarland	186	$'04 - '09$	292	$'98 - '09$
Saxony-Anhalt	319	$'04$	658	$'02 - '05$
Saxony	415	$'04$	834	$'02 - '05$
Total	18000		35194	
# within 60 % margin	13883		26407	
# within 5 % margin	1304		2410	
# within 2 % margin	555		950	

Source: Own calculations, based on the data provided by the federal election office.

Table 7: Design 1 - strongest party

	Dep. Var: Probability of becoming strongest party in the council					
	(1)	(2)	(3)	(4)	(5)	(6)
d	0.231***	0.176**	0.106	0.183	0.181***	0.204***
	(0.053)	(0.085)	(0.134)	(0.111)	(0.030)	(0.062)
N	514	219	95	514	5677	5677
R2	0.05	0.03	0.01	0.06	0.19	0.19
Sample	5 %	2 %	1%	5 %	full	full
Control function	none	none	none	linear	linear	3rd order

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable in all specifications is a dummy variable indicating whether the respective party became the strongest party in the council (based on the obtained vote share). The regressions in columns 1-3 are based on a limited sample within the respective margins and include only a constant and the treatment dummy. The estimations in columns 4-6 include a polynomial control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

Table 8: Design 1 - Effects by party

|	Center-Right			Center-Left		
	(1)	(2)	(3)	(4)	(5)	(6)
d	0.054***	0.058***	0.032**	0.064***	0.066**	0.080***
	(0.017)	(0.022)	(0.013)	(0.021)	(0.028)	(0.016)
N	140	335	3371	79	179	2306
R2	0.07	0.06	0.13	0.11	0.17	0.32
Sample	2 %	5 %	full	2 %	5 %	full
Control function	none	linear	3rd order	none	linear	3rd order

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of the respective party in a town council election. Columns 1-3 refer to results of the center-right party. Columns 4-6 are estimates for the center-left party. The regressions in columns 1 and 4 are based on a limited sample within a margin of victory of 2 percentage points and include only a constant and the treatment dummy. The estimations in columns 2,3 and 5,6 include a polynomial control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

Table 9: Design 1 - Placebo test (-5 percent, +5 percent)

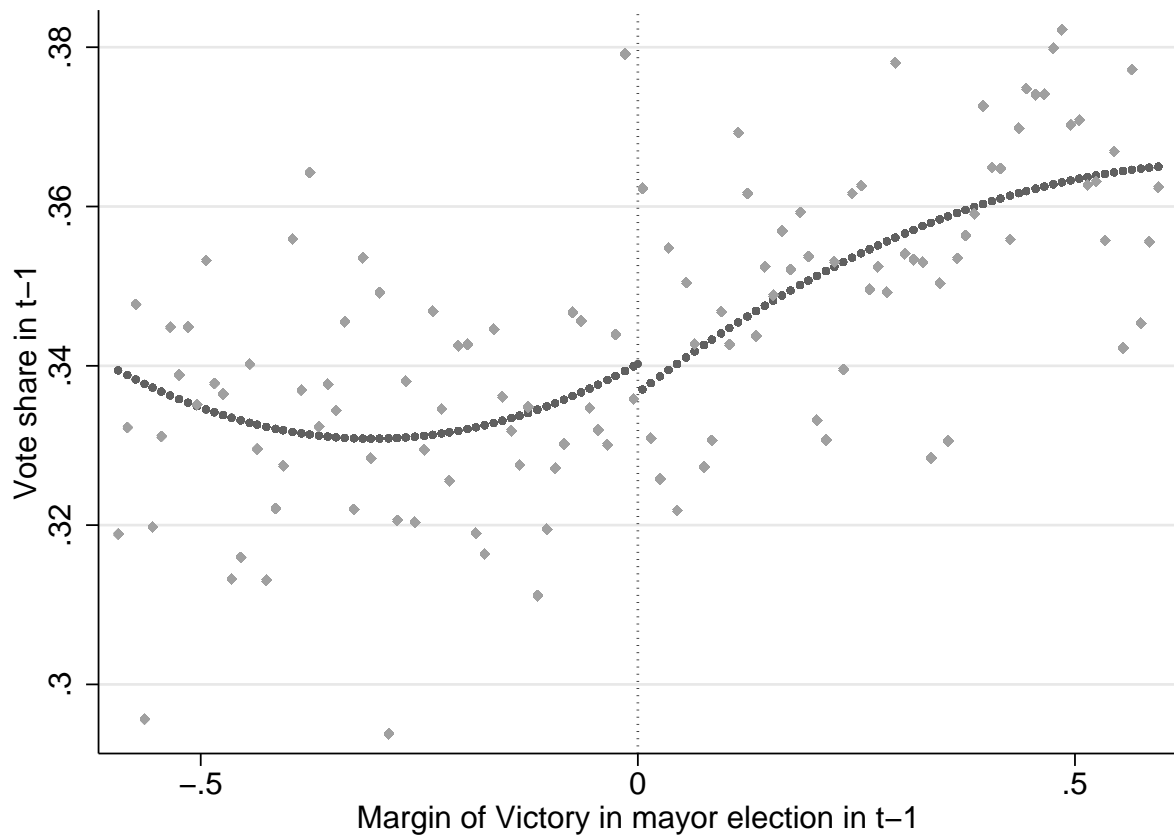|	- 5 percent			+ 5 percent		
	(1)	(2)	(3)	(4)	(5)	(6)
d	0.001	-0.013	-0.010	0.009	0.003	-0.000
	(0.014)	(0.019)	(0.011)	(0.014)	(0.020)	(0.009)
N	211	515	5677	189	478	5677
R2	0.00	0.01	0.26	0.00	0.00	0.26
Sample	2 %	5 %	full	2 %	5 %	full
Control function	none	linear	3rd order	none	linear	3rd order

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of the respective party in a town council election. Columns 1-3 refer to results of the placebo test in which we simulate that a party obtained the mayor incumbency status also if it lost the preceding mayor election with at most 5 percentage points. Columns 4-6 are estimates for the reverse placebo test in which a party needed more than 5 percentage points winning margin to gain the incumbency status. The regressions in columns 1 and 4 are based on a limited sample within a margin of victory of 2 percentage points and include only a constant and the treatment dummy. The estimations in columns 2,3 and 5,6 include a polynomial control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

Figure 3: RDD validity check - predetermined variables
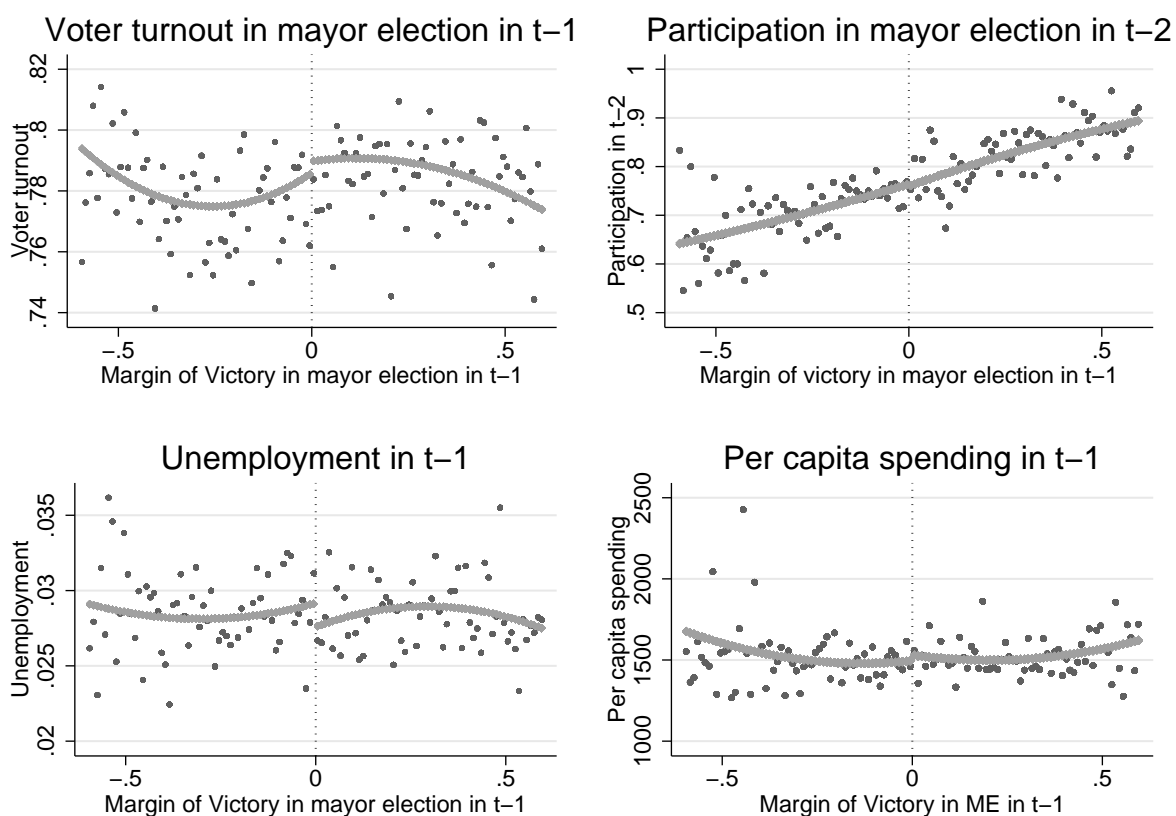
Notes: This figure illustrates that predetermined variables in $t-1$ are not affected by the treatment that occurs at $t-1$. Each bin in the graphs represents an interval of 1 percent in the margin of victory. The line fitted onto the scattered data is a polynomial function of degree two which is flexible on both sides of the threshold. The upper left figure depicts the relationship between the margin of victory in the mayor election and the voter turnout in the the very same election. The upper right figure illustrates the relationship between the mayor election in $t-1$ and the participation in the preceding mayor election (in $t-2$). The lower left graph shows the relationship between a measure of unemployment in the year of the election and the margin of victory in that election. The lower right figure illustrates the relationship between the margin of victory in the mayor election and the aggregate per capita spending of the respective community in the election year. *Source:* Own calculations.

Figure 4: RDD validity check - vote share in last town council

Notes: This figure illustrates that the predetermined variable vote share in the last council election is not affected by the treatment that occurs at t-1. Each bin in the graphs represents an interval of 1 percent in the margin of victory. The line fitted onto the scattered data is a polynomial function of degree two which is flexible on both sides of the threshold. *Source:* Own calculations.

Figure 5: RDD validity check - frequencies around the threshold
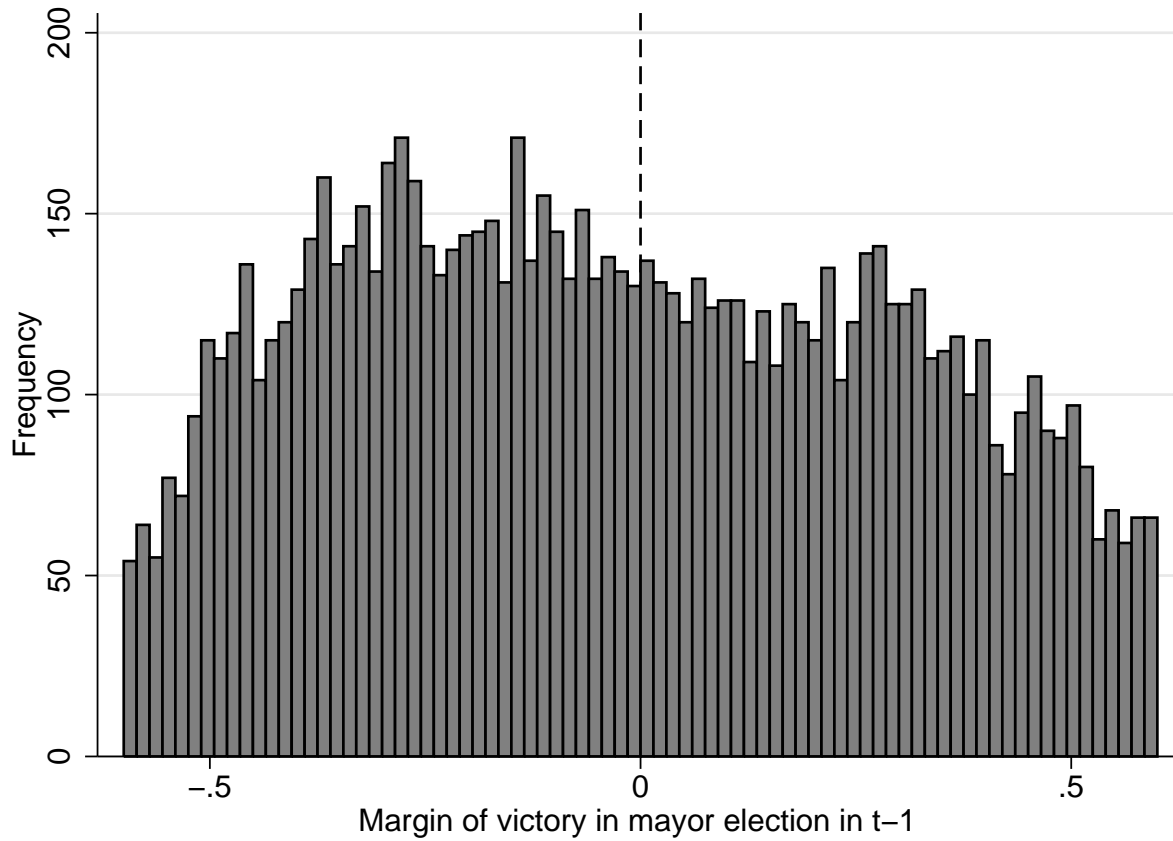
Notes: This figure presents the frequencies of observations in data with respect to the margin of victory. Each bin in the graphs represents an interval of 1.5 percent in the margin of victory. *Source:* Own calculations.