Incumbency, divided government, partisan politics and council size

–

Essays in local political economics

Ronny Freier

Acknowledgements

I have heard people say that writing a thesis is a lonely experience. My experience was the opposite. During my PhD studies, I got into contact with many interesting and inspiring people. I had the chance to work with the best and most passionate supervisors, I traveled the world and met some of the greatest scholars and teachers and I made lasting friendships with a number of amazing people. This is the time to thank them.

My academic interest in economics goes back to my time as an Erasmus student at Örebro University. My bachelor thesis advisor Hans Bylund, as well as my teachers Anders Lunander and Niklas Karlsson put the idea in my head to pursue a PhD. When I returned to my master program in Berlin, I found an inspiring academic atmosphere at the institute for statistics headed by Professor Jürgen Wolters and Professor Herbert Büning where I was teaching mathematics to undergraduates. With the support of Professor Irwin Collier, I was further given the opportunity to spend my senior year at Stanford University where I benefited from the teaching and support of Luigi Pistaferri, John Pencavel and Antonio Rangel.

In 2005, I took my first academic footsteps when I started working as research assistant for Professor Viktor Steiner at the DIW Berlin. Prof. Steiner taught me what it meant to be a researcher in economics, how to choose and go through with an empirical project, and the scrutiny that is needed in work with data. He also taught me that economics is only to be separated from the other social sciences by the introduction of a budget constraint. I am afraid, that I have ignored that advice in the papers in this thesis, but I look forward to coming back to it in future work. I am greatly indebted to Prof. Steiner not only for the two joined papers that we wrote, but for his continuous support in about every aspect of my professional life.

For guidance through my thesis experience my sincere gratitude goes to my two supervisors Magnus Johannesson and Henrik Jordahl. Magnus was in charge of my PhD fortunes in Stockholm from day one of the program. He was supportive and encouraging of all my research efforts, he trusted my judgments and granted me many freedoms. At the same time, he always

made it very clear, what is needed to finish a project, a paper and a thesis. For all your time, effort and commitment to my dissertation project, I thank you. Henrik came onto the project for the last one and a half years. During these decisive months, when all the papers must come together, he has been tremendously helpful. We had many hours of inspiring talks about our research topic, political economics, and I have benefited greatly from his comments and his advice.

Apart from my supervisors, I have received advice and help from many of the Stockholm program faculty. Tore Ellingsen and David Strömberg always had their doors open and their input to my work should not be underestimated. Moreover, I was happy to teach for Mark Voorneveld and Jörgen Weibull and I greatly appreciate their support. I am also indebted to Torbjörn Becker, Olle Folke, Richard Friberg, Emmanuel Frot, Björn Hinnerich-Tyrefors, Juanna Joensen, Chloe Le Coq, Erik Lindqvist, Per Petterson-Lidbom, Torsten Persson and Jesper Roine.

During my PhD studies I was given the opportunity to visit and stay at different institutions. I would like to thank Agar Brugiavini for inviting me to Ca' Foscari University Venice; Rainald Borck for hosting me at Ludwig-Maximilians Universität in Munich; and Orley Ashenfelter, Alan Krueger, Ilyana Kuziemko, Jesse Rothstein and Eytan Sheshinski for their teaching and support at Princeton University. At each institution, I also met intriguing fellow PhD students and made friends in Katja Maria Kaufmann, Tim Mennel and Giacomo Pasini at Stanford, Alessandro Tavoni at Venice, Martin Wimbersky und Alexander Mahle at LMU Munich and Jose Azar, Damien Ientile, Marta Lachowska, Jesus Puente, Martin Schmalz and Alexander Warkentin at Princeton.

Writing this thesis would not have been half as fun and rewarding without my superb co-authors. The papers with Florian Ade and Christian Odendahl made it into this dissertation. However, I have also greatly benefited from working with Damien Ientile, Marta Lachowska and Abel Schumann. You are the best! And I am looking forward to future collaborations.

I am further thankful to all colleagues and the administrative staff at the School for making my time in Stockholm such a pleasant experience. To Mark Bernard, Gökhan Buturak, Samuel Lee, Andreas Müller, Petra Persson, Merle Plassmann, Eva Ranehill, Mark Sanctuary, Nicholas Sheard, Pedro Soares Brinca, Daniel Spiro, Ignat Stepanok, Wieland Wermke, Claudia Wolff, I say thank you for your friendship. A big thanks goes to all my first-year colleagues and especially to Petra and Daniel for all the hours of joint studying. Also, I owe Daniel, Andy and Wieland a million times over for their repeated hospitality throughout the last years.

My academic career began at the DIW Berlin and it will continue there when I start a Postdoc position in the Department of Public Economics in September 2011. Throughout the years the institute has always been my second academic home. My special gratitude, thus, goes to all my colleagues and friends: Stefan Bach, Hans Baumgartner, Martin Beznoska, Hermann Buslei, Marco Caliendo, Frank Fossen, Johannes Geyer, Daniela Glocker, Peter Haan, Michal Myck, Maria Nieswand, Olga Nottmeyer, Richard Ochmann, Davud Rostam-Afschar, Nicole Scheremet, Helke Seitz, Martin Simmler, Johanna Storck, Florian Walch, Katharina Wrohlich and Petra Zloczysti.

None of this would have been possible without the support of my family and my girlfriend. My mother, my father, my sister, my grandparents and Irina have been there for me all through the PhD experience. You backed up my decisions, you carried me through bad times and you put things into perspective. I love you dearly. This is dedicated to you!

Contents

General introduction

"Political economics, like public choice, is defined as the economic analysis of politics."[1] It is concerned with studying *"government behaviour and the interaction between economics and politics."*[2]

This thesis comprises four empirical papers, each devoted to a specific topic in local political economics. Paper one and two evaluate the importance of the mayor position to the future electoral success of the mayor's party. In the first paper, the focus is on the party's electoral outcome in subsequent mayoral elections, while the second paper is concerned with the interdependencies between the mayor's office and elections on other levels of government. The third paper investigates the causal effect of individual parties on policy in the context of German town council politics. The objective is to measure the impact of political representation in a proportional election system on core fiscal decisions of the municipalities. The final paper studies the specific concerns when using population thresholds in regression discontinuity designs for causal inference (in the German case). The analysis reviews the German evidence on the link between the size of the legislation and government spending.

In the larger frame, political economics aims to contribute to a basic understanding of the functioning of the state. How should we design institutions and rules to organize the state activity in a democracy? How should we set up the election system and the structure of the political process? My dissertation touches on these important questions. The four papers evaluate details in the electoral system, study coalition formation of parties, the intrinsic advantage of holding office, and they consider the importance of constitutional rules for local policy making.

It is the *"common goal of public-choice* [and political economics[3]] *scholars to derive normative suggestions for improvements of political and economic institutions."*[4] The papers in this thesis

[1] Blankart and Koester (2005), see abstract.
[2] Alberto Alesina and Tabellini (2006), p.207.
[3] Addition by the author.
[4] Blankart and Koester (2006), p.172.

are mostly brief on the normative implications of the documented empirical findings. The focus generally is on the sound identification of causal relationships rather than a normative evaluation of the effects. Nevertheless, the empirical evidence reported may help to draw normative conclusions on, for example, the need to introduce term limits for mayors, the design of electoral systems with regard to the timing of different elections or the number of members that should comprise the town council.

All four papers are furthermore interconnected as they use similar statistical methods. The empirical methodology that is used borrows from statistical tools of the program evaluation literature such as instrumental variable estimation and the application of regression discontinuity designs. This thesis thus seeks to make very clear what variation is used to identify the relationship of cause and effect.

I say "*local*" political economics, because all papers focus on the municipal level in Germany. They are based on data for municipalities in Germany which are in charge of a significant part of the state activity within the German tier structure. In total, there are 12,500 municipalities in Germany, and their major responsibilities concern infrastructure investments, administration of social services and the provision of local public goods. The local municipal level is further free to finance part of its operation itself and it sets three local tax rates. To steer and administrate the local authorities, the voters elect both a mayor and a town council. Throughout, the papers discussed in this thesis are concerned with studying these local political institutions.

In the following, I introduce each of the four papers in more detail:

Chapter 1: Incumbency as the major advantage

This paper studies the empirical evidence on the scope and determinants of the electoral advantage of parties of incumbent mayors in Germany. The main interest is to identify the causal effect of holding the mayor's office on the party's electoral outcome in the next mayoral election.

I estimate the party incumbency effect to be in the order of 38-40 percentage points in the probability of winning the next mayoral election. The results are stable in various specifications and compare well with estimates from other western democracies.

I can furthermore investigate the estimate in a number of subgroups and therefore evaluate the heterogeneity of the effect. I find that the incumbency effect is larger for full-time mayors,

increasing with municipality size, independent of the particular party identity and stable over the last 60 years.

Finally, I also analyse the causal dynamic effect of incumbency on distant future elections and show that an electoral advantage also can be found in the mayoral election two periods ahead. However, the effect is relatively small and sometimes not significant. In particular, the observed dynamic effect falls short of the theoretical effect which is implied if one assumes that the local "one-period" effect is global. I conclude that the local average treatment effect is an exaggeration of the more general average treatment effect. I interpret this such that, despite the large size of the one-period effect, there is no dynasty building.

The paper relies on a regression discontinuity design to identify the causal effect. I focus on close mayoral elections and argue that within a close band around the winning 50% threshold the partisanship of the mayor is quasi-randomly assigned. In a detailed validation section, I thoroughly validate the identifying assumptions and find that the research design passes all relevant tests.

This paper has been put first in this dissertation for a number of reasons. Methodologically, the paper uses a "classic" regression discontinuity design, a method that has seen a lot of recent attention. Paper two and four will also directly rely on this statistical approach. The paper introduces the argument that the precise outcome of close elections is subject to randomness (building on a literature that has been started with papers by Pettersson-Lidbom (2008) as well as Lee (2008)). This argument is thoroughly tested in the validation section. This is particularly comforting as both paper two and three will elaborate on this identification strategy. Finally, this first paper establishes that there is a large intrinsic advantage to the party of the mayor in the next mayoral election. That finding is then used in paper two, to motivate why the mayor's position can have a positive spillover effect on elections for other political institutions within the same entity.

Chapter 2: Divided government versus incumbency externality

Chapter two is concerned with the interdependency of voting outcomes for different institutions at the local level. We ask whether it matters for the party's electoral success that the party holds the mayor's office. The other way around, we are also interested in whether voters condition their vote for the next mayor on the result of the last council election. Given that actual policy outcomes will depend on the interaction between the different institutions, we therefore try to shed light on the complex decision problem of the voter within this system of

political integration. We relate our empirical findings to two mechanisms from the theoretical literature: the theory of preferences for divided government and an incumbency externality effect.

Our main empirical findings are the following. We show that the party of the mayor receives a sizable and significant bonus in subsequent council elections, if and only if, the local elections are held simultaneously (meaning that on the actual election day, the voters have to decide both on a new council and a new mayor). This bonus is as large as 4-6 percentage points in vote share or up to 20 percentage points in the probability of becoming the strongest party in the next council. Similar to council elections that are *not* held simultaneously, elections on higher government tiers (federal and European elections) are *not* affected by the partisanship of the mayor.

Finally, we investigate how a party's outcome in a council election affects the next mayor election. We illustrate that a candidate loses about 2.5 percentage points in the run-off mayoral election for every 10 percentage points in additional vote share that her party received in the recent council election.

Methodologically, this paper uses a similar approach to the first paper and exploits a regression discontinuity design in which the assumption is that the partisanship of the mayor is exogenously shifted close to the 50% threshold in the mayoral election. Moreover, the paper makes use of an additional design, in which the results of second-ballot (run-off) mayoral elections are used as outcome variables. By controlling for the first round results, we can obtain consistent estimates of the effect of the recent town council election on the subsequent mayoral election.

The timing of elections is the most crucial institutional feature that this paper studies and exploits. We work through the theoretical implications of holding elections simultaneously, the information set of the voter under the different scenarios, and we directly use the specific timing of events in the run-off mayoral election approach. As we show, the timing of elections matters crucially for the possibilities of interacted voting decisions.

As mentioned above, the paper is directly linked to the first paper in this thesis both in content and in methodology. Together the two papers provide insights on the effects of holding the mayor's position. The third paper will take a similar identification argument using close elections and extent it to the outcome of the council elections. In a wider sense, paper four will in part again focus on timing in changes in the institutions of the local council and the mayor position and explore the impact of that precise institutional setting.

Chapter 3: Do parties matter?

The third paper asks whether parties in German municipal councils matter for economic policy. In particular, we study how political representation of a specific party under a proportional election system - measured either in seat shares or voting power in the local council - affects the local tax policy.

We find that OLS results relating seat shares or voting power to policy outcomes show clear and expected correlations. However, applying causal estimation the results are less clear. We report significant effects for only two parties. Increasing the representation of the social democrats, surprisingly, seems to lower the local tax burden, while increased representation of the Green party results in a sizable increase in the local property tax.

The main challenge in identifying the effect of representation is to isolate it from underlying voter preferences. In this paper, we identify seats in local councils that were at the margin of being won or lost. The main idea of identification is again that elections are subject to some degree of randomness and that the exact outcome (within small margins) can be considered random (here, we follow Folke (2010b)).

In terms of methodology, the paper has two contributions. First, we suggest a new algorithm to determine what a close election actually is. As the seat outcome of each party depends on the entire vote vector, clear thresholds are ill-defined. We use repeated perturbation of the vote vector within a small margin to determine how likely it is that the actual seat distribution changes. We then use this information in the construction of our instrument. Second, we translate seats into voting power, which enables us to both implicitly weight the importance of seats and to simulate what we think is a reasonable proxy for political bargaining within the council.

The specific institutional details used in this study are the precise seat allocation rules applied in the different German states. In order to determine for which observations the election outcome can be considered close, we need to understand and be able to apply the current legislative rules that translate votes into seats.

This paper extends the concept of close elections used in the two papers above and develops a new algorithm of how to define closeness in a setting of a proportional election system and multiple parties. In content, this paper is close to the fourth and last paper in the thesis, which also investigates the determinants of local fiscal policy making.

Chapter 4: When can we trust population thresholds?

The last paper reviews the potential to use regression discontinuity designs based on legislative rules with population thresholds. We highlight three key concerns for the valid causal inference of such designs: (1) the existence of simultaneous exogenous treatment effects, (2) the existence of simultaneous (endogenous) choices and (3) the precise manipulation of population figures. Using data from the state of Bavaria in Germany, we show that these concerns do indeed matter crucially.

The paper is a comment on the study by Egger and Koethenbuerger (2010) who study the relationship of council size and government spending. The authors apply a regression discontinuity design using population thresholds in the Bavarian legislation which determine shifts in the number of council members to identify the effect of legislature size on local expenditures.

Our new findings are the following. In Bavarian laws and bylaws, our legal research turned out 16 different legislative rules that use population thresholds and determine local institutions. Given that the Bavarian state tier has a need to induce differential rules by municipality size, many of the same population thresholds are used in multiple rules and laws.

Besides the problem of simultaneous exogenous treatments, we also show that simultaneous endogenous choices can matter for the estimation. Before each election cycle in Bavaria, the municipality has to make a choice on whether the mayor works part-time or full-time and whether the municipality will take over certain new responsibilities. These changes can interfere with the identification of effects in changes in local institutions which happen at the same time.

Finally, we show that manipulation of the precise population figure around specific thresholds is of concern in Bavaria. Given the importance of the population thresholds and the fact that the population number is a continuously observed and publicly known measure, we find that there is indeed evidence that municipalities sort at the thresholds. Causal inference based on the assumption that those thresholds identify exogenous variation in the treatment thus fails.

In contrast to the first papers, the methodological focus of this study is on how a seemingly reliable regression discontinuity design may fail. We conclude that regression discontinuity designs based on population thresholds should only be used with utter care to the precise institutional setting and requires sound checking of the identifying assumptions.

Chapter 1

Incumbency as the major advantage

The electoral advantage for parties of incumbent mayors[1]

Abstract: This paper provides empirical evidence on the party incumbency advantage in mayoral elections in Germany. Using a regression discontinuity design on a data set of about 25,000 elections, I estimate a causal incumbency effect of 38-40 percentage points in the probability of winning the next mayor election. The electoral advantage is larger for full-time mayors, increasing in municipality size, independent of the specific partisanship of the mayor and constant between 1945 and 2010. Moreover, it increases with local spending hikes and it is independent of municipal debt. I also illustrate the causal dynamic effects of the incumbent status on distant future elections and therefore evaluate the global properties of the LATE estimate. Finally, I show that the total effect is due to an effect on the probability that the party participates in the next election (about 40% of the total effect) and an effect on the vote share (about 60%).

[1]Acknowledgments: I would like to thank Florian Ade, Tore Ellingsen, Mikael Elinder, Magnus Johannesson, Juanna Joensen, Henrik Jordahl, Christian Odendahl, David Strömberg as well as participants at the Annual Meeting of the EPCS in Izmir and the seminars at IIES (Stockholm) and DIW (Berlin). Comments of colleagues at DIW Berlin and Princeton University are also gratefully acknowledged. Furthermore, I would like to thank Helke Seitz, Jenny Freier and Heike Hauswald who provided excellent research assistance to help me organize the data. Funding of the Hedelius foundation for a research visit to Princeton University and of the Jan Wallander and Tom Hedelius Foundation for generous financial support is gratefully acknowledged. I am further grateful for editorial support from Adam Lederer. The usual disclaimer applies.

1.1 Introduction

A central feature in most democracies is that the citizens elect members of a party into political positions. It is well known that these incumbents in the political office receive an incumbency advantage in future elections for the same office. Theorists in political science and political economics have long recognized the specific role of incumbents (see, e.g., Besley and Case (1995), Alesina and Rosenthal (1995), Persson and Tabellini (2000)). Starting with Erikson (1971), a large strand of empirical literature investigates the issue of incumbency and the intrinsic advantage that parties and candidates derive from holding a political office during reelection campaigns.[2]

Lee (2008) was first to examine the empirical facts of this phenomenon by applying rigorous identification techniques from the program evaluation literature. He investigates elections for the US House of Representatives and analyses the effect of party incumbency on the probability of winning the subsequent election. Lee estimates the intrinsic electoral advantage to be around 45 percentage points in probability of winning the seat.

Following Lee's analysis, a number of studies use similar designs to estimate the incumbency advantage of parties in different settings. Ferreira and Gyourko (2009) study the effects of incumbency in mayoral elections in the United States, finding an incumbency effect of about 32 percentage points for the probability of reelection.[3] For Germany, Hainmueller and Kern (2005) investigate the party incumbency advantage within districts for federal elections and Ade and Freier (2011) show that a positive incumbency advantage also persists in German state elections. In the context of developing countries, researchers find substantial negative effects (see Linden (2004) and Uppal (2005) for India, Miguel and Zahidi (2004) for Ghana or Titiunik (2009) for Brazil).

This paper studies the electoral advantage for the party of incumbent mayors in German municipalities. Estimation results from a regression discontinuity design (RDD) show an effect on the order of 38-40 percentage points in increasing the probability of winning the next election. The results remain stable over a range of different specifications and a number of tests validate the credibility of the research design.

Analysing the effect further, I find that about 15-16 percentage points (hence, 40% of the total effect) are due to differences in the probability that the party participates in the next

[2]See among others , Cover (1977), Krehbiel and Wright (1983), Jacobson (1987), Gelman and King (1990) King and Gelman (1991), Levitt and Wolfram (1997), Ansolabehere, Snyder, and Stewart (2000), Cox and Katz (2002), and Ansolabehere and Snyder (2004).

[3]See also the analysis by Folke and Snyder (2010), who investigate the incumbency advantage and its determinants for US governors.

election. Parties of close winners are on average more likely to participate in the subsequent mayor elections than parties of close losers. The remaining 23-24 percentage points (60% of the total effect) stem from an increase in the party's vote share.

Given that the data spans more than 60 years, I can identify the effect of party incumbency on distant future elections. This paper is first to show that an electoral advantage can be found not only in the next, but also in the following subsequent mayoral election, although the estimates are insignificant in some specifications. This analysis is of particular interest as it provides an idea about the implied global properties of the local "'one-period" effect (the local average treatment effect or LATE). I show that the observed dynamic causal effect falls short of the implied global effect, based on extrapolation of the local "one-period" estimate.

The data used in the analysis come from the state of Bavaria in Germany and comprise more than 25,000 individual mayoral elections. Due to the exceptional richness of the data, I can also explore the results for a number of subgroups. Although the findings from the subgroup analyses must be taken with caution, I highlight some interesting correlations that shed light on the underlying mechanisms at work. I find that the incumbency advantage is larger for full-time mayors, increasing in municipality size, independent of the particular partisanship of the mayor and constant over time.

Furthermore, I use key economic variables to evaluate whether the voting decision on the incumbent mayor is contingent on past economic performance. Voters seem to reward the incumbent mayor if spending increases over and above the general level, but fail to punish if municipal debt levels increase. Using additional data, I report no significant differences between estimates from different German states.

I use close elections to identify quasi-experimental variation in the partisanship of the mayor. I argue that the precise election outcome is subject to a random component when the election is exceptionally close. Under this assumption, treatment just above the defining threshold of 50% in vote share is assigned quasi-randomly. The party incumbency advantage effect is, then, consistently estimated from the regression discontinuity design, as the continuity assumption directly follows from local randomization. This identification strategy is used in a number of studies, not just to estimate the effect of incumbency, but also in other contexts.[4]

[4]The idea was first used in parallel discussion papers, which came out in the year 2001, by Pettersson-Lidbom (2008), who evaluates the economic policy consequences of party block mayorities in Swedish councils and by Lee (2008) (mentioned above). For further applications, see, e.g., Ferreira and Gyourko (2009) and Gerber and Hopkins (2011) on partisan effects of US mayors, Eggers and Hainmueller (2009) on financial benefits to MP's in Great Britain, Brollo and Nannicini (2011) on fiscal transfers in Brazil.

One critic of this approach is Caughey and Sekhon (2010), who reassess the study by Lee (2008) and point out that the identifying assumptions may not necessarily hold. They argue that the extent to which elections might be subject to a random event might differ and that careful testing of the underlying assumptions is necessary to validate the research design. I test and confirm that the identifying assumption is valid for the application at hand.

The paper is structured as follows. I describe the institutional background and the data set in section 2. In section 3, I lay out the empirical methodology. Section 4 discusses the main results and investigates the validity of the research design. Section 5 concludes the analysis.

1.2 Data and institutional setting

In Germany, the municipal level is the lowest of the four governmental tiers.[5] Local governments in Germany oversee local public firms, administrate mandated spending allocated by higher tiers and carry out a number of direct responsibilities such as child care provision, expenditures for culture and recreation as well as investments in local infrastructure. The average budget of a German municipality is about 1400 Euro per capita. In total, about $\frac{1}{3}$ of total government spending is allocated to the local level. The municipalities also oversee about 40% of all government personnel. The affairs of the municipality are in the joint responsibility of the mayor and the municipal council.[6]

In this study, I mainly use data on mayors from the German state of Bavaria. Here, the mayor is granted a strong and independent position by the South German Council Constitution (*Süddeutsche Ratsverfassung*). The constitution puts the mayor in charge of the local administration, municipality personnel and all day-to-day decisions. In local politics, the mayor heads the city council (with active voting rights), presides over all council committees and sets the local policy agenda. In many municipalities, she is often the only full-time working politician. Despite this powerful position of the mayor, it should be noted that the council remains the legislative body for all municipality decisions.

Mayors are directly elected by the voters.[7] Mayors are elected for a period of 6 years and the election is typically simultaneous with the local city council election. If the first-ballot is indecisive (no candidate above 50%) the law requires a second, run-off, election in which

[5] Beside the federal level, Germany is organized in 16 states (of which 3 are city states), about 450 counties and about 12500 municipalities.

[6] See Bundesbank (2007).

[7] Bavaria is in fact only one of two states in Germany (the other being Baden-Württemberg) where the mayor is directly elected since World War II. Most German states only instituted direct mayoral elections at the beginning of the 1990s.

the two leading candidates compete. The position of the mayor can be either full-time or part-time.[8]

In total, I observe the results of 25,180 elections since 1945 for 2,056 municipalities.[9] More than 45 percent of all elections are single-candidate elections, while about 38 percent had 2 candidates, 11 % had 3 candidates, 4% had 4 candidates, and less than 1% had 5 or more candidates. In the analysis, I use the observations of two parties: center-right (CSU - *'Christlich Soziale Union'*)[10] and center-left (SPD - *'Sozialdemokratische Partei Deuschlands'*). Results of other parties are not considered in the analysis.[11] If both the center-right and the center-left party field a candidate in a mayoral election, I will treat them as separate observations.[12] As the electoral rule in Bavaria allows for candidates to be supported by several parties, I recode joint nominations to be nominations of the party indicated to be the primary affiliation. Moreover, I use the result of the run-off election, if applicable.

Table 1.1 illustrates the scope of the data set. In total, I have 18,761 observations for which I can link the outcome of a center-left or center-right candidate in $t-1$ to the outcome of this party in a subsequent mayor election in t within the same municipality. As indicated in the table, about $\frac{2}{3}$ of those observations come from the center-right party that participates most often in local mayor races. The table also presents the number of observations within subsamples restricted by the size of the margin of victory. To effectively exclude single-candidate elections, I restrict the general sample to be within the 60% winning margin.[13] Moreover, I highlight the number of observations in the samples within the 5%, 2% and 1% margin of victory.

[8]The current law prescribes that a municipality is to have a part-time mayor if it has 5,000 or fewer residents, however communities may deviate from this rule. Communities with more than 5,000 and 10,000 or fewer residents are expected to have a full-time employed mayor, however, deviations from this rule are allowed. Communities with more than 10,000 citizens are required to have a full-time mayor. In practice, all municipalities with at least 7,500 inhabitants or more have full-time mayors.

[9]In Bavaria, there are 2,056 municipalities as of 2008. Note, that Bavaria saw changes in the municipality structure in the end of the 1960s and beginning of the 1970s.

[10]The CSU is an independent party existing only in Bavaria. At the federal level, the party forms a close alliance with the CDU (*'Christlich Demokratische Union'*) and, effectively, acts as one party.

[11]The reason is twofold. For smaller parties, such as the Greens (*'Die Grünen'*) or the liberal party (FDP - *'Freie Demokratische Partei'*), winning the mayor's office is very rare and those observations might be special in their own regards. In contrast, local independent parties are often very strong and regularly win mayor elections. However, the local party identity is often not fix over time and such party groups sometimes only form to support a specific candidate.

[12]As those observations might be subject to common shocks, I will cluster standard errors in the analysis at the level of each municipal election.

[13]Thus, in a two candidate election the winning candidate won with a maximum of 80 percent of the vote. In single-candidate elections, the participant runs against herself and usually receives 80-100% of all votes.

Table 1.1: Data set description

	All	Center-left	Center-right
Observations			
total	18761	6692	12069
within 60% margin	13797	5525	8272
within 5% margin	1231	817	414
within 2% margin	487	314	173
within 1% margin	220	142	78
Elected mayors		4006	9817

Source: Own calculations, based on the data provided by the federal election office. Center-left refers to the social-democrats (SPD) and center-right refers to the conservative Christian democrats (CDU).

In table 1.2, I show descriptive statistics of the main variables used in the analysis. For the full sample, the average number of voters is 3,454. Further, 45% of the elections in the data set have candidates competing for a full-time mayoral position. The vote share of the winner in $t-1$ is, by definition, restricted to be greater than 50% and is, on average, 72%. As outcome variables in t, I highlight the vote share of the center-left and center-right party respectively (conditional on participation in t and $t-1$). Panel 2 of table 1.2 shows the descriptive figures for the number of voters and the distribution of the mayor status for the sample within the 5% margin. I find that observations for which the election were close in $t-1$ are slightly larger on average and had a larger share of full-time mayors.

Table 1.2: Data set - descriptive statistics of the variables

Variable	Observations	Mean	Std. dev	Min	Max
	Panel 1 : Variables in the full sample				
Number of voters	18761	3454.7	3689.3	20	38461
Dummy for mayor status	18625	0.45	0.50	0	1
Vote share of the winner	18755	0.72	0.16	0.5001	1
Vote share of the center-right	10073	0.64	0.24	0.004	1
Vote share of the center-left	4758	0.49	0.25	0.028	1
	Panel 2 : Variables in the 5% margin sample				
Number of voters	1231	4143.1	4581.3	166	36876
Dummy for mayor status	1217	0.51	0.50	0	1

Notes: The table shows the descriptive statistics for the variables used in the analysis. The vote share of the winner refers to the election in t-1. The vote shares of the center-right and the center-left party refer to the outcome of these parties in the election at t, conditional on participation in t and t-1. Panel 2 highlights the descriptives for the number of voters and the dummy for the mayor status (0 = part-time and 1= full-time employed) within the sample of elections that was as close a 5% in the margin of victory (in t-1). *Source*: Own calculations, based on the data provided by the federal election office.

To investigate interesting subsamples, I supplement the data with a number of additional variables on fiscal characteristics for the period 1983-2007. In particular, I collected yearly per capita data on total municipality expenditures, local public debt, revenue from local trade tax and an indicator of the tax power (defined as own revenue from local taxes divided by total revenue from all sources).

1.3 Empirical model and methodology

In this section, I describe the empirical strategy pursued, which is based on a regression discontinuity design (RDD). The object of interest is the effect of party incumbency on subsequent mayoral election outcomes. I denote this treatment with the dummy variable $d^p_{i,t}$, where the superscript p refers to the party identity, i refers to the municipality and t denotes the election period. The treatment variable is uniquely determined by a score variable (also referred to as running variable), the margin of victory, $v^p_{i,t}$. The margin of victory (or loss) is defined as the distance in vote share of the party to the best opponent. If the margin of victory is positive in $t-1$, the party has its candidate in office during the election period t. I say, the party earned incumbency status during this election period.

$$d^p_{i,t} = 1\left[v^p_{i,t-1} > 0\right] \tag{1.1}$$

I consider election outcomes from the subsequent mayoral elections within the municipality i as outcomes variables and denote them $y^p_{i,t}$. Note that the outcome is indexed with t, as the outcome will be measured at the end of election period t. As outcome variables, I consider an indicator variable that takes the value one when the party wins office in t. Alternatively, I also look at the vote share of the party in the next election and a dummy indicating whether the party ran a candidate. Furthermore, I will also consider outcomes of mayor elections further apart into the future, at $t+1$, $t+2$ and $t+3$.

The problem of estimation is twofold. First, I am interested in obtaining consistent estimates for the causal effect of political incumbency of a party on future election outcomes. Second, I want to investigate the effect in various subgroups. For clarity, I will turn to each of these problems separately.

I follow Pettersson-Lidbom (2008) and Lee (2008)[14] in their use of RDD.[15] The basic model

[14] A similar framework is used by Hainmueller and Kern (2008) in the setting of German district candidates in federal election as well as by Ferreira and Gyourko (2009) for US mayors.

[15] The RDD intuition was first developed by Thistlethwaite and Campell (1960) to test the effect of scholarships on subsequent educational attainment of college students. Hahn, Todd, and van der Klaauw (2001) clarified the conditions needed to estimate treatment effects from a RDD.

is given as follows:

$$y^p_{i,t} = \delta_0 + \delta_1 d^p_{i,t} + \epsilon_{i,t} \quad for \quad |v^p_{i,t-1}| < \Delta \tag{1.2}$$

where δ_1 is the parameter measuring the incumbency effect. In this limited sample model, the regression analysis is based only on observations within a small margin, Δ, from the threshold (with is at $v_0 = 0$). Alternatively, I also consider a control function approach as follows:

$$y^p_{i,t} = \delta_0 + \delta_1 d^p_{i,t} + h(v^p_{i,t-1}, \theta) + \epsilon_{i,t} \tag{1.3}$$

Here, the analysis can rely on a larger set of observations. The flexible function $h(\cdot)$ then represents the influence of the margin of victory on the election outcome on election day t.

The intuition of the RDD is to focus on the observations just around the threshold. Assuming that the margin of victory $v^p_{i,t-1}$ contains a random component and cannot be precisely manipulated by the parties, observations just right and left of this decisive threshold should have the same characteristics both observable and unobservable. The main argument for the validity of the RDD used here is that election results in German mayoral elections are sufficiently prone to random factors that can shift the election outcome within a small margin.[16] Such factors could be things that affect, for example, participation (like weather or simultaneously held higher level elections, (see Knack (1994.)))[17] or direct shocks in party popularity (e.g. media coverage of political scandals).

The formal identifying assumption required for the RDD to hold is the continuity assumption (see Lee and Lemieux (2009)). Around the threshold, all characteristics, except treatment, must be distributed continuously for the identification to be valid. If no covariates besides treatment changed at the threshold, the effect of treatment on the outcome can be consistently estimated. Note, that continuity is a direct result of randomization at the threshold.

The limited sample approach, eq. 1.2, makes direct use of this idea. Here I compare election outcomes only for the samples within a small margin from the threshold. In the analysis, I rely on a margin of 2 percent in the margin of victory. Hence, I compare treatment between observations in which a mayoral race was decided by no larger margin than a 51-49 split. To obtain a consistent estimate of the treatment effect, I must assume that the winner of such a close race was effectively randomly chosen.

[16] The argument of randomness in elections is not just a point made by empiricists, also theoretical models in political economy often rely on a similar assumption. Lindbeck and Weibull (1987) introduced the thought of random chance in elections in their model of probabilistic voting which has come to be a standard model of voting theory.

[17] Participation shocks could shift election results especially when they apply to some groups of voters more than to others.

By including the control function $h(\cdot)$, the RDD can maintain the same identifying assumption, but allow for estimation with the entire sample. Introducing a flexible functional form of the margin of victory in the full sample regression, any correlation of treatment with omitted variables in the error term can be controlled for. However, the additional assumption then is that the control function is correctly specified. In practice, I use various parametric polynomial specifications of different degrees to illustrate that the effects found are not relying on precise functional form assumptions.

It is important to note that I estimate the incumbency of parties and not of specific candidates. While it might at first appear to be more intuitive to study the incumbency effects of specific candidates that run for consecutive elections, such an analysis requires significantly more data and econometric modeling. For detailed discussion of differences between the incumbency party estimator (as proposed by Lee (2008)), the incumbent advantage estimators by Gelman and King (1990) (also see Cox and Katz (2002)) as well as by Ansolabehere, Snyder, and Stewart (2000) can be found in Caughey and Sekhon (2010).

Due to the richness of the data, I can investigate the size of the incumbency party effect in various subgroups. To compare subgroup results, I split the sample and run separate regressions. In practice, the sample is divided into two distinct subsamples ($E = 0$ and $E = 1$) along a specific dimension. The dimensions considered are: status of the mayor (part-time or full-time employed), municipality size (above and equal 5,000 voters or below), party identity of the mayor (center-left or center-right), time of the election (before or after 1980), the change of expenditure (high or low changes), the change of debt (high or low changes) and the change of the local tax base (high or low changes). For the fiscal data, I split the samples at the median of the distributions.[18] The separate regressions then read:

$$y^p_{i,t} = \gamma^E_0 + \gamma^E_1 d^p_{i,t} + h(v^p_{i,t-1}, \vartheta^E) + \eta_{i,t} \quad for \quad E = \{0, 1\} \tag{1.4}$$

Subgroup estimations provide hints as to what are the contributing factors that generate the incumbency effect. However, they are limited in their potential to identify a specific causal effect of the subgroup characteristic. Say I find, for example, a large incumbency advantage effect for full-time mayors but only a small effect in the sample of part-time mayors. It is now correct to say that within each of those samples, I have estimated a consistent effect. However, I cannot be sure that the differences are solely driven by the status of the mayor and not by other sample characteristics that are also different.

For predetermined variables (like the mayor status, municipality size, party identity and time

[18] The way I split the sample is similar to Folke and Snyder (2010).

of the election) this problem is limited to uncontrolled subsample characteristics. When I also consider sample splits based on postdetermined variables (fiscal characteristics) there is the additional problem of sorting into the specific subgroups. Overall, these analyses can highlight interesting differences in subgroups, however, interpretation of these disparities is limited.

1.4 Results

The following section consists of three parts. In the first part, I present the main results of the RDD analysis, including the evidence from the dynamic analysis and a discussion of the two potential mechanisms. The estimation results in various subgroups are highlighted in part two. The third part focuses on a detailed validation of the RD design.

1.4.1 Main results

Table 1.3 presents the findings for the causal effects of the electoral advantage for German mayors. I evaluate the effect of having a mayor of a certain party in office during the election period in $t-1$ on the election outcome of this party in the next mayor election in t. The dependent variable is an indicator that takes the value one if the party obtained the mayor office in t. The columns show the basic effects under different specifications. The results are generally stable over all specifications and range within 35%-42% increase in the probability of holding office after the next election.

Columns 1-3 present the limited sample approach with varied closeness of the elections. In column 2, the estimation uses a margin from the threshold of below 2 percentage points. Each observation in this subsample is such that the party's candidate either won or lost with a distance to the best opponent of below 2 percentage points (hence at a maximum distance of 51% to 49% when there are two candidates). Within this margin, I consider the precise outcome of the local election to be a random event. I estimate the incumbency advantage to be 38.9 percentage points in the probability of winning the election.

Alternatively, I also specify limited sample regressions with a margin of 5 and 1 percentage points accordingly (see columns 1 and 3). The 5 percentage points margin might be hard to defend as still random in treatment. The 1 percent sample, however, is quite restrictive in the available sample. Using a control function, I present the results from three further specifications (see columns 4-6). The estimation in column 4 uses a linear control function with the 5% sample. Columns 5 and 6 use the full sample and a control function with a

Table 1.3: Main results - incumbency effect for mayors

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel 1 : Probability of winning in ME in t					
d	0.420***	0.389***	0.345***	0.360***	0.392***	0.410***
	(0.030)	(0.048)	(0.078)	(0.063)	(0.016)	(0.035)
N	1231	487	220	1231	13797	13797
R2	0.177	0.152	0.119	0.178	0.313	0.314
	Panel 2 : Probability of winning in ME in t+1					
d	0.100***	0.050	-0.033	0.008	0.064***	0.100**
	(0.031)	(0.047)	(0.069)	(0.061)	(0.019)	(0.045)
N	1049	424	200	1049	10600	10600
R2	0.01	0.00	0.00	0.01	0.05	0.05
Sample	5 %	2 %	1%	5 %	60%	60%
Control function	none	none	none	linear	linear	4rd 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 Panel 1 is the indicator variable whether the respective party obtains the mayor's office in t. In Panel 2, the dependent variable is a similar indicator variable, however, the time is one extra election period into the future (in $t+1$). 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.

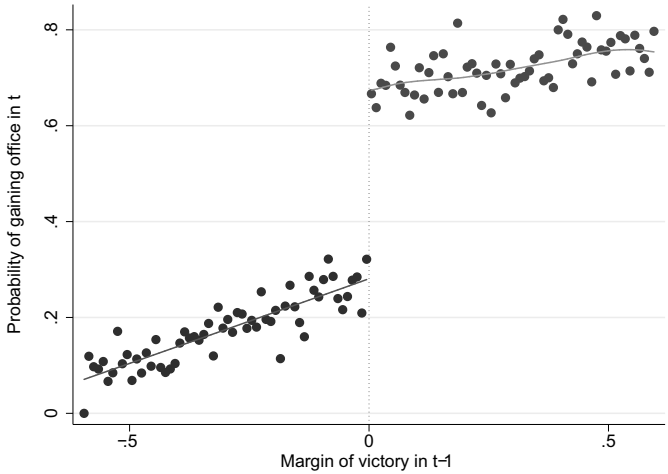
polynom of degree one and four respectively.[19] All results are within a similar range and none are statistically different from the others.

Graphical evidence on the large scope of the effect is presented in figure 1.1. The graph plots the raw data in bins of the margin of victory in $t-1$ (1 percentage point bins) against the bin average of the probability of gaining the office in t. To visualize the inherent discontinuous jump at the threshold, I superimpose a regression fit from a local linear regression. The graph illustrates that a party that just lost the last mayoral race only has a 28-30 percent chance to gain office (on average) compared to the winner of the last election who will win with a probability of just under 70 percent.

I conclude that there is a large and persistent incumbency effect in German mayoral elections. While the effect may appear to be large at the outset, it is only slightly larger than the results for US mayors. Ferreira and Gyourko (2009) report a discontinuous jump in the probability of winning of 33% for Democratic mayor candidates in US cities. In contrast, authors focusing

[19]The polynomial function is flexible on either side of the threshold. I choose a polynomial of degree four as the fifth order was no longer significant. All regressions use robust standard errors, clustered on the level of each individual municipality election.

Figure 1.1: Main result - incumbency effect



Notes: This figure graphically illustrates the discontinuous jump of the probability of winning in t at the zero threshold of the margin of victory in $t-1$. 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 data is based on a local kernel regression using endogenous Epanechnikov weights. *Source:* Own calculations.

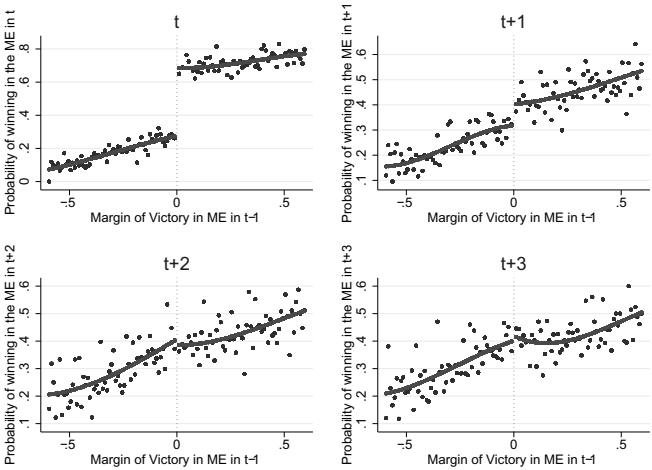
on developing countries have found zero or even negative incumbency effects (Titiunik (2009) for Brazilian mayors, Linden (2004) and Uppal (2005) for Indian parliament, and Miguel and Zahidi (2004) for the parliament in Ghana). They interpret their results to show that politicians engage in rent-seeking activities.

Interestingly the size of the incumbency advantage for German mayors is much larger than the electoral advantage of German politicians in higher tiers. Hainmueller and Kern (2005) report an incumbency effect of a mere 1.4-1.7 percentage points for the party of a district candidate in the election for the German parliament (*Deutscher Bundestag*). For German state elections, Ade and Freier (2011) show that a similar incumbency effect of 1.1-1.5 percentage points exists for district candidates.

In panel 2 of table 1.3, I present the incumbency effect on the outcome of the mayoral election in $t+1$. This realization is two election periods (12 years) after the initial treatment. Thereby, I ask whether party incumbency today affects election outcomes over and beyond the next mayor election. The empirical evidence is less clear. Point estimates range from insignificant zero to a significant 10 percentage point advantage. The results are not conclusive, but they

suggest that merely a small effect is persistent to also effect distant future elections.

Figure 1.2: Main result - dynamic incumbency effect



Notes: This figure presents the dynamic analysis graphically. The four panels show the relationship between the margin of victory in t-1 and the probability of obtaining the mayor's office in t, $t+1$, $t+2$ and $t+3$ with a special focus on the jump at the zero threshold. 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 a polynomial function of degree three which is flexible on both sides of the threshold. *Source:* Own calculations.

In figure 1.2, I highlight the effect on future mayor elections graphically. While the effect is clearly visible for the time period in t, the effect is much smaller in $t+1$ (about 8 percentage points). In election periods even further apart, I find no significant effects any longer, as indicated in the subgraphs at the bottom. These results highlight that, despite the enormous incumbency effect in the election directly following the treatment, the incumbency advantage does not lead to a dynasty manifestation. One normative conclusion from that is that term limits[20] are indeed unnecessary if the objective is to prevent dynasty formation based on the intrinsic electoral advantage of holding office.

Due to the nature of RDD, the estimated incumbency effect of about 37-40 percentage points can only be regarded as a local effect. It is a valid measure only for elections there were close races in $t-1$. The dynamic analysis gives rise to the conjecture that the results are indeed not valid over the entire distribution of election results. If it were true that every incumbent

[20]There are no direct term limits for mayors in Bavaria.

party received an electoral advantage of about 40 percentage points, then I should observe the dynamic effect on the election period in $t+1$ to be around 16 percentage points.[21]

The local estimate seems to overstate the general overall party incumbency effect. This could be true, for instance, if elections close to the threshold are different in nature to other elections. For example, it is quite likely, that the elections that were close (on either side of the threshold) at t are between candidates that have not been in office before and are therefore both younger and "unconsumed". For these elections, it is reasonable that the winner will also seek reelection, whereas the likelihood of continuation might be much smaller for candidates in elections that were not close in *t-1*.

While the probability of winning the next mayoral election is ultimately the most informative measure of the incumbency advantage, it is of interest to analyse the underlying mechanisms that comprise this total effect. The probability of winning is, in fact, influenced by two contributing factors: (a) a significantly lower vote share in election at t for the "marginal" losers in $t-1$ compared to the "marginal" winners, as well as (b) an effect on the probability that the party participates in the next election.

Table 1.4 presents the results of the RDD analysis along those dimensions. Panel 1 highlights the effect of just getting into office on the probability of running a candidate in the next mayor election. The estimates show that there is a causal treatment effect in the order of 14-16 percentage points. A party that narrowly obtained the mayor's office is therefore about 14-16 percentage points more likely to run a candidate in the next election.

In the second panel, I illustrate the electoral advantage in terms of the received vote share. The incumbent party receives a bonus of 15-17 percentage points in the next mayoral election. Those estimates are not to be interpreted as causal effects as these are potentially biased due to selection. I observe vote shares only for parties that participated in the race for mayor. The fact that parties can be expected to make a strategic choice in whether to even participate or not introduces a selection bias into the estimate of the vote share advantage. Moreover, the

[21] Back of an envelope calculations are as follows: The elected party in *t-1* can obtain office in *t+1* either by winning twice ($0.7 \cdot 0.7 = 0.49$) or by losing in$t$ and later winning in$t$+1 ($0.3 \cdot 0.3 = 0.09$). The observed probability of winning in *t+1* under constant incumbency advantage should then be $0.49 + 0.09 = 0.58$. Hence, given a global effect of 40 percentage points, I should observe a 58 to 42 split also in *t+1* (hence an estimate of 16 percentage points). By the same logic, the implied effect in *t+2* would be a 53.2 to 46.8 split.

Table 1.4: Decomposing the total effect -two mechanisms

	(1)	(2)	(3)
	Panel 1 : Prob of running in ME in t		
d	0.160***	0.158***	0.141***
	(0.034)	(0.043)	(0.020)
N	487	1231	18761
R2	0.04	0.04	0.15
	Panel 2 : Vote share in ME in t		
d	0.163***	0.147***	0.174***
	(0.020)	(0.026)	(0.013)
N	403	1029	14831
R2	0.14	0.17	0.33
Sample	2 %	5 %	full
Control function	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. In Panel 1, the dependent variable is the indicator variable whether the respective party ran a candidate in the mayor election in t. In Panel 2, the dependent variable is the vote share that the party received in the mayor election in t. The regression in columns 1 is 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 is within the 5% margin and a linear control function and 3 runs the entire sample including a polynom control function of third degree which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

bias cannot be determined in sign.[22]

Generally, I can conclude that both mechanism are of significant importance. The increase in the likelihood of the party participating in the mayor election can account for about 15 percentage points in the probability of winning. Given a total incumbency advantage of 38 percentage points in the reelection probability, the remaining 23 percentage are due to the electoral gain in terms of vote share.

[22]On the one hand, one could find it intuitive to argue that only the candidates with the better outlook for the elections, are prone to participate. Under this argument, the vote share (and the probability of winning) just right of the threshold would be the outcome of the better candidates that remained after the less strong politicians dropped out. Hence, the estimates would appear even larger once controlling for participation. On the other hand, it is equally sensible to assume that the better candidates (in the sense of them being more likely to win) have, in fact, better foresight into the outlook of the election and can (given the large incumbency advantage) decide to drop out earlier. Under this argument, it would in fact be the poorer candidates, for example, individuals with a strong ideological component, who remain in the race. Here, the estimates would exhibit an upward bias.

1.4.2 Subgroup analyses

Given the substantial number of elections in the analysis, I can also evaluate the RDD estimate of the incumbency effect on various subgroups. I consider two groups of variables for which I investigate subgroup outcomes. In the final part of this section, I further highlight evidence from additional states both from the former west and the former east of Germany.

Firstly, I analyse subgroup results along specific town characteristics. Here, I consider the position of the mayor (full-time versus part-time), size of the municipality, the specific party identity of the mayor and the time of the elections (before or after 1980).

The subgroup outcomes by position of the mayor and size of the town are of interest as both dimensions are likely to be linked with the amount of resources that the incumbent party can use to regain the mayor's post (see, e.g., Levitt and Wolfram (1997), Stroemberg and Snyder (2010), Serra (1994), Folke (2010a)). Full-time mayors might have a significantly higher electoral advantage as they spend more time on the job, have more contact with the constituency and are likely to have more coverage in the media. Also, parties in larger communities are mostly better organized, have more members and higher electoral campaign budgets.

The analysis of subgroups by the specific partisanship of the mayor is important to illustrate the general nature of the intrinsic advantage.[23] I expect that the findings are neutral with respect to the partisanship of the mayor. Similarly, it is interesting to document potential changes over time. Given that the period of observation goes back all the way to World War II, I can compare estimates from Germany from when the democracy was still young to a political system with more than 60 years of experience and stability (see, e.g, Titiunik (2009)).

The second group of variables comprises information on the fiscal and budgetary state of the municipality in the course of the election period. I ask whether the voters reward the incumbent mayor when the per capita expenditures or the revenue from the local business tax increase and whether they punish when the municipal debt rises. In essence, I ask whether the fiscal measures are of importance for the electoral choice of the voters. This analysis therefore relates to a large literature on economic voting (see, e.g., Folke and Snyder (2010), Jordahl (2006), Wolfers (2007)). Moreover, it also helps to evaluate the extend to which the inhabitants vote retrospectively in general (see, e.g., Berry and Howell (2007), Ferejohn (1986), Fiorina (1978)).

As described above, the analyses in subgroups may again lead to issues of endogeneity that the RDD initially attempts to avoid. What a subgroup analysis can do, is to highlight interesting

[23] Especially, given the fact that Bavaria is generally considered a conservative stronghold (it has seen no state government other then a conservative since World War II).

correlations and evaluate the heterogeneity of the party incumbency advantage along the respective dimension. It cannot, however, allow for consistent estimation of the causal impact of a specific characteristic on the size of the incumbency effect.

Table 1.5: Subgroup analysis

	Dependent variable: Probability of winning					
	E=0		E=1		Difference	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel 1 : Town characteristics					
Mayor status	0.211***	0.300***	0.546***	0.512***	0.336***	0.212***
(E=0: part-time, E=1: full-time)	(0.075)	(0.049)	(0.065)	(0.049)	(0.100)	(0.069)
Municipality size	0.332***	0.377***	0.540***	0.508***	0.208*	0.132*
(E=0: <5000, E=1: >=5000)	(0.057)	(0.040)	(0.093)	(0.069)	(0.109)	(0.079)
Party identity of the mayor	0.359***	0.381***	0.439***	0.452***	0.081	0.071
(E=0: CDU, E=1: SPD)	(0.052)	(0.038)	(0.066)	(0.048)	(0.064)	(0.048)
Decade of the elections	0.340***	0.341***	0.411***	0.445***	0.071	0.103
(E=0: <1980, E=1: >= 1980)	(0.089)	(0.060)	(0.060)	(0.042)	(0.109)	(0.072)
	Panel 2 : Fiscal characteristics					
Diff in total expenditures	0.283***	0.373***	0.551***	0.531***	0.267**	0.156*
(E=0: low, E=1: high)	(0.091)	(0.080)	(0.092)	(0.088)	(0.128)	(0.093)
Diff in debt	0.407***	0.419***	0.379***	0.433***	-0.028	0.014
(E=0: low, E=1: high)	(0.079)	(0.046)	(0.121)	(0.070)	(0.141)	(0.084)
Diff in revenue trade tax	0.366***	0.362***	0.456***	0.503***	0.091	0.141*
(E=0: low, E=1: high)	(0.084)	(0.049)	(0.113)	(0.064)	(0.144)	(0.083)
Sample	2 %	60%	2 %	60%	2 %	60%
Control function	none	3rd order	none	3rd order	none	4rd 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 variables is the indicator variable whether the respective party obtained the mayor's office in t in the respective subsample indicated in the left column. The Columns 1 and 3 are based on a limited sample within a margin of 2% for the subsample labeled with E=0 and E=1 respectively. Column 2 and 4 consider a 60% sample and a cubic control function within the subsamples. Columns 5 and 6 show the estimated difference between the respective two samples. *Source*: Own calculations.

Table 1.5 presents the results from the subgroup analyses. Panel 1 shows the results for subgroups of certain town characteristics. Firstly, I split the sample by the mayor status (full-time and part-time employed). I compare the estimates of the incumbency advantage in the sample of only part-time (columns 1 and 2) with the sample of full-time employed mayors (columns 3 and 4). The difference for the two alternative specifications is highlighted in the

last columns. I find that full-time mayors receive a significantly larger electoral bonus. The incumbency advantage is 21-34 percentage points larger than that of part-time mayors.

Next, I divide the sample by municipality size using a population threshold of 5000 voters.[24] The incumbency advantage is larger in municipalities with more individuals, however, the differences are only marginally significant.

The estimates for the above subgroups highlight the specific issue of interpretation of the subgroup estimates. The mayor status for example is, *ex ante,* given and hence exogenous to the incumbent mayor. It is therefore correct to argue that the incumbency advantage is larger in communities with a full-time mayor. However, I cannot claim causality. Towns with a full-time mayor may also have other features in which they differ, for example municipality size. I cannot rule out that it is another mechanism that makes the electoral advantage increase. However, it can be argued that the incumbency advantage estimate is a reliable measure for municipalities with this characteristic.[25]

I test whether the specific partisanship of the mayor matters for the incumbency effect in the third row. As described above, I use data on candidates for the center-right conservative party ("CDU") and the center-left social democratic party ("SPD"). Estimating the model on the separate sample, I find no significant effect of the particular party affiliation. Point estimates are slightly higher for SPD candidates, however, the differences is imprecisely measured even if I include further covariates (time dummies, mayor status, municipality size, economic variables).

Further, the electoral advantage is constant over time. I split the sample in half, using observations from before the year 1980 and after. The differences in the estimates are insignificant from zero. The same is the case when I estimate the model by each decade. This is interesting as it highlights that the incumbency advantage for German mayors was present and substantial even when Germany was just a young democracy shortly after World War II.

In panel 2 of table 1.5, I turn to subgroup analysis using economic variables that proxy the fiscal state of the municipality. At first, I look at the difference in total expenditures (per capita) from the year before the incumbent mayor took office to the year before the next mayoral election. This difference is informative on whether the incumbent mayor was successful in

[24]The choice of 5,000 eligible voters is arbitrary. The intention was to capture a notion of a larger local center versus a village. The precise devision is not of crucial importance and results remain comparable for alternative thresholds such as 3,000, 4,000 or 10,000.

[25]I also run estimations in which I include both the full-time versus part-time mayor variables and the measure of municipality size. From that horse race model, I find that it is the mayor's position that prevails (in the same magnitude) and the municipality effect turns insignificant. Results are available upon request.

increasing the expenditure level or not. I split the sample at the median of the distribution and compare the half that increased expenditures most to those that increased only slightly or lowered the expenditure level. I find that higher spending correlates with a larger incumbency estimate. The difference is a significant increase of 16-27 percentage points in the probability of winning. Voters reward the incumbent mayor for spending hikes.

In contrast, voters do not take local debt levels into account when voting for the next mayor. I again split the sample at the median of the distribution for new debt and find that the incumbency advantage is similar in both groups. Finally, I also compare municipalities with regard to the increase or decrease in the level of revenue from the local business (trade) tax. This measure is interesting as it proxies for a successful local business activity as opposed to expenditures, which can be increased even if the local economy is not doing well. Here, I find sizable and marginally significant differences in the order of 9-14 percentage points. Note that those effects are notably smaller than the findings for the expenditure level.

As argued above, the findings in the subgroup analysis for the economic variables should not be taken at face value. Issues of selection into the respective groups and problems of omitted variable bias are of large concern in the interpretation of those results. However, comparing the results for expenditures, debt and revenues are still quite informative. Changes in expenditures are clearly visible to the voters and will be associated with the work of the incumbent mayor. Here, I find clear and sizable effects. Voters know if the local economy is booming (proxied by the revenue from trade tax), although, this fact is not directly linked to the mayor's work. The effect reported is smaller, but still meaningful. The local debt level is not directly assessable by the voter and exerts no effect whatsoever.

To infer whether the results of the party incumbency advantage are specific to the state of Bavaria, I also collected data on mayoral elections in further German states both in the former western part and in states from the former east.[26] In particular, I obtained data on elections in *Hessen*, *Rheinland-Pfalz*, *Saarland* for the west, and *Brandenburg* and *Thüringen* for the east. For generalisation, comparing the incumbency advantage estimate from Bavaria with other states is of interest. Bavaria grants mayors a strong position in local politics, whereas other states in Germany have local constitutions with a less powerful mayor position. Also, German states from the former east might be fundamentally different as the democracy there is still young and political competition differs.

[26]For the general analysis I use only data for Bavaria. The data from the other states is limited as the time series is quite short and economic variables are not generally available.

Table 1.6: Evidence from additional states

	Dependent variable: Probability of winning				
	(1)	(2)	(3)	(4)	(5)
	Panel 1 : Western states				
d	0.538***	0.496***	0.573***	0.461***	0.469***
	(0.073)	(0.106)	(0.142)	(0.047)	(0.066)
N	131	66	131	1153	1153
R2	0.30	0.25	0.30	0.32	0.32
	Panel 2 : Eastern states				
d	0.248	0.192	0.066	0.337***	0.398**
	(0.152)	(0.229)	(0.284)	(0.080)	(0.164)
N	44	21	44	478	478
R2	0.06	0.04	0.08	0.32	0.32
Sample	5 %	2 %	5 %	60%	60%
Control function	none	none	linear	linear	4th 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 variables is the indicator variable whether the respective party obtained the mayor's office in t. Panel 1 highlights the results for the western states of Hessen (1993-2009), Rheinland-Pfalz (1994-2009) and Saarland (1994-2009). Panel 2 investigates the sample of former east states Thüringen (1999-2009) and Brandenburg (2001-2010). The Columns 1 and 2 are based on a limited sample within margins of 5% and 2%. The estimations in columns 3-5 include a polynom control function of the degree indicated which is specified to be flexible on both sides of the threshold. *Source*: Own calculations.

In table 1.6, I show the results of the RDD analysis in the additional German states. In panel 1, I present the estimates of different specification in the three western states. The party incumbency advantage in those states is a large and significant effect in the order of 46-57 percentage points in the probability of winning in the next mayoral election. Point estimates are slightly larger than the results for Bavaria (compare table 1.3), however, those differences are not statistically significant. In the states of the former east (panel 2), estimation results are more variable, partly due to the smaller sample size. The effects appear to be smaller and often insignificant from zero, although, also here no statistical significant difference to the results of Bavaria can be stated. Overall, the incumbency advantage effect in other German states are comparable to the estimates of Bavaria.

1.4.3 Validity of the RDD

The evidence above crucially relies on the validity of the identifying assumptions. While it is impossible to directly observe whether the election outcome and hence the treatment was subject to sufficient randomness, there are two implicit tests. If treatment is indeed randomly

assigned close to the threshold, the observations must be comparable in predetermined observables. Also, given random local assignment, I should not observe differences in the number of observation around the threshold.

Table 1.7 in the appendix highlights the first test for a broad range of predetermined variables. I distinguish between two groups of variables: political variables of past election periods and elections as well as economic variables. I show the results of the model in eq. 1.3 using the predetermined observables as the dependent variable. I thereby test, whether the distribution differs significantly at the threshold. The test confirms the validity of RDD if I find no significant differences. For all variables I use the predetermined observables both for the last and the next to last election (or election period) and I present estimation results for three different specifications.

As political variables (see panel 1) I use the following observables: Party incumbency status in prior election periods, the votes share of the party in past elections, the participation in the mayoral race, mayoral position (full-time or part-time), number of candidates in the first round of the mayoral election, number of eligible voters and the turnout rate. I find no significant differences in either of those variables for any specification. Specifically, the fact that the incumbent status and the vote share come out to be well balanced is assuring that the quasi-experiment is valid. Caughey and Sekhon (2010) argue that exactly these variables are of major concern for the validity of the design.[27]

In panel 2, I also check the distributions of the fiscal data just prior to the past elections (the observation refers to the year before the election). Again, I find no predetermined differences in any of the following variables: total municipal expenditures (per year and capita), the stock of debt (per capita), the revenue from trade tax (per capita) and a measure of taxing power (share of tax revenue to total revenue).

The argument of perfectly balanced predetermined variables not only includes the mean of the distribution but also higher moments of the predetermined variable distribution. Therefore, I group the data in bins (one percentage point in the margin of victory) and compute the standard deviation per bin both for past incumbency status and the prior vote shares of the party. Figure 1.3 in the appendix shows that also the 2nd moment of the distribution of these predetermined variable is well balanced around the threshold.

[27]They show that a similar design used by Lee (2008) fails exactly due to predetermined differences in the past incumbency status. The incumbent candidate is significantly more likely to win a close race, which invalidates the research design. The argument that they make, raises concern that the identification strategy of using close election might not be applicable in all instances.

As a second check, I also investigate the number of observations just around the threshold. Given that assignment is random close to the threshold, I should not observe any differences in the observed frequencies. I present two graphs in figure 1.4 in the appendix showing that this is indeed the case. In the left graph, I use bins of one percent in the margin of victory and the full sample. I superimpose a fitted line based on a linear local regression to allow for comparison directly at the threshold. Similarly, I also present a graph (on the right) in which I use a smaller bins size (0.25 percentage points) and observations that are no further from the threshold than 0.1 in the margin of victory. For both graphs, I observe no difference in the frequencies just around the thresholds.

Overall, the analysis of both the predetermined variables as the frequencies of observations gives a convincing argument for the validity of the research design. Close outcomes in German municipal mayor elections can be considered a valid research design to establish the causal effect of the incumbency status.

A final test that I estimate a valid treatment effect can be given by running placebo regressions. Table 1.8 in the appendix shows the results of simulating alternative thresholds. First, I simulate that a party already obtained access to the mayor's office if the margin of victory was only -0.05. This implies that the candidate of a party that just obtained above 47.5 percent (against an opponent with 52.5 percent) already (hypothetically) got the position. Alternatively, I simulate the opposite case, in which a party had to receive a winning margin of 0.05 to get hold of the mayor's office. Both tests show that there is no effect on those simulated thresholds.

1.5 Conclusions

In this study, I examine the electoral advantage for the party of the incumbent mayor. I use quasi-random variation in the partisanship of the mayor to identify the causal incumbency advantage effect. Estimation results are based on a regression discontinuity design relying on close elections. The data comprises about 25,000 mayoral elections from the German state of Bavaria between 1945 and 2010.

The main results show a party incumbency effect on the order of 38-40 percentage points in the probability of winning the next election. These results are stable over a range of different specifications. The identifying assumptions of the research design are supported in a variety of validity checks.

Decomposing the effect, I find that about 40% of the total effect comes from increasing the

probability of participation of that party in the next mayor election. The other 60% percentage points are increases in the vote share of the party of the incumbent mayor.

Using data from additional German states, I show that the results from Bavaria are comparable to the electoral advantage for mayors across all of Germany.

In several subgroup analyses, I highlight interesting differences of the incumbency advantage: the incumbency effect is larger for full-time mayors, increasing with municipality size, independent of the particular party identity of the mayor and constant over the last six decades.

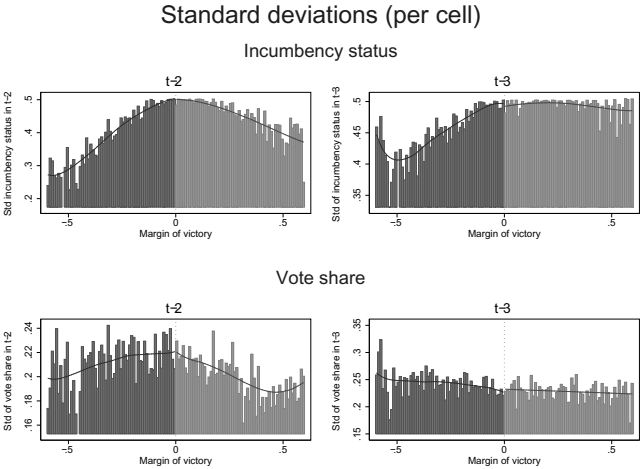
Moreover, I investigate how local fiscal measures are related to the incumbency effect. I find that the treatment effect is larger when local public spending increases, but independent of the municipal debt levels. These results indicate that voters are at least partly backward looking and take the election as an opportunity to hold a referendum on the past performance.

In the data, I observe repeated mayor elections within municipalities over more than 60 years. I can therefore also evaluate the impact of the incumbency status today on consecutive future elections. I report, that the current party incumbency will also have an effect on elections two periods ahead, but no further in time. Given that the score variable in one period is also the outcome variable in a different period, I can use the estimates of the dynamic analysis to evaluate the global properties of the LATE estimate. I find that the dynamic effect falls short of the implied "one-period" effect and interpret this to mean that the LATE estimate in this application is an overstatement of the more general ATE.

These estimates for Germany are large, but compare well with incumbency estimates reported for the seats in the US House of Representatives (Lee (2008)) and for US mayors (Ferreira and Gyourko (2009)). Compared to negative incumbency effects in developing countries, they further highlights the apparent differences between well established and developing democracies. The subgroup analyses and the reported dynamic effects provide some first insights into the determinants of the effect, however, further research is needed to understand the causal mechanisms behind the electoral advantage of incumbent parties.

Appendix

Figure 1.3: RDD validity - 2nd order moments of past political performance



Notes: This figure illustrates the second moments of the two most important predetermined variables in t-2 and t-3: the past incumbency status and the vote share in the last mayor elections. The graph highlights that the standard deviations within group bins of 1 percent in the margin of victory are not different at either side of the threshold. The line fitted onto the data is based on a local kernel regression using endogenous Epanechnikov weights. *Source:* Own calculations.

Table 1.7: RDD validity - predetermined variables

	Predetermined variables from election period in t-1			t-2		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel 1 : Political variables					
Incumbency status	-0.018	-0.068	-0.027	-0.013	-0.035	-0.043
	(0.054)	(0.071)	(0.050)	(0.056)	(0.073)	(0.052)
Votes shares	-0.031	-0.047	-0.036	-0.027	-0.051	-0.037
	(0.025)	(0.033)	(0.024)	(0.027)	(0.036)	(0.025)
Participation in mayor race	0.065*	0.046	0.046	0.044	0.037	0.004
	(0.039)	(0.049)	(0.037)	(0.044)	(0.057)	(0.041)
Mayor status	-0.015	-0.003	-0.035	0.015	0.028	-0.007
	(0.031)	(0.040)	(0.028)	(0.030)	(0.038)	(0.028)
# of candidates in the race	-0.029*	-0.013	-0.005	-0.035	-0.066	0.010
	(0.016)	(0.017)	(0.019)	(0.042)	(0.054)	(0.040)
# of voters	-36.799	-56.613	-114.680	-90.747	-125.410	-112.773
	(184.528)	(221.843)	(156.768)	(172.337)	(209.599)	(147.720)
Turnout rate	-0.001	-0.002	0.001	0.004	0.002	0.007
	(0.005)	(0.007)	(0.005)	(0.005)	(0.007)	(0.005)
	Panel 2 : Economic variables					
Expenditures (per capita)	34.870	59.877	2.222	41.210	50.113	44.046
	(43.471)	(55.487)	(43.900)	(61.528)	(81.028)	(56.295)
Debt (per capita)	0.248	0.385	0.772	-9.998	-8.056	-5.490
	(0.231)	(0.370)	(0.638)	(9.570)	(7.668)	(5.333)
Revenue from trade tax (per capita)	7.952	10.141	0.456	19.308	4.758	3.181
	(10.541)	(14.261)	(12.626)	(13.713)	(16.441)	(13.347)
Tax power	-0.009	-0.010	-0.005	-0.008	-0.009	-0.008
	(0.006)	(0.007)	(0.006)	(0.008)	(0.011)	(0.008)
Sample	2 %	5 %	60%	2 %	5%	60%
Control function	none	linear	4th order	none	linear	4rd order

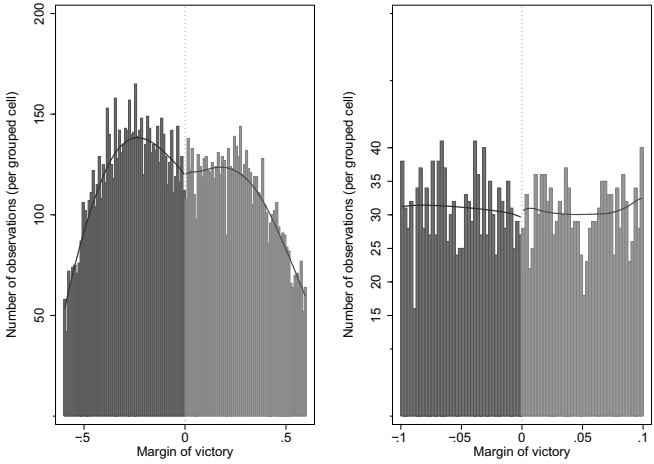
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variables are indicated in the left column. Each regression coefficient represents the result from a RDD analysis on the predetermined outcome. The Columns 1 and 4 are based on a limited sample within a margin of 2%. Column 2 and 5 consider a 5% limited sample and a linear control function and columns 3 and 6 include the full sample (60% margin) and a fourth-order polynomial control function which is flexible on both sides of the threshold. *Source*: Own calculations.

Table 1.8: RDD validity - Placebo test (-5 percent, +5 percent)

	- 5 percent			+ 5 percent		
	(1)	(2)	(3)	(4)	(5)	(6)
d	0.033	0.036	0.032	-0.018	-0.014	0.030
	(0.038)	(0.050)	(0.028)	(0.042)	(0.055)	(0.030)
N	516	1243	13797	471	1210	13797
R2	0.00	0.00	0.29	0.00	0.00	0.29
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 indicator variable whether the respective party obtained the mayor's office in t. The columns 1-3 highlight the 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 present the 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 1.4: RDD validity - frequency histograms



Notes: This figure presents the frequencies of observations in the data with respect to the margin of victory. Each bin in the left graph represents an interval of 1 percent in the margin of victory. In the right panel, the graph is zoomed in further and represents the frequencies within bins of 0.25 percent in the margin of victory. *Source:* Own calculations.

Chapter 2

Divided government versus incumbency externality effect

Quasi-experimental evidence on multiple voting decisions[1]

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.

[1] This paper is joint work with Florian Ade. Acknowledgments: We would like to thank Charles B. Blankart, Tore Ellingsen, Mikael Elinder, Olle Folke, Frank Fossen, Jens Hainmueller, Magnus Johannesson, Juanna Joensen, Henrik Jordahl, Holger Kern, Erik Lindqvist, Christian Odendahl, Yaniv Reingewertz, Carsten Schröder, David Strömberg, Marvin Süsse as well as participants at the Annual Meeting of the EPCS in Izmir and the seminars at IIES (Stockholm) and DIW (Berlin). Comments of colleagues at DIW Berlin and Princeton University are also gratefully acknowledged. Our particular gratitude goes to Viktor Steiner and Peter Haan who invited us to work at their Department for Public Economics at the DIW Berlin and gave valuable comments and advice. Furthermore, we would like to thank Helke Seitz, Jenny Freier and Heike Hauswald who provided excellent research assistance to help us organize the data. Ronny Freier would like to thank the Hedelius foundation for funding the research visit to Princeton University and the Jan Wallander and Tom Hedelius Foundation for generous financial support. We are further grateful for editorial support from Adam Lederer. The usual disclaimer applies.

2.1 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[2] as well as non-monetary aspects, such as media presence for the party and time spent with the electorate during the election period.[3] 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 (2010a) 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)).[4] 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

[2]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.

[3]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.

[4]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.[5] 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.[6]

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.[7] 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 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,

[5]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.

[6]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.

[7]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.

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.

2.2 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 2.1 for an overview of the states and elections included).

2.2.1 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[8] 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.[9] Another feature that is crucial for the correct implementation of our design is the presence of a second (or run-off) ballot. If no candidate reaches the majority of 50% in the first ballot, a second ballot

[8] 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).

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

Table 2.1: 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.

is held between the two leading candidates.[10] 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[11], 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 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

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

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.

2.2.2 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.[12] 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.[13]

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).

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

[12]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).

[13]See Bundesbank (2007).

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.

2.3 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 2.1 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.[14]

[14]The council election years of the other states in our data base were: Bavaria 1984, 1990, 1996, 2002, 2008, Thuringia: 1999, 2004, Brandenburg 2008, Rhineland-Palatinate 1999, 2004, Saarland 1999, 2004, Saxony-Anhalt 2004, Saxony 2008.

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).[15]

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').[16] Brandenburg is the only state where we have observations of both designs. 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 2.2: 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 2.2 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

[15]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.

[16]The reason for the non-alignment of the two elections in Hesse, Rhineland-Palatinate, and Saarland are historic administrative processes. When these states introduced the direct election of the mayor, the term of the administrators leading the local government before the first election was honored. As these terms were of differing length, the dates of mayoral and council elections are not aligned.

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 with the most votes and her best opponent is more than 60 percent.[17] Applying this limitation we lose 1569 observations, of which most are cases where only one candidate ran for office. Table 2.2 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 2.6 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.[18] 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.

2.4 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 $d^p_{i,t}$, where i refers to the community or town for which we observe elections, t refers to the election year[19] and p refers to the party under consideration. The variable that unambiguously determines treatment in our application is the margin of victory, $v^p_{i,t}$, which is

[17] 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.

[18] 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.

[19] 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 $d^p_{i,t}$. 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.

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 $v^p_{i,t}$ has to cross is $v_0 = 0$. The relationship between $d^p_{i,t}$ and $v^p_{i,t}$ is as follows:

$$d^p_{i,t} = 1\left[v^p_{i,t-1} > 0\right] \tag{2.1}$$

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 $y^p_{i,t}$. 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:

$$y^p_{i,t} = \delta_0 + \delta_1 d^p_{i,t} + \epsilon_{i,t} \quad for \quad |v^p_{i,t-1}| < \Delta \tag{2.2}$$

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:

$$y^p_{i,t} = \delta_0 + \delta_1 d^p_{i,t} + h(v^p_{i,t-1}, \theta) + \epsilon_{i,t} \tag{2.3}$$

where $h(\cdot)$ is some function that represents the influence of the margin of victory in the preceding mayor election on the voting outcome $y^p_{i,t}$. In this alternative specification, we run the estimation on the entire sample.[20]

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 $v^p_{i,t-1}$ must contain a random component. If it were true that a party had absolute certainty about

[20] 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.

the final vote outcome, the design would be invalid. Instead of 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.[21]

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 $d^p_{i,t}$ 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.

[21]Figure 2.5 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.

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 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.

2.5 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.

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

Table 2.3 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 2.3 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. 2.2 and the control function approach (columns 4-6) as in eq. 2.3. We observe a sizable and significant effect of around 4.3-5.9 percentage points throughout almost all specifications. This implies, that (on average)

Table 2.3: Interdependency between mayor's office and town council elections

	Dependent variable: Vote share in TCE in t					
	(1)	(2)	(3)	(4)	(5)	(6)
	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

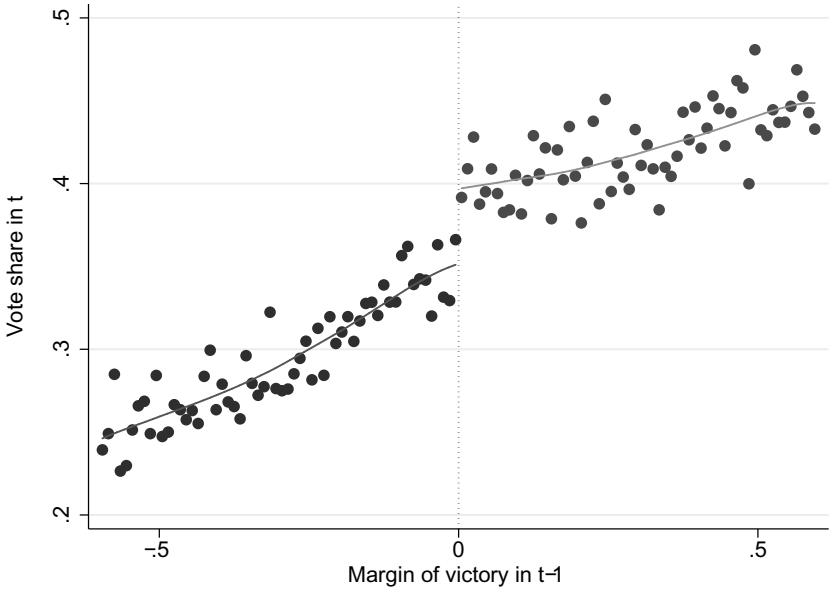
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of a party (which had a candidate in the preceding mayoral race) in a town council 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.

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 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 2.1, which illustrates a clear discontinuous jump at the threshold. Table 2.8 in the appendix shows that the effects are present irrespective of the individual parties.[22]

[22]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

Figure 2.1: 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.

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 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 2.7 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.[23]

higher when a social democrat narrowly makes it into office.

[23]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

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 2.3 (Panel 2) and figure 2.2. We find that when 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.[24] Figure 2.2 does not indicate a significant break of the overall trend at the discontinuity.

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 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.

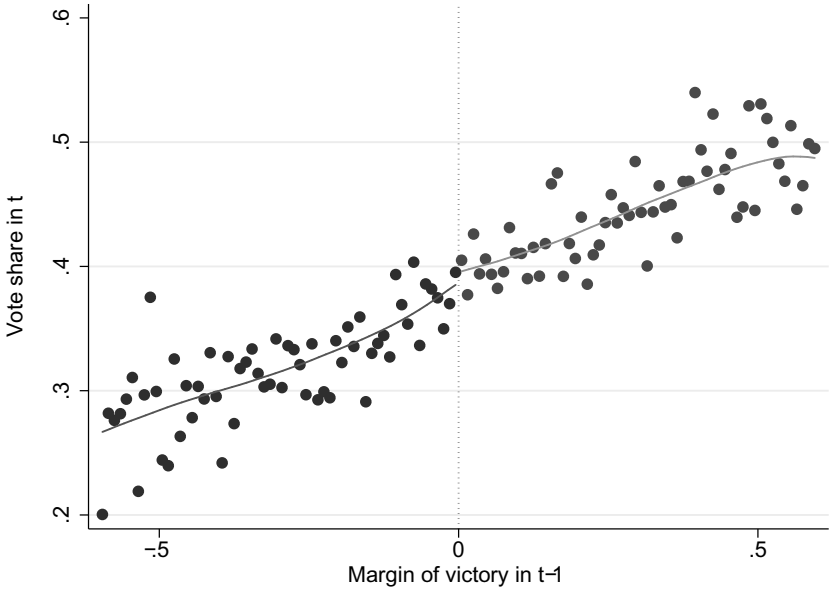
It is reasonable to assume that the incumbency externality effect could be larger (more positive) in the case of joint elections. In both designs the party that has access to the mayor's office can profit from the use of this resource (media presence, increased financial resources, time spent with the 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.

points (column 2), and a clearly significant effect of 20.4 percentage points in column 6.

[24]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.

Figure 2.2: 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.

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.

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

In table 2.4 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.[25]

The panels 1 and 2 in table 2.4 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.

[25]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 2.4: 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.

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 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.

2.5.3 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 2.5 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.[26] 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.[27] 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

[26]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.

[27]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 2.5: 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.

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

2.5.4 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 observations 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 2.3 and 2.4 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 2.5 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.[28] 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 2.9 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

[28]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.

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.

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.[29] 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

[29]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.

differences to the effects presented above.[30]

2.6 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 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.

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

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.

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 2.6: 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 2.7: 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 2.8: 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

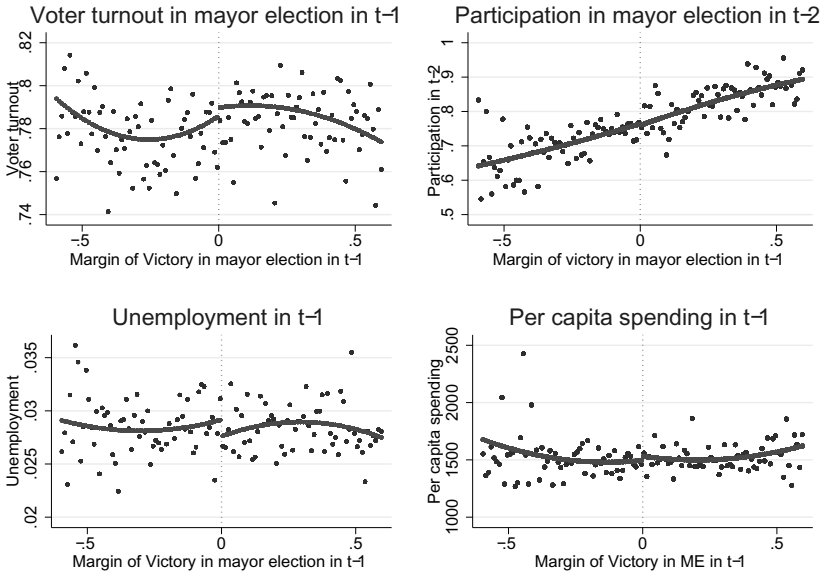
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of the respective party in a town council election. Columns 1-3 refer to results 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 2.9: 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

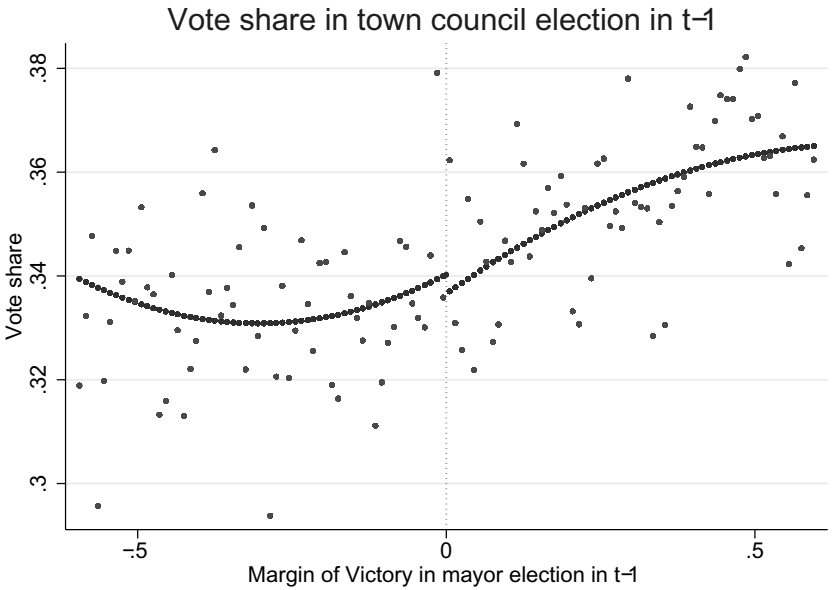
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust and clustered on the level of each individual municipality election. The dependent variable is the vote share of the respective party in a town council election. Columns 1-3 refer to results 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 2.3: RDD validity check - predetermined variables



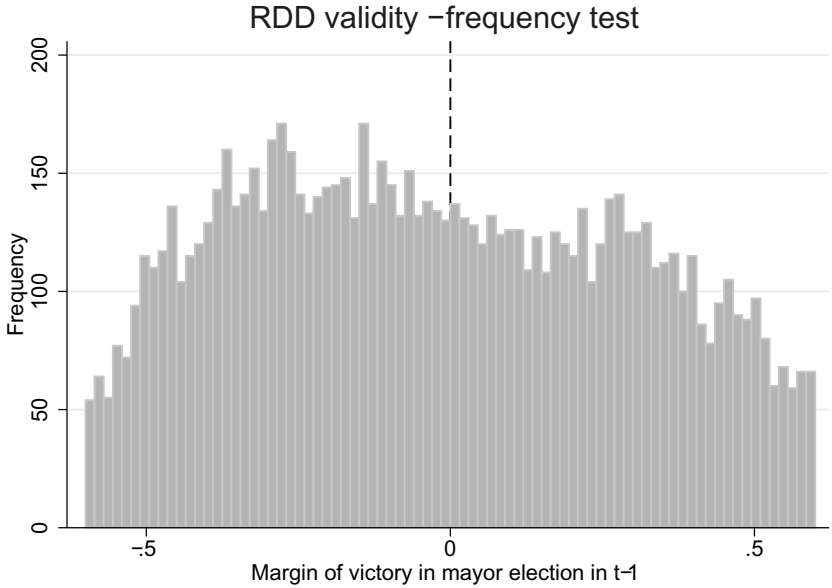
Notes: This figure illustrates that predetermined variables in $t-1$ are not affected by the treatment that occurs at $t-1$. Each bin in the graphs represents an interval of 1 percent in the margin of victory. The line fitted onto the scattered data is a polynomial function of degree two which is flexible 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 2.4: 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 2.5: 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.

Chapter 3

Do parties matter?

Estimating the effect of political representation in multi-party systems[1]

Abstract: This paper estimates the causal effect of political representation in local governments on tax policy in municipalities under a proportional election system. The main challenge in estimating the causal effect of parties on policy is to isolate the effect of representation from underlying voter preferences and the selection effect of parties. We use an instrumental variable approach where close elections provide the exogenous variation in our measures of representation: seat shares and voting power. Using data from German municipalities our estimation results suggest that representation does matter. The effects are mostly small, but statistically significant. Somewhat surprisingly, representation of the center-left party is found to lower the local taxes, whereas The Greens increase the property tax considerably. These effects remain robust to weighting voting power by the likelihood of coalitions and different definitions of close elections and the instrument.

[1] This paper is joint work with Christian Odendahl. Acknowledgments: We would like to thank Florian Ade, Tim Besley, Olle Folke, Magnus Johannesson, Juanna Joenson, Henrik Jordahl, Torsten Persson, Imran Rasul, Viktor Steiner and David Strömberg as well as seminar and conference participants at Tilburg University, Wissenschaftszentrum Berlin, DIW Berlin and UCL for helpful comments and suggestions. We would also like to thank Helke Seitz, Jenny Freier and Heike Hauswald who provided excellent research assistance. Ronny Freier would like to thank the Hedelius foundation for funding the research visit to Princeton University and the Jan Wallander and Tom Hedelius Foundation for generous financial support. Christian Odendahl would like to thank the Hedelius foundation for funding the research visit to the London School of Economics. The usual disclaimer applies.

3.1 Introduction

Does political representation matter for policy? In a majoritarian system, this question is usually equivalent to: does it matter who wins the election? The reason is that after Duverger's law a two-party system is likely to emerge. In a proportional election system, however, the question is often more complicated. We frequently observe a multitude of parties and winning an election is not as clear a concept as in a majoritarian system.

There are two reasonable measures of political representation for a multi-party system, both of which are equivalent in a two-party system. The first is seat shares. In a two-party system, the party with the higher seat share is the winner. In a multi-party system the seat shares of parties do not necessarily indicate the winner of an election: the party with the highest share may well find itself in the opposition to a coalition of smaller parties. Nevertheless, the seat share gives us a measure of representation in the legislature of the party in question.

The second measure is voting power.[2] Voting power can be investigated from two different angles, a policy-oriented and an office-motivated perspective. Since we are interested in the policy decisions of a legislature, we will use the purely probabilistic measure of voting power after Penrose (1946) and Banzhaf (1965).[3] In a two party system, voting power will either be one or zero, depending on who has the higher seat share. Thus, there is a simple mapping of seat shares on voting power. For multi-party systems on the other hand, voting power may differ substantially from seat shares. As an introductory example, consider three parties, where two big parties have 48% each, and the third party has only 4%. Their relative voting power (measured by the Banzhaf index) will be one third each because each party is equally useful in forming a winning coalition.

In the current paper, we estimate the causal effect of seat shares and relative voting power of parties on policy outcomes in a municipal council. Three key results emerge. First, there certainly are correlations between seat shares or voting power on the one hand and policy outcomes like taxes and spending on the other. However, these OLS results are misleading and the causal estimation shows different results. Second, political representation measured in seat shares and voting power does affect policy. And third, the political influence of small parties depends on measuring the voting power of these parties rather than approximating power with seat shares.

[2] See Felsenthal and Machover (1998) for details on voting power. The exposition here is based on their book.

[3] The office-motivated measure is the game-theoretic index after Shapley and Shubik (1954) which measures the power to acquire a share of a price, e.g. an office.

Estimating the causal effect of political representation on policy crucially depends on overcoming a fundamental endogeneity problem: the correlation of political representation with (unobserved) voter preferences. The observation that, say, a pro-business party obtains 3 seats or a certain amount of voting power in a town council and the local business trade tax is observed to be below average does not imply causality of political representation of the liberal party on tax rates. Ideally, we would like to run an experiment, where we assign political representation randomly to municipalities. In the perfect experiment, voter preferences are similar on average in treatment and control group and we could measure the causal effect of a higher seat share or more voting power for a specific party. This is (and should be) impossible.

In the absence of experimental data, a new literature has emerged in political economics that applies techniques from the program evaluation literature to draw valid inference from observational data. Pettersson-Lidbom (2008) and Lee, Moretti, and Butler (2004) were the first to use a regression discontinuity design (RDD) to estimate the effect of parties on policy. Lee, Moretti, and Butler (2004) estimate the effect of party affiliation in the US House of Representatives on policy. Under the assumption that vote outcomes are subject to some randomness, they argue that the party affiliation (Democrat or Republican) of a representative of a district is a quasi-random event if the margin of victory was very small. They show significant effects of party affiliation on the voting record of the representatives. Pettersson-Lidbom (2008) studies the policy effects of having a left-block majority (of one or more parties) in Swedish municipalities. The identification is also a RDD at the 50% vote margin. He finds significant effects of block majorities on tax rates and different spending categories. Ferreira and Gyourko (2009) investigate the effect of having a Democratic or a Republican mayor in office in US cities on fiscal outcomes in a similar RDD framework. They find effects of the mayor on policy only if Tiebout competition between municipalities is weak.

In all these studies, the threshold is the absolute majority of the votes that is needed to gain control of a certain political office or institution. The political setting in all studies is a two-party majoritarian system or an assumed two-block party structure in a proportional system. This has two consequences for the research design. First, there is a simple mapping of seat shares on voting power, as outlined above. Second, it is possible to use a simple RDD framework at a fixed threshold. While our research design also relies on close elections, we are specifically interested in different measures of political representation that are not only binary variables of winning or losing. Moreover, in proportional election systems blocks of parties are very difficult to define, especially in local politics. Finally, if we take the proportional system seriously, there are no fixed thresholds at which a party gains or loses a seat or voting power.

Folke (2010b) was the first to estimate the effect of seat shares in Swedish municipalities on

policy outcomes under a proportional election system. He finds interesting results for the green party, which has an effect on local environmental policy, and an anti-immigration party, which he shows to have an effect on local immigration policy. He finds no effect of seat shares on taxes. A study that uses a related identification design is Liang (2009). He estimates the effects of party representation on *political* outcomes in subsequent local elections. The two incumbency effects that he investigates are the effect of holding a seat and the effect for each member of the council to be part of the government coalition. He finds a positive incumbency effect of holding a seat, but no incumbency effect of being part of the government.

Identification in Folke (2010b) stems from observations in which a certain seat allocation was very close: he determines counterfactual observations as those in which, for example, party A was either very close to winning or very close to losing a seat. By assuming that such a close seat allocation is the result of an as-good-as-random event, he solves the endogeneity problem of political representation and is able to isolate a causal effect of seat shares on policy.

In our analysis, we generally build on the estimation strategy developed in Folke (2010b). However, we investigate both seat shares and voting power and show that seat shares can have no effect while voting power does. Second, we outline a conceptual framework of randomness in elections. This conceptual framework leads naturally to a different method to identify close seat allocations. We shortly describe this method in the following for the purpose of introduction.

In general, a certain seat allocation depends on both, the council size and the entire vector of votes of all parties. Both are arguments in the seat allocation function that translates votes into seats. Folke (2010b) uses an algorithm based on this seat allocation function to calculate the minimal vote change that is required for a seat to change. This is done for every party in the council. This minimal vote change is his measure of closeness of a seat allocation for a specific party.[4]

In our method, we repeatedly perturb the vote vector for each observation by adding a random variable to the votes of each party. Then we simulate the seat allocation and voting power and observe whether they change. Observations whose seat allocations or voting power change often are considered close and are used for identification for the respective parties. Those observations for party A with repeated seat *gains* during the perturbations will receive a *negative* treatment because they were close to gaining a seat but did not receive it in the actual seat allocation. Positive treatment is assigned accordingly. This treatment variable is then used as an instrument for seat shares. For voting power, we calculate the average gain

[4]We will discuss this in more detail in section 3.3.2.

or loss of voting power when voting power does change. The instrument for voting power is then given by combining this average with the number of perturbated simulations in which it did change.

We have compiled a new data set that combines information on both election outcomes and fiscal data at the municipal level in Germany. As of now, we have complete data for the two recent municipal elections in four out of 13 German states: Thuringia, Bavaria, Hesse and North Rhine-Westphalia.[5] In the current analysis, we use the results of about 7200 independent elections and subsequent fiscal outcomes. The election data contains information on total votes and seats for all parties. For the fiscal variables, we have yearly information on the three local tax rates (trade tax as well as property tax A and B) and the aggregate spending data for each municipality.

The paper is organized as follows. In section 2, we outline the electoral rules and give some background on the responsibilities of German municipalities. Section 3 presents our empirical strategy, including the methodological background on how to define close elections and how to calculate voting power. The data is presented in section 4, section 5 contains the results before the analysis is concluded in section 6.

3.2 Institutional background

In this section, we describe the seat allocation functions that are used in German municipalities and give some institutional background on local politics in Germany.

3.2.1 Elections and electoral rules

In Germany, there are four tiers of government: federal, state, county and municipality. Our focus is on municipal elections and policies. In municipalities, the legislative body is the council (*Gemeinderat* or *Stadtrat*, depending on the size). It is elected every 5 years[6] in a proportional election system, sometimes with elements of a mixed member proportional system.[7] The parties that participate are the 5 major parties in Germany: a center-right

[5]The three states Berlin, Bremen, and Hamburg are city-states where municipality, county and state coincide. In total there are 16 German states. We have to restrict the analysis to those four states due to limitation in data availability.

[6]Except in the two states Bremen and Bavaria where it is 4 and 6 years respectively.

[7]A mixed member proportional system combines elements of a majoritarian election in districts and proportional elections of the overall council. The resulting composition of the council, however, will be determined by the proportional election part of the election – unless a party won more direct seats than it should receive given the proportional election result. As we cannot identify such cases, we will have some measurement error in the specification for the state of North-Rhine Westphalia where such a system is used.

conservative party (CDU - Christlich Demokratische Union), a center-left social democratic party (SPD - Sozialdemokratische Partei Deutschlands), a pro-market liberal party (FDP - Freie Demokratische Partei), the Greens (DIE GRÜNEN) and the left socialist party (DIE LINKE), as well as some minor small parties and local parties.[8] The Free Voters (Freie Wähler), while purely local and independent, are often member of a collective of Free Voter parties, either regional or on a state level. The mayor of the municipality is elected by the public as well. Often, the timing of the elections coincide. The mayor is head of the administration and also member of the council in most states. Even though the mayor proposes the budget and generally sets the agenda, the council is free to change it and has the power over the legislation.

In every proportional election system, a seat allocation function is used to distribute (discrete) seats to parties based on their votes. This seat allocation function is by design a step-wise function since there are more votes than seats. The locations of these steps for party A, however, are not predetermined. They are jointly determined by all arguments of the seat allocation function: the votes of *all* parties, the sum of those votes and the council size. In order to clarify this issue, we first describe the seat allocation functions used in German municipalities before returning to the question of where the seat thresholds for a party A lie.

The states choose the allocation method for their municipalities such that all municipalities in a certain state have the same seat allocation function. There are two different seat allocation functions used in German municipalities.[9]

The first is the largest remainder method (or Hare-Niemeyer method). The first step in this method is to calculate the Hare quota: total votes divided by total seats. This gives the "price" in terms of votes that a party has to "pay" for one seat. Then the votes of all parties are divided by this price. The resulting quotient is the exact number of seats that each party should receive in case of perfect proportionality. However, it is rarely, if ever, an integer. Therefore, the largest remainder method allocates the seats according to the integer of this quotient. This results almost always in at least one remaining seat. The remaining seat(s) are then distributed according to the rank order of the remainders of each party.

[8]The CDU is the party of the current federal chancellor Angela Merkel, of Helmut Kohl and Konrad Adenauer, the SPD is the party of Gerhard Schröder, Helmut Schmidt and Willy Brandt, the FDP is the party of Guido Westerwelle and Hans-Dietrich Genscher, The Greens is the party of Joschka Fischer and The Left is the former PDS which was founded as the successor party of the Socialist Unity Party (SED) in former East Germany.

[9]The state of North Rhine-Westphalia recently reformed its voting legislation and is now using a third allocation function. However, under the ruling law in our time period we only observe the two functions explained in this section.

The second method is the highest averages method (or d'Hondt method).[10] This method proceeds just as the largest remainder method by calculating the price of a seat (the Hare quota), dividing each party's votes by this price and then distributing seats according to the integer of the resulting quotient.[11] As under Hare-Niemeyer, there will be at least one remaining seat. Under the d'Hondt method, however, the price of the seats is lowered in small increments in order to distribute the remaining seat(s). The procedure is repeated with a lowered price and seats are allocated according to the integer until the procedure results in a complete allocation of all seats.[12]

Now we return to the question of where a seat threshold lies for party A. Consider a council with 10 members, and 100 voters. The seat allocation function is the Hare-Niemeyer method. The Hare quota or price is 10 votes per seat. If party A has 56 votes, and parties B and C 22 each, party A receives 6 seats (5 seats from the integer, and the remaining seat for having the largest remainder) and parties B and C two each. If parties B and C have 27 and 17 votes respectively, party A receives only 5 seats for its 56 votes. However, it was very close to gaining the sixth seat. Therefore, the distance to a seat change depends on the whole composition of the vote vector, not only on the votes for a specific party. A similar argument would hold under the d'Hondt method.

3.2.2 Responsibilities of German municipalities

The local government is head of the administration of the municipalities, and it manages a yearly budget of about 1400 Euro per capita on average. This amounts to roughly one third of total per capita government spending in Germany. Moreover, two thirds of all investment spending is allocated by municipalities, and they employ around 40% of all state employees.[13]

The municipalities set three tax rates whose revenues completely accrue to them: a trade tax for businesses and two types of property taxes. The local budget also contains a share of the income tax revenue raised in the municipality and a part of the VAT revenue. Setting the rates for those two taxes is however not a municipal responsibility. Another part of their budget is

[10]There are several highest averages methods, for instance the method of Sainte-Lague. Since only the method of d'Hondt is used in Germany, we will describe only this method.

[11]There are several different ways of reaching the final seat allocation in the highest average method. The other common form is the use of a divisor series. Both yield the exact same result.

[12]The two approaches may lead to different seat allocations since the d'Hondt method slightly favors larger parties. The intuition is as follows: the "price" of a seat in the d'Hondt method is lowered until the distribution of seats according to the integer leads to a full distribution of seats. If we regard the votes of a party as its budget, lowering the price is best for the party that already has the most seats – compared to the Hare-Niemeyer method where just the remainders are compared.

[13]See Bundesbank (2007).

federal or state allocated funds that the municipalities administer, e.g. for public schools or social services.

The municipalities spend their revenues in the following areas:

- general administration
- public order
- public schools, cultural centers and services, social services (elderly care, immigration housing, child care, youth services)
- sport and recreation
- infrastructure investments (housing projects, roads), public transport, business development, management of public firms

However, the division of tasks between the tiers is often complex: which tier of government pays for the service or investment, or for part of it, which tier enacts the law or by-law, which tier administers the service and so forth. More often than not, each of these tasks is itself divided between the tiers of government. Thus, the degree of discretion for the municipality varies by field. While municipalities are, for example, completely free to decide about cultural or recreational institutions, most social services have to be administered within clearly defined laws and by-laws. Those rules are to a large extent enacted by the federal or state government and the municipalities use mandated funds from higher tiers. However, even there the municipalities can choose to apply rules strictly or not. What is more, they can for example supplement social services, expand or restrict child care or start an initiative for public transport.

3.3 Empirical strategy

In this section, we will outline the empirical strategy for both, seat shares and voting power. We will start with seat shares and the problems of identifying a causal effect. Then we will turn to using close election results only and describe in subsection 3.3.2 how we define close elections. In the last subsection 3.3.3, we will discuss voting power and show how we can construct instruments for voting power in our setting using close elections.

3.3.1 Estimation setup for seat shares

In general, the effect of parties on policy is hard to distinguish from the underlying preferences of the voters. Assume that we have parties A, B and C, a vote vector $\mathbf{v}_i$, seat shares s_i^j and outcome Y_i in municipality i. We would like to estimate the effects of the parties on policy outcome Y_i. Experimental data is impossible to obtain and OLS of the outcome on seat shares or voting power is biased by underlying voter preferences if we estimate

$$Y_i = \alpha + \beta_B s_i^B + \beta_C s_i^C + e_i. \tag{3.1}$$

The error term in this equation contains not only a random component w_i but also unknown voter preferences ϕ_i:

$$e_i = w_i + \phi_i.$$

These voter preferences affect both the outcome and the seat shares, as voter preferences are a main determinant of the election results. Therefore, $E(s_i^j e_i) \neq 0$ and the coefficients will be biased.

However, we could use the fact that the seat allocation function is a discontinuous function of the vote vector: a party can only gain a full seat, not a fraction of a seat. That in turn means, vote shares and seat shares do not necessarily coincide. If we interpret vote shares as reflecting voter preferences and take seat shares as our measure of political representation, we could estimate:

$$Y_i = \alpha + \beta_B s_i^B + \beta_C s_i^c + f(\mathbf{v}_i) + u_i. \tag{3.2}$$

The problem with this approach is that the function $f(\cdot)$ needs to capture the voter preferences ϕ_i correctly over the whole range of possible values for $\mathbf{v}_i$ – and for the interaction of parties – in order for the seat shares to be uncorrelated with the error term. Formally, the condition is:

$$E(s_i^j e_i \mid f(\mathbf{v}_i)) = 0.$$

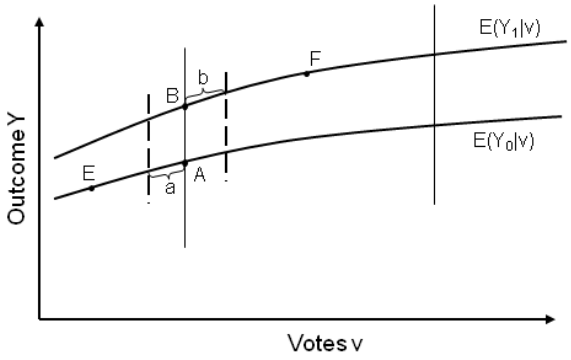
It is hard to argue that we will be able to accomplish that.

However, there is a different way to use the step-wise nature of the seat allocation function: only use seat allocations that were close to being changed for identification. This relaxes the need to specify the whole function $f(\cdot)$ correctly. The two assumptions that we need to make is that $f(\cdot)$ is continuous at the steps, and that it is correctly specified close to these steps.

As explained in section 3.2.1, these steps are not predetermined, but depend on the votes of *all* parties, the total votes and the council size. This in turn means that the steps could be anywhere in the vote vector space. Does that imply that we need to specify $f(\cdot)$ correctly over the whole range, just as in equation 3.2, if we use close seats only?

In order to answer this question, consider an extreme case where we have almost unlimited amounts of data. This implies that we can get arbitrarily close to seat thresholds. If we are just one vote away from a seat change for each close observation that we identify, it is safe to argue that the averages that we calculate on both sides (for close seat gains and close seat losses) will be unbiased and any difference is the result of the seat change. In figure 3.1, we draw the potential outcome functions and a hypothetical seat threshold (the solid vertical line), using the votes for this party on the horizontal axis.[14] Our interest lies in comparing A and B, because the difference between the two is the causal effect of representation on policy outcome Y. If we can get arbitrarily close to the threshold, we get unbiased estimates of A and B.

Figure 3.1: Using the discontinuity



The figure is based on Lee and Lemieux (2009).

Of course, we need to relax this assumption since we do not have unlimited amounts of data.

[14] As outlined above, this is not entirely correct because thresholds depend on the whole vote vector, not only the votes of one party. The figure is therefore just for illustrative purposes.

However, we stay very close to the seat thresholds. In figure 3.1, we stay between the dotted vertical lines. In these ranges we calculate averages on both sides, in the segments a and b. Without any control for the underlying functional relationship between the vote vector (the forcing variable) and the outcome variable, those averages will be biased: the average over all observations in b will be too high, and too low over all in a.[15]

Therefore, we have to use the function of the vote vector to account for this distance to a threshold when calculating averages. This distance on the other hand is small, and so is our reliance on the functional form of $f(\mathbf{v}_i)$: we only need to make sure that the averages on both sides of the thresholds correspond to the points A and B that we are ultimately interested in. In contrast to the specification in equation 3.2, we do not use $f(\mathbf{v}_i)$ to compare observations in E and F.[16]

For the close seats in a council of size Z_i we will define a treatment variable t_i^j. This treatment variable takes on the value $+\frac{1}{2}/Z_i$ if the close seat was just above a threshold, that is, the party was close to *losing* this seat, and $-\frac{1}{2}/Z_i$ in the opposite case. We call the two values positive and negative treatment respectively:

$$t_i^j = \begin{cases} +\frac{1}{2}/Z_i & \text{for positive treatment,} \\ -\frac{1}{2}/Z_i & \text{for negative treatment,} \\ 0 & \text{otherwise.} \end{cases}$$

The values ensure that the difference between gaining and losing a seat is one divided by the council size, that is, the difference between positive and negative treatment is measured in terms of seat shares.[17] This facilitates the interpretation of the coefficients and makes them comparable to OLS results, as we will discuss later.

Moreover, we define the variable c_i^j:

$$c_i^j = \begin{cases} 1 & \text{for positive and negative treatment,} \\ 0 & \text{otherwise.} \end{cases} .$$

These closeness dummies c_i^j ensure that our treatment variables t_i^j only act as instruments for those observations for which party j was close to gaining or losing a seat, that is, where the instrument is unequal to zero. The observations for which it was not close – where the instrument is zero – will only be used in order to add precision to the estimation of the control function $f(\mathbf{v}_i)$ and increase overall efficiency of the estimates.

[15] See Hahn, Todd, and van der Klaauw (2001) for a formal treatment of this bias.

[16] In other words, we rely on $f(\mathbf{v}_i)$ only for very short distances, for which even an linear approximation might suffice.

[17] Below, we will describe in detail how we assign positive and negative treatment for close elections.

We could use these variables in a regression of the outcome y_i without instrumenting. If we had just one forcing variable and could get arbitrarily close to the thresholds, the difference between negative and positive treatment observations would be equal to one seat and we could just use it in a regression.[18] However, the forcing variable is in fact a vector and we cannot get arbitrarily close to the thresholds. Therefore, our treatment variable is only approximately correct and we therefore use an instrumental variable strategy, where we instrument for seat shares with our treatment variable. This ensures that our regression coefficients represent the effect we want to estimate: the effect of 1% additional seat share.

The regression that we are going to estimate therefore takes the following form:

$$y_i = \alpha + \sum_j \beta_j s_i^j + f(\mathbf{v}_i) + \mathbf{X}_i\gamma + \mu_i + e_i, \tag{3.3}$$

where $f(\cdot)$ is the flexible function of the vote share, $\mathbf{X}_i$ is a set of control variables and μ_i is a municipality fixed effect. We instrument for the seat shares s_i^j with our treatment variables t_i^j. The set of control variables include population of the municipality, dummies for each state-wide municipal election and the closeness dummies c_i^j.

Note that we leave out one party, the center-right CDU, in all specifications. The reason is that seat shares add up to one by definition. This is also true for relative voting power, as we discuss below. The CDU is the biggest party and takes part in almost all of the elections in the western states, and most of the elections in Thuringia. The interpretation of the β's is therefore: the effect of an increase in party j's seat share by 1% at the expense of the CDU on the policy outcome.[19]

There are two important assumptions for the validity of our research design: individuals (here: parties or voters) cannot manipulate the vote vector such that a party ends up just above or just below such a seat threshold. And second, parties cannot manipulate policy in anticipation of a close election.

Election manipulation is (hopefully) impossible in Germany, and voters have no precise information about which side of a seat threshold parties are when making their voting decision (they neither know the voting behavior of others, nor do they understand the seat allocation functions), so we can safely argue that the first condition is satisfied.[20] The other issue is more difficult to dismiss a priori. However, we show evidence that there are no party effects

[18] We would correct for different council sizes, of course.

[19] We measure seat shares from 0 to 100 in order to allow for this interpretation.

[20] Freier (2011) shows for mayor elections, that are easier to understand and manipulate, that there are no signs of manipulation by the voters.

for policies enacted *before* the election. Moreover, we include municipality fixed effects and a dummy for close elections in order to control for some of this variation.

The next step in the empirical setup is to define which seats can be considered close and to assign positive and negative treatment accordingly. In contrast to standard RDD settings, the cutoff points in our setting are not immediately obvious. Our forcing variable is a whole vector, and the votes for one party alone tell us nothing about those cutoffs. Therefore, we need to find a way to determine which seat allocations are close.

3.3.2 Defining close elections

In this section, we discuss how we define closeness of a seat allocation. We intend to use close elections as a source of exogenous variation because we assume that there is some random component in elections (e.g. the weather, errors in polls or popularity shocks). Thus, randomness and closeness are linked. In fact, the way we think about the randomness in elections determines the way we define which elections or seats are close. In the following, we outline a conceptual framework of randomness in elections. From this conceptual framework, we deduce our definition of close elections and the method that we use.

To begin the discussion of a conceptual framework, we may consider sincere voting and assume a continuum of voters' bliss points along one policy dimension. Let us further assume that parties are exogenously given and their allocation along this dimension is known and fixed. However, the distribution of voting decisions along this dimension changes, according to popularity shocks to the parties. In other words, voters take policy *and* party popularity into account when making a voting decision.

Under the assumption that everyone votes, parties may gain or lose votes only at the expense of other parties (and not at the expense of non-voters). In our simple one dimensional setting, the parties next to each other are gaining or losing from one another. For higher policy dimensions, that is, a policy vector, the voter migration pattern will be more complex. Whether a seat allocation was close for party A then depends on how much voter migration between parties is necessary for party A to gain or lose a seat. Alternatively, a close seat is one that changes often in repeated elections with the same preferences and party locations, but newly drawn popularity shocks.

As a second step, we can introduce the voters' decision to participate in the election. The participation shocks may be independent across parties (mobilization shock), or related (weather

shock).[21] Thus, in this setting parties not only gain votes at the expense of other parties (their neighbours) but also from non-voters. Whether a seat of a party is close depends as before on the voter migration between parties, but it is now combined with the participation shocks. Alternatively, a close seat is one that changes often in repeated elections with the same preferences and party locations, but newly drawn popularity and participation shocks. The difference between popularity and participation shocks is not of much practical relevance, and is for the most part a purely theoretical concept. One could think of just a popularity shock for parties that not only leads to voter migration between parties, but equally to changes at the participation margin.

Another story assumes strategic voting. Here, each voter has a preferred seat allocation in parliament that best matches her own policy preferences. Polls and other information provide a background against which the voters can assess and formulate their voting strategy. However, voters need to interpret the information. In interpreting the information, voters make mistakes and thus their voting is a combination of a strategic choice and a random error term. The error might consist of two parts: a common error that may be caused by erroneous news reports or some other correlated errors in interpreting the information, and an individual interpretation error. A close seat is therefore one that changes often in repeated elections with the same preferences and information, that is, the same strategically optimal decision, but newly drawn errors. The migration patterns are less clear here: strategic voters may prefer to vote for a rather "distant" party, to prevent an unwanted coalition or party from winning.

Based on these notions of election randomness, we propose the following definition of closeness of elections. For each observation i with vote vector $\mathbf{v}_i$ and the resulting (known) seat allocation we add a vector of random variables to the vector of votes. We then calculate the resulting seat allocation from this perturbed vote vector and track whether the seat allocation has changed. This procedure is repeated multiple times. In practice we add a vector $\mathbf{r}_i$ of independently normally distributed random variables to the vote vector of observation i with expectation zero and variance $\left(kv_i^j\right)^2$. The standard deviation of these random variables is thus k percent of the votes of party j. This ensures that for a small party the perturbation is small. Seat allocations for party j in municipality i are considered close, if in repeated perturbations of the vote vector, the seat allocation for this party j changes more than q% of the time.

Next we discuss our choice of q and its interpretation. For normally distributed random variables, roughly $\frac{1}{3}$ of the probability mass lies outside the interval of the mean plus/minus

[21] It is hard to imagine a participation shock that is the same for all parties because any reason for such a shock may affect the voters of different parties differently.

one standard deviation. Moreover, we observe from our perturbations that almost all seat changes go in one direction only, not both.[22] It follows that if we observe one additional seat for party j in municipality i in $\frac{1}{6}$th of our perturbations, we know that roughly one standard deviation in vote change for this party was required for this change in seat allocation.[23] When we vary the degree of closeness in later specifications, we will only change k in the standard deviation of the random variables. The share q will always be $\frac{1}{6}$ in order to allow for this interpretation.[24]

In accordance to the specification above, we define close seats as those that change more than $\frac{1}{6}$th of the time, and assign treatment as follows. We assign negative treatment to those that were close to gaining an extra seat and therefore, had one extra seat in more than $\frac{1}{6}$th of the perturbations. And we give positive treatment to those that were just above a threshold and therefore close to losing one seat.

In future work, we would like to constrain the vector of random variables to reflect a certain covariance structure in the shocks of parties (instead if assuming independence). This covariance matrix should reflect the actual migration patterns between parties and at the participation margin. We intend to estimate this migration from the municipal election data. As a result, the closeness would be reflected more realistically compared to using independently distributed random variables.

In contrast to our approach, Folke (2010b) identifies close elections by analytically computing the vote change from the realized vote vector to the closest point in the vote space where the seat allocation changes. This *minimal* vote change requires that there is no voter migration except the migration that leads to exactly this vote change. This approach runs into problems if this migration pattern is utterly unrealistic (a green party gaining votes at the expense of a right-wing party for instance) and may therefore incorrectly identify close seats.

In our approach, especially when we add the estimated migration between parties, randomness in the elections translates naturally into determining which seats were close. Besides this conceptional advantage of the perturbed simulations there are two additional benefits to our approach. First, the implementation is substantially easier. Folke (2010b) develops a complex

[22] It is possible, that a seat is close in both directions: for instance, if three parties have very similar remainders in the Hare-Niemeyer method. However, such situations occur only rarely for very small perturbations like the ones in the present paper.

[23] That does not mean that for every seat change one standard deviation vote gain or loss was necessary for the party in question. A seat change for party A can be the result of vote changes for other parties, too. The interpretation given here only offers an idea of the magnitudes involved.

[24] If we use $k = 2\%$, for a party that received 100 votes, we will perturb its votes such that the vote count is between [98,102] in about 66% of the cases and between [96,104] in 95% of the cases.

algorithm to compute the minimal vote change. This algorithm is specific to the electoral system in use in Sweden and is not easily adjustable to variations in the electoral system. Specifically, it cannot be used without substantial adjustments in the electoral system in some of the German states. Our approach does not depend on the specifics of the electoral system, but can be applied to any system. This is especially valuable for us, as each German state does in fact have its distinct electoral rules. Secondly, it is simple to implement different structures of randomness, for example a voter migration matrix or conditioning the randomness on the party's size.

3.3.3 Estimation setup for voting power

In this section, we will first define voting power and then discuss our instrumental variables strategy. As outlined in the introduction, we use the policy-oriented measure of voting power after Penrose (1946) and Banzhaf (1965). This a purely probabilistic measure which is defined as follows.[25] In the set N there are n different parties with weights equal to their respective seat shares. The quota is one half which means that every coalition of parties whose seat shares sum up to more than one half is a winning coalition. The power set 2^N with its 2^n elements consists of all possible coalitions plus the empty set. For a party A, there are $2^{(n-1)}$ possible coalitions that it could be a part of. Party A is said to be critical in a coalition, if this coalition (with A) is a winning coalition, but without A it is not.

Then the (absolute) voting power of party A is defined as

$$\beta_a = \frac{\eta_a}{2^{(n-1)}}, \tag{3.4}$$

where η_a is the number of times party A is critical. If we assume that all coalitions are equally likely to form, then β_a measures the a priori probability of party A to be in a position to change the fate of the decision. It is this interpretation that makes this measure of voting power a measure of influence on policy.

Voting power measured in this way does not necessarily add up to one, as indicated in the introduction. To construct an index of voting power that adds up to one, we divide the η_a not by $2^{(n-1)}$, but by the sum of the η_j's:

$$\beta'_a = \frac{\eta_a}{\sum_{j \in N} \eta_j}. \tag{3.5}$$

Which one is the appropriate measure of voting power in our context? Note first that in case of an absolute majority for a party A, the absolute and the relative measure of voting power will

[25] The definition and discussion of voting power is based on Felsenthal and Machover (1998).

be equal to one for this party, and zero for all other parties. If there is no absolute majority there will be negotiations about coalitions. And in these negotiations, the parties with more *relative* voting power will be more likely to influence policy. The absolute measure of voting power is less relevant because even if party C has some absolute voting power, the resulting policy will be more in line with those of party A and B if their absolute voting power is even higher. Therefore, we will use relative voting power as our measure of representation.

As indicated above, the (a priori) measure of voting power assumes that all coalitions are equally likely to form. This might be unlikely to hold in practice.[26] In robustness tests, we will add weights to the coalitions when we calculate voting power. In practice, we will locate parties in a two-dimensional policy space. These dimensions capture unknown policy dimensions (often called left-right or liberal-conservative). The distance between two parties is then a measure of the likelihood of a coalition between the parties. This likelihood will then be the weight of this coalition:

$$w_{kj} = 1 - (d_{kj})^s, \tag{3.6}$$

where d_{kj} measures the distance between party k and party j, and $s > 1$ affects the curvature of our distance measure. It only measures the curvature because we normalize the distance between the most extreme parties to be unity. For those parties (like a strongly right-wing party and a communist party) the weight on their coalition will be zero which implies that they do not receive any voting power from this coalition, even if it were a winning coalition and each of them were critical. If we have a coalition of three or more parties, the distance within this coalition will be equal to the distance between the two parties within that coalition that are farthest away from each other.

Relative voting power with weights will then be calculated as:

$$\beta_a'^w = \frac{\eta_a^w}{\sum_{j \in N} \eta_j^w}, \tag{3.7}$$

where

$$\eta_j^w = \sum_{j:critical} w_{kj}. \tag{3.8}$$

In words, η_j^w is the sum of all the weights of those coalitions in which party j is critical.[27] So

[26] The a priori voting power measure is used for example to analyse different voting rules, where it is very useful. In our context, where we want to analyse the power of known parties with a certain policy position in a simple majority rule council, the a priori measure is less useful.

[27] Bilal, Albuquerque, and Hosli (2001) propose a similar approach to weighted relative voting power.

far, we have presumed specific party locations since we are still in the process of estimating them. However, already these guesses illustrate how voting power and subsequent results can change, once voting power is weighted by coalition likelihoods. We return to the comparison between a priori and actual voting power in the results section.

Now we turn to the specification of the regression and how we define our instrument. The regression that we would like to estimate is

$$y_i = \alpha + \sum_j \beta_j p_i^j + f(\mathbf{v}_i) + \mathbf{X}_i\gamma + \mu_i + e_i, \tag{3.9}$$

where p_i^j is relative voting power of the parties, $f(\cdot)$ is the flexible function of the vote share, $\mathbf{X}_i$ is a set of control variables and μ_i is a municipality fixed effect. Note that we have to leave out one party as relative voting power adds up to one. However, voting power is endogenous so we instrument for voting power using close election outcomes in the following way.

We again perturb the vote vector of each observation, simulate the new seat allocation, but this time also calculate the voting power of the parties under this new seat allocation. If one seat switches there are three possible consequences for voting power. First, nothing changes because the seat change was not crucial for whether a coalition is a winning coalition or whether a party is critical. As an example, consider an absolute (super)majority for party A where this party A loses one seat but maintains its absolute majority. Second, the voting power of those parties changes that had a seat change. For instance, if party A has 6 and party B has 7 seats, and after the perturbation party A receives 7 and party B 6 seats. For the other parties, nothing changes. With weighted voting power, this scenario is almost impossible. And third, the voting power of more than two or even all parties changes. This is far from uncommon: when a seat change leads to different winning coalitions, the voting power of all parties is likely to change, especially with weighted voting power.

To construct our instrument, we first count the number of changes in voting power for each party in a municipality i during the perturbations. However, the size of the jump may also contain useful information that we would like to keep. Therefore, we also calculate the average change in voting power for the times that it did in fact change. Our instrument is then

$$z_i^j = \begin{cases} \frac{1}{2}(p_i^j - \bar{p}_{i,perturb}^j) & \text{if it changes more than } q\% \text{ of the time,} \\ 0 & \text{otherwise.} \end{cases}$$

where $\bar{p}_{i,perturb}^j$ is the average voting power of party j in municipality i during the perturbations when it in fact changed. The reason for dividing the instrument by two is the same as for the

seat shares: we compare observations that had a positive difference to those with a negative difference in the instrument specification. If we take the full difference for each observations, we in fact double the difference.

One might wonder how we can use z_i^j as an instrument for p_i^j when the former contains the latter. However, the difference between p_i^j and $\bar{p}_{i,perturb}^j$ is in fact unrelated to p_i^j when the election outcome was close – under the assumption that some form of randomness in elections determines whether you are close and above or close and below a threshold. We again let q be $\frac{1}{6}$ because this allows for the interpretation that roughly one standard deviation (of our random variable) was necessary to induce this shift.

3.4 Data

We have compiled a new data set that combines information on both election outcomes and fiscal data on the municipality level in Germany. We use data from four German states: Bavaria, North Rhine-Westphalia and Hesse from the western part and Thuringia from the eastern part.

Table 3.1: Data set - descriptive statistics for the election data

Election	Observations	Participation rate in the elections for				
State/Year		CDU	SPD	FDP	Greens	Left
Thur 1999	766	0.69	0.38	0.18	0.01	0.32
Thur 2004	716	0.66	0.33	0.14	0.01	0.29
Hes 1997	417	0.98	1.00	0.24	0.55	0.00
Hes 2001	414	0.98	0.99	0.50	0.49	0.00
NRW 1999	396	1.00	1.00	0.82	0.84	0.05
NRW 2004	396	1.00	1.00	0.88	0.82	0.06
Bay 1996	2056	0.84	0.66	0.08	0.18	0.00
Bay 2002	2056	0.83	0.65	0.09	0.16	0.00
Total	7217					

Notes: The table shows the number of observations for each election in the first column. The remaining columns show how often the different parties participated (in shares) in the respective elections. The abbreviations for the states are: Thur - Thuringia, Hes - Hesse, NRW - North-Rhine Westphalia, Bay - Bavaria. The abbreviations for the party are: CDU - conservative center-right (*Christlich Demokratische Union*), SPD - socialdemocrats center-left (*Sozialdemokratische Partei Deutschlands*, FDP - pro-market (*Freie Demokratische Partei*), The Greens - green party (*Buendnis 90/Die Gruenen*), The Left - communists (*Die Linke* former *Partei des Demokratischen Sozialismus*). *Source*: Own calculations.

Table 3.1 shows the descriptive statistics for the political variables. For each state we have election data on two municipal elections. There are between 400 and 2050 municipalities in

each state. The center-right party (CDU - Christlich Demokratische Union) and the center-left party (SPD - Sozialdemokratische Partei Deutschlands) participate in almost all elections in western states and in many communities in the eastern state. The green party (DIE GRÜNEN) and the pro-market party (FDP - Freie Demokratische Partei) participate in roughly half the elections in the western states (except in Bavaria), but in considerably less elections in the east. For the socialist party (DIE LINKE or PDS), which had a strong focus on eastern Germany until recently, the pattern is the reverse.

Table 3.2 shows the descriptive statistics for the fiscal outcomes. For all these outcomes, we took the average over the period between two elections. We left out the data from the election years because we are unable to assign them to a government term.[28]

Table 3.2: Data set - descriptive statistics for the fiscal data

Fiscal Category	Observations	Mean	Std. dev	Min	Max
Trade tax multiplier	7217	329.4	38.3	170	490
Property tax A multiplier	7217	287.5	68.2	0	800
Property tax B multiplier	7217	318.0	52.8	140	800
Total Spending	7217	1479.6	591.0	504.0	17746.7
Spending on Personel	7217	262.1	111.9	17.9	1268.2
Investment Spending	7217	303.2	222.7	2.8	3088.5
Total Debt (stock)	4035	710.7	588.4	0.7	13340.2
Revenue from Fees	7217	142.4	124.0	0	2653.4
Revenue from Invest Grants	7217	100.6	108.4	−6.9	2381.1

Notes: The table shows the descriptive statistics for the fiscal data used as outcome variables in the analysis. The information on the local taxes refer to the respective multipliers in the tax formula. These multipliers are bounded between 0 - 800 (in the period of observation). The data on the aggregate spending categories as well as stock in debt and revenues from fees and investment grants are per capita. Note that we have data on debt only for Bavaria. *Source*: Own calculations.

The tax multipliers require some explanation. The trade tax is a tax on business income, where "business" includes all companies and firms as well as self-employed that do not belong to the "free professions" (*Freiberufler*). These include for example artists, lawyers, scientists, teachers, accountants, doctors, all medical therapists, architects, journalists, fotographers and engineers. The tax payment is calculated based on federal tax law and then multiplied by the trade tax multiplier that the municipality sets. This trade tax is separate from the federal business income tax. The multipliers range from 0 to 800,[29] and the actual tax rates for this

[28] In cases where we do not have data for the whole period – for instance for the election in 2004, the term just ended recently – we took the average over all years that we had data for. Our data on fiscal outcomes starts in 1998.

[29] During a reform in 2003 a statutory federal minimum of 200 for the trade tax multipliers was implemented. However, there was only a handful of communities that were directly affected by this reform.

trade tax is in the range of about 7% to 28%. The property taxes in Germany are *ad valorem* taxes where tax A is applied to agricultural property and tax B to all other property. Again, the tax payment is calculated based on federal law and then multiplied by the municipal tax multiplier.

All spending data is measured in Euros per capita. The data display considerable variation in all spending categories and, to a lesser extent, in tax multipliers. We use fixed effects in order to reduce the large variation across municipalities. However, the large variation – especially in spending – is a problem for our estimation since we are estimating the effect of small changes in seat shares and voting power.

3.5 Results

The results section is organized in three parts: first, we shortly present the results for seat shares, second we show the main results for voting power and finally we discuss instrument validity and robustness.

3.5.1 Seat shares

In the left part of table 3.3, we present the OLS results corresponding to equation 3.1. These results give us a first indication of whether political parties are at all statistically associated with policy outcomes. The coefficients for each party give the effect of a 1% change in seats from the CDU to the respective party.[30] Column (1) shows that the seat share of parties is indeed a predictor of the trade tax rate with the expected sign: the "left" parties (SPD, Left/PDS, Greens) show a positive sign relative to the CDU, and the pro-market party (FDP) shows a negative sign. For the other taxes, the relationship is less clear. Interpreting the size of the coefficient, the Left/PDS is associated with 5.7 points increase in the trade tax multiplier if it gains 10% of the seat share from the CDU. 200 points roughly equal 7% in the final tax rate, so 5.7 points correspond to 0.2% increase in the actual tax rate.

Since the OLS results contain only statistical associations rather than causal effects, we turn to the instrumental variable (IV) approach. The results of estimating equation 3.3 with our treatment variables as instruments are presented in columns (4) - (6). Only a seat change of the social democrats (SPD) relative to the benchmark center-right party (CDU) has a significant effect on taxes, with the FDP and the Others being marginally significant for some taxes.

[30]The coefficients in equation 3.1 measure the effect of a switch of 100% seat share of the CDU to 100% seat share of the respective party. In the results, we divided this coefficient by 100.

Table 3.3: OLS and IV results for seat shares - average tax rate multipliers

	OLS			IV		
	Average Taxes			Average Taxes		
	Trade Tax	Property A	Property B	Trade Tax	Property A	Property B
	(1)	(2)	(3)	(4)	(5)	(6)
SPD	0.21***	0.17**	0.03	−0.70**	−0.69	−1.01**
	(0.03)	(0.07)	(0.06)	(0.27)	(0.44)	(0.43)
The Greens	0.53***	0.25	0.33	0.06	0.37	0.57
	(0.15)	(0.31)	(0.26)	(0.59)	(0.95)	(0.94)
FDP	−0.23***	0.22*	0.01	−0.40	−1.66*	−1.03
	(0.06)	(0.13)	(0.10)	(0.59)	(0.95)	(0.94)
The Left	0.57***	0.26	0.27	−0.11	−0.19	−0.36
	(0.11)	(0.23)	(0.19)	(0.62)	(1.01)	(1.00)
Others	0.03	0.00	−0.02	−0.25	−0.43	−0.49*
	(0.02)	(0.05)	(0.04)	(0.19)	(0.30)	(0.30)
N	7217	7217	7217	7217	7217	7217

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. Columns 1-3 refer to the OLS regressions of seat shares for the respective parties on the three direct policy instruments (tax rate multipliers). Each regression contains a population control as well as dummies for each party if it did not receive any votes and state-election dummies. Columns 4-6 refer to the estimates of the IV regressions, in which seat shares are instrumented by the seat share shifts around a threshold in close elections. The instrument is based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. *Source*: Own calculations.

The first stage results of the IV approach for seat shares are presented in the first two columns of table 3.5 in the appendix. Each coefficient is the result of a separate regression: equation 3.3 with the actual seat share of one party as outcome variable. We expectedly observe a highly significant effect of about 100 for all parties: our coefficient measures the effect of a 100% change in the seat share of the council towards the party in question on the outcome, in this case the seat share.[31]

The aim of analysing seat shares is to identify *average* effects of an additional seat. However, not every seat is of equal importance. The effect of a seat that increases a party representation from, say, 65% to 70% might not matter at all, likewise a seat that does not change the coalition options in a council. The effect of a seat that has an impact on the identity of the governing coalition on the other hand, probably has the largest impact. Therefore, we will explore the effects of voting power on policy outcomes in the following section.

[31]The very high values for the F-Statistics (in squared brackets) indicate that the instrumental variable strategy identifies higher seat shares for parties almost by definition.

3.5.2 Voting power

The OLS results for voting power are shown in table 3.4 in columns (1) - (3). Again, the coefficient estimates for the trade tax go in the expected direction but are much smaller than for seat shares. For the property taxes, both SPD and Left/PDS have positive and significant coefficients. As before, those results are no proof of voting power to causally affect policy outcomes. Rather they illustrate the link between voter preferences and local policy as well as a selection effect of parties not to run in specific municipalities.

Table 3.4: OLS and IV results for voting power - average tax rate multipliers

	OLS			IV		
	Average Taxes			Average Taxes		
	Trade Tax	Property A	Property B	Trade Tax	Property A	Property B
	(1)	(2)	(3)	(4)	(5)	(6)
SPD	0.10***	0.16***	0.10***	−0.08	−0.19*	−0.25**
	(0.02)	(0.03)	(0.03)	(0.07)	(0.11)	(0.11)
The Greens	0.20***	−0.12	0.10	−0.03	0.46	0.93***
	(0.07)	(0.16)	(0.13)	(0.19)	(0.32)	(0.32)
FDP	−0.15***	0.01	−0.03	−0.05	−0.21	−0.09
	(0.04)	(0.08)	(0.07)	(0.30)	(0.49)	(0.49)
The Left	0.17**	0.32**	0.26**	0.15	−0.06	−0.23
	(0.07)	(0.15)	(0.13)	(0.21)	(0.35)	(0.35)
Others	0.00	−0.04	−0.03	−0.11*	−0.14	−0.11
	(0.01)	(0.03)	(0.02)	(0.06)	(0.09)	(0.09)
N	7103	7103	7103	7103	7103	7103

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. Columns 1-3 refer to the OLS regressions of voting power for the respective parties on the three direct policy instruments (tax rate multipliers). Each regression contains a population control as well as dummies for each party if it did not receive any votes and state-election dummies. Columns 4-6 refer to the estimates of the IV regressions, in which the voting power measures are instrumented by the shifts in voting power around a threshold in close elections. The instrument is based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. *Source*: Own calculations.

To identify causal effects of voting power, we again turn to the instrumental variable estimation of equation 3.9. Only for property tax B, the non-agricultural tax, do we find effects: for the SPD a 2.5 points decline in the tax multiplier for an increase in (relative) voting power by 10 percentage points relative to the CDU, for the Greens the effect is an increase of 9.3 points. Since the multiplier ranges from 200 to 800, this may seem like a small effect. However, the average change in the multiplier of the property tax B over an election period is around 7 points.

So far we have taken the average value during a government term as the outcome, excluding the election years since we are unable to pin down changes during a year. Alternatively, we

could look at the outcomes in the last year before the election. The top panel of table 3.6 in the appendix contains the result for this IV regression. The effects of the SPD and The Greens on the property tax B are the same, whereas the effect of the SPD on the agricultural property tax increases slightly in magnitude and is now significant. The effect of The Greens on property tax A increases as well and is now marginally significant. These results for the end period suggest that the effects found for the average are not driven by a political business cycle.

For the voting power approach, we also look at spending and other sources of revenue. The first three columns of table 3.7 in the appendix suggest that parties' voting power has no impact on spending: neither total nor personnel or investment spending display any significant effect. However, spending is highly variable and only in part the result of a local policy decision. Since we only measure the effect of small increases in voting power, these results are not surprising. Also, the results for other sources of revenue – debt, fees and investment grants – display a similar pattern: no effects of voting power. One explanation for the lack of an effect on fees could be that we look at revenues here, not fee rates. The direct policy instrument, the rates, could be affected but the market reaction (for instance fees for house construction, for sewage or waste collection) may counteract those changes.

How do we interpret our findings? The effect of the Green party on the major local property tax is substantial and follows the a priori expected sign. It is notable that we can isolate the policy effect only when we account for voting power of the Green party. In the seat share specification the Green party is not shown to have a significant impact. If one believes that seat shares are a noisy measure of actual voting power, this difference is to be expected especially for small parties. The fact that we find an effect only for the Green party and not for other small parties is likely to have two reasons. First, the Greens participate more often (compared to FDP and The Lefts), thus generating more observations. Second, the Green party has seen an increase in representation during the period of observation.

Our results for the center-left party are surprising. With increased representation (relative to the center-right party) local taxes are documented to decrease. Why would representation of the social democrats have that effect? We can rule out two possible explanations. We showed that it is not the case, that the social democrats keep the taxes down while increasing other sources of revenues (debt, fees, grants). Also, our results indicate that the social democrats are not trying to be more conservative than the conservatives. That is because the results for voting power are clearly smaller in size than the seat share effects. If the social democrat position would be to decrease the local tax rates, we should observe that more in the voting power specification than in the seat share regressions.

We consider a third explanation to be likely. The tax deductions could be induced by the leading conservative party as a response to increased representation of the social democrats. It is when the CDU faces a larger number of SPD members in the opposition that they decrease taxes to make for distinct policy platforms. The observed differences between seat shares and voting power then indicate that the CDU can pursue that strategy better if only the seat share of the SPD changed and not the relative voting power. In the next section, we also provide evidence from regressions by each state. This evidence supports this interpretation, as the observed negative effect of the SPD is only to be found in Bavaria, a state where the SPD is mostly second to the center-right CDU. Moreover, in Thuringia, a state where the SPD is often the dominant party, we indeed observe a positive effect of the SPD on one of the local tax rates.

While that interpretation is in line with some of our findings, it also means bad news for our voting power measure. If that interpretation is correct, it indicates that we have only partially succeeded in identifying a voting power index that takes adequate account of the actual party influence.

The first stage results of the IV approach for voting power are presented in columns (3) and (4) of table 3.5 (see appendix). As before, each coefficient is the result of a separate regression: equation 3.9 with the actual voting power as the outcome variable. The effects are high and statistically significant. We also expect effects around 100 as the coefficient should measure the effect of a one percent increase in relative voting power on the outcome variable. The coefficients show larger variation around 100 than the seat shares, indicating that our instrument is not as precise. This is also reflected in the much lower F-statistics. Nevertheless, the first stage works well overall.

3.5.3 Validity and robustness

Apart from examining the first stage of our instrumental variable estimation, we conduct tests on the validity of the instrument. The second part of this section analyses several robustness tests.

Our assumption is that close election outcomes provide us with exogenous variation in seats and voting power that we can use to analyze local public policy making. If this assumption is correct, we should *not* observe any of the following:

- an effect of instrumented voting power on variables *before* the government term in question;

- different distributions of vote shares for positive and negative instrument values;
- skewed distributions of the instrument for each party: the difference between the actual voting power and the voting power during the perturbations.

We will discuss these aspects in turn.

The lower panel of table 3.6 in the appendix shows the IV regression of equation 3.9 using the outcome variable in the last year *before* the election. The results are clearly insignificant. Moreover, the effects for the Greens and the SPD that we found for the end of the government term (see the upper panel) are economically sizable whereas the same effects on the policy measure before the government term are economically negligible. For spending and revenues (not included in the table), there are also no effects on outcomes before the period in question.

As another validity check, we look at the distributions of vote shares and of the instrument for each party. Figure 3.2 (see appendix) shows the frequency of a positive or negative value of the instrument by the vote share of that party. All graphs are reasonably well balanced, even though the CDU seems to be getting slightly more positive than negative treatments, contrary to the SPD. In figure 3.3 (see appendix) we plot the distribution of the instrument, which is also reasonably well balanced for each party.

We conduct four robustness tests to check whether our results hold up to changes in the specification and data: changes in the IV specification, changes in the degree of closeness, estimates by state, and finally weighted voting power. The results are highlighted in the appendix.

The specification of the instrument can be changed in two opposite ways: extract more or less information out of our perturbations. First, we extract more information by adding the squared instrument to the specification. The reason is that large changes in voting power during the perturbations have a different impact on voting power than small changes: large parties have disproportionately high voting power and therefore relatively high differences, too. The squared instrument allows for such a non-linear relationship. The top panel of table 3.8 shows the results for this specification. The results do not change.

The other way is to extract less information from our perturbations, for instance by using only the direction of the difference, not the the size of the difference. The lower panel of table 3.8 contains the estimates from this IV regression. The results change slightly, but remain robust.

When we change our parameter k in the perturbation from 1% to 2% or .5%, we expand or restrict the sample we use for our estimation. The results of doing so are presented in table

3.9 in the appendix. In the top panel, we have expanded the sample to also include less close elections. The effects of the SPD and The Greens on property tax B are in the same range, albeit less significant. If we restrict the sample to include only very close elections (shown in the lower panel), we find that the effects remain in the same range but lose significance. We interpret these tests as confirmation that our results are robust to a change in the definition of closeness.

The estimates are so far based on the data in four German states. If we estimate each state separately, we find that results indeed change (table 3.10 in the appendix). For Bavaria alone, the results of the SPD and The Greens are consistent with the overall results, even though the SPD seems to affect all tax rates. In Hesse, only The Greens affect a tax, namely the property tax B. In North Rhine-Westphalia (NRW) we do not find any effects.[32] For Thueringia, there are no effects, except that the SPD and The Left seem to affect the trade tax, even though this effect is only marginally significant.

As a further check, we excluded all observations with a council size below twelve, thus restricting the sample to larger communities. We find that the signs and significance of coefficients do not change although point estimates vary slightly (not included in the tables).

Finally, we use weighted voting power as explained in the estimation strategy. Weighted voting power takes into account that not all coalitions are equally likely to form. In the absence of better information, we used a rough approximation for party locations in Germany as a starting point. Figure 3.4 (see appendix) shows our presumption for two-dimensional party positions in local German politics. The results are shown in table 3.11 in the appendix. The results are almost unchanged.

3.6 Conclusions

The importance of political representation in a proportional election system has long been a topic of the theoretical literature in political economy. However, the empirical literature has been mostly unsuccessful in reliably identifying the effects of representation. This paper attempts to estimate the causal effects of political representation on fiscal policy in local governments.

Our analysis builds on the estimation strategy developed in Folke (2010b) to estimate *causal* effects of representation. We use seat shares *and* relative voting power as our measures of party

[32]Note that NRW was the state in which the mixed member proportional system poses some problems for our estimation, adding noise to the instrument.

representation in the municipal council. The empirical analysis is based on an instrumental variable approach in which we identify the effects of individual parties by focusing on narrow seat allocations. Given that election outcomes are subject to some randomness, the exact seat allocation on the margin can be regarded as an as-good-as random event. To determine whether a party was close to lose or win a seat, we develop an algorithm based on random perturbations that is both simple to implement and easy to adjust to any electoral system. Moreover, we show that this algorithm can be directly linked to a conceptional framework for randomness in elections.

We use data on about 7200 elections and subsequent fiscal outcomes in German municipalities from 4 different states. While OLS results hint at a clear correlation pattern between political representation and fiscal policy, the results of the causal analysis are less definite. However, they indicate that representation does matter. For both seat shares and voting power, the effects are mostly small, but statistically significant. Representation of the center-left party is, somewhat surprisingly, found to *lower* the local tax rates. The Greens on the other hand increase the property tax considerably. This effect is only present in the estimation setup with voting power and remains robust to modifications in the specification. This suggests that small parties can be shown to influence major policy decisions if their voting power is sufficiently taken into account.

Appendix

Table 3.5: First stage results and # of close elections

	Seat shares		Voting power	
	# of close elections	FS coef	# of close elections	FS coef
	(1)	(2)	(3)	(4)
CDU	744	113.5***	569	107.7***
	(12.3%)	(46.9)	(9.6%)	(10.7)
		[1289.3]		[129.5]
SPD	566	110.3***	504	107.8***
	(11.7%)	(65.4)	(10.6%)	(16.6)
		[1170.0]		[82.1]
The Greens	148	105.7***	221	106.4***
	(8.2%)	(38.9)	(12.7%)	(17.0)
		[627.5]		[51.7]
FDP	126	105.7***	169	54.3***
	(8.1%)	(88.0)	(11.3%)	(8.63)
		[1322.7]		[100.2]
The Lefts	47	98.6***	53	150.3***
	(9.2%)	(60.3)	(10.7%)	(19.0)
		[1543.4]		[129.6]
Others	600	97.6***	585	113.3***
	(8.9%)	(61.8)	(8.2%)	(10.5)
		[1570.1]		[140.8]

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns 1 and 3 present the number of close elections in which seats or voting power was at the margin. The percentage points in round parentheses refer to the share of these close elections in the total number of elections in which the party participated. Columns 2 and 4 show the results of the first stage regressions for each party's representation measure. T-values (based on robust standard errors) in round parentheses are the respective values on each coefficient. F-Stat values are in squared brackets and refer to the F- Statistic for the join test of all variables in the first stage. Each coefficient represents a separate regression. Included controls: 2nd order vote share function including squared vote shares of all parties, population, dummies if the parties received no votes,the c_i^j dummies for each party if it was close, as well as state-election dummies. *Source*: Own calculations.

Table 3.6: IV results voting power - end versus start of the election period

	Average Taxes		
	Trade Tax	Property Tax A	Property Tax B
	(1)	(2)	(3)
	Panel 1: Tax rates at the END of the election period		
SPD	−0.14	−0.24**	−0.27**
	(0.09)	(0.13)	(0.14)
The Greens	−0.24	0.63*	0.97**
	(0.25)	(0.38)	(0.38)
FDP	0.23	−0.11	−0.01
	(0.38)	(0.58)	(0.58)
The Left	0.01	−0.22	−0.24
	(0.27)	(0.42)	(0.42)
Others	−0.07	−0.04	−0.03
	(0.07)	(0.11)	(0.11)
	Panel 2: Tax rates BEFORE the election period		
SPD	0.03	−0.10	−0.15
	(0.10)	(0.14)	(0.14)
The Greens	0.04	0.04	0.06
	(0.26)	(0.37)	(0.38)
FDP	−0.46	−0.84	0.18
	(0.38)	(0.55)	(0.57)
The Left	0.11	0.04	−0.26
	(0.25)	(0.36)	(0.38)
Others	−0.10	−0.10	−0.04
	(0.07)	(0.10)	(0.10)

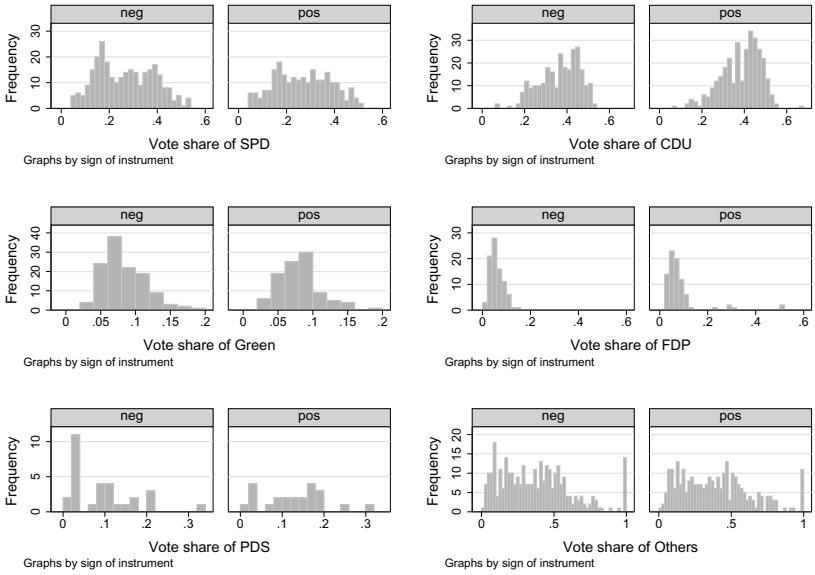
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The table shows the estimates of IV regressions for voting power, in which the the tax rate outcomes are measured at the end of the election period (panel 1) and just before the start of the election period (panel 2). The instrument is based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. The number of observations in each regression is 7103. *Source*: Own calculations.

Table 3.7: IV results voting power - aggregate spending and revenues sources

	Aver spending (per cap)			Aver revenue (per cap)		
	Total	Personnel	Investment	Stock in Debt	Fees	Invest Grants
	(1)	(2)	(3)	(4)	(5)	(6)
SPD	1.50	0.13	1.84	0.00	0.23	−0.09
	(2.12)	(0.27)	(1.31)	(0.00)	(0.27)	(0.69)
The Greens	4.26	−0.44	−0.01	0.00	−0.23	−1.15
	(5.94)	(0.76)	(3.66)	(0.01)	(0.75)	(1.94)
FDP	0.47	−0.30	−2.41	0.00	−0.12	−3.10
	(9.12)	(1.16)	(5.62)	(0.03)	(1.16)	(2.98)
The Left	−5.27	0.16	−5.06		0.12	−3.46
	(6.56)	(0.84)	(4.05)		(0.83)	(2.15)
Others	−2.45	−0.12	−0.87	0.00	0.16	0.13
	(1.75)	(0.22)	(1.08)	(0.00)	(0.22)	(0.57)

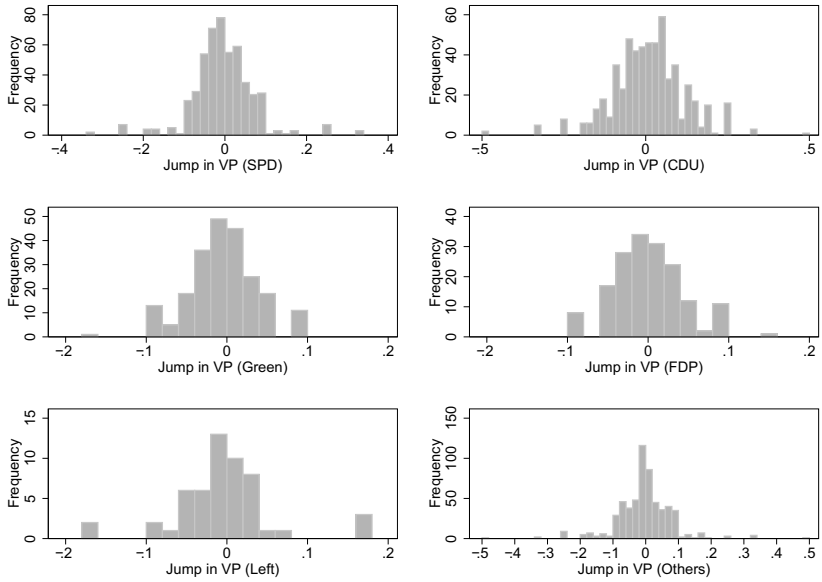
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The table presents IV voting power estimates for the following additional fiscal outcome variables (per capita): total spending, spending on personnel, investment spending, stock of debt, revenue from fees for communal services and revenue from higher level grants for investment. The instrument is again based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. The number of observations in each regression is 7103 (except for column 4 for which we only have data for Bavaria, N=3951). *Source*: Own calculations.

Figure 3.2: Distribution of vote share by sign of instrument



Notes: This figure graphically illustrates the distribution of the vote share of the respective parties separated into the distinct subsample when the instrument for voting power was positive or negative. A positive instruments indicates that an observation for this party has narrow obtained the extra voting power (through one narrow seat), but was close to lose that seat. A negative instruments relates to the reverse accordingly. The graph excludes all observations that were not coded as close elections. The six different panels show the respective graphs for each party. Note that, the graph for the "Others" (lower right) is special as it shows the aggregate vote share of all other parties. *Source:* Own calculations.

Figure 3.3: Distribution of the instrument - voting power jump



Notes: This figure graphically illustrates the distribution of the voting power jump that is used as instrument in the voting power analysis. The graph excludes all observations that were not coded as close elections. The six graphs show the distribution of each party respectively. Note that, the graph for the "Others" (lower right) is special as it shows the aggregate mean of the instrument over the Others parties. *Source:* Own calculations.

Table 3.8: IV results voting power - varying the instrument

	Average Taxes		
	Trade Tax	Property Tax A	Property Tax B
	(1)	(2)	(3)
	Panel 1: Adding the squared jumps in vp		
SPD	−0.06	−0.18	−0.25**
	(0.07)	(0.11)	(0.11)
The Greens	−0.06	0.42	0.90***
	(0.19)	(0.31)	(0.31)
FDP	−0.10	−0.27	−0.11
	(0.25)	(0.40)	(0.40)
The Left	0.12	−0.04	−0.21
	(0.20)	(0.34)	(0.33)
Others	−0.12**	−0.14	−0.11
	(0.06)	(0.09)	(0.09)
	Panel 2: Using only 1/ -1 information		
SPD	−0.22**	−0.38**	−0.41**
	(0.10)	(0.17)	(0.17)
The Greens	−0.05	0.39	0.79**
	(0.22)	(0.36)	(0.36)
FDP	−0.26	−0.27	−0.26
	(0.30)	(0.50)	(0.50)
The Left	0.04	−0.39	−0.54
	(0.30)	(0.50)	(0.50)
Others	−0.06	−0.16	−0.10
	(0.10)	(0.16)	(0.16)

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The table shows a robustness test of the voting power analysis on tax rate multipliers in which we have altered the instrument that we use. In panel 1, we additionally use the squares of the jump in the first stage regression. In panel 2, we limit the information of the instrument to the sign of the jump and use indicator variable for a negative or a positive jump. The instrument is again based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. The number of observations in each regression is 7103. *Source*: Own calculations.

Table 3.9: IV results voting power - alternative close samples

	Average Taxes		
	Trade Tax	Property Tax A	Property Tax B
	(1)	(2)	(3)
	Panel 1: 2% closeness definition		
SPD	−0.07	−0.13	−0.16*
	(0.05)	(0.09)	(0.09)
The Greens	−0.11	0.15	0.53**
	(0.15)	(0.24)	(0.24)
FDP	0.19	−0.15	0.03
	(0.20)	(0.33)	(0.32)
The Left	0.14	0.10	−0.08
	(0.19)	(0.32)	(0.31)
Others	−0.05	−0.03	−0.09
	(0.05)	(0.08)	(0.08)
	Panel 2: 0.5% closeness definition		
SPD	0.01	−0.19	−0.32*
	(0.11)	(0.18)	(0.18)
The Greens	−0.29	0.14	0.91
	(0.38)	(0.62)	(0.62)
FDP	0.16	−0.47	−0.53
	(0.35)	(0.57)	(0.57)
The Left	0.28	−0.19	−0.49
	(0.37)	(0.60)	(0.60)
Others	−0.12	−0.18	−0.18
	(0.08)	(0.13)	(0.13)

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The table shows a robustness test of the voting power analysis on tax rate multipliers in which we use a different definition of closeness in the perturbation used to calculate close elections. In the upper panel 1, we perturb the vote vector using a variance of 2% of the vote count. In panel 2, we use a narrow band for the variance of 0.5% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. The number of observations in each regression is 7103. The number of close elections in those specifications is as follows: CDU - 1080 (2% sample) and 284 (0.5% sample), SPD - 955 and 261, The Greens - 419 and 113, FDP - 319 and 84, PDS - 89 and 26, Others - 1108 and 294. *Source*: Own calculations.

Table 3.10: IV results voting power - by states

	Average Taxes		
	Trade Tax	Property Tax A	Property Tax B
	(1)	(2)	(3)
	Panel 1: Bavaria		
SPD	−0.36**	−0.66**	−0.68**
	(0.15)	(0.26)	(0.27)
The Greens	−0.32	1.00*	1.06*
	(0.34)	(0.58)	(0.60)
Others	−0.06	−0.04	−0.04
	(0.10)	(0.17)	(0.17)
N	4028	4028	4028
	Panel 2: Hesse		
SPD	−0.00	0.02	0.17
	(0.18)	(0.44)	(0.39)
The Greens	0.59	0.66	1.90**
	(0.42)	(1.03)	(0.92)
FDP	1.09	0.45	2.37
	(0.69)	(1.69)	(1.52)
Others	−0.04	0.14	0.29
	(0.23)	(0.57)	(0.51)
N	823	823	823
	Panel 3: NRW		
SPD	0.12	0.13	0.16
	(0.15)	(0.25)	(0.24)
The Greens	−0.11	−0.23	0.31
	(0.27)	(0.46)	(0.43)
FDP	−0.34	−0.34	−0.36
	(0.25)	(0.42)	(0.40)
Others	−0.30	−0.33	0.06
	(0.23)	(0.39)	(0.36)
N	736	736	736
	Panel 4: Thuringia		
SPD	0.26*	0.20	−0.02
	(0.15)	(0.14)	(0.15)
The Left	0.54*	0.34	0.17
	(0.30)	(0.29)	(0.30)
Others	−0.07	−0.05	−0.07
	(0.11)	(0.10)	(0.10)
N	1455	1455	1455

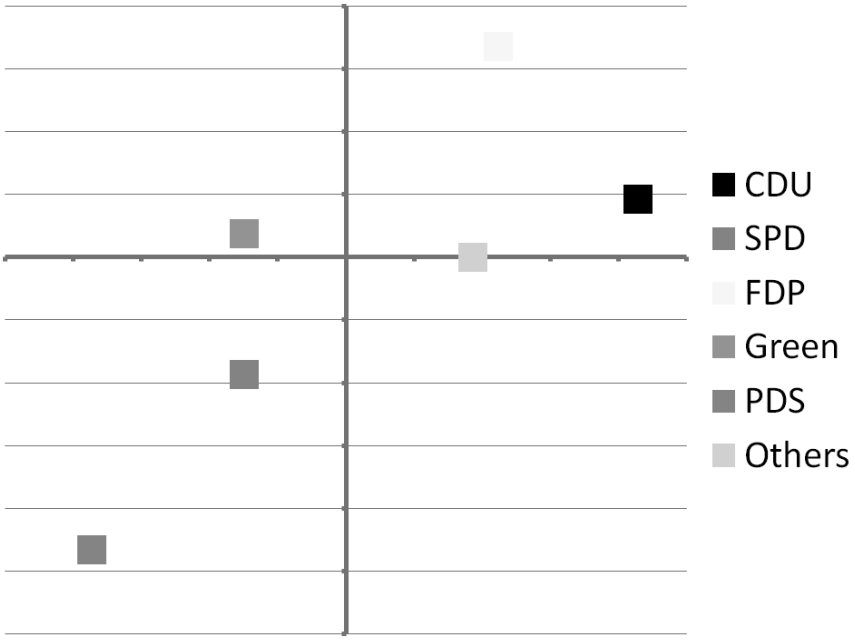
Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. In this table we repeat the analysis of the effect of voting power on tax decisions by each state for Bavaria, Hessen, NRW and Thuringia respectively. We excluded specific parties if there were to few cases within the state that close elections occurred. The instrument is again based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. *Source*: Own calculations.

Table 3.11: IV results weighted voting power

	Average Taxes		
	Trade Tax	Property Tax A	Property Tax B
	(1)	(2)	(3)
	Panel 1: Weighted by distance in *one* dimensional policy space		
SPD	−0.08	−0.18	−0.26**
	(0.07)	(0.11)	(0.11)
The Greens	−0.02	0.28	0.74***
	(0.17)	(0.28)	(0.27)
FDP	−0.17	−0.14	−0.01
	(0.19)	(0.31)	(0.31)
The Left	0.19	0.20	−0.22
	(0.29)	(0.47)	(0.47)
Others	−0.11*	−0.18*	−0.13
	(0.06)	(0.09)	(0.09)
	Panel 2: Weighted by distance in *two* dimensional policy space		
SPD	−0.08	−0.19**	−0.26**
	(0.07)	(0.11)	(0.11)
The Greens	−0.01	0.42	0.90***
	(0.19)	(0.31)	(0.30)
FDP	−0.14	−0.18	−0.05
	(0.26)	(0.42)	(0.42)
The Left	0.24	0.16	−0.29
	(0.33)	(0.54)	(0.54)
Others	−0.11*	−0.17*	−0.13
	(0.06)	(0.10)	(0.09)

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. In this table we use weight voting power in coalitions with an approximated party distance. Panel 1 shows the results on tax making when we use a traditional one-dimensional distance measure. In panel 2, we instead use a two dimensional policy space. The instrument is again based on 1000 perturbations of the vote vector using a variance of 1% of the vote count. Each regression contains population controls and council size, dummies for each party if it did not receive any votes, state-election dummies, and a polynomial control function that is quadratic in each party's vote share. The specification is estimated using municipality fixed effects. *Source*: Own calculations.

Figure 3.4: Party locations in policy space



Notes: This figure presents the assumed party positions in a two dimensional policy space. The dimensions can be thought of as pro-market versus state ruled (vertical axis) and conservative versus progressive (horizontal axis). *Source:* Own graphic.

Chapter 4

When can we trust population thresholds in regression discontinuity designs?

A comment on Egger and Koethenbuerger (2010)[1]

Abstract: A recent literature has used variation just around deterministic legislative population thresholds to identify the causal effects of institutional changes. This paper reviews the use of regression discontinuity designs using such population thresholds. Our concern involves three arguments: (1) simultaneous exogenous (co-)treatment, (2) simultaneous endogenous choices and (3) manipulation and precise control over population measures. Revisiting the study by Egger and Koethenbuerger (2010), who analyse the relationship between council size and government spending, we present new evidence that these three concerns do matter for causal analysis. Our results suggest that empirical designs using population thresholds are only to be used with utmost care and confidence in the precise institutional setting.

[1] This paper is joint work with Florian Ade. Acknowledgments: We would like to thank Charles B. Blankart, Peter Haan, Magnus Johannesson, Juanna Joensen, Henrik Jordahl, Erik Lindqvist, Christian Odendahl, Viktor Steiner and David Strömberg as well as colleagues at Stockholm School of Economics and DIW Berlin. Comments at the Applied Microeconomics Seminar at Stockholm School of Economics are also gratefully acknowledged. We would like to thank Peter Egger and Marko Koethenbuerger for making their data available online and Viktor Steiner and Peter Haan who invited us to work at their department for public economics at the DIW Berlin. Ronny Freier would like to thank the Jan Wallander and Tom Hedelius scholarship for continued funding. We are further grateful for editorial support from Adam Lederer. The usual disclaimer applies.

4.1 Introduction

Scholars in political economy devote much attention to the causal identification of the effects of fundamental rules and features of governmental organization (e.g. Persson and Tabellini (2002)). Estimating the causal effects of institutional designs and constitutional rules, however, is generally a difficult task for a number of reasons (see Acemoglu (2005)). The main challenges are that institutional rules are usually endogenous and seldom change, that different aspects of constitutional designs often correlate and change simultaneously and that data analysis is often limited by small samples. The new interest has therefore turned to subnational levels and statistical methods from the program evaluation literature in the attempt to use quasi-random variation in specific rules to estimate their impact.

One specific class of designs being used in a range of different applications are regression discontinuity designs that focus on population thresholds. Pettersson-Lidbom (2006) and Egger and Koethenbuerger (2010) investigate whether the municipality's council size, which changes at deterministic population thresholds, affects local government spending.[2] Regression discontinuity designs, based on population thresholds, are also used to study performance of politicians when salaries increase (Gagliarducci and Nannicini (2009) and Ferraz and Finan (2009)), the effect of fiscal transfers on local elections (Litschig and Morrison (2010)), corruption and the quality of politicians (Brollo, Nannicini, Perotti, and Tabellini (2009)), as well as the impact of representative democracy versus direct democracy on government spending (Hinnerich and Pettersson-Lidbom (2010)).

In this paper, we devote attention to the specific challenges of using population thresholds for reliable causal inference. Our concern is threefold: (1) The first challenge is that the population threshold used may not only define the treatment considered but also additional simultaneous exogenous co-treatments. (2) The identification and interpretation of the treatment effect is further complicated as additional endogenous choices on other institutions are often taken simultaneously. The timing of events is likely to coincide as it is often at specific times that changes are implemented (e.g. at the beginning of an election period). (3) Given that the official population count is usually observable at any moment, the specific concern is that the precise number can be manipulated. In the empirical analysis, we revisit the study by Egger and Koethenbuerger (2010) and present new evidence on the importance of these concerns.

[2]Both studies use rules (given by the federal or state law) that set the number of members of legislature based on population count of the locality. Those contributions are perceived to be of great importance as they focus on causal identification. For further studies on this topic compare e.g. Baqir (2002) for cities and counties in the US, Gilligan and Matsusaka (2001) for state and local governments in the US, Schaltegger and Feld (2009) for Swiss cantons.

Researchers using quasi-experimental designs recognize the importance of ensuring that a particular treatment effect stands in isolation to other confounding simultaneous treatments. In applications using difference-in-difference designs it is common practice to make an explicit argument that the treatment group is not simultaneously affected by additional treatments. Similarly, for the use of instrumental variables, special care is taken in arguing that the instrument is only of importance to the specific treatment under investigation (see Acemoglu (2005)).[3] We argue that this problem is critical when using regression discontinuity designs with population thresholds. Population count is an intuitive and easy way for higher level governments to impose differential rules on lower tier structures. Hence, it is very likely that the same thresholds might be used in different dimensions of rules and institutions.

According to Bavarian state law the number of town council members must increase at 10 different population thresholds. However, researching the details of Bavarian laws and bylaws we find that the same thresholds also determine large changes in both local institutions (e.g. referendum quotas, politicians wages) and communal finances (e.g. additional funds from the state government). In total, we find 14 additional legislative laws or bylaws that induce differential rules by population size at the same thresholds. Applying these rules, we show that a part of the spending effect at population thresholds is financed by increases in state grants and additional revenue from fees.

The second challenge in using population thresholds for identification is the timing of changes in the treatment and changes of other endogenous choices. Local institutions are often changed simultaneously at the beginning of new electoral cycles. In Bavaria, e.g., apart from the change in council size set by the exogenous rule, communities must make various (endogenous) decisions just before the next election cycle starts. Those decisions involve, for example, the nature of the mayor position (full-time or part-time) as well as what new tasks the community wants to be responsible for. For the identification of a specific treatment effect, this problem becomes twofold: First, because timing coincides it is harder for the researchers to disentangle the effect in question and, in small samples, pure statistical variation is more likely to introduce bias in the estimation. Second, the interpretation of the effect might crucially hinge on the timing and interdependency of the treatment changes and the choice making. We supplement the data of Egger and Koethenbuerger (2010) with information on one such endogenous decision (work status of the mayor) and no longer find a significant spending effect at the council size thresholds.

[3] Acemoglu (2005) discusses the fundamental problem for the standard instrumental variable approach: that the instrument used for e.g. an institutional setting may not only be an instrument for the institution under investigation but also for a different institution. In this case it is not clear what effect is estimated.

Our final concern is that the population count may be precisely manipulated. As population thresholds are critical for the remuneration of government personnel, for allocation of finances from other government levels, as well as for the council size, there are large incentives for sorting around the respective population thresholds. The concern of sorting on the precise threshold is well-recognized in the literature on regression discontinuity design (RDD) (see Lee and Lemieux (2009), McCrary (2008)). For population thresholds this is of particular concern as in Germany the number of inhabitants is by no means a surprise, but rather a regularly updated and publicly known number. Hence, whenever a community comes close to a threshold the administration can purposefully manipulate the number and precisely sort on the desired population count. For Bavarian municipalities during the period of observation we find evidence that there is sorting around the population thresholds. Thus, the spending effects found by Egger and Koethenbuerger (2010) could be potentially biased by selection effects.

Given the results of our empirical analysis, we conclude that researchers must be cautious when using population thresholds for identification in regression discontinuity designs. To be covered against the threats of simultaneous exogenous co-treatment and simultaneous endogenous decisions, researches must acquire deep institutional knowledge while checking legal norms and customs thoroughly. Furthermore, testing of the key identifying assumptions of RDD is required.[4]

This study proceeds as follows. Section 2 introduces the empirical model and the underlying identification assumptions. Section 3 presents the institutional setting and the specific use of population thresholds in Bavarian law, before Section 4 highlights our results and draws comparisons to the findings reported by Egger and Koethenbuerger (2010). Section 5 concludes the analysis.

4.2 Empirical model and methodology

In this section, we consider the empirical methodology that is involved in the use of deterministic population thresholds in regression discontinuity designs (RDD). After discussing the basic empirical setup, we clarify the identifying assumptions needed to allow for causal inference of the treatment effect in question.

[4]Caughey and Sekhon (2010) revisit the empirical evidence presented in Lee (2008), who uses a RDD on close election outcomes. The former rigorously investigate the RDD assumptions and find that, for the US House of Representatives, observations in close elections still exhibit crucially unbalanced predetermined variables that are likely to invalidate the research design.

4.2.1 Basic model

Assume, for simplicity, that we consider a case in which a treatment is determined at a single population threshold. Define $v_{i,t}$ as the distance of the number of inhabitants in location i at time t from the threshold. Assume further that the treatment (e.g. number of council members) is discontinuously determined at the threshold (sharp discontinuity design). The relationship between treatment $d_{1_{i,t}}$ and $v_{i,t}$ is as follows:

$$d_{1_{i,t}} = 1\left[v_{i,t} > 0\right] \tag{4.1}$$

The estimation equation specifying the RDD then reads as follows:

$$y_{i,t} = \delta_0 + \delta_1 d_{1_{i,t}} + h(v_{i,t}, \theta) + \epsilon_{i,t} \tag{4.2}$$

where $y_{i,t}$ is the outcome in question (e.g. local government spending), δ_1 is the parameter of interest and $h(\cdot)$ is a flexible function that represents the underlying general relationship between the distance to the threshold (hence population size) and the outcome variable. This simple framework can also be easily adjusted to accommodate the fact that treatment is changing at multiple thresholds (see Egger and Koethenbuerger (2010)).

4.2.2 Identifying assumptions

The parameter δ_1 is an unbiased estimate of the treatment effect under the critical assumption of continuity(see Lee and Lemieux (2009)). It is only if all observable and unobservable covariates, except treatment, are distributed continuously around the threshold that one can assume to have valid counterfactual observations on either side. If observations just right from the required population count are systematically different from the ones just to the left, then identification fails.

The first implication of the continuity assumption is that we must ensure that there exist no further co-treatments. It is obvious that causal inference of an individual effect cannot be upheld when other treatments are simultaneously determined at the same threshold. Formally, the existence of co-treatment implies the following. Let a second treatment, $d_{2_{i,t}}$, be determined at the same threshold v_0:

$$d_{2_{i,t}} = 1\left[v_{i,t} > 0\right] \tag{4.3}$$

The individual effects of $d_{1_{i,t}}$ and $d_{2_{i,t}}$ cannot be identified. The second treatment is omitted in eq. (4.2) and the flexible function in v cannot control for it as the threshold coincides.

Even if the second treatment is observable, one cannot include it in the regression due to multi-collinearity. The researcher can only identify a joint effect. If one has several thresholds and differences in when certain treatments apply, one may hope to disentangle the isolated treatments. However, the major concern is when and if we fail to recognize the existence of additional (co-)treatments. Although an outcome was in fact induced by several changing factors, we instead falsely attribute the effect to only one treatment.

It is important to note that co-treatment is of particular concern because standard RDD tests are not very likely to detect it. If a second treatment is implemented at the same threshold, distribution tests of predetermined covariates cannot be expected to detect such differences. Further, there is no reason to believe that a direct test of the distribution of the score variable (McCrary (2008)) will be of help.

For population thresholds the problem of co-treatments has to be critically reviewed on a case by case basis. Population thresholds are tools for legislatures to induce differences in laws and bylaws by the population size of local entities. Thus, the same thresholds are likely used in several legislative rules.

Apart from exogenous co-treatments, we consider the precise sorting around thresholds to be of particular concern when using population count in RDD. Lee (2008) shows that it is sufficient to show that there is a random component in the scoring variable to uphold the continuity assumption. If agents cannot precisely control the variable that determines the treatment, then in some neighborhood of the threshold, assignment of treatment is effectively random. This random assignment close to the threshold directly implies continuity.

Thus, any RDD application relying on this argument must investigate whether agents can precisely sort around the threshold. For population count this issue is indeed crucial. The official population of a municipality in Germany is known at any given time. Day to day changes in the population number are very small and even well ahead of the defining deadline, the precise population number can be well anticipated by the authorities of the municipality.[5]

If the agent can indeed ensure that she ends up just to the right of a certain threshold then the continuity assumption is likely to fail. Instead of valid counterfactuals, observations on either side are likely to differ. Variables that drive the selection will be systematically different on both sides and confound the treatment effect estimate.

[5]Caughey and Sekhon (2010) make the point that it is important to assess the magnitude of the random component in the score variable as compared to the precision with which agents can actively manipulate this variable.

One advantage of the RDD is that the sorting argument can be tested in a number of different ways. McCrary (2008) suggests a direct test of the distribution of the score variable. If it is profitable for agents to sort on a threshold and the possibility of doing so exists, then one should observe a higher frequency of observations on one side rather than on the other. Also, we can look at the distribution of predetermined variables which - given local randomization - should not systematically differ around the thresholds.

Our last concern is a particularity of the use of population thresholds. Typically, a legal rule that applies a population threshold defines a definite point in time when the population count is taken. This point in time is often distinct from the actual implementation of the new rule. For example, for Bavarian municipalities council size is determined by a population count about a year before the council election.[6] This year, however, is also the time during which important institutional decisions are made. The municipal council, for example, decides whether the next mayor will be part- or full-time and / or whether the local authorities will be responsible for specific services.

These endogenous decisions can be problematic in two different ways. First, interpretation of the effect can be complex. Say, a municipality knows that it will increase the council size in one year. It must decide whether the future council will become responsible for certain tasks (e.g. water management). Their decision to do so might depend on the anticipated council size. A larger council might, for example, be able to support additional committees that can oversee these responsibilities. Those kinds of effects, however, are not included in the definition of the treatment effect, as treatment, *per se*, has not yet started. The researcher might be willing to redefine treatment, however, specific care has to be given to exactly what the object of interest is.

Apart from problems in interpretation, these endogenous decisions might also pose a threat to the validity of the identification, particularly in small samples. Assume that the simultaneous endogenous decision is in fact independent from the treatment. As with any postdetermined variable the researcher can only hope to not pick a strangely unbalanced sample in which the observations just right and just left are different by random selection. However, as opposed to other variables the timing of those endogenous institutional choices coincides exactly with the determination of treatment. This makes it much more likely that an unfavorable sample is picked. In Bavaria, for example, there is a trend to have more full-time mayors and to locate

[6]In preparation for a new election, the new council size must be known well ahead of time as certain preparations directly depend on it. For example, each party will name a number of candidates that is (typically) exactly as many as there are council seats.

additional responsibilities at the local level. Naturally, these decisions are made at exactly the same time as council size increases: just before the new election cycle.

4.3 Institutional setting

The German federal system comprises of four tiers. Apart from the federal, state and county level, some key decisions are taken by local municipalities. This local authority decides, among other issues, on local roads, theaters, cultural events, local business development, as well as school buildings and administers social welfare programs, kindergarten spending, etc. Furthermore, they often own and control key parts of the local economy such as waste disposal, public transport, as well as the energy and water supply. As the municipalities are the lowest tier of the federal system, most laws regulating their decision mechanisms, freedoms, and duties are exogenous to them: in Bavaria alone there are hundreds of state laws and by-laws that refer to municipalities.[7] As the state laws apply to all communities, the state legislature often uses population thresholds to adjust the rules to different requirements of smaller and larger communities.

We worked through the most relevant parts of the applicable state legislation to detect rules that use population thresholds. We found 14 (excluding the changes in council size) rules that cover nine of the ten thresholds determining the council size (see table 4.1 for an overview of the thresholds relevant to the council size and corresponding further rules).[8]

We found rules that define both local institutions and affect budget size. With regards to budgeting, we found five rules (see table 4.6 in the appendix for a detailed description).[9] For example, communities with more than 5,000 inhabitants receive 7.6 percent of the vehicle tax collected in their territory, while smaller communities do not receive anything - at the same threshold the council increases from 16 to 20 members. Another rule states that a city with more than 50,000 inhabitants may apply to become a county free city. If choosing to do so, the city takes over all the duties previously provided by the county and receives additional transfers accordingly. We conclude that larger cities receive more transfers and provide more

[7] All laws and by-laws are available online at a website provided by the government of Bavaria: http://www.gesetze-bayern.de/jportal/portal/page/bsbayprod.psml (retrieved in March 2011).

[8] Given the amount of laws and bylaws that regulate municipal decision making, it is still possible that there are further undetected thresholds.

[9] Following the state development program (*Landesentwicklungsprogramm Bayern*, December 22, 2009 version) communities may also be grouped into one of five levels that indicate regional relevance for public services. While there are population thresholds among the criteria, these thresholds do not refer to the community itself but rather to the total population of the respective community plus the population of the communities it serves.

Table 4.1: Changes at council size population thresholds

	Population thresholds at # of inhabitants (in tsd)									
	1	2	3	5	10	20	30	50	100	200
	Panel 1: *Thresholds used in Egger and Koethenbuerger (2010)*									
Council size	x	x	x	x	x	x	x	x	x	x
Wage of elected civil servants					x		x	x	x	
	Panel 2: *Further thresholds defining local institutions*									
Wage of full-time mayors		x	x	x	x		x	x	x	
Wage of part-time mayors	x		x	x						
Full-time council members					x					
Petition for referendum					x	x	x	x	x	
Referendum quota								x	x	
City districts									x	
Open council					x				x	
Accounting committee				x						
Mayor status				x	x					
	Panel 3: *Further thresholds defining budgeting rules*									
County free city								x		
Status of larger city							x			
Vehicle Tax				x						
Fiscal equalization				x	x			x	x	

Notes: We have included only those rules and thresholds that correspond to the thresholds relevant for the council size. For a more detailed description the changes at each of those population thresholds see tables 4.6 and 4.7. *Source*: Own research.

services than smaller cities and that the thresholds typically coincide with the community council size thresholds.

Regarding local institutions, there are even more rules that depend on population thresholds (we found nine, see table 4.7 in the appendix for details). These prescriptions range from relatively minor directives (e.g. that the council in cities with more that 5,000 inhabitants must include an accounting committee) to setting key rules of the local game. The latter include stipulations for the remuneration of the mayor[10] and a detailed regulation concerning

[10]This is likely to be of significant importance for their performance and may hence affect government spending. Compare e.g. Besley (2004), Gagliarducci and Nannicini (2009), Messner and Polborn (2004), Ferraz and Finan (2009)

local referendums. The requirements to bring a proposed referendum to the ballot are much higher in small communities (where 10 percent have to sign the petition) than in very large cities (where only three percent need to sign) with multiple steps in the signature requirement in between. The same general logic applies to the participation quota in the referendum itself.[11]

We conclude that while the number of seats in the council is an important feature of local institutions, there are several other important local institutions that also change at the same population thresholds. This leads to double or multiple simultaneous treatment at any given threshold. The same applies when population thresholds also change the financial endowment of the community (see paragraph above).

Apart from setting population thresholds with budgetary and institutional relevance, the state laws also provide the communities with many choices in those dimensions. For example, the communities may decide to take over tasks from the county, which brings along new responsibilities, funds and administrative work. Tasks that communities may take over from higher government levels include: maintaining certain types of roads (compare *Finanzausgleichsgesetz*, Article 13a, version of 2010/06/03), construction supervision (compare Article 53, *Bayerische Bauordnung*, version of 2010/02/25), and waste disposal (compare Articles 5, 7 *Bayerisches Abfallwirtschaftsgesetz*, version of 2010/03/24). While some of the tasks that communities acquire from higher levels of government only induce small changes in revenue (e.g. construction supervision), others trigger large increases in revenues and spending (e.g. waste disposal).

Moreover, the municipality must regularly make decisions about institutional questions such as: Should we have a full-time mayor?, How many full-time working council members do we want? As mayors are elected at the same time that the council is elected and because additional responsibilities are also more likely to be taken over at the beginning of the electoral cycle it is probable that such decisions also coincide with increases in council size. The econometric implications of the resulting possible simultaneous endogenous decision making are discussed in section 4.2.

[11] e.g. Romer, Rosenthal, and Munley (1992) and Holcombe and Kenny (2008) show that referendum requirements affect spending.

4.4 Results

Our empirical analysis generally builds on the strategy and data presented by Eggers and Koethenbuerger (2010).[12] Similar to their work, we use data on 2,056 municipalities from the German state of Bavaria over the 1983 to 2004 period. To investigate our specific points of interest, we complement their data base with information on the revenue side of the municipality budget and the status of the town mayor (part-time or full-time).[13] The result section is split into three parts: (1) the importance of simultaneously determined finances, (2) the need to control for endogenous (but simultaneous) decisions of the municipality, and (3) the potential manipulation by local authorities of population numbers around thresholds.

4.4.1 Importance of simultaneous exogenous co-treatment

In the above section, we illustrate that legislative population thresholds in Bavaria not only affect council size but also a number of other important institutional features. Among those features are legislative rules that affect the budgeting of communities. Naturally, those direct fiscal consequences for the community are crucial in the identification of the effect of legislative size and governmental spending.

To understand the structure of municipality budgets in more detail, we illustrate some of the basic figures in table 4.8 (see appendix). The average municipality budget in Bavaria (for all 2056 municipalities in the period 1983-2004) reaches 1909 Euro per capita in total expenditures.[14] On the expenditure side, we highlight the shares of three major budget components that are (partly) in the discretion of the municipality: expenditures on personnel, on materials, as well as investments.[15]

Below, we present important categories from the revenue side. Under the full discretion of the local authority are three tax rates[16]: (1) property tax A on agricultural land, (2) property tax B on all non-agricultural property and (3) a trade tax on local businesses. In column (2) we highlight the share that each category has in the overall budget. We find that local taxes account for 20.6 percent of the total revenue with the major part of that income generated by the trade tax. Local property taxes contribute only about 4.4 percent to local finances.

[12] We obtained the data and Stata-dofiles from the authors through the journal website.

[13] This additional data is publicly available and can be obtained free of charge from the Bavarian statistical service.

[14] The same figure has to appear on the revenue side, as all expenditures have to be refinanced either by tax income, grants or increases in new debt.

[15] Other important expenditures that are not included are expenditures for debt repayments and expenditures for mandated social services. These are usually not directly in the control of the municipality.

[16] Only the tax rates are at the discretion of the municipalities - the rules defining the tax base are set by higher government levels.

Equally important are grants from other tiers of government and fees for communal services. We document that about 12.7 percent of revenues come from grants and 9.4 percent are raised through fees on local services.[17] It is important to note that the communities have only very limited decision power when it comes to influencing grants. Of course, they can lobby to receive grants and they have some discretion in setting fees, however, for the most part grants underlie legislative rules and many fees are regulated by the state.[18]

Table 4.2: Replication of the results by Egger and Koethenbuerger (2010)

	Outcomes (15 percent window)				
	Log Total Expenditures	Log Debt	Log Prop Tax Rate A	Log Prop Tax Rate B	Log Trade Tax Rate
	(1)	(2)	(3)	(4)	(5)
Treatment	0.109**	0.137*	0.055***	0.058***	0.009**
	(0.052)	(0.074)	(0.009)	(0.008)	(0.004)
N	22631	22162	22631	22631	22631
R2	0.00	0.00	0.00	0.00	0.00

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We use the program provided by Egger and Koethenbuerger (2010) to replicate their main results (see their tables 3 and 6). Standard errors in parentheses are robust. The dependent variable are indicated above. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. The estimation is done within the 15 percent window of population around the thresholds. *Source*: Own calculations based on the program provided by Egger and Koethenbuerger (2010).

Table 4.2 replicates the results reported by Egger and Koethenbuerger (2010).[19] Egger and Koethenbuerger estimate a version of eq. 4.2 for the thresholds at which population size changes and document the effect on total expenditures to be an immense 10.9 percent increase (see column 1, table 4.2) or 4.2 percent for every additional seat.[20] Given the average total expenditure (per capita) of 1909 Euro, the effect is hence argued to increase expenditures by the order of about 200 Euro (per capita and year). The size of this effect is to be evaluated even larger considering that only part of the local expenditures is effectively under local control.

Looking at the revenue side of the effect, Egger and Koethenbuerger (2010) report the estimates on debt, both types of property tax and the local trade tax (see columns 2-5 in table 4.2).

[17] Those fees are levied on services such as water supply, sewage and waste management, kindergartens, etc.

[18] The remaining part of the revenue comes from higher levels authorities and are raised, in part, through income tax and VAT. The community gets a fixed share of these revenues and has no control over the tax rate. Moreover, revenues can also be generated through new debt.

[19] We use the same dofiles and data and get the exact some results from their table 3 and table 6.

[20] The average increase in the number of seats at the thresholds is 2.56.

They find sizable and significant effects primarily for the two property tax rates A and B of 5.5 percent and 5.8 percent increases respectively. They argue that municipalities mainly rely on those sources of revenue to finance the additional expenditures (see Egger and Koethenbuerger (2010, p. 211)). However, given that the property taxes only account for 4.4 percent of total revenue, an increase by about 5.8 percent in the those taxes is likely to increase revenue by only 5 Euro (per capita and year).[21]

Table 4.3: The results of council size on state grants and municipality fees

	Outcome (15 percent window)		
	Log General Grants	Log Invest Grants	Log Muni Fees
	(1)	(2)	(3)
Treatment	0.089*	0.273***	0.236***
	(0.047)	(0.085)	(0.070)
N	21513	21383	21511
R2	0.00	0.00	0.00

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We use the program provided by Egger and Koethenbuerger (2010) to estimate the regressions using their setup, however, with the outcome variables that we have added to the data. Standard errors in parentheses are robust. The dependent variable are indicated above. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. *Source*: Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

Given the exploration on the institutional changes at population thresholds, we suspect that the additional spending is largely driven by automated changes in the rules that affect grants as well as the services communities provide (and hence fees collected). In table 4.3 we apply the design to estimate the increase in grants and fees at the thresholds. We show that there are substantial, significant increases both in grants (27.3 percent in additional investment grants) and communal fees (23.6 percent additional revenue). In total, the increases in grants and fees account for about 81 Euro per capita. Hence, we can explain a substantial part of the expenditure increase (although not the entire effect) by automated changes in revenues provided by higher levels of government to larger communities.

[21]The effect that they report on the local trade tax is very small with less than 1 percent increase, which implies that this source of revenue does not play an important part in financing the expenditure increase. The point estimate on the debt is in fact quite large and could explain a substantial part of the expenditure increase. However, it is insignificant at the 5 percent level.

We document that the identification of the council size effect may suffer from co-treatment effects that are simultaneously determined by the same population thresholds. Without the detailed institutional knowledge and more specific data on all aspects of the revenue sources[22] it is impossible to distinguish between the "true" effect of legislative size and other changes at the same thresholds.

4.4.2 Simultaneous endogenous decision making

Causal analysis is further complicated by the fact that municipalities must make important decisions on the structure of the local institutions and the services provided by the community. It is in the nature of political cycles that the timing of those endogenous choices coincides with adjustments in council size. As explained above those decisions can have substantial consequences for municipal fiscal situations. We investigate one important feature of local institutions, namely whether the mayor is working part-time or full-time.

In figure 4.2 in the appendix we highlight the share of full-time mayors over the population distribution. The vertical lines indicate the population thresholds used in the council size analysis. Significant differences in the share of full-time majors can be observed around the 2000 and 3000 inhabitants thresholds.[23] To further investigate the differences, we present the distribution of full-time majors just around those two marks in figure 4.3 in the appendix. As indicated by the local kernel regression fitted onto the data, we observe substantial differences right of the cut-off points.

Moreover, we show that the differences in mayor status also prevail in the entire sample. Using the same estimation setup, we use the analysis to predict the mayor status (see table 4.4). If mayor status is independent of council size, the estimates should be insignificant from zero. Our results indicate that there is a sizable and significant effect of the population thresholds on the probability of choosing a full-time mayor.[24]

The above findings indicate that mayor status differs significantly between the observations just to the left and just to the right of the thresholds. The important question is how that affects the estimates of the council size effect. In table 4.5 we repeat the estimation including

[22] To rule out all automated changes in revenues / expenditures one would have to get detailed data on very specific grants that are given from the state or federal level to the communities.

[23] As described above, there are two legal thresholds at 5,000 and 10,000 inhabitants that prescribe consequences for the mayor status. However, those rules are non-binding and only suggestions. We find that the probability of choosing a full-time mayor does not seem to be altered systematically at those thresholds. Rather, the important changes are at thresholds below 5,000.

[24] We present the results within the 4 window sizes used by Egger and Koethenbuerger (2010). A window size of 15 implies that only municipalities within 15 percent of the population threshold are used in the analysis.

Table 4.4: The results of council size on predetermined mayor status

	Probability of having a full-time mayor			
Window size	15 Percent	20 Percent	25 Percent	30 Percent
	(1)	(2)	(3)	(4)
Treatment	0.068***	0.058**	0.051**	0.064***
	(0.026)	(0.023)	(0.020)	(0.019)
N	22611	29783	37177	44295
R2	0.00	0.00	0.00	0.00

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In this regression, we estimate the effect of council size on the mayor status defined as a dummy variable which takes the value 1 if the mayor is full-time employed. Standard errors in parentheses are robust. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. In general, the setup is similar to the regressions run by Egger and Koethenbuerger (2010). *Source*: Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

a dummy for each observation that takes the value one when the municipality employs a full-time mayor during that year. For comparison, we highlight the results from Egger and Koethenbuerger (2010) in panel 1 of table 4.5. In the panels below we show the results controlling for mayor status (panel 2) and additionally controlling for year fixed effects (panel 3).

Including a control for mayor status, we find that the point estimates reported in Egger and Koethenbuerger (2010) drop significantly and are no longer statistically significant in the preferred specifications. The specification using a 15 percent window is shown to exhibit no more effects of the council size on governmental spending.[25] If the mayor status is a predetermined variable, this result would immediately suggest that the RDD is invalid. However, as argued above, the timing is specific. The mayor status is determined after the new council size is fixed, yet before this new council takes office. One could argue that it is in anticipation of the new council members that the old council decides to employ a full-time mayor and that this in turn increased spending. This would mean to change the interpretation of the effect drastically. However, we believe that the choice of the mayor's status is in fact unrelated to the future council size. Under this assumption, results imply that the RDD fails to identify

[25]In table 4.9 in the appendix, we show that this is also true for the effect of council size on the disaggregate spending categories (referring to table 5 in Egger and Koethenbuerger (2010)). However, when we include the mayor dummy in the analysis on revenue sources, we can not reject their estimates for the local tax rates. In table 4.10 in the appendix, we show that the effects on property taxes and trade tax remain stable when we include the mayor dummy. As these categories are true choice variables of the community, they might reflect a different dynamic than the categories steered by choices of responsibilities and grants from other tiers.

Table 4.5: Main results - council size effect with controlling for mayor status

	Log Total Expenditures			
Window size	15 Percent	20 Percent	25 Percent	30 Percent
	(1)	(2)	(3)	(4)
	Panel 1 : Replication of Egger and Koethenbuerger (2010)			
Treatment	0.109**	0.121***	0.149***	0.202***
	(0.052)	(0.045)	(0.041)	(0.037)
N	22631	29803	37197	44324
R2	0.00	0.00	0.00	0.00
	Panel 2 : Controlling for mayor status			
Treatment	0.024	0.049	0.087***	0.124***
	(0.040)	(0.034)	(0.031)	(0.028)
Mayor status	1.240***	1.233***	1.227***	1.217***
	(0.010)	(0.009)	(0.008)	(0.007)
N	22611	29783	37177	44295
R2	0.42	0.42	0.41	0.41
	Panel 3 : Controlling for mayor status and year effects			
Treatment	-0.006	0.019	0.044	0.063***
	(0.034)	(0.030)	(0.027)	(0.024)
Mayor status	1.157***	1.146***	1.129***	1.118***
	(0.009)	(0.008)	(0.007)	(0.006)
N	22404	29483	36766	43809
R2	0.47	0.47	0.47	0.46

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust. The dependent variable is the log of total expenditures. Panel 1 replicates the results in Egger and Koethenbueger (2010) (see their table 3). In panel 2, we control for the mayor status in the estimation by including a dummy variable that takes the value 1 if the mayor is full-time employed. Panel 3, additionally also controls for year fixed effects. In those last estimations we also excluded bigger communities that had the status of a county free city. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. *Source*: Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

the causal effect of an increase in council size.

Interestingly, we observe a large estimate for the mayor status dummy, implying a 124 percent increase in total spending when the mayor is full-time. We do not intend to interpret this to be a causal effect. Rather, we believe that is it very likely that this estimate reflects the importance of the complete choice set of the community at the time of elections. It is precisely when the council decides to take over more responsibilities from other tiers (like child care, water supply, etc.) that they also choose to have a full-time mayor that administers the local

administration.

4.4.3 Manipulation of the population numbers around the thresholds

We show above that simultaneous co-treatment as well as simultaneous endogenous choices are posing a threat to the validity of the causal inference. As argued in the section on the empirical methodology, the important identifying assumption is the argument that municipalities can not sort around the population thresholds (continuity assumption). This is of specific concern here as the population figure is generally observable at all times and action can be taken to manipulate the precise figures before the deadlines in question. Municipalities could act to manipulate the statistical numbers directly within the administration or (more likely) they could start programs designed to attract new residents.[26] Given that the thresholds involve a multitude of consequences both for the political institutions and the fiscal budgeting of the municipality, the concern of manipulation needs to be taken seriously.

Egger and Koethenbuerger (2010) indicate the total number of observations just left (10914 obs) and just right (11690 obs) within the window of the limited sample.[27] However, it is standard praxis in applications of RDD to show a histogram of the frequencies just around the thresholds. In figure 4.1, we present such a histogram for the data used throughout this analysis (within the 15 percent window).

The graph indicates that the frequencies of observations just right and left of the thresholds differ systematically. There is a definite jump in the number of observations if one compares the two groups just below the threshold (about 700 obs each) compared to the one group just above (about 900 obs). The difference is further indicated by the gap in the local kernel regression fitted onto the data.[28] In figure 4.4 in the appendix, we investigate the frequency histogram for each of the thresholds individually. With the exception of the 2,000 threshold, the frequency tend to be always higher just to the right of the threshold. Of considerable difference are the jumps in the graphs for the 1,000, 3,000, 5,000 and 50,000 thresholds.[29]

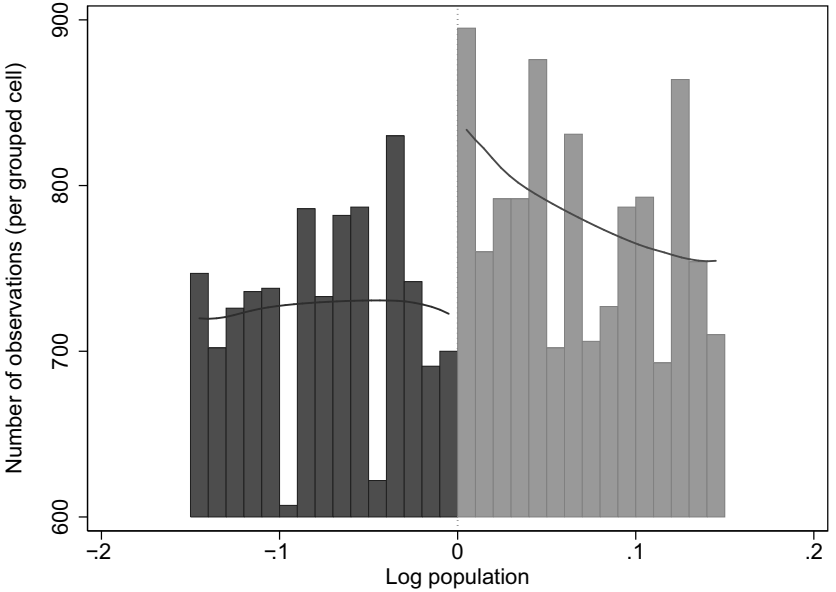
[26] Municipalities could, for example, open new community areas for housing projects, guarantee kindergarten spots to newcomers or give direct financial incentives to move to the town.

[27] Their results are presented in their table 3.

[28] We also ran flexible polynomial regressions on the binned data to obtain standard errors on the observed difference. As shown in the histogram, the data are quite variable between the bins. Nevertheless, the jump is significant at the 10 percent significance level when we run linear and quadratic specifications and at the 5 percent level using a third order polynomial that is flexible on either side.

[29] We also tested differences of important predetermined variables such as total expenditures the year before the election and mayor status during the last election period. However, we found no significant differences there, which implies that sorting is not along those dimensions.

Figure 4.1: RDD validity - frequency check



Notes: The figure shows the frequencies of observations in the analysis in grouped bins within the 15 % window (using 30 bins with a size of one percent each). The thresholds in the analysis have been normalized to zero. The line fitted onto the data is based on a local kernel regression using Epanechnikov weights. A regression analysis using those bin averages and a flexible polynomial specification in the log of population shows that the difference at the threshold is both sizable and significant at least at the 10 percent level (at 5 percent level when one uses a third order polynomial specification). *Source:* Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

This observation has direct consequences to the validity of the estimation design. If communities have the capability to manipulate their population figure at the margin, the continuity assumption for the RDD does not necessarily hold and the estimates are potentially biased due to selection effects. It might be the fast growing communities that seek to take on more responsibilities which manage to end up just to the right of the thresholds.

4.5 Conclusions

In this paper, we identify and discuss three main challenges requiring careful attention when using population thresholds in regression discontinuity designs. First, the population threshold

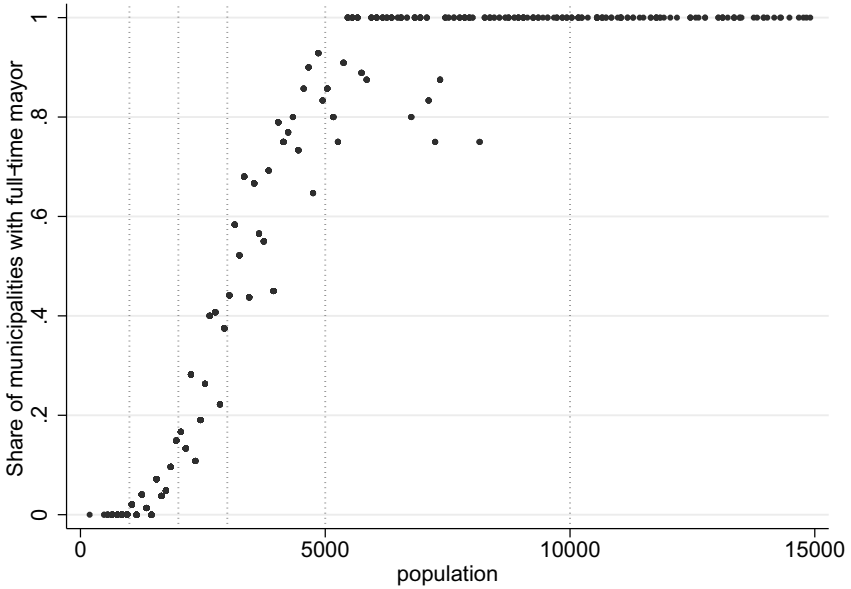
may define changes in multiple rules and not only the treatment considered. Second, when changes in population trigger the observed treatment it is likely that endogenous institutional choices occur simultaneously. Third, political entities may seek to manipulate the official number of inhabitants knowing that it affects institutions at certain thresholds.

Revisiting Egger and Koethenbuerger (2010), we find evidence that our three concerns are of practical relevance. (1) The population thresholds used by the authors not only trigger changes in council size but also affect many other budgetary and institutional rules. Bigger communities have more responsibilities, receive more transfers and differ in institutions. We show that a large share of the spending increase at population thresholds relevant for the seat count stems from increases in grants from higher government levels. (2) We observe that when city councils grow in seats, the communities often decide endogenously to have their mayor work full instead of part time. When we include the information on the status of the mayor the spending effects of additional seats found by Egger and Koethenbuerger (2010) become insignificant. (3) We find evidence for manipulation of the population count, again indicating that the reported estimates are likely to be invalid. We conclude that the causal effect of council size on government expenditures in Germany is still an open question.

From our results we find that it is crucially important to thoroughly verify the identifying assumptions of the RDD when using population thresholds. Taking into account the complexity of institutions and exogenous rules, notably at the local level, we recommend that researchers carefully document the legislative setting throughout the period of observation. In particular, they should discuss in detail whether population thresholds are used for other institutions (including the applicable laws and thresholds) and what other endogenous decisions might be taken simultaneously. The responsibility lies entirely on the researcher as the precise institutional setting is hard to assess for any outsider and near to impossible to judge from the pure data analysis. Further, the underlying continuity assumption needs to be discussed with great detail and all available tests have to be carried out such as to illustrate that the assumption is indeed supported by the data.

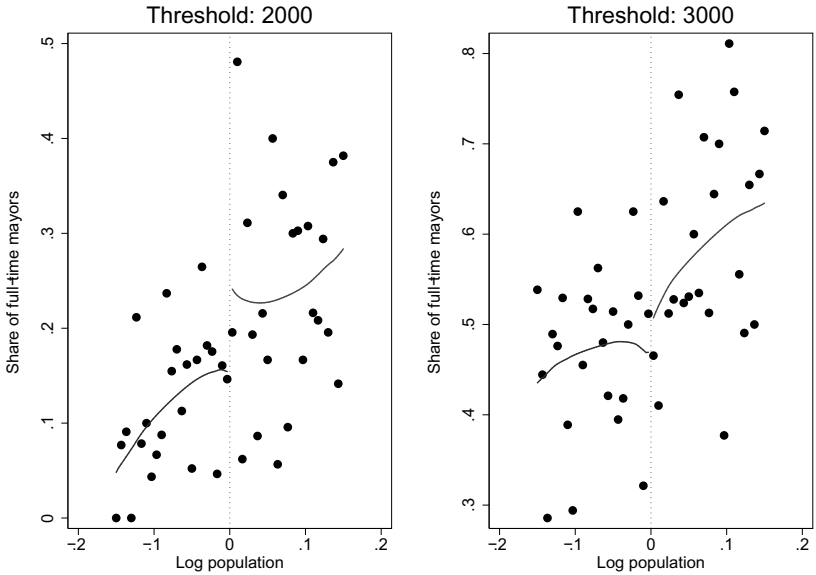
Appendix

Figure 4.2: Share of full-time mayors over the population distribution (in 1984)



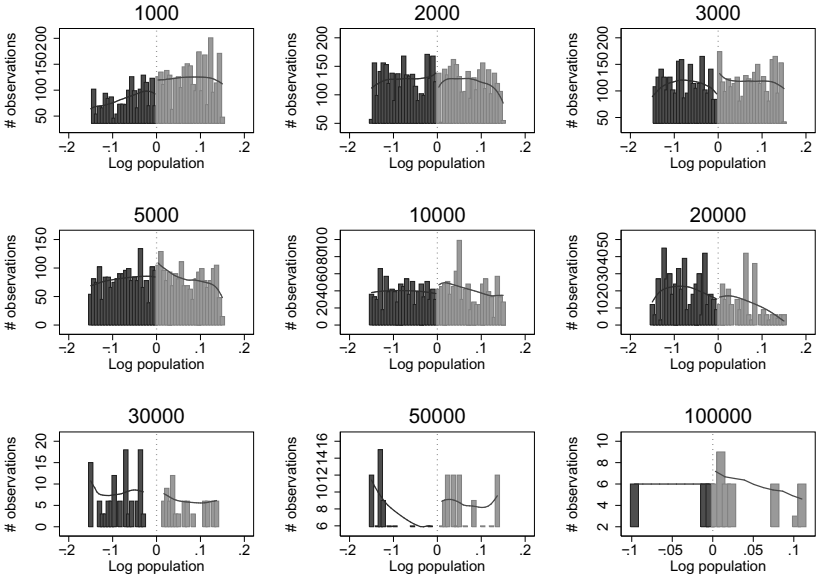
Notes: This figure illustrates the share of full-time mayors over the distribution of the population below 15000 inhabitants (there are no changes above). Each point represents the share within a bin (bandwidth equals 100 inhabitants) including all municipalities of that size. The vertical lines illustrate where the thresholds at with the council size changes are. *Source:* Own calculations.

Figure 4.3: Distribution of full-time mayors around thresholds



Notes: The figure investigates the share of full-time mayors just around the thresholds of 2000 and 3000 inhabitants. Similar to Egger and Koethenbuerger (2010), we present the results within the 15 % window (using 46 bins representing about 2/3 of one percent). The lines fitted onto the data is based on a local kernel regression using Epanechnikov weights. *Source:* Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

Figure 4.4: RDD validity - frequency check



Notes: The figure shows the frequencies of observations at all thresholds below 200,000 in grouped bins within the 15 % window (using 30 bins with a size of one percent each). The lines fitted onto the data is based on a local kernel regression using Epanechnikov weights. *Source:* Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

Table 4.6: Defining the budget: population thresholds in Bavarian law

Rule	Population thresholds	Legal source	Description
County free city	50,000	e.g. FAG Art. 2, 9, 12	Cities with more than 50,000 inhabitants can apply to become a county free city ('Kreisfreie Stadt'). That means that the city legally leaves the county it belonged to before and takes over all tasks that were executed by the county until then. These tasks include such important and costly assignments as social welfare, vocational schools, and car registration. The transfer of tasks also includes a transfer of the budget associated with these tasks. The additional funds that the council free city enjoys stem both from additional state transfers and a county levy that the city now no longer needs to pay.
Larger city	30,000	GO, Art. 5a	Communities with more than 30,000 inhabitants can apply for the larger city status ('Große Kreisstadt'). Once awarded the larger city status the community takes over some of the duties that are usually carried out by the county to which the community belongs. The additional tasks include construction supervision and road traffic authority. Some of these tasks also create additional revenue for the cities (e.g. fees of the construction supervision authority).
Vehicle Tax	5,000	FAG, Art. 13a	Communities with more than 5,000 inhabitants receive 7.6 percent of the vehicle tax collected on their territory, the rest of the tax revenue remains with the state government. Smaller communities do not receive a share of the vehicle tax. Whenever communities - regardless of their size - take over the responsibilities for certain types of roads they receive an even larger share of the vehicle tax.
Water management	90,000; 143,750; 300,000; 600,000	FAG, Art. 9	County free cities are in charge of managing all issues related to water in their city. To help them cover the associated cost county free cities receive a per capita allowance of 80 cents per inhabitant but not more than 115,000 EUR which is equivalent to to a population threshold of 143,750. Beyond they also receive a supplemental lump sum allowance that depends on their size. This lump-sum is 25,000 EUR for cities below 90,000 inhabitants, 35,000 EUR for cities with more than 90,000 and up to 300,000 inhabitants, 50,000 EUR for cities with more than 300,000 and up to 600,000 inhabitants and 100,000 EUR for all cities with more than 600,000 inhabitants.
Fiscal equalization	5,000; 10,000; 25,000; 50,000; 100,000; 250,000; 500,000	FAG, Art. 4	Next to own revenue the communities in Bavaria rely on transfers from the state government to finance their expenditures. One of the most important transfers is the so called 'key allocation' (Schluesselzuweisung). The larger the community is the more it receives in key transfers per capita. Communities with more than 5,000 inhabitants receive a per capita bonus for every additional inhabitant. The increase is 8 percent per inhabitants once the community reaches 5,000 inhabitants and is then a linear function on the population with several points a which the slope changes. For example at exactly 50,000 inhabitants the bonus is 35 percent compared to a community with less than 5,000 inhabitants. The population thresholds mentioned in this tables do not change the level of the key transfers discontinuously, however, the slope of the key transfer schedule (as a function of population size) changes. This change in the slope might in turn induce discontinuous changes in the optimizing strategy of the economic agent.

Legal sources: FAG - Finanzausgleichsgesetz fuer Bayern (version: June 2, 2010), GO - Gemeindeordnung für den Freistaat Bayern (version: August 17, 2009). *Source*: Own research.

Table 4.7: Defining local institutions: population thresholds in Bavarian law

Rule	Population thresholds	Legal source	Description
Full time council members	10,000	GO Art. 40	In communities with more than 10,000 inhabitants the council can elect additional full time council members that take over management functions in the local administration.
Wage elected civil servants [a]	2,000; 3,000; 5,000; 10,000; 15,000; 30,000; 50,000; 100,000.	BKBV, Art. 1; KWBG, supp. 1	The full-time mayor, the full-time deputy mayor(s), and the full time council members have the status of elected civil servants. Their remuneration depends on the size of the community. At each of the threshold indicated their salaries increase by one or two steps in the civil servants salary scale. For example in ordinary communities with more than 15,000 and up to 30,000 inhabitants the mayor is at the level B2 or B3 while in communities with more than 30,000 inhabitants she is at the level B3 or B4. An increase in one step in this system (from B3 to B4) implies about six percent wage increase (or EUR 400 per month) in Bavaria in 2010. Furthermore, elected civil servants receive a monthly expense allowance that depends on the population if the city is either a community free city or has the large city status. Then the allowance thresholds are 50,000, and 100,000 inhabitants. For example the mayor receives between 305.68 and 833.76 EUR in such communities with less than 50,000 inhabitants and between 437.72 and 965.79 EUR in communities with more than 50,000 but less than 100,000 inhabitants.
Cost allowance part time mayors	1,000; 3,000; 5,000	GO, supp. 1	Part time mayors receive a cost allowance that depends on the population size. The law sets overlapping allowance ranges for the different community size classes. For examples mayors in communities with up to 1,000 inhabitants receive a monthly allowance between 430.48 and 1,907.88 EUR and while this amount increases to between 1,830.10 and 3,307.51 EUR in the size class above.
Petition for referendum	10,000; 20,000; 30,000; 50,000; 100,000; 500,000	GO, Art. 18a	The inhabitants in a community can request a referendum. To be considered, the organizers of the request need to gather signatures from the inhabitants. What share of the inhabitants they need to sign depends on the size of the community. Ten percent have to sign if the community has up to 10,000 inhabitants, 9 percent up to 20,000, 8 percent up to 30,000, 7 percent up to 50,000, 5 percent up to 100,000 and 3 percent have to sign if the community has more than 500,000 inhabitants.
Referendum quota	50,000; 100,000	GO, Art. 18a	Once a petition for referendum has reached the necessary number of signatures the voters are called to the ballot. Then the referendum needs to fulfill two conditions to pass. First, it has to get more then fifty percent of the votes. Second, these votes must represent at least 20 percent of the eligible electorate in communities with up to 50,000 inhabitants. In communities with up to 100,000 inhabitants the supporting votes need to represent at least 15 percent of the electorate and in communities with more than 100,000 inhabitants this share has to be at least 10 percent.
City districts	100,000	GO, Art. 60	Communities with more than 100,000 inhabitants are subdivided into several city districts. These city districts take over part of the administration of the city. The city council can decide to hand over decision powers to the districts. In this case they also have a district councils whose members are elected.
Open council	10,000; 100,000	GO, Art. 18a	In an open council ('Buergerversammlung') the constituents meet to discuss topics of local relevance. The meetings are chaired by the local administration. The open council can pass recommendations that then have to be discussed in the community council. In communities with less than 10,000 inhabitants five percent of the voters have to request the open council meeting, in communities with more than 10,000 inhabitants only two and a half percent need to request the open council. In cities with more than 100,000 inhabitants open council can also be held within a city district.
Accounting committee	5,000	GO, Art. 103	In communities with more than 5,000 inhabitants the council has to establish an accounting committee composed of council members to conduct an examination of accounts.
Mayor status	5,000; 10,000	GO, Art. 34	Communities with more than 10,000 inhabitants must have full-time mayors. Communities with more than 5,000 but not more than 10,000 inhabitants have a full time mayor by default. However, they can choose to have a part-time mayor instead if the council decides this not later than 67 days before the mayor election. Communities with 5,000 inhabitants or less are governed by part-time mayors by default. However, they can also choose to have a full-time mayor if the council decides so not later than 67 days before the mayor election.

Legal sources: BKBV - Bayerische Kommunalbesoldungsverordnung (version: May 29, 2009), FAG - Finanzausgleichsgesetz fuer Bayern (version: June 2, 2010), GO - Gemeindeordnung für den Freistaat Bayern (version: August 17, 2009), KWBG - Gesetz ueber kommunale Wahlbeamte (version: August 17, 2009) *Source*: Own research.
a: Egger and Koethenbuerger (2010) state that the remuneration of the full time council members depends on the population thresholds but do not mention that this also applies to mayors and deputy mayors. As full time council members can only be elected in communities with more than 10,000 inhabitants they conduct a robustness that only includes communities with less than 10,000 inhabitants to make sure that remuneration steps for full time council do not play a role for their identification.

Table 4.8: Communal budgeting - shares of expenditure and revenues

	Mean	Share in %
Expenditures (Euro per capita)		
Total	1909.5	100.0
on Personnel	369.6	19.4
on Materials	251.2	13.2
in Investment	368.3	19.3
Revenues (Euro per capita)		
Total	1909.5	100.0
from Property Tax A	6.3	0.3
from Property Tax B	77.9	4.1
from Trade Tax	310.1	16.2
from General Grants	146.4	7.7
from Investment Grants	96.3	5.0
from Fees	178.6	9.4
Average debt (Euro per capita)	788.4	

Notes: This table illustrates basic figures of an average municipality budget. Column 2 presents the population weighted average number of Euro per capita in each category over all 2056 municipalities during the period of observation (1984-2004). Column 2 highlights the share that the individual item has on total expenditures or revenues respectively. The categories named are not exclusive, rather, we refer to all categories used in the analysis. *Source*: Own calculations.

Table 4.9: The results on expenditure categories controlling for mayor status

	Expenditure Categories (Window size 15 percent)		
	Invest Expend	Material Expend	Personnel Expend
	(1)	(2)	(3)
	Panel 1 : Replication of Egger and Koethenbuerger (2010)		
Treatment	0.103*	0.143**	0.169***
	(0.061)	(0.059)	(0.059)
N	22623	22626	22631
R2	0.001	0.001	0.001
	Panel 2 : Controlling for mayor status and year effects		
Treatment	-0.047	0.026	0.052
	(0.048)	(0.039)	(0.038)
Mayor status	1.157***	1.373***	1.367***
	(0.012)	(0.010)	(0.010)
N	22396	22399	22404
R2	0.310	0.488	0.505

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are robust. The dependent variables are indicated above. Panel 1 replicates the results in Egger and Koethenbueger (2010) (see their table 5). In panel 2, we control for the mayor status in the estimation by including a dummy variable that takes the value 1 if the mayor is full-time employed. Additionally, we also control for year fixed effects and we excluded bigger communities that had the status of a county free city. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. *Source*: Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

Table 4.10: The results on revenues controlling for mayor status

	Revenues in Logs (Window size 15 percent)					
	Debt	Prop Tax A	Prop Tax B	Trade Tax	Gen. Grants	Invest Grants
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.031	0.066***	0.066***	0.009**	0.136*	0.106**
	(0.060)	(0.009)	(0.008)	(0.004)	(0.077)	(0.049)
N	21935	22404	22404	22404	21186	21314
R2	0.256	0.050	0.052	0.019	0.181	0.454

Notes: Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are indicated above. In this regression, we estimate the effect of council size on the mayor status defined as a dummy variable which takes the value 1 if the mayor is full-time employed. Standard errors in parentheses are robust. Treatment is defined by being right to the population thresholds. All regressions include a third order polynomial in population that is flexible on both sides of the threshold. In general, the setup is similar to the regressions run by Egger and Koethenbuerger (2010). *Source*: Own calculations partly based on the program provided by Egger and Koethenbuerger (2010).

Bibliography

Acemoglu, D. (2005): "Constitutions, Politics, and Economics: A Review Essay on Persson and Tabellini's The Economic Effects of Constitutions," *Journal of Economic Literature*, 43, 1025–1048.

Ade, F., and R. Freier (2011): "Incumbency, contamination and economic voting. Evidence from a quasi-experiment for Germany," *Mimeo*.

Alberto Alesina, T. P., and G. Tabellini (2006): "Reply to Blankart and Koester's Political Economics versus Public Choice: Two Views of Political Economy in Competition," *Kyklios*, 59, 201–208.

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.

Ansolabehere, S., and J. M. Snyder (2004): "Using Term Limits to Estimate Incumbency Advantages When Officeholders Retire Strategically," *Legislative Office Quarterly*, XXIX, 487–515.

Ansolabehere, S., J. M. Snyder, and C. Stewart (2000): "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage," *American Journal of Political Science*, 44(1), 17–34.

Banzhaf, J. F. (1965): "Weighted Voting Doesn't Work: A Mathematical Analysis," *Rutgers Law Review*, 19, 317–343.

Baqir, R. (2002): "Districting and Government Overspending," *The Journal of Political Economy*, 110(6), 1318–1354.

Berry, C. R., and W. G. Howell (2007): "Accountability and Local Elections: Rethinking Retrospective Voting," *Journal of Politics*, 69(3), 844–858.

Besley, T. (2004): "Joseph Schumpeter Lecture: Paying Politicians: Theory and Evidence," *Journal of the European Economic Association*, 2(2-3), 193–215.

Besley, T., and A. Case (1995): "Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition," *American Economic Review*, 85(1), 25–45.

Bilal, S., P. Albuquerque, and M. . Hosli (2001): "The Probability of Coalition Formation: Spatial Voting Power Indices," Paper to be presented at the ECSA Seventh Biennial International Conference, Maddison, Wisconsin, USA.

Blankart, C. B., and G. B. Koester (2005): "Political Economics versus Public Choice: Two views of political economy in competition," *Humboldt University Discussion Paper*.

——— (2006): "Political Economics versus Public Choice: Two views of political economy in competition," *Kyklios*, 59, 171–200.

Brollo, F., and T. Nannicini (2011): "Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil," *IZA Discussion Paper*.

Brollo, F., T. Nannicini, R. Perotti, and G. Tabellini (2009): "Federal Transfers, Corruption, and Political Selection: Evidence from Brazil," *University of Bocconi Working Paper*.

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.

Cover, A. D. (1977): "One Good Term Deserves Another: The Advantage of Incumbency in Congressional Elections," *American Journal of Political Science*, 21(3), 523–541.

Cox, G. W., and J. N. Katz (2002): *Elbrigde Gerrys Salamander The Electoral Consequences of the Reappointment Revolution*. Cambridge University Press.

Egger, P., and M. Koethenbuerger (2010): "Government Spending and Legislative Organization: Quasi-experimental Evidence from Germany," *American Economic Journal: Applied Economics*, 2(4), 200–212.

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. (1971): "The Advantage of Incumbency in Congressional Elections," *Polity*, 3(3), 395–405.

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

Felsenthal, D. S., and M. Machover (1998): *The Measurement of Voting Power*. Edward Elgar Publishing, Cheltenham.

Ferejohn (1986): "Incumbent performance and electoral control," *Public Choice*, 50(1-3), 5–25.

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

Ferraz, C., and F. Finan (2009): "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance," *NBER Working Paper*, w14906.

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

Fiorina, M. P. (1978): "Economic Retrospective Voting in American National Elections: A Micro-Analysis," *American Journal of Political Science*, 22(2), 426–443.

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

——— (2010b): "Shades of Brown and Green: Party Effects in Proportional Election Systems," Working Paper, IIES.

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," *available as DIW working paper soon*.

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

Gelman, A., and G. King (1990): "Estimating Incumbency Advantage without Bias," *American Journal of Political Science*, 34(4), 1142–1164.

Gerber, E. R., and D. J. Hopkins (2011): "When Mayors Matter: Estimating the Impact of Mayoral Partisanship on City Policy," *American Journal of Political Science*, 55(2), 326–339.

Gilligan, T. W., and J. G. Matsusaka (2001): "Fiscal Policy, Legislature Size, and Political Parties: Evidence from State and Local Governments in the First Half of the 20th Century," *National Tax Journal*, 54, 1, 57–82.

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 (2005): "Incumbency Effects in German and British Elections: A Quasi-Experimental Approach," *Working Paper*.

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

Hinnerich, B. T., and P. Pettersson-Lidbom (2010): "Democracy, Redistribution, and Political Participation: Evidence from Sweden 1919-1950," *Stockholm University Working Paper*.

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

Holcombe, R., and L. Kenny (2008): "Does restricting choice in referenda enable governments to spend more?," *Public Choice*, 136, 87–101.

Jacobson, G. C. (1987): "The Marginals Never Vanished: Incumbency and Competition in Elections to the U.S. House of Representatives, 1952-82," *American Journal of Political Science*, 31(1), 126–141.

Jordahl, H. (2006): "An economic analysis of voting in Sweden," *Public Choice*, 127, 251–273.

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.

King, G., and A. Gelman (1991): "Systemic Consequences of Incumbency Advantage in U.S. House Elections," *American Journal of Political Science*, 35(1), 110–138.

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

Krehbiel, K., and J. R. Wright (1983): "The Incumbency Effect in Congressional Elections: A Test of Two Explanations," *American Journal of Political Science*, 27(1), 140–157.

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

Lee, D. S., and T. Lemieux (2009): "Regression Discontinuity Designs in Economics," Working Paper 14723, National Bureau of Economic Research.

Lee, D. S., E. Moretti, and M. Butler (2004): "Do Voters Affect or Elect Policies? Evidence from the U.S. House," *Quarterly Journal of Economics*, 119(3), 807–859.

Levitt, S. D., and C. D. Wolfram (1997): "Decomposing the Sources of Incumbency Advantage in the U. S. House," *Legislative Studies Quarterly*, 22(1), 45–60.

Liang, C. (2009): "Is There an Incumbency Advantage or Cost of Ruling in Proportional Election Systems," Working paper, Uppsala University.

Lindbeck, A., and J. W. Weibull (1987): "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, 52, 273–297.

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

Litschig, S., and K. Morrison (2010): "Government Spending and Re-election: Quasi-Experimental Evidence from Brazilian Municipalities," *University of Pompeu Fabra Working Paper*.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2), 698 – 714.

Messner, M., and M. K. Polborn (2004): "Paying politicians," *Journal of Public Economics*, 88(12), 2423 – 2445.

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.

Penrose, L. S. (1946): "The Elementary Statistics of Majority Voting," *Journal of the Royal Statistical Society*, 109(1), 53–57.

Persson, T., and G. Tabellini (2000): *Political Economics*. The MIT Press.

——— (2002): *The economic effects of constitutions*. MIT Press.

Pettersson-Lidbom, P. (2006): "Does the Size of the Legislature Affect the Size of Government? Evidence from Two Natural Experiments," *Stockholm University Working Paper*.

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

Romer, T., H. Rosenthal, and V. G. Munley (1992): "Economic incentives and political institutions: Spending and voting in school budget referenda," *Journal of Public Economics*, 49(1), 1 – 33.

Schaltegger, C. A., and L. P. Feld (2009): "Do large cabinets favor large governments? Evidence on the fiscal commons problem for Swiss Cantons," *Journal of Public Economics*, 93(1-2), 35 – 47.

Serra, G. (1994): "What's in it for me? The Impact of Congressional Casework on Incumbent Evaluation.," *American Politics Quaterly*, 22, 403–420.

Shapley, L. S., and M. Shubik (1954): "A Method for Evaluating the Distribution of Power in a Committee System," *The American Political Science Review*, 48(3), 787–792.

Stroemberg, D., and J. M. Snyder (2010): "Press Coverage and Political Accountability," *Journal of Political Economy*, 118, 355–408.

Thistlethwaite, D., and D. Campell (1960): "Regression-Discontinuity analysis: an alternative to the ex post facto experiment," *Journal of Educational Psychology*, 51 (6), 309–317.

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.

Wolfers, J. (2007): "Are Voters Rational? Evidence from Gubernatorial Elections.," *Unpublished manuscript*, University of Pennsylvania.

The Stockholm School of Economics

A complete publication list can be found at *www.hhs.se/research/publications*. Books and dissertations are published in the language indicated by the title and can be ordered via e-mail: *efi.publications@hhs.se*.

A selection of our recent publications

Books

Barinaga, Ester (2010). *Powerful dichotomies.*
Ericsson, Daniel (2010). *Den odöda musiken.*
Ericsson, Daniel (2010). *Scripting Creativity.*
Holmquist, Carin (2011). *Kvinnors företagande – kan och bör det öka?*
Lundeberg, Mats (2011). *Improving business performance: a first introduction.*
Melén, Sara (2010). *Globala från start: småföretag med världen som marknad.* Forskning i Fickformat.
Mårtensson, Pär & Mähring, Magnus (2010). *Mönster som ger avtryck: perspektiv på verksamhetsutveckling.*.
Sjöström, Emma (2010). *Ansiktslösa men ansvarsfulla.* Forskning i fickformat.
Wijkström, Filip (2010). *Civilsamhällets många ansikten.*
Engwall, Lars (2009). *Mercury meets Minerva: business studies and higher education: the Swedish case.*

Dissertations

Alexandersson, Gunnar (2010). *The accidental deregulation : essays on reforms in the Swedish bus and railway industries.*
Bjerhammar, Lena (2011). *Produktutvecklingssamarbete mellan detaljhandelsföretag och deras varuleverantörer.*
Bohman, Claes (2010). *Attraction: a new driver of learning and innovation.*

Bottero, Margherita (2011). *Slave trades, credit records and strategic reasoning: four essays in microeconomics.*

Buturak, Gökhan (2010). *Choice deferral, status quo bias, and matching.*

Bång, Joakim (2011). *Essays in empirical corporate finance.*

Carlsson-Wall, Martin (2011). *Targeting target costing: cost management and inter-organizational product development of multi-technology products.*

Engvall, Anders (2010). *Poverty and conflict in Southeast Asia.*

Fries, Liv (2011). *Att organisera tjänstesektorns röst: en teori om hur organisationer formar aktörer.*

Glassér, Charlotte (2010). *The fountainhead of innovation health: a conceptualization & investigation.*

Hemrit, Maetinee (2011). *Beyond the Bamboo network: the internationalization process of Thai family business groups.*

Khachatryan, Karen (2011). *Essays on managerial overconfidence and gender differences in preferences.*

Lundvall, Henrik (2010). *Poverty and the dynamics of equilibrium unemployment: [essays on the economics of job search, skills, and savings].*

Lychnell, Lars-Olof (2010). *IT- relaterad verksamhetsförändring : processer som formar växelspelat mellan utveckling och användning.*

Monsenego, Jérôme (2011). *Taxation of foreign business income within the European internal market: an analysis of the conflict between the objective of achievement of the European internal market and the principles of territoriality and worldwide taxation.*

Nelson, Maria (2010). *Utflyttning av aktiebolag: en analys i ljuset av den internationella skatterätten och EU-rätten.*

Ranehill, Eva (2011). *Essays on gender, competition and status.*

Stepanok, Ignat (2011). *Essays on international trade and foreign direct investment.*

Strid, Ingvar (2010). *Computational methods for Bayesian inference in macroeconomic models.*

Söderblom, Anna (2011). *Private equity fund investing: investment strategies, entry order and performance.*

Wallace, Björn (2011). *Genes, history and economics.*