

THE POLITICAL ECONOMY OF ELECTORAL REFORMS:
A TALE OF TWO COUNTRIES

Dissertation zur Erlangung des wirtschaftswissenschaftlichen
Doktorgrades der Wirtschaftswissenschaftlichen Fakultät der
Universität Göttingen

vorgelegt von
Mariana Lopes da Fonseca
aus Lissabon

Göttingen, 2016

Erstgutachter: Prof. Dr. Thushyanthan Baskaran

Zweitgutachter: Prof. Dr. Robert Schwager

Tag der mündlichen Prüfung: 14. Juli 2016

Acknowledgements

I benefited from discussions, suggestions, comments and encouragements from more people than I can enumerate during the three-plus years I spent in Göttingen working on my doctoral degree. I am eternally grateful to all of you. There are a few people I wish to name here.

My advisor, Thushyanthan Baskaran, who is a constant source of inspiration. Thank you for the trust and opportunity to work with you. I learn more from our collaborations and reading your work than I have learned in many years of formal education. This dissertation would not have been possible without everything you taught me. Your work is eye-opening, rigorous and always at the highest standard. You taught me the means to aim as high. I feel I learned from the best.

My second supervisor, Robert Schwager, who opened the doors of the Public Economics “Family” for me. My academic experience would have been very different without this feeling of belonging to your Chair, together with fellow PhD students, Kamila, Katharina, and Johannes, and Post-Docs, Christian and Emre. It has been a pleasure. I thank you all for the continuous support and the good times.

Monika Jackmann, thank you for all your help and availability. I could always count on you to solve so many of the challenges of living abroad and not mastering (to say the least) the language. I also want to thank Zohal Hessami for so promptly accepting to be my third supervisor and Michael Wolff for being a member of my examination committee.

In addition, many other people have played important roles during my time as a PhD candidate. I have met extraordinary people in the conferences, seminars and workshops I had the opportunity to participate in. I look forward to see you in conferences to come and to continue the exchange, contributing to and benefiting from, the body of knowledge of this community.

Contents

1	Introduction	1
2	Electoral Competition and Endogenous Political Institutions: Quasi-experimental Evidence from Germany	5
2.1	Introduction	5
2.2	Institutional Details	8
2.2.1	Political Setting	8
2.2.2	Electoral Rules and the <i>Kommunalwahlreform</i>	10
2.2.3	Council Size	12
2.3	Empirical Strategy	16
2.3.1	Difference-in-discontinuities Design	16
2.3.2	Difference-in-discontinuities Plots	20
2.4	Baseline Results	20
2.4.1	Graphical Evidence	20
2.4.2	Regression Results	22
2.5	Robustness Tests and Extensions	23
2.5.1	Placebo Tests	23
2.5.2	Event-study Plot	24
2.5.3	Individual Cutoffs	26
2.5.4	Percent Changes	27
2.5.5	Effect Heterogeneity	27
2.6	Validity of the Identification Strategy	29
2.6.1	Direct Election of the Mayor	29
2.6.2	Other Elements of the <i>Kommunalwahlreform</i>	29
2.7	Conclusion	31
3	Electoral Thresholds and Political Outcomes	33
3.1	Introduction	33

3.2	Background	34
3.2.1	Local Politics	34
3.2.2	The <i>Kommunalwahlreform</i>	34
3.2.3	Hypotheses	35
3.3	Empirical Design	37
3.3.1	Difference-in-discontinuities Design	37
3.3.2	Difference-in-discontinuities Plots	38
3.4	Results	38
3.4.1	Vote and Seat Shares	38
3.4.2	Seat Shares of Individual Parties	42
3.4.3	Election-specific Vote and Seat Shares	42
3.5	Extensions	43
3.5.1	Council Fragmentation	43
3.5.2	Turnout	45
3.6	Conclusion	47
	Appendix to Chapter 3	49
4	Candid Lane Ducks	57
4.1	Introduction	57
4.2	Term Limits and Political Incentives	58
4.3	Institutional Details	60
4.3.1	Local Politics and Finances	60
4.3.2	Electoral Reform	61
4.4	Data	62
4.5	Methodology	63
4.5.1	Empirical Model	63
4.5.2	The Common Trends Assumption	64
4.6	Empirical Evidence	67
4.6.1	Electoral Term and Yearly Effects	67
4.6.2	Heterogeneous Effects	69
4.7	Extensions	72
4.7.1	Political Business Cycle	72
4.7.2	Mayoral Resignation	72
4.8	Robustness Checks	75
4.8.1	Selection Bias	75
4.8.2	Controlling for Anticipation Effects	76

4.8.3	Controlling for Pre-existing Trends	78
4.9	Conclusion	80
5	Identifying the Source of Incumbency Advantage through a Constitutional Reform	82
5.1	Introduction	82
5.2	Incumbency Advantage and the Electoral Process	84
5.3	Institutional Background	86
5.3.1	Local Politics	86
5.3.2	Constitutional Reform	86
5.3.3	Data	87
5.4	Methodology	87
5.4.1	Incumbency Advantage and the RDD	87
5.4.2	Identification Strategy	90
5.5	Empirical Results	91
5.5.1	RDD and Diff-in-disc Estimates	91
5.5.2	Partisan and Personal Incumbency Advantages	93
5.5.3	Explanatory Hypotheses	94
5.6	Internal Validity	95
5.7	Conclusion	98
	Appendix to Chapter 5	100
	References	106

List of Tables

2.1	Population Cutoffs and the Number of Seats	13
2.2	Number of Council Seats	22
2.3	Individual Population Cutoffs	26
2.4	Number of Council Seats (log)	27
2.5	Heterogeneous effects	28
3.1	Seat and Vote Shares	40
3.2	Seat and Vote Shares of Individual Parties	41
3.3	Small Parties	43
3.4	Medium Parties	44
3.5	Large Parties	45
3.6	Council Fragmentation	46
3.7	Turnout	47
A.3.1	Smaller Bandwidths	51
A.3.2	Individual Population Cutoffs	54
A.3.3	Summary Statistics	56
4.1	Term Treatment Effects	68
4.2	Annual Treatment Effects	69
4.3	Heterogeneous Term Treatment Effects	70
4.4	Heterogeneous Annual Treatment Effects	71
4.5	Political Business Cycle	73
4.6	Resigning Mayors – Term Treatment Effects	74
4.7	Resigning Mayors – Annual Treatment Effects	75
4.8	Placebo Treatment	76
4.9	Quasi-myopic model	77
4.10	Group-specific Trends	79
4.11	Summary Statistics.	81

5.1	Rerunning & Success Rates	87
5.2	RDD and Diff-in-disc Estimates	93
5.3	Personal and Partisan Incumbency Advantage	94
5.4	Hypotheses Test	95
5.5	Diff-in-disc Assumption	98
A.5.1	Control Variables	100
A.5.2	Municipality Fixed Effects	101
A.5.3	Different Bandwidths and Polynomials	102
A.5.4	Internal Validity: Losers Self Selection	102
A.5.5	Different Specifications	103
A.5.6	Heterogeneous Effects	104

List of Figures

2.1	Average Population Size in Hesse	9
2.2	Average Council and Population Size in Hesse	13
2.3	Average Council Size in Different Population Brackets	15
2.4	Manipulation of the Running Variable	19
2.5	Council Size	21
2.6	Placebo Treatments with Fake Cutoffs	23
2.7	Placebo Treatment for Fake Treatment Year	24
2.8	Event-study Plot	25
3.1	Seat and Vote Shares	39
A.3.1	Average Council Size Before and After 2001	50
A.3.2	Placebo Treatments with Fake Cutoffs	52
A.3.3	Placebo Treatment for Fake Treatment Year	53
4.1	Mean Plots	65
4.2	Yearly Dummies	66
4.3	Group-specific Trends	67
4.4	Yearly Effects	78
5.1	Vote Share and Winning Probability	92
5.2	Manipulation of the Running Variable	96
5.3	Selection Bias	97

Chapter 1

Introduction*

“If patterns of political outcomes are acknowledged to be rule dependent, and if rules that describe the inclusive political process are themselves considered to be variables, subject to change by deliberative reform, more attention to the causal chains of relationships between rules and outcomes is surely warranted.”

– James Buchanan

Institutions shape society as a whole and play a crucial role in the general economic environment through the structuring of economic incentives. Acemoglu et al. (2005) place institutions, along with geography and culture, as one of the main factors explaining differences in economic performance. This book focuses on a specific set of institutions, political institutions, to study their relevance for a number of economic and political outcomes and behaviors. Despite the large amount of literature on political institutions that take these as predetermined and exogenous when assessing their impact on policy choices, recent contributions emphasize the endogeneity of institutions (Aghion et al., 2004). Studies show that institutions themselves evolve in response to societies’ drives and needs. Hence, evidence obtained on the premise of historical and static institutional contexts may misrepresent the mapping of institutional rules to policy outcomes (Besley and Case, 2003).

This book starts precisely by contributing to this emerging literature on endogenous political institutions by providing evidence on the manipulation of local political instruments by the established political class. It then continues in pursue of the causal relationship between rules and outcomes called for in Buchanan (2008). For the purpose, all chapters rely on electoral reforms used as natural experiments that induce exogenous variation on the institutional context and incentives coming therefrom, allowing for the identification of causal treatment effects. In particular, the identification strategies rely at least partially on one of the two following electoral reforms: the Hessian *Kommunalwahlreform* from 1999 and the recent introduction of term limits in Portuguese municipalities. By combining these institutional reforms and state of the art quasi-experimental methodologies, it

*This chapter heavily borrows from the abstracts of the paper versions of Chapters 2-5, Baskaran and Lopes da Fonseca (2016), Baskaran and Lopes da Fonseca (2014), Lopes da Fonseca (2015b) and Lopes da Fonseca (2015a), respectively.

is possible to measure the causal impact of political institutions not only on electoral but also economic outcomes. Additionally, in the last chapter, the exogenous variation induced by the electoral reform in Portugal is used to disentangle two important determinants of voting behavior.

The empirical analyses rely on a total of three different methodologies: the difference-in-differences (diff-in-diff) method, the regression discontinuity design (RDD) and the difference-in-discontinuities (diff-in-disc) design. The latter being a combination of the previous two. In general, the diff-in-diff methodology compares the period before and after a reform in identifying its causal effects. Given a treatment year, the diff-in-diff estimator calculates the difference in the change from the pre- to the post-treatment period between affected and unaffected individuals to assess the impact of a reform. The RDD in turn, relies on a discontinuity in treatment assignment to identify the treatment effect. By focusing on locally randomized samples of individuals differing only on treatment status, the RDD estimator measures the causal treatment effects. Finally, the diff-in-disc effectively combines diff-in-diff and RDD by relying on both changes from the pre- to the post-treatment period and a discontinuous treatment assignment. All following chapters provide comprehensive information on the empirical methods employed as well as the precise mechanics and identification strategies intrinsic to the different contexts under study. Furthermore, the assumptions underlying each of the methodologies are discussed and tested in order to establish the internal validity of the different approaches.

The first reform under analysis, in Chapters 2 and 3, is the *Kommunalwahlreform* dating back to 1999 that changed the rules governing local elections in the German state of Hesse as of 2001. Among the different elements of the reform, the abolishment of a five percent electoral threshold for the municipal council provides the core for the identification strategy. Until 2001 an explicit five percent electoral threshold was in place preventing parties that failed to overcome the threshold from gaining a seat in the municipal council. Since 2001, every party can gain a seat in the local council provided they receive the sufficient amount of votes. Given the absence of an obvious control group against which the abolishment of the electoral threshold can be assessed, the empirical analysis makes use of a state law mapping population size to council size in order to exploit heterogeneity in the intensity of treatment. The abolishment of the explicit electoral threshold had, on average, stronger effects on municipalities with larger council sizes since implicit electoral thresholds are inversely correlated to council size. Relying on a diff-in-disc design, the state law mapping population to council sizes, and a comprehensive dataset including all 426 Hessian municipalities for the period 1989-2011, the results capture the causal effect of the reform on the outcomes under study.

The abolishment of a five percent electoral threshold is likely to benefit smaller parties perhaps affecting the prevailing balance of powers among political representatives and threaten the dominance of the more established parties. Chapter 2 investigates whether the presumed increase in electoral competition coming from the non-existence of an explicit threshold led to a counteracting reaction by the political forces in favor of the status quo. In particular, the question is whether established parties change political institutions to the disadvantage of new political actors if the latter's electoral prospects improve. In the

face of the electoral reform, local politicians from the large mainstream parties could adjust municipal political institutions in such a way as to counteract unwanted effects of the reform. One such adjustment could be to reduce the size of the local council since a reduction in council size raises the implicit electoral threshold and thus disadvantages smaller parties. Indeed, municipalities where the electoral competitiveness of smaller parties improved more after the reform saw a larger reduction in council size. Hence, established parties appear to erect entry barriers by adjusting political institutions once new political actors become viable electoral alternatives.

Chapter 3 in turn, studies the impact of the reform on the following political outcomes: seat and vote shares, council fragmentation and turnout. *Ceteris paribus*, the abolishment of an explicit electoral threshold mechanically increases the seat share of smaller parties for which the threshold used to be binding. Parties that generally enter the council would therefore experience a reduction in the seat shares. Moreover, with more parties entering the municipal council, council fragmentation is likely to increase. However, all else must not be equal, as also indicated by the evidence in Chapter 2, and different responses to the abolishment of the explicit electoral threshold are conceivable. Chapter 3 discusses different scenarios and identifies the actual impact of the electoral reform on the different political outcomes. Results are interpreted in light of the evidence in Chapter 2 with the less than proportional increase in the seat share of small parties in relation to the gain in vote share as indicative of the successful manipulation of council sizes by the more established parties. In the long-run though, the success of smaller parties is apparent as is an increase in council fragmentation. Very weak evidence also suggests a slight increase in turnout, which was one of the objectives of the reform under study.

The second reform, explored in Chapters 4 and 5 of this book, involves a revision of the Portuguese Constitution introducing term limits at the municipal level. Law no. 49/2005, from August 29th 2005, establishing term limits for local officeholders was approved in Parliament on July 25th 2005 and entered into force on January 1st 2006. According to the law, mayoral positions at the municipal level are limited to a tenure of three consecutive terms. Upon entering into force it established a stand-by period of one election, the 2009 local elections, in which all incumbent mayors were allowed to seek re-election. Term limits were thus only first binding in the 2013 local elections for all mayors serving their at least third consecutive term in office.

Term limits introduce an exogenous variation in eligibility for office that may have an effect on short term electoral incentives. *Lame ducks*, or non-eligible officeholders, have arguably different motivations and abide by different incentives in their last term in office vis-a-vis re-eligible incumbents. The theoretical literature on the impact of term limits on fiscal policy is divided into two different types of models dealing either with problems of moral hazard or adverse selection. On the one hand, elections may act as a disciplining device that keeps officeholders from opportunistic behavior but ceases to have an effect on term-limited incumbents. These do not face re-election incentives and deviate from the optimal policy choice (Barro, 1973). On the other hand, elections may distort incentives instead, leading politicians to pursue distortive fiscal policies in an attempt to be re-elected. Term limits can thus lead officeholders to pursue more truthful fiscal policies by eliminating

re-election concerns and reducing the value of holding office (Smart and Sturm, 2013). As a result, the electoral process would be more transparent enabling voters to better select their representatives.

Empirical evidence on the impact of term limits on fiscal policy is sparse and almost limited to the U.S. experience. With one exception, List and Sturm (2006), who provide evidence of distortive policy choices by re-eligible incumbents, most of the empirical evidence is interpreted bearing in mind the disciplining role of elections and opportunistic behavior of term limited incumbents. Chapter 4 contributes to this literature by exploring how the exogenous variation in eligibility for re-election affects fiscal policy choices at the Portuguese local level. Relying on electoral and fiscal outcomes for the past three complete electoral terms, the empirical analysis relies on a diff-in-diff quasi-experimental approach to estimate how fiscal policy differs on average between re-eligible and term-limited incumbents. Results indicate that rather than engaging in opportunistic behavior, lame ducks pursue more conservative fiscal policies. Term limited officeholders choose lower property tax rates and reduced levels of current expenditure relative to re-eligible incumbents. Heterogeneous effects further suggest that ineligible mayors behave more truthfully and do not engage in political business cycles, challenging previous results in the literature.

Finally, Chapter 5 relies on the reform introducing mayoral term limits as a natural experiment that creates exogenous variation on the incumbency status of officeholders while holding the incumbency status of the party constant to provide one of the first causal estimates of both the personal and the partisan incumbency advantages. For the purpose it uses data on six local elections, taking place during the last 20 years in 278 municipalities in Portugal, and a new methodology that combines two quasi-experimental methods, the RDD and diff-in-disc design. Estimates from the two models are inputs in a system of equations relating personal to partisan advantage in the spirit of Fowler and Hall (2014). Results provide evidence of a significant personal effect and an insignificant partisan incumbency advantage in a context of a proportional representation system for the chief-executive position in a municipality. Hence, to the extent that incumbency affects voting behavior, candidates appear to have more political power than the parties they represent.

Taken together, this book offers a study on the effect of political institutions on economic and political incentives. The principal-agent nature of representative politics provides the background for the mapping of rules to policy outcomes resulting from the interplay between differently motivated groups in the society. All in all, results are a robust evidence of the impact of institutions on several aspects of the local political and economic environment. Moreover, by providing supportive evidence of the endogeneity of institutions this book reinforces that taking institutional contexts for granted when analyzing different societal questions may produce biased conclusions not suitable of causal interpretation. Instead, relying on institutional changes introducing exogenous variation and appropriate quasi-experimental empirical methodologies constitutes a better approach to identifying causal relations between rules and outcomes.

Chapter 2

Electoral Competition and Endogenous Political Institutions: Quasi-experimental Evidence from Germany*

2.1 Introduction

Even if all constitutional authority is supposed to derive from the will of the people, elites wield significant influence over the political process. In many countries, it is the political elite that decides how electoral districts are shaped, how votes are translated into seats, and how many seats the legislature has. Since the rules that govern the constitutional life of a country, its political institutions, can be reshaped by those who currently hold the reigns of power, it seems plausible that they would adjust these rules in an ad hoc fashion to maintain their position at the top of the political order. The political elite, in short, may change institutions in such a way as to erect additional barriers to entry for new political actors once they become a credible electoral threat (Doron and Maor, 1991).

Anecdotal evidence indeed suggests that mainstream parties adopt measures to disadvantage new political actors if the latter become too successful. After the success of Ross Perot's Campaign in the 1992 U.S. presidential race, for example, third party candidates were prevented by the Commission on Presidential Debates – which is controlled by both the Republican and Democratic parties – from participating in future presidential debates, a measure that likely diminished the electoral prospects of non-mainstream parties.

While such anecdotal evidence can be found easily, the question is whether it is indicative of a general phenomenon. This paper is one of the first to offer quasi-experimental evidence on whether established parties adjust political institutions to disadvantage new political actors. To do so, we make use of a natural experiment in the German State of Hesse: an electoral reform for local elections passed by the Hessian state parliament in 1999 and implemented in 2001 (when the first election after the reform was held). The professed purpose of the reform was to increase the degree of political competition at the

*This chapter originates from joint work with Thushyanthan Baskaran published in *Journal of Economic Behavior and Organization*, Volume 122, February 2016, Pages 43-61.

local level. Its important aspects were (i) the abolishment of an explicit electoral threshold, the so called “five-percent hurdle” and (ii) the introduction of a new voting system called *Kumulieren und Panaschieren* (KUP), which entailed in particular a switch from closed to open lists.

Before the reform political competition, both between and within parties, was limited in various ways. Between-party competition was limited because the electoral threshold prevented many smaller parties from receiving seats to which they were entitled to given their vote share. Furthermore, voters might have been reluctant to vote for smaller parties in the first place given the non-negligible chance that their vote would be “wasted” if their preferred party did not overcome the electoral threshold. Within-party competition was limited because voters could vote only for closed party lists, with the position of candidates on the lists essentially determined by the local party leadership. Consequently, candidates not closely affiliated with the local party leadership (for simplicity we refer to such candidates as party rebels in the following) had few chances to gain a seat in the council.

After the reform, both smaller parties and party rebels became a more viable electoral alternative. First, smaller parties required, in general, a substantially lower vote share than five percent to gain their first council seat and thereby legislative representation. Second, party rebels, even if they had been placed at a low-ranked position on the party list, could enter the council if they received sufficient personal votes to overcome the party leaderships’ pre-ordering.

The question we ask in this paper is how the established political parties reacted to this exogenous increase in the electoral competitiveness of smaller parties and party rebels: did they change the prevailing local political institutions to put these new political actors back at a disadvantage? We show that at least one municipal political institution was indeed adjusted after the reform: the size of the local council. By reducing the number of council seats, which can be done through a two-third council majority, the mainstream parties raised implicit thresholds – the minimum vote share that a party has to gain to receive at least one seat – and thus made it harder for smaller parties to gain their first council seat. Furthermore, party leaders may have also used council size reductions to prevent party rebels from receiving a seat following the switch to KUP. However, as we discuss below in more detail, reductions in council size are arguably less effective in preventing party rebels from entering the council than in preventing smaller parties. We therefore focus in the following on the link between the abolishment of the explicit threshold, council size reductions, and the competitiveness of smaller parties.

Since all Hessian municipalities were subject to the abolishment of the explicit electoral threshold, there is no obvious control group against which changes in council size from the pre- to the post-treatment period could be evaluated. Our identification strategy to uncover the causal effect of the abolishment on council size relies therefore on differences in the intensity of treatment. Municipalities that had a larger council were affected more strongly by the abolishment of the explicit electoral threshold than those with smaller councils because municipalities with smaller councils have higher implicit thresholds to begin with. That is, the competitiveness of smaller parties is higher the larger the council

after the reform. Therefore, the mainstream parties had a stronger incentive to reduce the size of the council if the current council size in their municipality was large. They are, however, not completely free in doing so. A state law relates municipal population sizes discontinuously to minimum and maximum council sizes. We exploit the discontinuous nature of the link between population size and council size to implement a variant of the regression discontinuity design (RDD), the difference-in-discontinuities design (diff-in-disc).

This paper contributes to the literature on political institutions, and specifically to the literature on their determinants. A strand of this literature assumes a historical perspective and analyzes how secular changes in political regimes, most notably the shift to popular democracy, were determined by conscious decisions of the elite (Acemoglu and Robinson, 2000; Aghion et al., 2004; Acemoglu and Robinson, 2006). On the other hand, the literature on the determinants of contemporaneous and arguably less decisive adjustments of political institutions within generally democratic societies is scarce. One of the few studies are Hayo and Voigt (2010, 2013) who analyze with cross-country regressions why countries witness constitutional change. They find that political factors, e.g. whether there is an internal armed conflict, determine how countries transition from a parliamentary to a presidential form of government (or vice versa). Another study is Ticchi and Vindigni (2010) who find with cross-section regressions that countries with higher income inequality are more likely to have majoritarian electoral rules.

Even rarer are studies on how political institutions are adjusted by established elites to counter threats from new political actors. Drometer and Rincke (2014) find that in the U.S., states which were affected more strongly by the Voting Rights Act of 1965 tightened ballot access restrictions to hinder new entrants into the political market.¹ Similarly, Trebbi et al. (2008) show that U.S. cities systematically changed electoral rules after the adoption of the Voting Rights Act to limit minority representation.² Our paper primarily contributes to this small literature and is the first that looks at a setting other than the U.S.. It documents that political institutions are indeed endogenous and thereby suggests that evidence from studies that take political institutions as exogenous is questionable (Trebbi et al., 2008). From a policy perspective, our results also suggest that adjustments of the wider political system by established elites should be taken into account when designing reforms aimed at increasing political competition.

Our paper is furthermore relevant for the literature that studies the consequences of legislature size and legislative fragmentation for fiscal outcomes. Relying on common pool theories, Gilligan and Matsusaka (2001), for example, find that larger U.S. state legislatures spend more. Similar evidence for U.S. local governments is offered by Baqir (2001). The related literature on legislative fragmentation and fiscal outcomes finds that increased fragmentation leads to higher spending and deficits (Borge, 2005). However, most studies in this literature operate under the assumption that the size of the legislature is essentially

¹ Lee (2013) shows that U.S. House incumbents respond to the threat of third party candidates within their constituency by adjusting how they vote in roll call votes. However, he does not study adjustments of political institutions.

²Doron and Maor (1991) and Bellettini et al. (2014) offer relevant theoretical analyses, showing for example how incumbent politicians can use bureaucratic red tape to disadvantage new entrants.

exogenous. We show that legislature size responds to political developments. Our results hence vindicate recent contributions on the effect of council size on fiscal outcomes that attempt to exploit quasi-experimental variation in council size for identification (Egger and Koethenbueger, 2010; Pettersson-Lidbom, 2012).

2.2 Institutional Details

2.2.1 Political Setting

The setting for our analysis is the German federal State of Hesse. This state has about six million inhabitants who live in 426 municipalities. Municipal population sizes vary: there are, on the one hand, municipalities with less than 1000 inhabitants and, on the other hand, the city of Frankfurt with more than 600,000 inhabitants. Figure 2.1 shows a map of Hesse and indicates the average population sizes of municipalities during the sample period.

Inhabitants in every municipality elect a local council in elections held at the same date throughout the state. The council is the most important political institution in a Hessian municipality. It decides, *inter alia*, on various municipal taxes, user fees, and on the provision of municipal public goods and services. The other important political institutions is the mayor. The mayor used to be appointed by the council (council-manager system), but as of 1993 she is directly elected by municipal inhabitants. Yet even after the reform, Hessian municipalities do not employ a full fledged mayor-council system as the competencies of the mayor remain limited (Hessami, 2014).³ We discuss the implications of the reform in the electoral rule for mayoral elections for our empirical strategy in Section 2.6.1.

Several parties contest the local council elections. First, the center-right CDU and the center-left SPD. These two parties typically receive 30 percent or more of the votes in both national and state elections. Second, the Green Party and the FDP. The characteristic feature of the Green Party is its emphasis on environmental issues. It is considered to be left of center regarding economic and liberal regarding social issues (e.g. immigration) and tends to form coalitions with the SPD. The FDP, on the other hand, emphasizes free market economics. It is considered to be right of center with respect to economic issues and liberal with respect to social issues. It tends to form coalitions with the CDU. The

³Specifically, the switch in how mayors were chosen in Hesse was part of a broader trend in the 1990s leading several German States to switch from a council-manager (*Norddeutsche Ratsverfassung*) to a mayor-council system (*Süddeutsche Ratsverfassung*). However, unlike other German States, Hesse made only a partial switch to the mayor-council system. The *Süddeutsche Ratsverfassung* not only requires that mayors are directly elected by the voters, it also gives the mayor a strong political position within the municipality, making him the head of the administration and providing him with veto rights against council decisions. Hesse, however, only introduced direct elections but left the political authority of the mayor otherwise unchanged. The Hessian local constitution is therefore labeled *Unechte Magistratsverfassung* rather than *Süddeutsche Ratsverfassung*.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

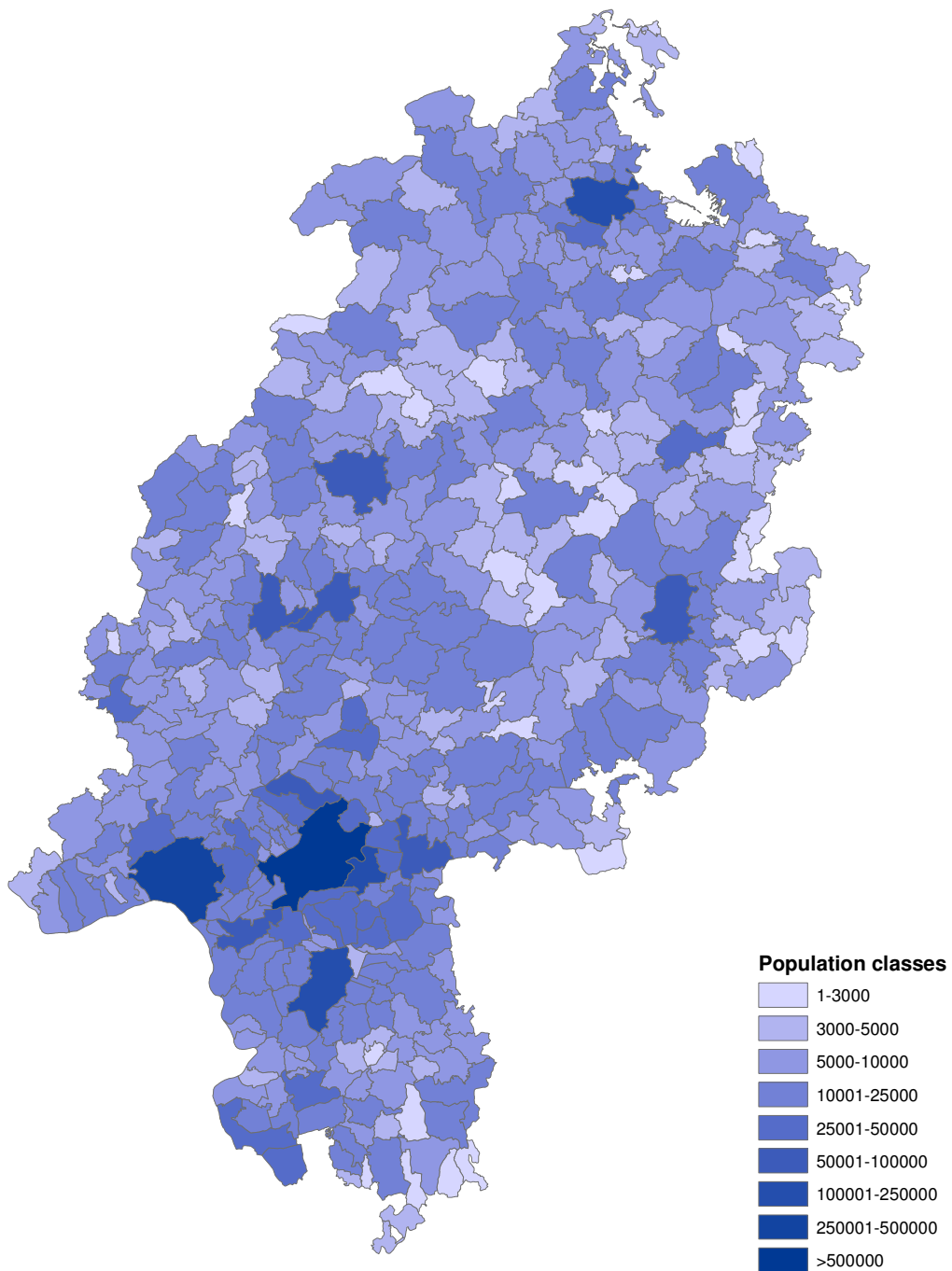


Fig 2.1: Average Population Size in Hesse for Sample Period.

Green Party and the FDP used to receive up to ten percent of the votes during the sample period.

In addition to the four large national mainstream parties, there are a number of smaller parties that contest local elections. First, small national parties which can be either centrist, far-left, or far-right with respect to economic and social issues. Second, municipal specific free voters' associations (*Wahlvereinigungen*) which often contest local elections by fielding a list of candidates. These free voters' associations can be very influential, particularly in smaller municipalities, and often receive more than 20% of the seats. Thus, the free voters' associations are not always a "small" party as such. However, they typically lack formal party structures and are largely a loose collection of citizens who agree on certain issues that are specific to their municipality.

2.2.2 Electoral Rules and the *Kommunalwahlreform*

The rules governing local council elections in Hesse differed before and after 2001. Until 2001, local elections took place every four years. Each of the parties would field one closed list and citizens were allowed to cast one vote for their favored party list. Parties would then be allocated seats in the council according to the Hare-Niemeyer procedure. All candidates placed sufficiently high on their respective lists would receive a seat. However, even if a party had a sufficiently large vote share to gain one or more seats in the council, it would not receive a seat if its vote share was below five percent, i.e. there was an explicit electoral threshold.

In 1999, the state parliament passed a law that fundamentally changed the rules that governed local elections from 2001 onward (*Kommunalwahlreform*). First, the length of the legislative period was extended from four to five years. Second, the law introduced a new voting system called *Kumulieren und Panaschieren*. In this system, voters may cast as many votes as there are seats available in the council. Up to three votes can be cumulated and given to individual candidates. Alternatively, voters are allowed to give all their votes to a certain party list, but they can also drop individual candidates from the list. Seats to which a party is entitled according to the number of total list votes are allocated to candidates according to their number of personal votes. That is, the original ranking of candidates on the party lists can be changed by voters through personal votes. While KUP consists of many individual aspects, its main feature was arguably this switch from closed to open party lists. Third, the five percent electoral threshold was abolished. Parties could enter the parliament if they surpassed the implicit threshold, i.e. if they had sufficient votes to gain at least one seat.⁴

⁴It may be surprising that the state government implemented the *Kommunalwahlreform* as it ostensibly disadvantaged the established parties, including those forming the state government. The main reason for the reform was that voter turnout at local elections was continuously declining in Hesse during the 1990s, threatening the legitimacy of the political process at this tier of government. The *Kommunalwahlreform* was intended to reverse this trend by giving voters more choices and thus motivating them to participate in the local elections.

The abolishment of the five percent threshold had decidedly heterogeneous effects across municipalities. It affected municipalities that had large municipal councils more strongly because they have smaller implicit thresholds. First, the abolishment made it easier for smaller parties to enter the council. In a council with e.g. 100 seats a vote share of around 0.5 percent would be sufficient for a party to gain its first council seat if there was no five percent threshold.⁵ But if the council has for example only 20 seats, a party has to receive around 2 to 3 percent of the votes to get a seat even if there is no explicit five percent threshold. Therefore, the abolishment of the five percent threshold mattered less for municipalities with smaller councils. This heterogeneity in treatment intensity forms the core of our identification strategy below.

In addition to the effects of the electoral threshold's abolishment, the effects of the switch to KUP may also vary at the population cutoffs. In particular, party rebels may be more likely to gain a seat in municipalities with larger councils, i.e. smaller implicit thresholds, because they require fewer personal votes to change the parties' pre-ordering such that they receive a council seat. However, while this is a possibility, it seems unlikely that the leadership of the mainstream parties would adjust council size to keep out party rebels. The link between council size and the propensity of party rebels entering the council is tenuous. For example, highly popular party rebels will end up very high on the final list and thus gain a seat irrespective of the size of the council. Adjustments in council size may thus only help to keep out party rebels with limited popularity in their municipality. Furthermore, council size reductions are an imprecise instrument to keep party rebels out. As the final ranking is uncertain due to KUP, a reduction in council size may end up preventing a candidate aligned with the party leadership from entering the council rather than a party rebel. For these reasons, it is likely that council size reductions are primarily a response to the abolishment of the explicit electoral threshold and intended to keep out smaller parties.

Finally, the lengthening of the legislative period had in all likelihood relatively limited effects across municipalities. That is, it is plausible that any change in local officials' strategic incentives due to the lengthening of the legislative period did not vary discontinuously with council size.

With the reform of 1999, the state legislature intended to improve the competitiveness of smaller parties and new candidates and thereby foster political competition. Indeed, given fixed voting patterns and fixed council sizes, having no legal electoral threshold should mechanically increase the seat shares of smaller and decrease the seat share of larger parties. In addition, voting patterns must not remain fixed (Moser and Scheiner, 2004). They will likely change such as to increase the seat shares of small parties even further. Prior to the abolishment, supporters of small parties might have chosen to vote for one of the more established parties if there was a non-negligible chance that their preferred small party would fail to overcome the five percent threshold. Once the threshold was

⁵The actual value of the implicit threshold for a given party is endogenous and depends inter alia on the vote shares of all other parties. Typically, a vote share that is sufficiently large for half a seat entitles a party to a full seat in the council. See <http://www.wahlrecht.de/kommunal/hessen.html>.

abolished, voting for their preferred small party became more worthwhile for this subset of the electorate (Perea, 2002). In short, the abolishment had in all likelihood, in addition to any mechanical effects, what is called psychological effects in the political science literature, making the abolishment even more beneficial for smaller parties (Duverger, 1954; Fiva and Folke, 2014).⁶

The switch to KUP, on the other hand, may have reduced incentives to vote for smaller parties because of the increased within-party competition in the established parties. If voters can vote for candidates that are not closely aligned with the local party elite, they may have fewer incentives to switch to smaller parties. However, it is unclear to what extent increased within-party competition prevents a switch of voters to smaller parties. Even candidates that are not closely aligned with the party leadership and thus receive a low rank on the list still have to be voted onto the list by party members in primaries. Thus, there is a limit to the extent a party rebel can move away from the core positions of the party if she wants a spot on the list. Voters who hold unorthodox political views may thus prefer to vote for a small party that has become politically viable due to the reform than to vote for a candidate from the established parties even if she is not part of the traditional party elite.⁷

2.2.3 Council Size

In each legislative period, municipalities can choose the size of their council for the next legislative period within certain ranges depending on population cutoffs that are defined in the Hessian local government code (HGO).⁸ These cutoffs were not affected by the electoral reform in 1999. Table 2.1 notes the minimum and maximum council sizes for each population bracket according to the HGO.⁹ Municipalities must choose a council size within the bracket that is relevant for them. Hence, a municipality with e.g. 1-3000

⁶In general, voters may vote for either instrumental or expressive reasons (Fiorina, 1976). The above discussion implicitly assumes that voters vote at least in part for instrumental reasons. If voters voted for expressive reasons only, the abolishment of the electoral threshold would arguably have no psychological effects since whether a particular party is represented in the local council should not matter to expressive voters. However, at the local level at least, it is plausible that a large share of voters engage in instrumental voting. Thus, the abolishment of the electoral threshold likely motivated many supporters of smaller parties to actually vote for these parties. The switch to KUP, on the other hand, may have had substantive effects also on expressive voters. In particular, since parties likely became less important with the introduction of open list, the benefits from voting for expressive reasons may decline. It is, however, unclear how such changes to the incentives of expressive voters will affect electoral outcomes.

⁷We study the consequences of the reform for vote and seat outcomes in a companion paper (Baskaran and Lopes da Fonseca, 2014). The results suggest that smaller parties indeed benefited from the reform (see Chapter 3).

⁸See Art. 38 of Hessian Local Government Code (*Hessische Gemeindeordnung*, HGO)

⁹The law states that council size brackets are determined by the latest available population data when the date for next local election is fixed. This population data is not the same as the annual data published by the state statistical office. For the elections of 2006 and 2011, we obtained the relevant data from the homepage of the statistical office. For the previous elections, we collected the data by hand from various issues of the Hessian government gazette.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

inhabitants must have at least 11 and at most 15 council seats while a municipality with 3001-5000 inhabitants must have at least 15 and at most 23 council seats

Table 2.1: Population Cutoffs and the Number of Seats.

Population	Council size	Observations
1-3000	11-15	219
3001-5000	15-23	473
5001-10,000	23-31	874
10,001-25,000	31-37	779
25,001-50,000	37-45	137
50,001-100,000	45-59	42
100,001-250,000	59-71	18
250,001-500,000	71-81	6
500,001-1,000,000	81-93	6
> 1,000,000	93-105	-

Notes: This table collects the population cutoffs at which municipalities may increase their council size. Municipalities must choose a council size within the respective bracket. For example, municipalities between 5001 and 10,000 inhabitants may have up to 31 council seats but must have at least 23 seats.

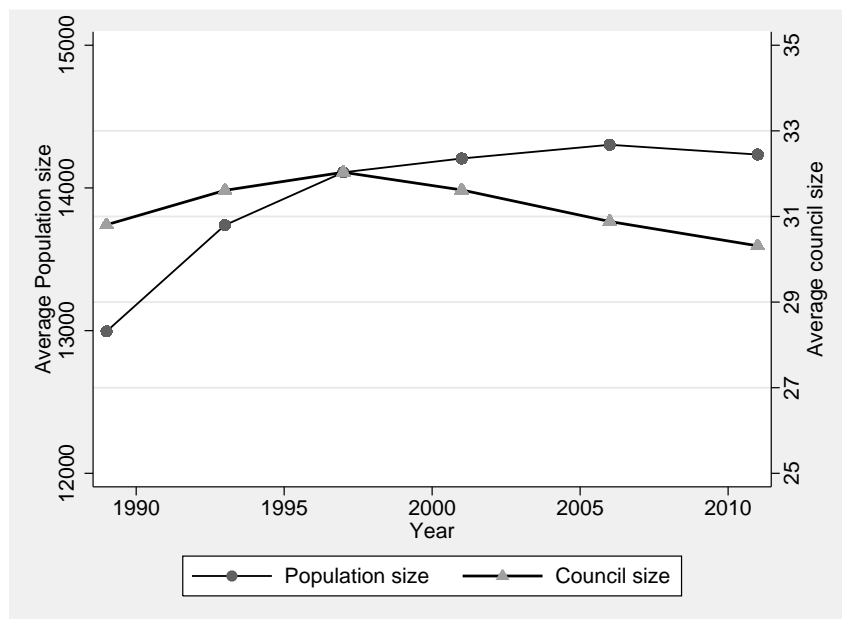


Fig 2.2: Average Council and Population size in Hesse. This figure depicts average council and population sizes in Hessian municipalities during the sample period.

Up until one year before the next local election, the current council is allowed to change the size of the next council with a two-third majority as long as final council size remains within the ranges listed in Table 2.1. One way through which the established parties can

therefore counteract the abolishment of the five percent threshold is to lower the size of their council within the limits set by the HGO and thereby raise implicit thresholds for the next council. Figure 2.2 plots the unweighted average council size in Hessian municipalities over time. There is a clear break in the trend for council size in 1997. Before 1997, average council sizes increase continuously. The reason for this continuous increase is presumably that Hessian municipalities were witnessing positive population growth, as depicted in the second plot in the graph, so that several municipalities were crossing the council size cutoffs from below and therefore were allowed to adopt larger councils.¹⁰ From 1997 onward, however, council size has been continuously declining even as population size has been continuing to grow.¹¹

The coinciding of the reform with the start of the decline in council size is suggestive. However, it cannot be interpreted causally. For example, it is possible that even as average population size grew after 1997, those municipalities that reduced their council size witnessed shrinking populations. One way to control for the effect of population growth on changes in council size is to plot the development of average council size within each of the population size brackets defined by the HGO. Figure 2.3 shows that before 2001, essentially all municipalities chose the largest possible council size for their bracket, either because local politicians wanted to maximize their chances of receiving a seat in the council or because the wording in the relevant article of the local government code suggests the highest possible council size as the default¹², or both.¹³ There were also virtually no adjustments from 1989 to 1997, which in turn suggests that the increasing average council size in Hesse as documented in Figure 2.2 was indeed due to increasing population size. The only exception is a small decline in average council size from 1993 to 1997 in the lowest bracket. After 2001, however, average council sizes began to decline for all but the three highest brackets (into which altogether only five large cities fall). For example, average council sizes in municipalities that have between 50,001 and 100,000 inhabitants is about 3.5 seats smaller in 2011 than in 1997.

¹⁰Apart from “natural” demographic developments, the increase in population size of Hessian municipalities in the early nineties might be due to immigration from Eastern Germany following the German Reunification.

¹¹A plot for the weighted average council size, with population size as weights, shows the same pattern.

¹²Paragraph 1 of the relevant article states that the number of representatives in a municipalities should be the maximum for each bracket as listed in Table 2.1. Paragraph 2 states that municipalities *may* choose the council size listed for the previous bracket, or some uneven number in between.

¹³Two municipalities had larger council sizes than permissible given their population size (one had 4999 inhabitants and a council size of 31 and the other 9754 and a council size of 37) in 1989. We drop these two observations from the sample. While we have no definite explanation, we suspect that these two municipalities made use of an exception defined in the electoral law that allows municipalities that crossed any of the thresholds from above to keep the council size intended for municipalities in the next higher threshold for another legislative period.

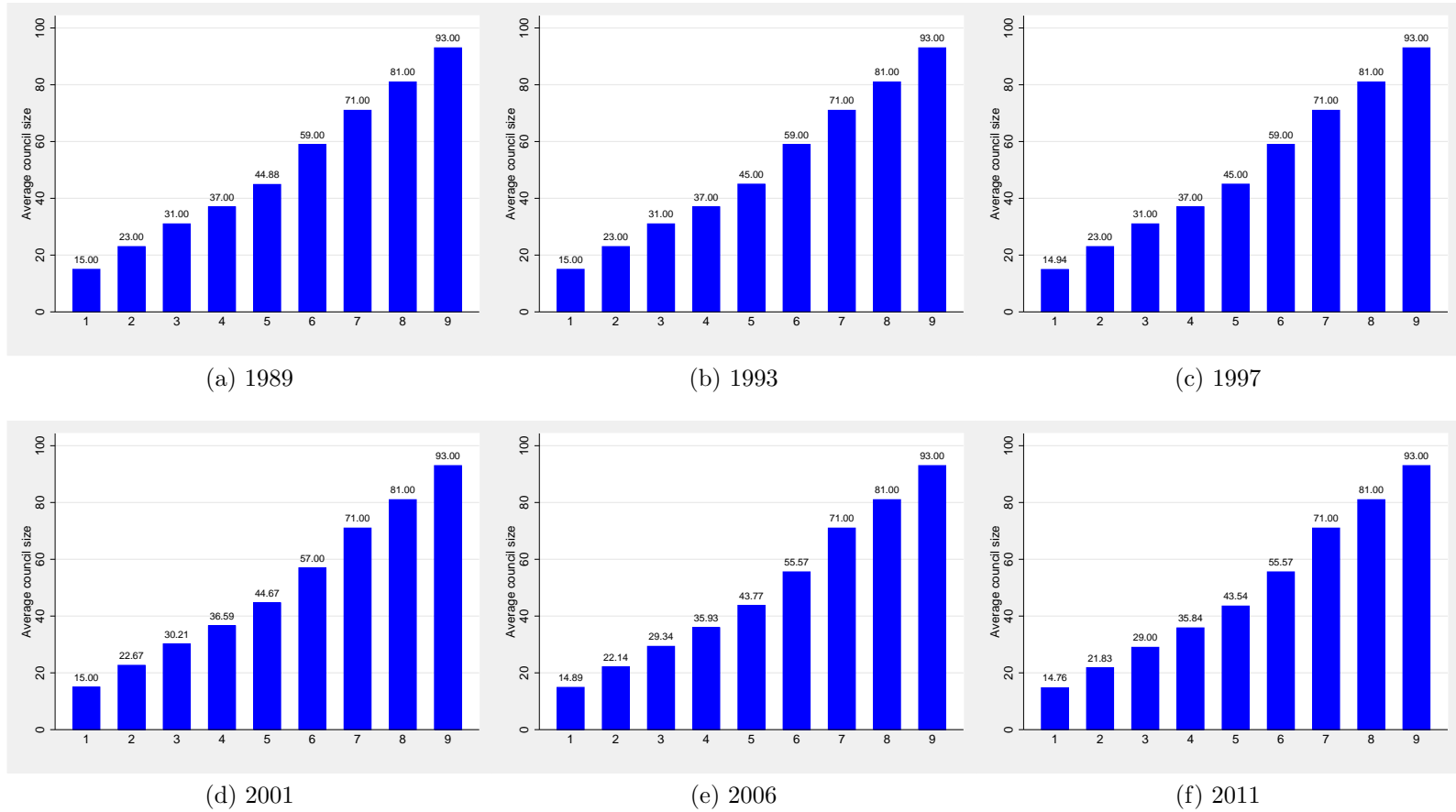


Fig 2.3: Average Council Size in Different Population Brackets. This figure shows the average council size of municipalities in population brackets 1-3000 (1), 3001-5000 (2), 5001-10,000 (3), 10,001-25,000 (4), 25,001-50,000 (5), 50,001-100,000 (6), 100,001-250,000 (7), 250,001-500,000 (8), 500,001-1,000,000 (9) at the beginning of each legislative period between 1989-2011.

Yet, it would still be premature to relate the decline in council sizes causally to the electoral reform. There might have been unobserved trends in Hesse that have caused municipalities to reduce their council size after 1997 but which were not systematically related to the electoral reform of 1999. For example, Hessian municipalities may have encountered fiscal difficulties and cut council sizes to reduce costs. Alternatively, the belief that smaller councils are more efficient might have gained traction throughout the state for some unobserved reason after 1997. In short, the presence of unobserved trends cannot be ruled out. Since all Hessian municipalities were affected by the reform, there is also no obvious control group that would offer a counterfactual with which the causal effect of the reform on council size could be easily identified. The following section is therefore concerned with identifying treatment and control municipalities within a quasi-experimental framework.

2.3 Empirical Strategy

2.3.1 Difference-in-discontinuities Design

To causally relate the reform of the electoral law to reductions in council size, we focus on the abolishment of the five percent threshold. In particular, we rely on the fact that the abolishment did not affect all Hessian municipalities equally. Municipalities that were affected less offer a reasonable counterfactual for those that were affected more strongly, as long as it can be ensured that both sets of municipalities were subject to similar trends. As indicated by our previous discussion on implicit thresholds, municipalities with larger councils were exposed to a stronger treatment because their implicit thresholds are lower and thus smaller parties could enter the council more easily. Therefore, after 2001 municipalities with larger councils should have reduced their council size more than those with small councils if the local political elite uses adjustments in council size to limit competition from smaller parties.

In principle, we could compare changes in council size in municipalities that had in the pre-treatment period large councils with changes in municipalities that had small councils to uncover how local politicians react to an increase in the competitiveness of small parties. There are, however, two problems with this approach. The first is that council size is under the direct control of municipalities and therefore an endogenous variable. It is hence questionable whether council size in the pre-treatment period can be used as an exogenous explanatory variable. For example, municipalities where local politicians dislike political competition might have chosen small councils already in the pre-treatment period (even though this does not seem to be the case given the evidence in Figure 2.3), making it more difficult to reduce council size further in the post-treatment period. Another problem is that municipalities with different pre-treatment council sizes differed in other characteristics as well, most clearly in their population size. It is hence possible that they were subject to different trends, either with respect to their council size or with respect to the political competitiveness of smaller parties. Municipalities that had larger councils

in the pre-treatment period hence might have lowered their council size more strongly for reasons unrelated to the reform.

Given these difficulties when using pre-treatment council size to determine treatment intensity, we adopt the following research design. We rely on the fact that the probability of a larger council increases discontinuously at the population cutoffs as defined in Table 2.1. Figure 2.3 shows that mean council size is indeed increasing between the different brackets. Typically, municipalities choose the largest possible council size in the pre-reform period. Even though the relationship between population size and council size is fuzzy at the cutoffs, since the council size that municipalities to the left and the right of each cutoff are allowed to overlap to some extent, it is clear that there is a positive and discontinuous relationship between both variables. Therefore, the treatment intensity of abolishing the five percent threshold will, in expectation, increase discontinuously at the population cutoffs. More specifically, a given municipality with e.g. 3000 inhabitants will be affected less by the abolishment of the electoral threshold than municipalities with 3001 inhabitants because the former will choose on average smaller councils and thus have larger implicit thresholds in both the pre- and the post-treatment period.

Since treatment intensity increases in council size, and the probability of a larger council increases discontinuously in population size, we can implement a fuzzy RDD (Lee and Lemieux, 2010) using pre- and post-treatment data to identify the causal effect of the abolishment of the electoral threshold. The basic idea is to estimate the effects of the population cutoffs on council size for the pre- and post-treatment periods, and then to observe whether there is a difference in the effect of the cutoffs between the two periods. If the effect of the cutoffs has changed between the pre- and the post-treatment period, we can reasonably ascribe the change to the reform of 1999.

In other words, we study whether municipalities that have population sizes barely above either of the cutoffs defined in Table 2.1, and therefore in expectation discontinuously larger councils, reduce their council size more strongly than municipalities with population sizes barely below the cutoffs, and therefore discontinuously smaller councils. Since the RDD relies on local randomization, this design implicitly accounts for differential trends between municipalities that were exposed to stronger and weaker treatments.

More formally, the effect of a stronger treatment at M , the natural log of a given threshold, in period t can be defined as follows:

$$\hat{\gamma}_t^M = \lim_{NLPOP_i \downarrow 0} E[y_{i,t} | NLPOP, t] - \lim_{NLPOP_i \uparrow 0} E[y_{i,t} | NLPOP, t], \quad (2.1)$$

where $y_{i,t}$ is council size in period t and $\hat{\gamma}_t^M$ is the estimate for the treatment effect. $NLPOP$ is the normalized value of the natural log of the relevant population figure such that $NLPOP = LPOP - M$.

The treatment effect $\hat{\gamma}_t^M$ can be obtained with the following general RDD model in a regression framework:

$$y_{i,t}^M = \gamma_t^M D_i + f(NLPOP) + D_i f(NLPOP) + \epsilon_i \text{ if } |NLPOP| < h, \quad (2.2)$$

where D_i is a dummy that is 1 if $NLPOP \geq 0$ and 0 else. $f(NLPOP)$ is a flexible polynomial of normalized population size which is allowed to have different slopes to the

left and right of a normalized cutoff M . This type of RDD model can be estimated by local polynomial regression using different polynomials and bandwidths h .

An estimate for the effect of the abolishment of the electoral threshold on council size is the difference in the treatment effects in the pre- and post-treatment period:

$$\begin{aligned} \hat{\delta}^M \equiv \hat{\gamma}_{t \geq T}^M - \hat{\gamma}_{t < T}^M = & \lim_{NLPOP_i \downarrow 0} E[y_{i,t} | NLPOP, t \geq T] - \lim_{NLPOP_i \uparrow 0} E[y_{i,t} | NLPOP, t \geq T] \\ & - \left(\lim_{NLPOP_i \downarrow 0} E[y_{i,t} | NLPOP, t < T] - \lim_{NLPOP_i \uparrow 0} E[y_{i,t} | NLPOP, t < T] \right), \end{aligned} \quad (2.3)$$

where $t \geq T$ denotes the post- and $t < T$ the pre-treatment period.

The corresponding model in a regression framework is:

$$\begin{aligned} y_{i,t}^M = & f(NLPOP) + D_i(\gamma_{t \geq T}^M + f(NLPOP)) + I_t(\alpha + f(NLPOP)) \\ & + D_i(\delta^M I_t + I_t f(NLPOP)) + \epsilon_{i,t} \text{ if } |NLPOP| < h, \end{aligned} \quad (2.4)$$

where I_t is a dummy indicating the post-treatment period.

This specification is an extension of the standard RDD specified in Equation (2.2) and is labeled diff-in-disc design by Grembi et al. (2015).¹⁴ It allows the control function to vary both to the left and the right of the cutoff M , between the pre- and post-treatment periods, and within treated municipalities in the pre- and post-reform periods. We are interested in the estimate for δ^M which captures the change in the effect of the discontinuity at M between the pre- and the post-treatment periods.

We motivated the RDD and diff-in-disc models above by referring to a single population cutoff M . In our case, there are multiple cutoffs. Rather than analyzing all cutoffs individually, we follow in the baseline regressions the previous literature that uses the RDD methodology with multiple population cutoffs, and normalize all observations such that they are around a single cutoff (Egger and Koethenbueger, 2010). This approach has the advantage of a larger sample size. In addition, the results can be presented more compactly. In robustness tests, however, we also report results for individual cutoffs to show that these results are consistent with our more compact baseline analysis.

We report results for various bandwidths around the normalized cutoff and polynomials of the control function. More specifically, we use the following bandwidths: 0.25, 0.2, 0.15, 0.1, and 0.05.¹⁵ With respect to polynomials of normalized population size, we use up to a cubic specification. To control for common trends in the pre- and post-treatment period, we include legislative term fixed effects (rather than only dummies for the pre-

¹⁴Several studies have adopted a diff-in-disc design recently, for example Asatryan et al. (2015) or Gulino (2014).

¹⁵We experimented with the data driven bandwidth selection procedure by Imbens and Kalyanaraman (2012). However, this procedure was developed for cross-sectional RDD and may result in inappropriate suggestions if the arbitrary initial bandwidth is wrong. In our application, it typically suggested unreasonably large optimal bandwidths. We therefore establish the robustness of the results by reporting estimates for different bandwidths, ranging from relatively large (0.25) to relatively small (0.05).

and post-treatment periods). We also follow the previous literature and include in all regressions municipality fixed effects to improve efficiency and reduce finite sample bias (Hoxby, 2000).¹⁶

The identifying assumptions in the diff-in-disc design are arguably less strict than in the RDD. Notably, we do not require that there is no co-treatment at the cutoffs (Eggers et al. (2015) show that co-treatment can be a major problem in standard RDD with population thresholds). Instead, we only require that the effect of any co-treatments remains constant between the pre- and post-treatment periods.¹⁷ For example, one particular reason for potential co-treatment is that several of the cutoffs in Table 2.1 are relevant for equalization transfers (Baskaran, 2015). However, the stipulations surrounding the intergovernmental transfer scheme were not changed from the pre- to the post-treatment period, and thus the effect of different transfers receipts at the population thresholds on council size should remain constant in both periods.

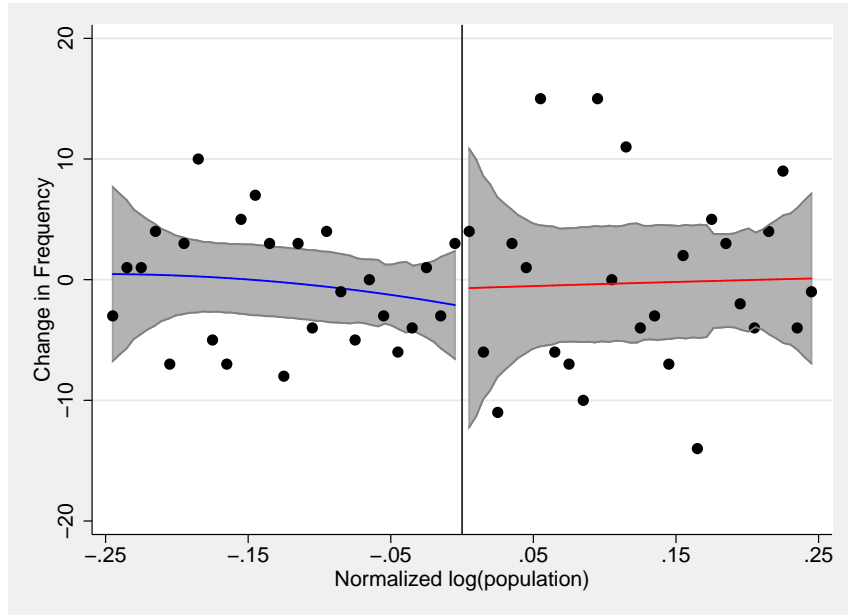


Fig 2.4: Manipulation of the Running Variable. This figure presents a density plot for the diff-in-disc design in the spirit of McCrary plots (McCrary, 2008). We first divide normalized log population size in bins of width 0.01. Then we calculate the change in the total number of observations within each bin from the pre- to the post-treatment period. Finally, we fit local polynomial plots using a bandwidth of 0.05, a degree of 2, and a rectangular kernel to the number of changes within bins. 95% confidence intervals are indicated in gray.

¹⁶Since the diff-in-disc design relies on local randomization, fixed effects are not necessary to obtain consistent estimates. We indeed find that omitting fixed effects does not change the results substantively. The estimates are, however, less significant and less stable across bandwidths, suggesting that including fixed effects reduces their variability. Our preferred specification therefore includes fixed effects and we focus on these results in the following.

¹⁷A different approach to deal with co-treatment when multiple cutoffs are available would be to rely on a standard RDD while implementing various robustness tests that deal with potential co-treatment, e.g. establishing that the results remain robust at those cutoffs where co-treatment can be credibly ruled out or conducting appropriate placebo regressions (Arnold and Freier, 2015).

The other crucial assumption is that the ability or incentives of municipalities to manipulate population size at the cutoff did not change from the pre- to the post-treatment periods. This assumption appears plausible as it is unlikely that municipalities would persistently misrepresent their population sizes only to avoid being forced to change their council sizes. Moreover, given that most municipalities chose the highest possible council size in the pre-treatment period, they were allowed to reduce their council sizes anyway and therefore had no incentives for manipulation. Finally, a McCrary (2008) style density plot¹⁸ reported in Figure 2.4 also fails to indicate that incentives for manipulation changed from the pre- to the post-treatment period at the normalized cutoff.

2.3.2 Difference-in-discontinuities Plots

In addition to regression results, we also present graphical evidence on the treatment effect. We construct RDD plots for the pre- and post-treatment periods and a diff-in-disc plot by first dividing the control function, $NLPOP$, into bins of size 0.01 within a window of 0.25. Then we calculate average council size y within each bin for the pre- and post-treatment period, i.e. $\bar{y}_{b,t}$ with the index $b = 1, \dots, 50$ denoting the bin and $t = 0, 1$ denoting the pre- and the post-treatment period.

For the RDD plots, we smooth the observations with a local polynomial plot of quadratic degree and a bandwidth of 0.25 at both sides of the normalized cutoff, using a rectangular kernel. For the diff-in-disc plot, we first obtain the difference within each bin in the pre- and post-treatment period $\Delta y_b = (\bar{y}_{b,1} - \bar{y}_{b,0})$, then plot the differences in y against $NLPOP$ to the left- and the right of the normalized cutoff, and finally smooth the differences with a local polynomial plot of quadratic degree and a bandwidth of 0.25 at both sides of the normalized cutoff, using a rectangular kernel and the number of observations within each bin as frequency weights.

2.4 Baseline Results

2.4.1 Graphical Evidence

Figure 2.5 presents the graphical evidence. Subfigure (a) shows the effect of the normalized threshold on council size in the pre-treatment period. There is a large positive discontinuity. Municipalities to the left of the normalized threshold have councils that are about

¹⁸Standard McCrary plots are inappropriate for our design as we are interested in the difference of the effect at the normalized population cutoff in the pre- and the post-treatment period. The idea underlying our diff-in-disc variant of the McCrary plot is that if either the ability or the incentives for manipulation changed at the cutoff from the pre- to the post-treatment period, we should observe a discontinuity in the changes in the number of observations close to the cutoff. More specifically, assume that because of the treatment, municipalities systematically start to (mis-) report lower population sizes in order to be able to reduce their council size. Then the increase in observations just below the normalized cutoff from the pre- to the post-treatment period should be significantly higher than the increase in observations just above the threshold.

8.4 seats larger than municipalities to the left. Subfigure (b) shows the effect of the normalized threshold in the post-treatment period. There is still a positive discontinuity, but it is markedly smaller. Average council size to the right of the threshold is only about 4.7 seats higher than to the left. Finally, subfigure (c) presents the diff-in-disc plot, which explicitly plots the change in the effect of the normalized threshold from the pre- to the post-treatment period. There is a negative discontinuity of about 4.1 seats. Hence, municipalities that were exposed to a stronger treatment reduced the number of seats more strongly.

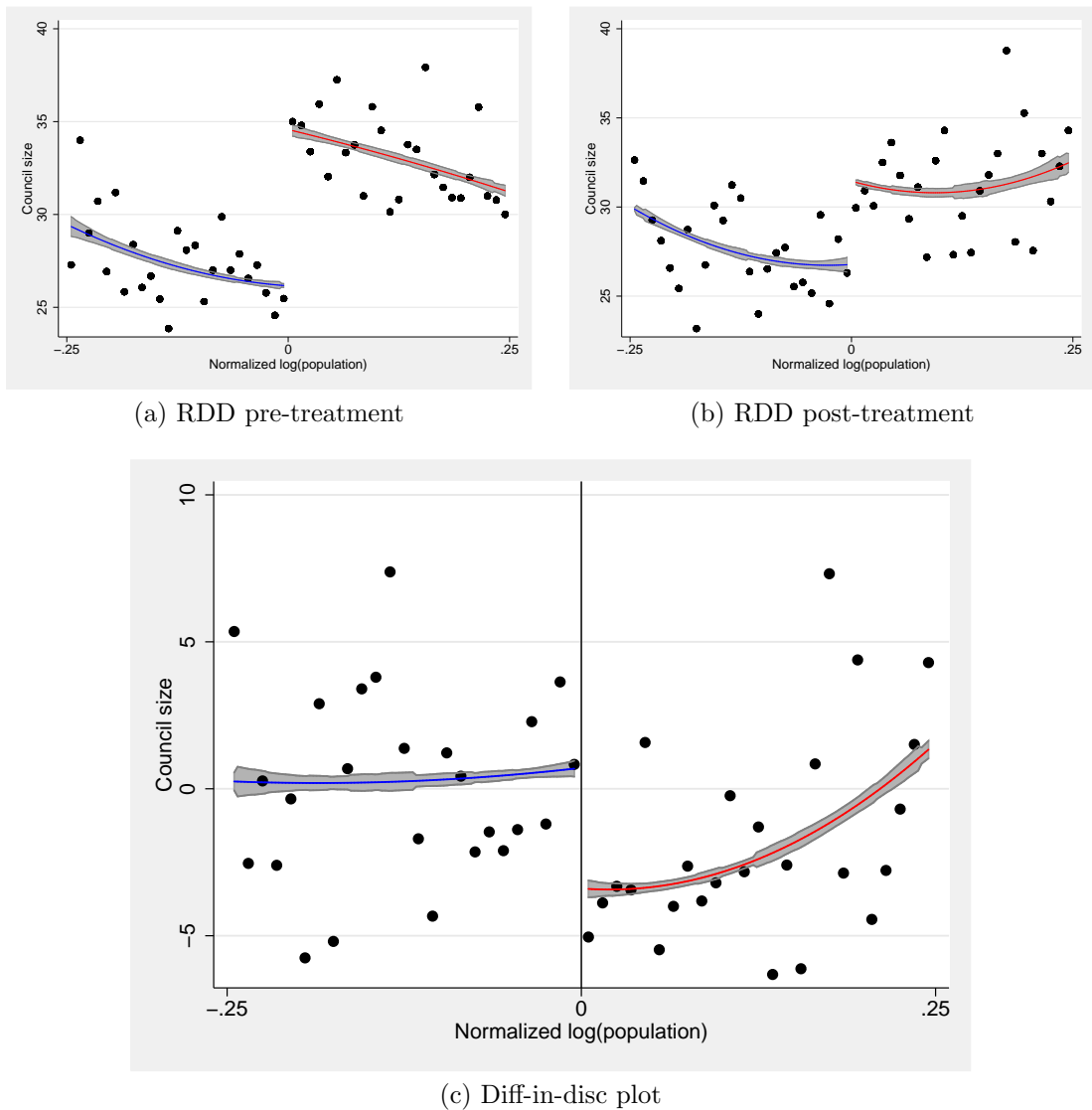


Fig 2.5: Council Size. This figure shows standard RDD plots for the pre- and post-treatment periods, subfigure (a) and (b), and a diff-in-disc plot, subfigure (c), for council size. Observations are averaged within bins of size 0.01. The polynomial plots are constructed using a rectangular kernel, a degree of 2, a bandwidth of 0.25, and the number of observations within bins as frequency weights. 95% confidence intervals are indicated in gray.

2.4.2 Regression Results

In Table 2.2 we collect the regression results. Consistent with the graphical evidence, the effect of abolishment is consistently negative and significant for almost all bandwidths and polynomials. Council size after the abolishment decreases in municipalities that were exposed to a stronger treatment by about 3 to 4 seats. The weighted average increase in treatment strength, i.e. the decline in the implicit electoral threshold, at the normalized population threshold is around 0.46.¹⁹ Scaled by the average increase in treatment intensity, these estimates imply that municipalities reduce the size of their council by about 7 to 8 seats after the abolishment of the five percent hurdle for every percent point lower implicit threshold. According to back of the envelope calculations, reducing the council size from e.g. 23 to 15 members increases implicit thresholds from about 2.2 to about 3.3 percentage points²⁰, i.e. by about 50%. This increase in implicit thresholds is presumably sufficiently large to keep several small parties out of the council that would otherwise have been entitled to a seat.

Table 2.2: Number of Council Seats.

	BW=0.25	BW=0.2	BW=0.15	BW=0.1	BW=0.05
Linear	-2.361*** (0.389)	-2.607*** (0.453)	-3.227*** (0.472)	-3.937*** (0.592)	-4.309*** (0.847)
Quadratic	-3.130*** (0.581)	-3.848*** (0.634)	-4.185*** (0.710)	-4.266*** (0.873)	-3.531*** (1.239)
Cubic	-4.323*** (0.762)	-4.526*** (0.823)	-4.552*** (0.931)	-4.450*** (1.150)	-3.020* (1.656)
N	1770	1416	1079	723	369

Notes: This table presents diff-in-disc regressions for the size of the council (number of seats). All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.25, 0.2, 0.15, 0.1, 0.05) and increasingly flexible polynomials (linear to cubic) of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***)

¹⁹Crossing the population threshold at 3001 from below implies on average a reduction in the implicit threshold from about 3.33 to about 2.17 percentage points, assuming that all municipalities choose the highest possible council size. Hence, the intensity of treatment from abolishing the explicit threshold increases by around 1.16 percentage points at the 3001 threshold (recall that the implicit threshold for the first seat is a sufficiently large vote share to gain half a seat) At the next threshold of 5001, the implicit threshold decreases from around 2.17 to around 1.61 percentage points. The intensity of treatment increases by around 0.56 percentage points. The same argument applies for all further thresholds. We weight the increase in treatment strength at each threshold with the number of observations within each population bracket when calculating the average size of the treatment.

²⁰Note that there is no exact formula for implicit thresholds. For the above calculations, we assume that parties are entitled to a seat if they have enough votes for “half a seat”.

2.5 Robustness Tests and Extensions

2.5.1 Placebo Tests

As a first set of placebo tests, we let the treatment set in at fake cutoffs and compare the estimated effects with those estimated at the correct cutoff. If the coefficient estimates in these placebo tests are significant, it is possible that municipalities adjusted their council size for reasons other than that the abolishment of the electoral threshold. We define D_i in Equation (2.4) such that it is 1 if $NLPOP = -2, -1, 0, 1, 2$ ($NLPOP = 0$ indicates the true threshold). We estimate the regressions for all combinations of bandwidths and polynomials reported in the baseline estimates. To save space, we summarize the results in graphs.

The structure of the plot in Figure 2.6 is as follows. For each fake cutoff, we plot the 15 coefficient estimates obtained by combining the bandwidths and bin sizes used in the baseline regressions and indicate their median value. We find that the placebo tests confirm the baseline estimates. The average coefficient estimate at the fake cutoffs revolves around 0, but decreases noticeably at the true cutoff.

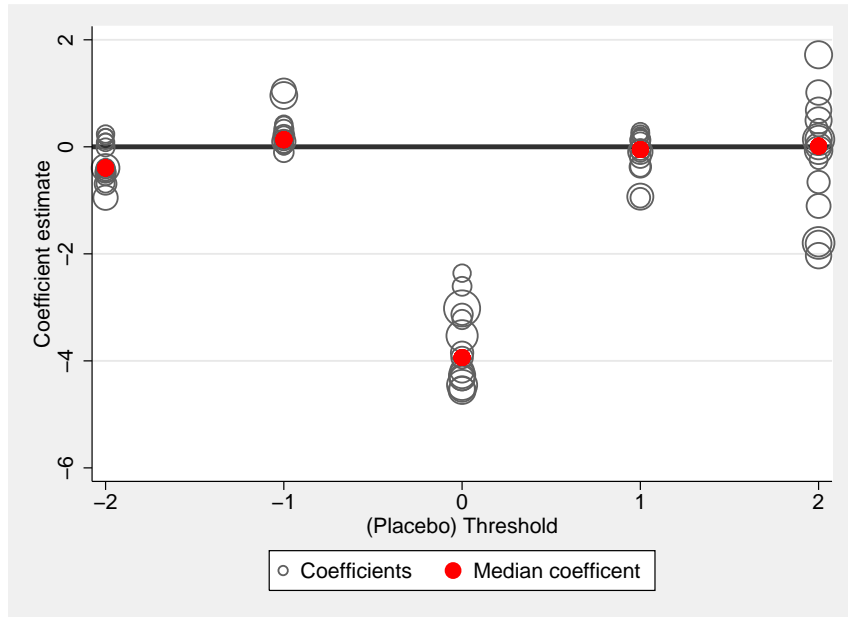


Fig 2.6: Placebo Treatments with Fake Cutoffs. This figure shows coefficient estimates of the diff-in-disc model for council size with placebo treatments. The size of the dots indicates the standard error of each estimate. The median estimates are indicated with red dots. The cutoffs are redefined such that treatment sets in at $NLPOP = -2, -1, 1, 2$. For comparison, the coefficient estimates at the true threshold of 0 are also indicated.

As a second set of placebo tests, we let the treatment set in at a fake treatment year. More specifically, we limit the sample to the pre-treatment period (1989-1997) and let the treatment set in 1993. If the treatment effect in the baseline regressions is not due to the reform but due to pre-existing trends, these placebo estimates should be significantly

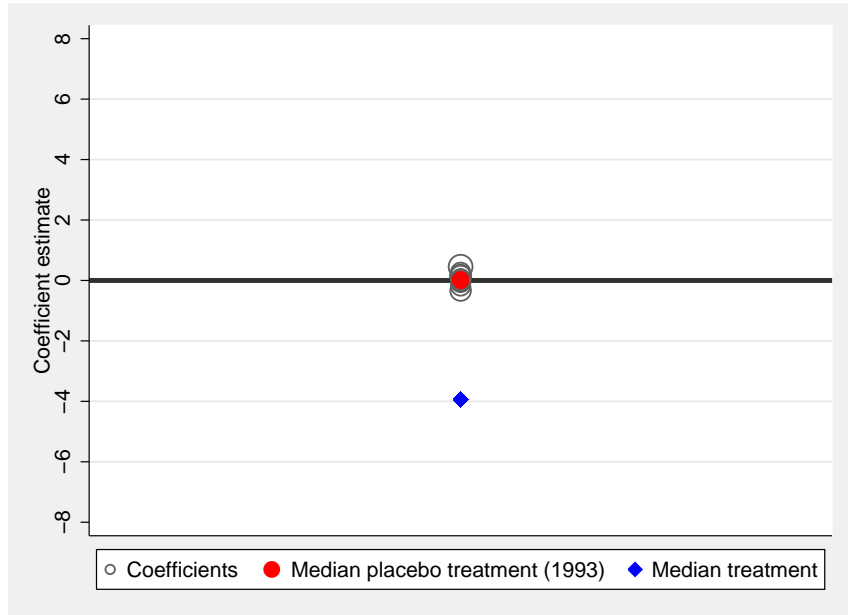


Fig 2.7: Placebo Treatment for Fake Treatment Year. This figure shows coefficient estimates of the diff-in-disc model with a placebo treatment defined to set in 1993. The sample covers the period 1989-1997. The size of the dots indicates the standard error of each estimate. The median estimate at the fake treatment year is indicated with a red dot. The median estimate at the true treatment year is indicated with a blue dot.

negative. We collect the results in Figure 2.7. The median coefficient at the fake cutoffs is indicated in red. For comparison, we also indicate the median estimate at the true cutoff with a blue dot. The estimates at the fake treatment year are close to 0, in contrast to the median estimate at the true threshold. Hence, these placebo regressions, too, confirm the baseline results.

2.5.2 Event-study Plot

Another way to evaluate whether the treatment effect found in the baseline regressions is due to pre-existing trends is to conduct an analysis in the spirit of an event-study. Consequently, we run traditional RDD regressions for each election year separately and estimate how the normalized cutoff is related to council size in each year. Given our baseline results, we expect the RDD coefficients for the 2001 election to be substantially smaller than the RDD coefficients for the 1997 election. If there were pre-existing trends, however, we should also observe declining RDD coefficients from 1989 to 1993 and from 1993 to 1997. A persistent decline in the RDD coefficients up to the 2001 election would suggest that at least part of the estimated treatment effect for the reform in 2001 is due pre-existing trends.

More formally, we estimate Equation (2.2) separately for $t = 1989, 1993, 1997, 2001, 2006, 2011$. We estimate the model for the bandwidths and polynomials used for the baseline estimates. To save space, we again collect the results in a graph. Figure 2.8 reports for each election year the 15 coefficients estimates for the various bandwidths

and polynomials. The size of circles indicates the standard errors. We indicate the median coefficient estimate with a red dot. We observe a small decline in the median RDD estimate from 1989 to 1993. A large number of estimated coefficients for 1989 are also quite large, but they are estimated imprecisely. The small decline from 1989 to 1993 in the RDD coefficients is presumably a statistical artifact due to a changing composition of treatment and control groups between the two elections. Given the substantial change in population size during this time period in Hesse (see Figure 2.2), some municipalities that were located below the normalized threshold in 1989 probably transitioned above the threshold in 1993, which in turn might affect the RDD coefficient estimates. Other explanations, especially deliberate responses of the established parties to some political events, are less likely because the descriptive evidence in Figure 2.3 indicates that almost all municipalities chose the highest possible council sizes both in 1989 and in 1993. That local politicians deliberately reduced council sizes would be inconsistent with this descriptive evidence.

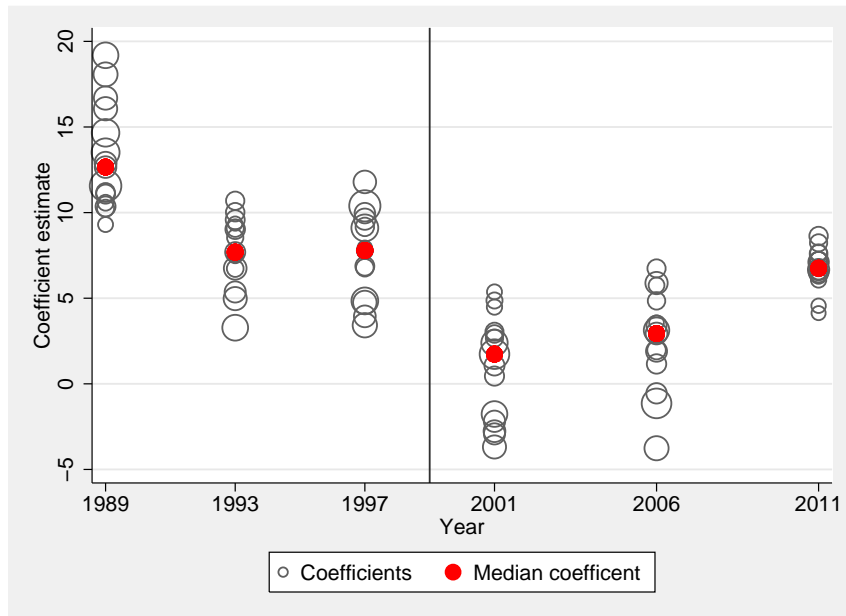


Fig 2.8: Event-study Plot. This figure shows coefficient estimates of traditional RDD regressions for all elections from 1989 to 2011. The size of the dots indicates the standard error of each estimate. The median estimate for each election is indicated with a red dot.

We find no decline in the RDD coefficient from 1993 to 1997. In 2001, there is, as expected, a substantial drop in the coefficient estimates. This drop persists to the 2006 election. By 2011, however, the RDD coefficient is as large as in 1993 and 1997. Chapter 3, shows that even though selected smaller parties benefited from the abolishment of the electoral threshold – i.e. they saw a significant increase in their vote and seat shares – overall council fragmentation did not increase significantly (Baskaran and Lopes da Fonseca, 2014). It is hence possible that the established parties, having observed for two consecutive

elections that consequences of the electoral reform were manageable, decided to expand the council size again before the 2011 election.²¹

Overall, there was a significant decline in the effect of the normalized discontinuity on council size just around the “event” of the reform in 2001. Furthermore, in the election immediately before and after the 2001 election, we observe no significant decline in the RDD coefficients. These results hence indicate that the treatment effect found in the baseline regressions was not due to pre-existing trends.

2.5.3 Individual Cutoffs

Are the baseline estimates driven by only selected cutoffs? Does the effect of the various cutoffs on council size differ? To answer these questions, we report results for the four smallest individual cutoffs. For compactness and since sample sizes are smaller, we only report results for specifications with a relatively large bandwidth of 0.25.

Table 2.3: Individual Population Cutoffs.

	T=3001	T=5001	T=10,001	T=25,001
Linear	-0.219 (0.472)	-3.529*** (0.707)	-2.358*** (0.637)	-3.853*** (1.026)
Quadratic	-0.484 (0.572)	-4.147*** (1.013)	-3.459*** (0.886)	-6.367*** (1.311)
Cubic	0.847 (0.574)	-5.712*** (1.326)	-3.303*** (1.176)	-8.199*** (1.864)
N	375	571	554	200

Notes: This table presents diff-in-disc regressions at the following individual population cutoffs: 3001, 5001, 10,001, and 25,001. Estimates for the average treatment effect of abolishing the election threshold are reported for a bandwidth of 0.25 and linear to cubic polynomials of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the level of the municipality and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

The results are collected in Table 2.3. We find for all cutoffs except the first that the coefficient estimates are consistently negative and significant. That we observe no effects at the lowest cutoff is plausible if politicians in smaller municipalities have little leeway or interest to reduce the council size. In particular, local politicians in small municipalities might feel that they may harm their own chances of winning a seat too much if they reduce the size of the council further. At the 5001 and 10,001 cutoffs, we observe significant but fairly moderate reductions: the coefficient estimates are largely between 2 to 5. At the 25,001 inhabitants threshold, we observe somewhat stronger effects. The treatment effect is between 4 to 8.

²¹A much more thorough description of the electoral consequences of the reform can be found in Chapter 3 or Baskaran and Lopes da Fonseca (2014).

2.5.4 Percent Changes

One further concern with the baseline estimates is that municipalities to the right of the normalized cutoff might appear to cut more seats in absolute terms than those to the left even if the abolishment of the electoral threshold was irrelevant when there are common trends that lead to the same relative reductions at the left and the right of the cutoffs. That is, the relative decline in council size to the left and the right of the normalized cutoff might be the same even if the decline in absolute terms is larger. To account for this concern, we estimate the baseline models with the log of council size as dependent variable, which allows for a percentage interpretation. The results are collected in Table 2.4. They show that the percentage decline at the normalized cutoff, too, is significantly negative. More specifically, council sizes declined by about 10 to 15 percent more to the right of the normalized threshold than to the left after the abolishment of the five percent threshold.

Table 2.4: Number of Council Seats (log).

	BW=0.25	BW=0.2	BW=0.15	BW=0.1	BW=0.05
Linear	-0.073*** (0.015)	-0.083*** (0.017)	-0.100*** (0.018)	-0.124*** (0.022)	-0.141*** (0.031)
Quadratic	-0.097*** (0.022)	-0.119*** (0.024)	-0.133*** (0.027)	-0.146*** (0.034)	-0.105** (0.051)
Cubic	-0.139*** (0.028)	-0.153*** (0.031)	-0.160*** (0.036)	-0.148*** (0.046)	-0.079 (0.065)
N	1770	1416	1079	723	369

Notes: This table presents diff-in-disc regressions for the log of the size of the council (log number of seats). All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.25, 0.2, 0.15, 0.1, 0.05) and increasingly flexible polynomials (linear to cubic) of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

2.5.5 Effect Heterogeneity

An interesting question is whether the treatment effects vary according to the composition of the council. Do councils where the two large mainstream parties (the CDU and SPD) together have a majority, either absolute or two-thirds, decrease council sizes more? Does the difference in the seat share between the two largest parties – whichever they may be – in the council matter? Do municipalities where the free voters’ associations have a substantial presence in the council react differently than other municipalities? To explore these questions, we interact the treatment dummy in Equation (2.4) with various variables that measure the composition of the current council. The basic model we estimate is as follows:

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

$$y_{i,t}^M = f(NLPOP) + \zeta V_{i,t} + D_i(\gamma_{t \geq T}^M + f(NLPOP)) + I_t(\alpha + f(NLPOP)) + D_i(I_t(\delta^M + \kappa V_{i,t}) + I_t f(NLPOP)) + \epsilon_{i,t} \text{ if } |NLPOP| < h, \quad (2.5)$$

where $V_{i,t}$ indicates various measures for the composition of the council before the election held in t . All other variables are defined as in Equation (2.4). If the effect varies with the composition of the council, the estimate for κ should be significant.

Table 2.5: Heterogeneous Effects.

	BW=0.25	BW=0.2	BW=0.15	BW=0.1	BW=0.05
Linear	-1.302*** (0.331)	-1.532*** (0.364)	-1.720*** (0.366)	-1.768*** (0.456)	-2.175** (1.090)
Quadratic	-1.209*** (0.330)	-1.393*** (0.330)	-1.581*** (0.361)	-1.772*** (0.446)	-2.234** (1.123)
Cubic	-1.198*** (0.298)	-1.369*** (0.324)	-1.597*** (0.353)	-1.765*** (0.453)	-2.487** (1.207)
N	1770	1416	1079	723	369

Notes: This table presents estimates from diff-in-disc regressions for an interaction variable between the dummy indicating the normalized threshold and a dummy indicating whether the CDU and SPD have jointly an absolute majority (i.e. whether CDU seat share + SPD seat share $\geq 50\%$) in the council. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the interaction effect are reported for different bandwidths (0.25, 0.2, 0.15, 0.1, 0.05) and increasingly flexible polynomials (linear to cubic) of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

As proxies for $V_{i,t}$, we have experimented with various variables, including a dummy for whether the CDU and SPD together have an absolute majority, a dummy for whether these two parties together have a two-third majority, and whether the free voters' associations have a substantial presence. We have also experimented with variables that account for the difference in the seat share of the two largest parties in the council, whichever they may be, and variables that capture whether the largest party has an absolute majority. We found a significant interaction effect only for one specification: when we used a dummy for whether the CDU and SPD together have an absolute majority. Table 2.5 reports the estimates for interaction effects for this specification. The results suggest that the decline in council size is more pronounced when the two large mainstream parties have an absolute majority.²²

It is curious that we find significant interaction effects for joint CDU and SPD absolute majorities but insignificant ones for two-third majorities, even though a two-third majority is required to change council size. If there are differences in the size of the treatment effect, we should expect them to be particularly pronounced at the two-third majority threshold. It is possible, however, that an absolute majority is already sufficient for the large parties

²²Results for the other interaction models are omitted.

to push through a change in council size. That is, an absolute majority gives the large parties a veto over council decisions. It is hence possible that the remaining parties are willing to trade their support for a reduction in council size against concessions in other areas.²³

2.6 Validity of the Identification Strategy

2.6.1 Direct Election of the Mayor

As mentioned in Section 2.2.1, another important reform occurred in Hesse at the local level in the early nineties: the introduction of direct elections for the mayoral office. Before 1993, mayors in all municipalities were appointed by the council. After 1993, municipalities switched to direct elections of the mayor. The transition to the new system for mayoral elections was staggered, with some municipalities switching immediately while others retaining their appointed mayors for a few years. It is possible that the switch may have affected municipalities differently along the population thresholds for council size. Hence, it may not be possible to separately identify the effect of the abolishment of the electoral threshold from the effect of the electoral rule for the mayor.

However, there is no good reason why the switch in how to choose the mayor should affect council size, and even less so why any effect should vary along the population thresholds. Second, the mayoral office, as described above, continued to be relatively unimportant even after the switch to direct elections, and thus any effects of the switch on council size or other political variables should be limited. Finally, all municipalities had switched to direct elections of the mayor by 1998. It is therefore unlikely that the effects of the switch to direct elections, even if they differed between municipalities at the population thresholds, persisted to the period after the implementation of the *Kommunalwahlreform*.

2.6.2 Other Elements of the *Kommunalwahlreform*

Lengthening of the legislative period

We interpret the negative treatment effect as evidence that the mainstream parties reduced council size to offset a perceived increase in the competitiveness of smaller parties. There might be other explanations. One possible alternative explanation is that voters demanded a reduction in council size following the reform because of the lengthening of the legislative

²³Another interesting question, as indicated above, is whether municipalities where the free voters' associations have a substantial presence react differently to the electoral reform than municipalities where they have only a marginal presence. We explore this issue with the same specification as above. We use various thresholds to define "substantial presence", i.e. a seat share of 5%, 10%, 15%, or 20% in the council. We also estimated models where we simply interacted the seat share of the voter initiatives in the previous legislative period with the treatment dummy. However, none of these specifications suggest that municipalities respond differently to the treatment depending on whether the free voters' associations have a substantial presence.

period from four to five years. That is, the lengthening of the representatives' terms of office may have led to demands for fewer representatives. Politicians, in turn, may have found it easier to reduce council size if councils were larger to begin with.

While we cannot explicitly test whether the reductions in council sizes are ultimately due to an attempt by the large parties to limit electoral competition or due to voter demands since we do not observe voters' preferences regarding their preferred council size nor the politicians' true motive to reduce the number of seats in the council, it is more plausible that council size was reduced to limit political competition. It is unlikely that the lengthening of the legislative period, a minor aspect of the reform, would have led to so much more vociferous calls from voters for council size reductions to the right of a given population cutoff than to the left, that it would force representatives to implement these reductions. Second, a two-third majority in the council is necessary for a change in council size. It seems unlikely that even if the voters demanded a reduction in council size, these demands were so strong to enforce coalitions across party lines. Third, we observe that council size already declines in 2001, i.e. the first election after the reform. Given that the reform law passed the state parliament in 1999, there were at most two years in which council sizes could be changed. Even if voters demanded smaller councils, it is unlikely their demands would be sufficiently pronounced and well articulated to effect concrete actions by the representatives within this relatively short time period.

Further Effects of *Kumulieren* and *Panaschieren*

Another, and arguably the most important, element of the *Kommunalwahlreform* was the switch to KUP. As discussed above, this switch likely increased the competitiveness of party rebels and may provide further incentives for the established parties to reduce council size to limit political competition. However, while the leadership of the established parties may indeed adopt measures to disadvantage party rebels following the switch to KUP, it is unlikely that reductions in council size would be one of these measures for reasons outlined above.

Nevertheless, the switch to KUP entails further aspects which may suggest a different interpretation for the regression results. First, besides the introduction of open lists, KUP allowed voters to cast as many votes as seats are available in the council. While the substantive effect of this second aspect of KUP should be small as the votes *per seat* remain constant at the population cutoffs, it may bring the size of their councils to the voters' attention. Before the reform, voters may have been unaware how large their council actually is. Realizing the size of their council due to this second aspect of KUP, voters may then demand a smaller council, particularly in those municipalities that have large councils, i.e. that are to the right of the cutoffs.

While this would provide an alternative interpretation for our regression results, it is unlikely that this is the main reason why council sizes decline more in municipalities with larger councils. First, established parties could always counter voter complaints about too large councils by pointing out that they are merely following the HGO. As discussed above, the wording in the HGO suggests the maximum council size as the default. Second,

any possible reduction in council size is limited. For example, a municipality with a population size between 10,001 and 25,000 could at most reduce its council sizes from 37 seats to 31 seats. It is unlikely that such a reduction will mollify those voters who are sufficiently outraged about the size of their council to have their votes swayed. Finally, as also mentioned in the previous subsection, it is unlikely that the various parties in the council agree with a two-third majority to reduce council sizes, thereby reducing the number of seats they can gain, only due to a diffuse concern that voters may favor smaller councils.

The second competing explanation for the diff-in-disc results is the following. Given that voters can cast as many votes as seats are available in the council after the switch to KUP, municipalities may have tried to provide voters with a “nice” number of votes, i.e. a number divisible by 10 or 5. But since practically all municipalities chose the highest possible council size in the pre-treatment period, municipalities had only the option of reducing council size to get at such “nice” numbers. However, an attempt to achieve a “nice” number of votes is unlikely the reason for the council size reductions. We find that there is no clustering at “nice” numbers after the reform. This absence of clustering at such numbers is in fact expected as according to the HGO, the number of seats in the council must be uneven.

2.7 Conclusion

We ask in this paper whether politicians from established parties adopt measures to disadvantage new political actors once they become a viable political alternative. To study this question, we exploit a reform in the German federal state of Hesse that improved the competitiveness of smaller parties and party rebels for exogenous reasons. The impact of one aspect of the reform, the abolishment of an electoral threshold, varied discontinuously between municipalities with different population sizes. Thus, we can implement a Diff-in-Disc design to identify how established politicians responded to the abolishment. One straightforward measure that established politicians can adopt is to decrease the size of the council. In smaller councils, implicit thresholds are lower, and hence particularly smaller parties will find it more difficult to gain legislative representation.²⁴ The results suggest that municipalities which had on average smaller implicit thresholds because they were to the right of a certain population cutoff indeed reduced their council size more.

The substantial effects on council size we find are particularly remarkable because local politicians should typically prefer, in line with the literature on pork barrel politics (Shepsle and Weingast, 1981), to have large councils in order to maximize their number of seats. That the established parties in the Hessian setting in principle share this preference is reflected by the fact that in the pre-treatment period, municipalities consistently chose

²⁴As mentioned above, it is also possible that party rebels are less competitive in smaller councils, thus potentially providing further incentives for the leadership of the established parties to reduce council size. However, for reasons also outlined above we believe this is a less plausible reason why council sizes would decline after the reform.

the largest possible council size. But it appears that preventing new political actors from entering the council is even more important to them than maximizing their number of council seats. There may be various reasons why they may be willing to forgo some seats only to keep out new entrants. First, they may have genuinely believed that increased legislative fragmentation due to the entry of new parties may make it harder to govern a municipality as it becomes more difficult to form stable coalitions. Alternately, they may have been concerned that any new entrants may disturb “old boys” networks and complicate rent seeking since the new entrants may become privy to any backroom deals or other questionable behavior. The established local politicians could have sacrificed a few seats in order to avoid this type of outside scrutiny. Third, established politicians could also have been worried that once some small party enters the council with a few seats, it may use these seats as a platform for future election campaigns, developing in the long-run into a potent electoral threat. The established parties may have been so worried about such long-term threats that they were willing to forgo some seats in the short-run.

From a policy perspective, our findings imply that the chances of new political actors to graduate into the political mainstream are possibly even lower than commonly believed. Established parties appear to erect additional barriers once new political actors improve their electoral prospects. Given the potentially positive implications of political competition, endogenous adjustments of political institutions to disadvantage new actors might be perceived as problematic.²⁵ On the other hand, of course, it is also possible to make the case that erecting barriers to entry into the mainstream increases the stability of the political system. The measures taken by the established parties in Hesse in the aftermath of the *Kommunalwahlreform* and, more generally, attempts of established parties to limit the electoral prospects of new political actors, should ultimately be viewed through the lens of how political competition affects economic and social outcomes. Theoretically, the effect seems ambiguous (Lizzeri and Persico, 2005). To further explore this question is an interesting and important avenue for future research.

²⁵However, we document in the following chapter that the reduction in council sizes did at least not harm the smaller parties too much (Baskaran and Lopes da Fonseca, 2014). Municipalities where implicit thresholds were lower saw a significant increase in the vote and seat shares of smaller parties.

Chapter 3

Electoral Thresholds and Political Outcomes*

3.1 Introduction

At the core of any electoral system is a trade-off between political representation and legislative cohesion. While majoritarian systems often ensure that only a small number of parties are able to achieve legislative representation, countries with proportional electoral rules (PR) are by default more prone to fragmented party systems (Duverger, 1954; Rae, 1971). PR systems thus often call for alternative electoral rules to limit the effective number of parties (Rokkan, 1968). Many countries rely on legal electoral thresholds for this purpose. Legal thresholds entail that a party may only receive seats in parliament if its overall vote share is above some fixed and relatively high value.

Recently, many countries have held acrimonious debates about the benefits and drawbacks of electoral thresholds.¹ Against this backdrop, it is important to obtain credible causal evidence on their political effects. However, existing studies overwhelmingly use cross-country variation and employ empirical methodologies that rely on selection on observables which may lead to biased estimates due to omitted variables.²

To fill this gap in the literature, we revisit the question of whether electoral thresholds have a causal effect on political outcomes relying on credible sources of exogenous variation: an institutional reform in the German state of Hesse that involved the abolishment of the five percent legal electoral threshold for local elections as of 2001 combined with a state law

*This chapter originates from joint work with Thushyanthan Baskaran (see Baskaran and Lopes da Fonseca (2014)).

¹In fact, several countries recently adjusted their thresholds. In 2014, Germany's federal constitutional court ruled the three percent electoral threshold for European Union elections unconstitutional. In New Zealand, the Electoral Commission recommended the party vote share threshold to be lowered from five to four percent. In Israel, on the other hand, the electoral threshold for the Knesset (national parliament) continually increased ever since the country was founded; the highest increase occurred as recently as April 2014.

²See e.g. Calvo and Micozzi (2005) Gebethner (1997) or Carey and Hix (2011). An exception is Vatter (2003) who uses subnational data at the level of the Swiss Cantons and finds that legal electoral thresholds have no effect on the number of parties represented in parliament. Second, Pellicer and Wegner (2014) find significant mechanical and psychological effects of local electoral thresholds in Morocco. However, a limitation of their analysis is that they have only data on seats but not on votes.

that discontinuously maps population to council size. As implicit electoral thresholds vary with council size, we can use the heterogeneity in how strongly the abolishment affected individual municipalities to identify the causal effect of electoral thresholds on party seat and vote shares with a difference-in-discontinuities design (diff-in-disc) (Grembi et al., 2015). Our estimates suggest that the vote and seat shares of small parties increased in municipalities that were affected more strongly by the abolishment while the seat share of large parties decreased accordingly. These results imply that the reform had strong psychological effects on voters.

3.2 Background

3.2.1 Local Politics

Hesse is divided into 426 municipalities. The most important local political institutions is the municipal council.³ Municipal council elections take place at the same date statewide with the electoral districts encompassing municipalities as a whole.

Several parties participate in local elections. We divide the contesting parties into three categories: large, medium, and small. We define the seat and vote share of large parties as the aggregate seat and vote share, respectively, of the two largest national parties, the center-right CDU and the center-left SPD. These two parties usually received around 30% of the votes in national and state elections during the sample period. The seat and vote shares of medium parties is the aggregate seat and vote share of the smaller national parties: the market liberal FPD and the environmentalist Green Party. These two parties usually received up to 10% of the votes in national and state elections during the sample period and were well established in the political mainstream.

Finally, many local and several very small national parties run in municipal elections. We refer to this group of parties as small since they often struggled with the five percent threshold. However, among these small parties, municipal specific voter initiatives (*Wahlvereinigungen*), are popular in some municipalities and often receive a substantial fraction of the votes.⁴ As there are many small parties, we define in the following the seat and vote share of small parties simply as 100% minus the seat or vote shares of the large and medium parties, respectively.

3.2.2 The *Kommunalwahlreform*

In 1999 the Hessian State Parliament passed a reform that changed the rules governing local elections from 2001 onwards, i.e. the *Kommunalwahlreform*. One of the main features of

³The other important institution is the mayor. Even though Hessian mayors are directly elected since the mid-nineties, the most important municipal decisions are still made by the local council. See Hessami (2014) for details on the mayoral office in Hesse.

⁴Note that the label “small” is therefore not always accurate. We use this label for simplicity, but in some cases the “small” parties, especially the voter initiatives, can be quite large. In fact, “small” should be understood as a shorthand for “party that is not important at the national level”.

the reform was the abolishment of the five percent explicit electoral threshold for municipal council elections. Parties can now enter the council as long as their vote share is sufficient to gain at least one seat. In addition, the reform entailed the extension of the legislative period from four to five years and the introduction of a new voting system called *Kumulieren und Panaschieren*, which was essentially a shift from closed to open lists.⁵

Unlike the last two elements of the reform, which arguably affected all municipalities equally, the abolishment of the threshold had decidedly heterogeneous effects.⁶ This heterogeneity results from the existence of intrinsic implicit thresholds in municipal councils, as parties still need to gain a minimum vote share in order to gain their first seat. Implicit thresholds vary throughout municipalities depending on the total number of seats in the municipal council.⁷ Council size, in turn, is by law a positive and discontinuous (albeit “fuzzy”) function of municipal population size (see Table 2.1).⁸ Consequently, implicit thresholds decrease and treatment intensity increases in expectation exogenously and discontinuously at the population cutoffs. This heterogeneity in treatment intensity forms the core of our identification strategy.

3.2.3 Hypotheses

It is plausible that the electoral reform benefited the smaller parties, particularly in municipalities with lower implicit thresholds. First, abolishing a threshold should have a

⁵Specifically, in the old system each voter cast one vote for their preferred party list. In the new system, each voter can cast as many votes as there are seats in the council for individual candidates and thus affect their within-party list ranking.

⁶As mentioned previously, the absolute number of votes that can be cast increases with the number of seats in the council. One important assumption for our empirical design is that the increase in the number of available votes as council sizes increase has no significantly heterogeneous effects on voters’ incentives. This is a plausible assumption as the number of votes *per seat*, which is arguably a more important variable for voters than the total number of votes, remains constant across municipalities with different council sizes. Moreover, any heterogeneous effect of the different number of votes is arguably less important than the heterogeneous effect arising from the different implicit thresholds.

⁷The actual value of an implicit threshold for a given party is endogenous and depends inter alia on the vote shares of all other parties and the seat allocation formula. In this context, the seat allocation formula is the largest remainder method (Hare-Niemeyer). Thus, a vote share that is sufficiently large for half a seat typically entitles a party to a full seat in the council. See <http://www.wahlrecht.de/kommunal/hessen.html>. Taagepera (2007, p. 243) presents formulas for implicit thresholds, deriving in particular inclusion (minimum vote share necessary to gain the first seat) and exclusion thresholds (maximum vote share until which the first seat can be denied), for various seat allocation formulas. He also presents a formula for an overall average estimate of an implicit threshold, i.e. : $T = 75\% / (M + 1)$, with M being the number of seats in a district (Taagepera, 2007, p. 241). Accordingly, as the whole municipality is a district in Hessian municipalities, the implicit threshold in municipalities with councils that have 15 seats is about 4.7% and in municipalities with 93 seats about 0.8%.

⁸The fuzziness of the relationship between population size and council size is due to the legal stipulation that municipalities may choose the (smaller) council size intended for the previous population bracket. However, in practice we observe discontinuously increasing council sizes at the cutoffs (see Section 3.A.1 of the Appendix to Chapter 3 for details).

mechanical effect on the seat share of smaller parties. That is, assuming that voting patterns remain the same, the absence of an explicit threshold should allow more marginal parties to enter the council in municipalities with more council seats and thus lower implicit thresholds. Taagepera (2007, 2008), for example, shows theoretically that implicit thresholds and the number of effective parties depends systematically on the size of the legislature.

Yet, there may be further effects. With a threshold, supporters of smaller parties may have preferred to vote for a larger party as a second best option, i.e. they may have anticipated that their vote would be effectively lost if they cast it for a smaller party that was unlikely to pass the electoral threshold (Perea, 2002). Once the legal threshold was abolished, smaller parties may have become a more viable electoral alternative in municipalities with lower implicit thresholds. Thus, a section of voters may have switched allegiance from a less-preferred large to their preferred small party for psychological reasons (Moser and Scheiner, 2004; Fiva and Folke, 2014), thereby amplifying any mechanical effects. In addition, individuals that previously abstained, anticipating that their preferred small party would not enter the council, might now participate in the election, increasing the vote and seat shares of smaller parties further (Perea, 2002). Overall, such psychological effects may be very important – indeed, legal thresholds and other electoral rules that increase the disproportionality of a voting system and thereby mechanically disadvantage smaller parties are introduced not only for their mechanical effects as such, but also because they are expected to possibly have even larger psychological effects (Moser and Scheiner, 2012).⁹

On the other hand, countervailing psychological effects may be also at work. With respect to voter participation, some supports of smaller parties might now lack the incentive to go to the polls, as smaller parties no longer need to overcome the five percent threshold to enter the council. Such a response would reduce the vote shares of smaller parties. Furthermore, voters may be less likely to engage in certain types of strategic voting in the absence of an electoral threshold. For example, before the reform, if supporters of a large party anticipated that it would not receive an absolute majority and would thus have to form a coalition with a smaller party, they might have voted for their preferred small coalition partner in order to help it overcome the electoral threshold. Without a threshold, the need to engage in such strategic voting is smaller, and thus the vote share of smaller parties may decline. Another example relates to how explicit thresholds might affect the structure of the party system (Cox, 1997). For instance, when there is a legal threshold, voters who are not satisfied with the established parties have an incentive to implicitly coordinate on one or a few smaller parties in order to gain legislative representation. Thus,

⁹In fact, it is difficult to disentangle mechanical and psychological effects. Moser and Scheiner (2012), for example, point out that the psychological effect may mask the size of the *perceptible* mechanical effect if a substantial number of voters respond to the disproportionality caused by a electoral rule by switching to a less favored larger party. That is, if voters who actually support smaller parties react to e.g. a high electoral threshold by switching their vote to larger parties, the disproportionality induced by the threshold may seem low as the large parties will have both a large seat *and* a large vote share. Thus, ultimately mechanical and psychological effects work in concert.

those small parties that benefit from any implicit coordination will receive a relatively large vote and a correspondingly large seat share. Without a legal threshold, the need for such coordination is lower. Thus, more non-mainstream formations may enter the council, but their aggregated seat share may decline because some non-mainstream formations may not receive sufficient votes to gain legislative representation. More generally, as Moser and Scheiner (2012) argue, the effects of electoral rules can be context-specific, and thus the strategic effects of the abolishment of a legal threshold on the party system are difficult to predict and arguably depend on many idiosyncrasies of a specific context, such as the extent of accurate information voters have on the electoral chances of smaller parties or the strategic calculus of local political elites.

3.3 Empirical Design

3.3.1 Difference-in-discontinuities Design

To identify the causal effect of abolishing the explicit electoral threshold, we implement a diff-in-disc design using a sample consisting of all 426 Hessian municipalities from 1989 to 2011 that encompasses three local elections prior and three after the electoral reform. This design combines regression discontinuity (RDD) and difference-in-differences (diff-in-diff) methodologies. Due to the RDD aspect of the design, it is robust against violations of the common trend assumption of standard diff-in-diff approaches. The diff-in-diff aspect in turn, makes it robust against specific violations of the RD assumptions, such as co-treatment and manipulation of the running variable.¹⁰

In our implementation of the diff-in-disc design, we exploit the fact that the abolishment of the electoral threshold had heterogeneous effects across municipalities with different population sizes. Specifically, as noted in Table 2.1, the number of seats in the council is a discontinuous function of population size. Implicit thresholds, in turn, vary with council size. The larger a council, the lower the implicit threshold: in a council with 10 seats, the implicit threshold is about 5% and thus as high as the previous explicit threshold. In a council with 100 seats, the implicit threshold is only about 0.5%, and thus much smaller than previously. As a result, at the council size cutoffs, the effect of abolishing the explicit threshold changes discontinuously. We exploit this discontinuity to identify the causal effect of the abolishment on vote and seat shares.

Following previous RDD with multiple population cutoffs, we normalize all observations such that they are around a single population cutoff denoted M (Egger and Koethenbueger, 2010). Specifically, we normalize the running variable, i.e. population size, by subtracting the log of each cutoff from the log of population $lpop$, i.e. $npop = lpop - M$. The empirical model we estimate is:

$$y_{i,t}^M = npop^\kappa + D_i(\gamma + npop^\kappa) + I_t(\alpha + npop^\kappa + D_i(\delta + npop^\kappa)) + \epsilon_{i,t} \quad (3.1)$$

¹⁰See Grembi et al. (2015) or Section 3.A.1 of the Appendix to Chapter 3 for a more comprehensive discussion of the diff-in-disc design .

where $y_{i,p}$ is either the vote or seat share of a small, medium, or large parties and D_i and I_t are dummy variables indicating the assignment to treatment and the post-treatment period, respectively. We include a flexible polynomial of the running variable, $npop^k$, and estimate the model for various bandwidths around the normalized cutoff $|npop| < h$, with $h = 0.4, 0.5, 0.6, 0.7$. We also follow the previous literature and include in all regressions municipality and election year fixed effects to improve efficiency and reduce finite sample bias (Hoxby, 2000). We are interested in the estimate for δ , which captures the change in the effect of the discontinuity at the normalized population cutoff between the pre- and the post-treatment periods.

3.3.2 Difference-in-discontinuities Plots

As a complement to our regression results, we present graphical evidence on the treatment effect. We construct the diff-in-disc plots by first dividing normalized population size into bins of size 0.001 within a window of 0.2. Then we calculate the average of the relevant outcome variable y within each bin for the pre- and post-treatment period, i.e. $\bar{y}_{b,t}$ with the index $b = 1, \dots, 200$ denoting the bin and $t = 0, 1$ denoting the pre- and the post-treatment period. Then we obtain the difference within each bin in the pre- and post-treatment period $\Delta y_b = (\bar{y}_{b,1} - \bar{y}_{b,0})$. Finally, we plot Δy_b against $npop$ to the left- and the right of the normalized cutoff.

To observe whether there is a discontinuity at the normalized cutoff, we smooth Δy_b with a local polynomial plot of quadratic degree and a bandwidth of 0.1 at both sides of the threshold, using a rectangular kernel and the number of observations within each bin as frequency weights. For presentational purposes, we use the average of the differences for bins of width 0.01 (rather than for the original bin widths). The smooth is constructed based on the original Δy_b .

3.4 Results

3.4.1 Vote and Seat Shares

Figure 3.1 shows the diff-in-disc plots for the aggregate seat and vote shares of the small, medium, and large parties. There is a noticeable discontinuity at the normalized population cutoff in subfigures (a) and (b) suggesting a causal increase in the seat and vote shares of the small parties due to the abolishment. Subfigures (c) and (d) show no discontinuity at the normalized population cutoff for the medium parties while subfigures (e) and (f) show a negative discontinuity at the normalized cutoff for the large parties. The graphical evidence suggests that small parties gained seat and vote shares mainly at the expense of large parties.

The coefficient estimates are collected in Table 3.1. For both seat and vote shares of small parties, we find that these are consistently positive and always significant across bandwidths and of a magnitude between 3 and 4. Accordingly, coefficient estimates

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

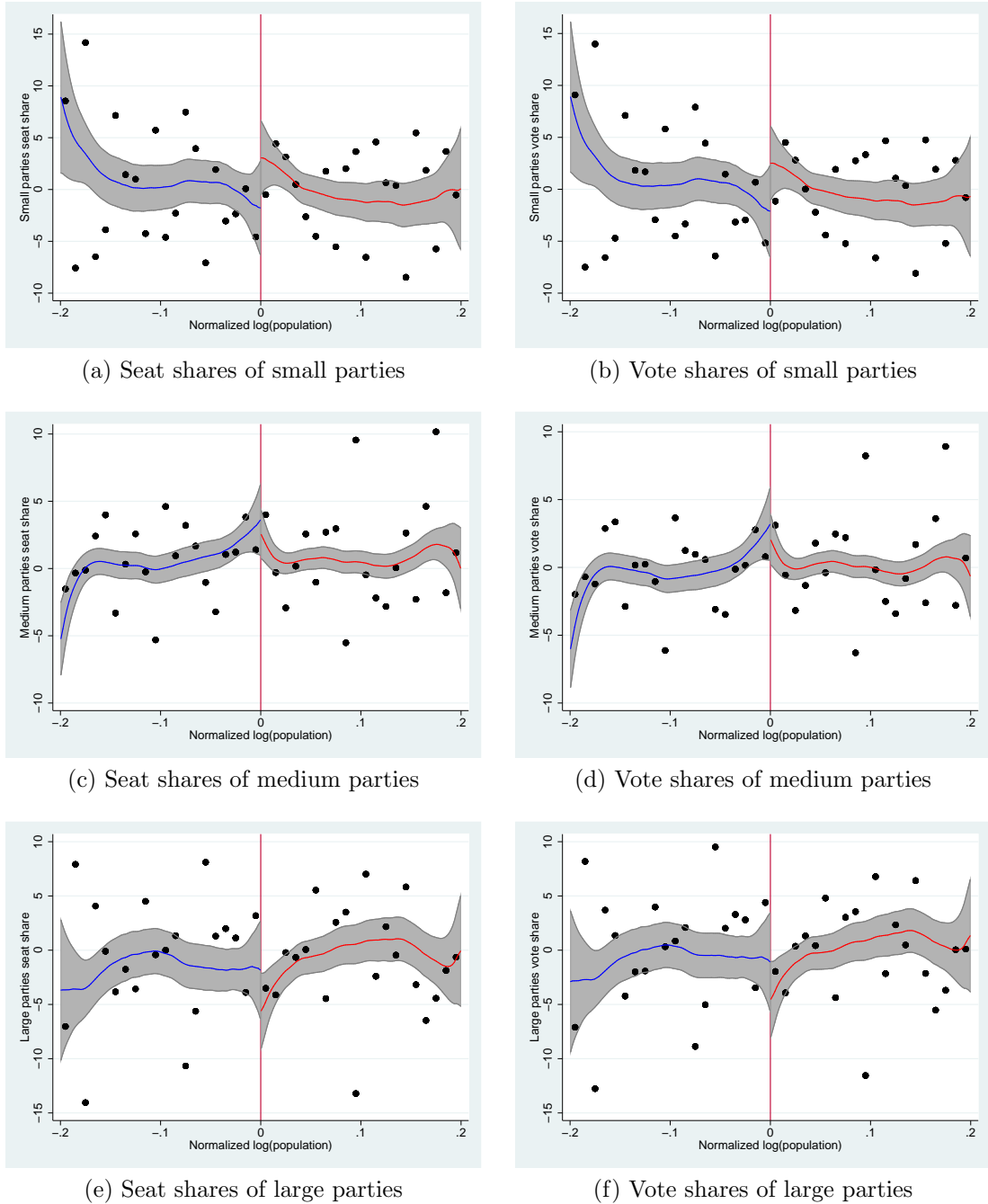


Fig 3.1: Seat and Vote Shares. This figure shows diff-in-disc plots for the change in the seat and vote shares of small, medium, and large parties from the pre- to the post-treatment periods. Observations are averaged within bins of size 0.001. The polynomial plots are constructed using a rectangular kernel, a degree of 2, a bandwidth of 0.1, and the number of observations within bins as frequency weights. 95% confidence intervals are indicated in gray.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

are consistently negative for both the seat and vote shares of medium as well as large parties, even if not always significant. Coefficient estimates are generally larger for the large than the medium parties. These results suggest that small parties benefited from the abolishment of the electoral threshold primarily at the expense of the large parties. As shown in Section 3.A.2 of the supplementary material, these results survive various robustness tests.¹¹

Table 3.1: Seat and Vote Shares.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Small parties				
Seat Share	4.913** (2.387)	3.711* (2.250)	3.682** (1.807)	3.027* (1.742)
Vote Share	5.028** (2.333)	3.834* (2.206)	3.690** (1.773)	3.143* (1.706)
Medium parties				
Seat Share	-2.152 (1.539)	-1.315 (1.369)	-0.589 (1.212)	-0.504 (1.189)
Vote Share	-1.493 (1.418)	-0.490 (1.257)	-0.022 (1.118)	0.070 (1.092)
Large Parties				
Seat Share	-2.761 (2.253)	-2.395 (2.138)	-3.093* (1.742)	-2.523 (1.673)
Vote Share	-3.536 (2.208)	-3.345 (2.092)	-3.669** (1.700)	-3.213** (1.623)
N	2797	3496	4208	4917

Notes: This table presents diff-in-disc regressions for the seat and vote shares of the small, medium and large parties in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipality level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

A follow-up question is whether the change in seat and votes shares is mainly due to the mechanical effects of the reform or if psychological effects are also at play. While the mechanical effect should automatically increase the seat share of small parties at a given vote share, the total effect also depends on whether voting patterns changed to the benefit or detriment of the smaller parties.¹² The evidence presented above shows that the

¹¹In particular, we report regression results using different bandwidths, various placebo tests, and estimates for individual cutoffs.

¹²As discussed above, mechanical and psychological effects are closely related. We interpret in the following any changes in vote shares as psychological effect and any changes in seat shares *at given vote shares* as mechanical effect. However, in view of the previous discussion, it is possible to argue that changes

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

vote share of smaller parties increased significantly. Hence, the abolishment of the legal threshold had indeed strong psychological effects in favor of smaller parties.

Furthermore, note that the estimated treatment effects for the small parties' vote shares are slightly larger than for their seat shares. This suggests that the purely mechanical effects of the reform, i.e. any increase in seat shares of smaller parties at given vote shares, in voting patterns, too, are ultimately a consequence of the mechanical effect, i.e. the reduction in the disproportionality of the electoral system.

Table 3.2: Seat and Vote Shares of Individual Parties.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
CDU				
Seat share	-2.134 (1.453)	-1.824 (1.372)	-2.199* (1.166)	-1.746 (1.117)
Vote share	-3.022** (1.400)	-2.724** (1.319)	-2.945*** (1.134)	-2.450** (1.087)
SPD				
Seat share	-0.627 (1.627)	-0.571 (1.547)	-0.893 (1.275)	-0.777 (1.223)
Vote share	-0.514 (1.571)	-0.621 (1.482)	-0.725 (1.215)	-0.764 (1.159)
FDP				
Seat share	-1.312 (0.879)	-0.995 (0.786)	-0.709 (0.681)	-0.416 (0.675)
Vote share	-0.657 (0.732)	-0.289 (0.649)	-0.187 (0.549)	0.122 (0.553)
The Greens				
Seat share	-0.840 (1.104)	-0.320 (0.984)	0.120 (0.870)	-0.088 (0.842)
Vote share	-0.836 (1.054)	-0.202 (0.941)	0.165 (0.834)	-0.052 (0.808)
Voter Initiatives				
Seat share	4.266* (2.432)	2.993 (2.258)	2.888 (1.828)	2.296 (1.763)
Vote share	4.250* (2.376)	3.033 (2.213)	2.829 (1.789)	2.300 (1.722)
N	2797	3496	4208	4917

Notes: This table presents diff-in-disc regressions for the seat and vote shares of individual parties. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

were small. As we show in a companion paper, one reason for these limited mechanical effects may be that municipalities exposed to stronger treatments decreased council sizes and thus increased implicit thresholds, thereby counteracting the mechanical effects of the abolishment of the explicit threshold (Baskaran and Lopes da Fonseca, 2016).

3.4.2 Seat Shares of Individual Parties

We report in Table 3.2 the effect of the abolishment of the electoral threshold on individual parties' seat shares. As expected, the coefficient estimates for the large parties, CDU and SPD, are consistently negative. Yet, they are only significant and of a large magnitude for the CDU. The coefficient estimates for the medium parties, FDP and the Greens, are also negative, but of a much smaller magnitude than for the CDU and never significant. It appears that the CDU was the main loser from this reform.

Finally, we look at municipality-specific voter initiatives. The estimated coefficients for the vote shares of voter initiatives are consistently positive and of a similar magnitude as those obtained for the vote shares of the small parties in the baseline regressions. Voter initiatives, already popular in Hesse, appear thus to have gained the most from the abolishment of the legal electoral threshold, even if the estimated coefficient is only significant for a relatively narrow bandwidth of 0.4.

3.4.3 Election-specific Vote and Seat Shares

In this section, we report election-specific estimates for the aggregated vote and seat shares of the small, medium, and large parties. That is, we estimate how the vote and seat shares of the three groups of parties compares to the pre-reform period in each post-reform elections separately. The results are collected in Tables 3.3 to 3.5. They suggests that the vote and seat shares of the large parties increased in municipalities exposed to stronger treatments in the 2001 elections relative to the pre-reform elections. Correspondingly, the vote and seat shares of small parties decreased. In the 2006 elections, however, the effect was reversed – the vote and seat shares of small parties increased and those of the larger parties decreased. This pattern prevails in the 2011 elections, causing the “average effect” estimated in Table 3.1 for the small parties to be positive.

These results suggest that smaller parties benefited from the reform, but only with a certain delay. A possible explanation for this pattern, following Cox (1997) and Moser and Scheiner (2012), is that in an environment where it is unclear which non-mainstream parties will manage to enter the legislature, i.e. when there is little information about electoral prospects of smaller parties, it makes sense for voters to vote for their most preferred large party as that party will definitely gain legislative representation. As it becomes clearer which of the smaller parties are competitive, strategic voters are likely to switch to one of the smaller parties.

Table 3.3: Small Parties.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Panel A: 1989-1997 & 2001				
Seat Share	-3.024 (3.133)	-3.090 (2.780)	-1.895 (2.292)	-1.725 (2.204)
Vote Share	-3.445 (2.984)	-3.353 (2.665)	-2.239 (2.184)	-2.013 (2.102)
N	1861	2345	2802	3269
Panel B: 1989-1997 & 2006				
Seat Share	8.930** (4.519)	7.876* (4.196)	7.800** (3.113)	6.277** (2.782)
Vote Share	9.570** (4.377)	8.432** (4.054)	8.121*** (3.019)	6.675** (2.699)
N	1865	2347	2803	3267
Panel C: 1989-1997 & 2011				
Seat Share	6.632* (3.849)	5.541 (3.579)	4.667* (2.833)	3.827 (2.797)
Vote Share	6.976* (3.801)	5.862 (3.571)	4.863* (2.811)	4.166 (2.780)
N	1853	2339	2802	3264

Notes: This table presents diff-in-disc regressions for the seat and vote shares of the small parties in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

3.5 Extensions

3.5.1 Council Fragmentation

To complement the analysis on the impact of electoral thresholds on the seat and vote share of the different parties, we study the effects of the treatment on council fragmentation. For the purpose we rely on three different fragmentation measures: the number of parties in the municipal council, the seat share held by the largest party in the council, and the inverse Herfindahl index.

The regression results are collected in Table 3.6. Panel A collects the results for a sample that covers the entire sample period, i.e. 1989-2011. The results suggest no significant impact of the abolishment of the legal electoral threshold on council fragmentation. In other words, council fragmentation in municipalities exposed to a stronger treatment was on average not higher in the three post-reform elections than in the three pre-reform ones.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table 3.4: Medium Parties.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Panel A: 1989-1997 & 2001				
Seat Share	-2.727** (1.351)	-1.448 (1.220)	-0.416 (1.092)	-0.633 (1.026)
Vote Share	-1.401 (1.174)	-0.242 (1.051)	0.412 (0.963)	0.130 (0.895)
N	1861	2345	2802	3269
Panel B: 1989-1997 & 2006				
Seat Share	-1.760 (2.104)	-1.543 (1.851)	-0.774 (1.582)	-0.449 (1.480)
Vote Share	-0.672 (1.926)	0.009 (1.690)	0.338 (1.452)	0.649 (1.359)
N	1865	2347	2803	3267
Panel C: 1989-1997 & 2011				
Seat Share	-1.476 (2.829)	-1.021 (2.465)	-0.585 (2.049)	-0.601 (2.014)
Vote Share	-1.055 (2.586)	-0.361 (2.263)	-0.321 (1.876)	-0.246 (1.835)
N	1853	2339	2802	3264

Notes: This table presents diff-in-disc regressions for the seat and vote shares of the medium parties in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

However, it is possible that the effect of the reform differed across the post-reform elections. The uncertainty associated with the reform, especially the difficulty in knowing which parties had a realistic chance of entering the council, may have led voters to coordinate at first on a few established parties, maybe even leading to a decline in the degree of fragmentation in the 2001 election. As experience with the new voting system increased, voters were arguably able to determine which smaller parties were competitive and may have switched their vote accordingly (Cox, 1997; Moser and Scheiner, 2012). Consequently, council fragmentation may have increased in the long-run while remaining unaffected or even declining in the short-run. Panel B to D explore this possibility.

In Panel B, we report regressions where the sample is restricted to the period 1989-2001, i.e. we only study the effect of the reform on the three measures of council fragmentation in the first post-reform elections. The results indicate that compared to the pre-reform period, council fragmentation decreased slightly in 2001, even if the estimated coefficients are mostly insignificant.

Table 3.5: Large Parties.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Panel A: 1989-1997 & 2001				
Seat Share	5.752** (2.894)	4.538* (2.625)	2.311 (2.179)	2.358 (2.065)
Vote Share	4.842* (2.817)	3.590 (2.563)	1.820 (2.112)	1.875 (2.006)
N	1861	2345	2802	3269
Panel B: 1989-1997 & 2006				
Seat Share	-7.169 (4.435)	-6.333 (4.102)	-7.026** (3.080)	-5.828** (2.832)
Vote Share	-8.894** (4.271)	-8.438** (3.942)	-8.455*** (2.953)	-7.321*** (2.712)
N	1865	2347	2803	3267
Panel C: 1989-1997 & 2011				
Seat Share	-5.156 (3.628)	-4.520 (3.428)	-4.082 (2.768)	-3.225 (2.734)
Vote Share	-5.928* (3.482)	-5.506* (3.308)	-4.545* (2.672)	-3.922 (2.650)
N	1853	2339	2802	3264

Notes: This table presents diff-in-disc regressions for the seat and vote shares of the large parties in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

Panel C reports results where the sample is restricted to the three pre-reform elections (1989-1997) and the 2006 elections, i.e. the second post-reform elections. The results suggest that compared to the pre-reform period, council fragmentation has noticeably increased in 2006, though the coefficients are not significant. Finally, Panel D reports results for a sample restricted to the three pre-reform elections and the 2011 elections. We observe that the degree of council fragmentation increased even further in the 2011 elections compared to the 2006 ones. The coefficients are positive and typically significant for the number of parties and Herfindahl index regressions. Overall, it appears that the abolishment of the electoral threshold increased council fragmentation in the long-run.

3.5.2 Turnout

As a final extension, we explore the effect of the abolishment of the explicit threshold on turnout as one of the aims of the reform was to increase voter participation in local elections. The expectation was that the abolishment of the electoral threshold would increase the

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table 3.6: Council Fragmentation.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Panel A: Full sample				
Nr. of parties	-0.031 (0.226)	0.025 (0.201)	0.147 (0.187)	0.093 (0.170)
Max. share	-0.035 (1.925)	0.660 (1.714)	-0.129 (1.413)	-0.281 (1.389)
Herfindahl index	0.461 (1.414)	0.268 (1.277)	1.255 (1.073)	1.142 (1.081)
N	2323	2909	3510	4112
Panel B: 1993-1997 & 2001				
Nr. of parties	-0.501** (0.254)	-0.441* (0.226)	-0.223 (0.208)	-0.227 (0.194)
Max. share	2.461 (2.148)	2.897 (1.968)	1.687 (1.588)	1.047 (1.549)
Herfindahl index	-2.060 (1.662)	-2.335 (1.525)	-0.749 (1.272)	-0.456 (1.254)
N	1385	1758	2104	2464
Panel C: 1993-1997 & 2006				
Nr. of parties	0.077 (0.322)	0.043 (0.279)	0.177 (0.245)	0.080 (0.231)
Max. share	-0.903 (3.115)	0.153 (2.632)	-1.276 (2.166)	-1.672 (2.104)
Herfindahl index	0.956 (2.333)	0.696 (1.979)	2.024 (1.604)	1.774 (1.595)
N	1389	1760	2105	2462
Panel D: 1993-1997 & 2011				
Nr. of parties	0.637* (0.373)	0.691** (0.322)	0.708*** (0.273)	0.607** (0.247)
Max. share	-2.886 (2.665)	-2.258 (2.380)	-2.399 (1.904)	-2.107 (1.879)
Herfindahl index	2.967 (2.124)	3.359* (1.915)	3.234** (1.543)	3.039** (1.540)
N	1381	1752	2104	2459

Notes: This table presents diff-in-disc regressions for the different measures of fragmentation in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

ability of voters to vote according to their true preference and thus make it worthwhile to go to the polls.

Table 3.7: Turnout.

	BW=0.4	BW=0.5	BW=0.6	BW=0.7
Panel A: Full sample	-0.141	-0.271	0.119	1.029
	(0.802)	(0.740)	(0.657)	(0.668)
N	2797	3496	4208	4917
Panel B: 1993-1997 & 2001	-0.636	-0.887	-0.259	0.754
	(1.288)	(1.166)	(0.951)	(0.965)
N	1861	2345	2802	3269
Panel C: 1993-1997 & 2006	-0.351	-0.370	-0.050	1.296
	(1.546)	(1.370)	(1.198)	(1.134)
N	1865	2347	2803	3267
Panel D: 1993-1997 & 2011	1.400	1.476	1.444	2.124*
	(1.547)	(1.366)	(1.137)	(1.139)
N	1853	2339	2802	3264

Notes: This table presents diff-in-disc regressions for turnout in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.4, 0.5, 0.6, 0.7) and a cubic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

The results are collected in Table 3.7. As for council fragmentation, we differentiate between the average effect over the entire post-reform period (Panel A), and effects in each of the post-reform elections separately (Panel B to D). We find the average effect to be insignificant. Similarly, the election specific estimates in Panels B and C also suggest that turnout did not increase in the 2001 and 2006 elections compared to the pre-reform elections. However, the estimated coefficient is larger in Panel D, where we focus on the 2011 election, and once significant. Therefore, there is some, albeit relatively weak, evidence that the reform increased turnout in the long-run.

3.6 Conclusion

We study the effect of abolishing an electoral threshold on vote and seat shares of small parties, thereby drawing conclusions regarding the mechanical and psychological effects of explicit electoral thresholds. Our results indicate that abolishing an explicit electoral threshold increases the seat shares of smaller parties at the expense of more established larger parties. We also find that the increase in the small parties' seat shares is due to changes in voting patterns, which suggests that the reform had substantial psychological effects.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Overall, electoral thresholds appear to have a substantial causal effect on political outcomes. In particular, they seem to be a suitable means to achieve legislative cohesion by ensuring that non-mainstream parties receive only a relatively small share of the seats in the legislature. By the same token, however, electoral thresholds reduce the legislative voice of minorities. Policy makers and voters should therefore be aware that electoral thresholds entail a strong trade off between legislative cohesion and political representation.

Appendix to Chapter 3

3.A.1 Empirical Methodology

This section collects further details on the empirical methodology. Two features of our setting make the identification particularly credible. First, according to a state law, council size in Hesse is a positive discontinuous fuzzy function of municipal population size. As can be seen in Figure 2.1 our sample of 426 municipalities exhibits substantial variation in population sizes.¹³ The law states that council size brackets are determined by the latest available population data when the date for next local election is fixed.¹⁴ Figure A.3.1 plots mean council size for each of the population brackets in the pre- and post-treatment period. It is obvious that mean council size is increasing between the different brackets, with municipalities typically choosing the largest possible council size in the pre-treatment period.¹⁵ More generally, the probability of a larger council increases discontinuously at the population cutoffs. Consequently, the treatment intensity of abolishing the five percent legal threshold also increases discontinuously. Thus, a given municipality with e.g. 3000 inhabitants will be less affected by the abolishment of the legal electoral threshold than municipalities with 3001 inhabitants because the former will choose on average smaller councils and thus have larger implicit thresholds.

The second feature that we use for identification is that given the nature of the electoral reform, we have a pre- and post-treatment period. Since the legal electoral threshold existed only before 2001, we can rely on within- in addition to between-variation along the population brackets for identification.

The presence of these features in our setting, the discontinuities at population cutoffs combined with the electoral reform, enables us to focus on changes in political outcomes at the population cutoffs between the pre- and post-treatment periods, thereby effectively combining RDD and diff-in-diff methods. The RDD aspect of this approach allows us to control for observable as well as unobservable characteristics of municipalities that may result in differential trends and thereby invalidate a traditional diff-in-diff design. The diff-in-diff aspect addresses concerns regarding co-treatment and manipulation at the relevant population cutoffs which may invalidate a standard RDD (Eggers et al., 2015). Combining the RDD and diff-in-diff designs results in a diff-in-disc design, that leads to a “as good

¹³In fact, the smallest municipality in terms of population has as few as 638 inhabitants and the largest 669,992 inhabitants (see summary statistics in Table A.3.3).

¹⁴This population data is not the same as the annual data published by the state statistical office. For the 2006 and 2011 elections, we obtained the relevant data from the homepage of the statistical office. For the previous elections, we collected the data by hand from various issues of the Hessian government gazette.

¹⁵Two municipalities in 1989 have larger council sizes than permissible given their population size (one had 4999 inhabitants and a council size of 31 and the other 9754 and a council size of 37). We dropped these two observations from the sample. While we have no definite explanation, we suspect that these two municipalities made use of an exception defined in the Hessian law for local elections that allows municipalities that crossed either of the thresholds from above to keep the council size intended for municipalities in the next threshold for another legislative period.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

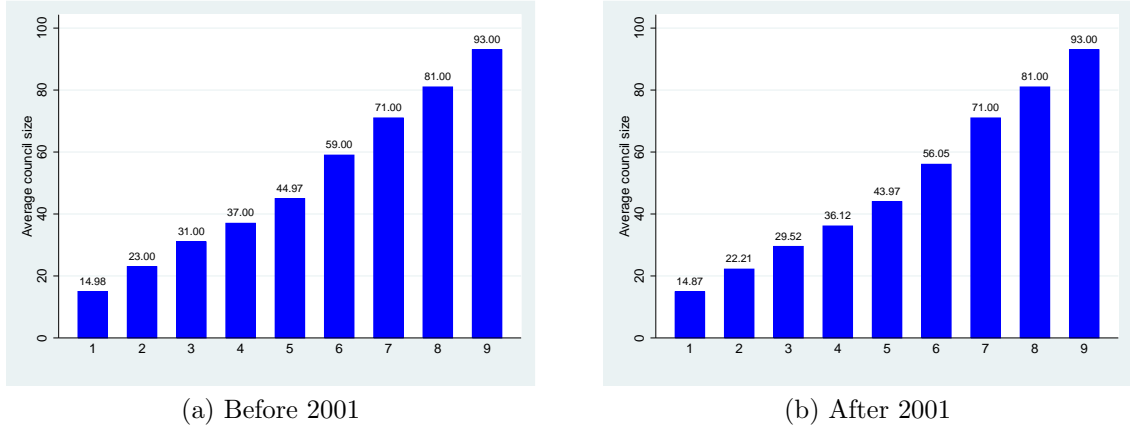


Fig A.3.1: Average Council Size Before and After 2001. This figure shows the average council size of municipalities in population brackets 1-3000 (1), 3001-5000 (2), 5001-10,000 (3), 10,001-25,000 (4), 25,001-50,000 (5), 50,001-100,000 (6), 100,001-250,000 (7), 250,001-500,000 (8), 500,001-1,000,000 (9) prior and after the 2001 elections.

as random” analysis even in settings where either differential trends or co-treatment and manipulation at the population cutoffs cannot be ruled out (Grembi et al., 2015).

In addition, as the abolishment of the legal electoral threshold was the only aspect of the reform having a different impact on municipalities with different population sizes, we are able to isolate its causal effect from the impact of the other changes included in the *Kommunalwahlreform*. Specifically, the estimates are robust to heterogeneous responses to the other changes in the electoral law. Since the analysis is confined to municipalities that are above and below the respective cutoffs, the response to the other changes in “treatment” and “control” municipalities will be identical in expectation.

Finally, a crucial assumption that must hold is that the ability or incentives of municipalities to manipulate population size at the cutoff did not change over the pre- to post-treatment period. This assumption is plausible as it is unlikely that municipalities would persistently misrepresent their population sizes only to avoid being forced to change their council sizes. This is tested in the McCrary (2008) style density plot reported in Figure 2.4. The idea underlying this plot is that if either the ability or the incentives for manipulation changed at the cutoff from the pre- to the post-treatment period, we should observe a discontinuity in the changes in the number of observations close to the cutoff. More specifically, assume that because of the treatment, municipalities systematically start to (mis-) report lower population sizes in order to be able to reduce their council size. Then the increase in observations just before the normalized cutoff from the pre- to the post-treatment period should be significantly higher than the increase in observations just after the cutoff. As the plot shows, there is no evidence that incentives for manipulation changed from the pre- to the post-treatment period at the normalized cutoff.

3.A.2 Robustness Tests

3.A.2.1 Polynomials and Bandwidths

In order to test the robustness of the baseline results we re-estimate our model relying on smaller bandwidths h (0.06, 0.07, 0.08, 0.09, 0.1, 0.2) and a linear polynomial. Regression results are grouped in Table A.3.1. Coefficient estimates for the seat and vote shares of small parties remain consistently positive and are always significant. Conversely, coefficient estimates for the seat and vote shares of medium and large parties are always negative. This robustness test hence supports and reinforces our baseline results. Not only we find significant effects for the average change in vote and seat shares of small and large parties, but also the results remain stronger and of a higher magnitude for vote rather than seat shares.

Table A.3.1: Smaller Bandwidths.

	BW=0.06	BW=0.07	BW=0.08	BW=0.09	BW=0.1	BW=0.2
Small Parties						
Seat Share	4.060 (2.982)	4.817* (2.766)	4.592* (2.396)	5.154** (2.220)	3.634* (2.101)	3.546** (1.664)
Vote Share	4.686 (2.931)	5.303** (2.700)	5.093** (2.356)	5.368** (2.188)	3.843* (2.091)	3.580** (1.639)
Medium Parties						
Seat Share	-0.665 (1.986)	-1.138 (1.816)	-1.646 (1.669)	-1.422 (1.569)	-1.201 (1.513)	-1.485 (1.105)
Vote Share	-0.995 (1.832)	-1.013 (1.655)	-1.551 (1.531)	-1.103 (1.457)	-0.601 (1.408)	-0.777 (1.007)
Large Parties						
Seat Share	-3.395 (2.936)	-3.679 (2.634)	-2.946 (2.336)	-3.731* (2.105)	-2.433 (1.988)	-2.060 (1.617)
Vote Share	-3.687 (2.819)	-4.290* (2.554)	-3.541 (2.275)	-4.265** (2.045)	-3.243* (1.952)	-2.803* (1.562)
N	442	523	591	655	723	1416

Notes: This table presents diff-in-disc regressions for the seat and vote shares of the different groups of parties in Hessian municipal councils. All population cutoffs at which council size is allowed to change are analyzed simultaneously by normalizing population size. Estimates for the average treatment effect of abolishing the election threshold are reported for different bandwidths (0.06, 0.07, 0.08, 0.09, 0.1, 0.2) and a linear polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***).

3.A.2.2 Placebo Tests

As a first set of placebo tests, we let the treatment set in at fake cutoffs and compare the estimated effects with the ones obtained for the correct cutoff. More specifically, we

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

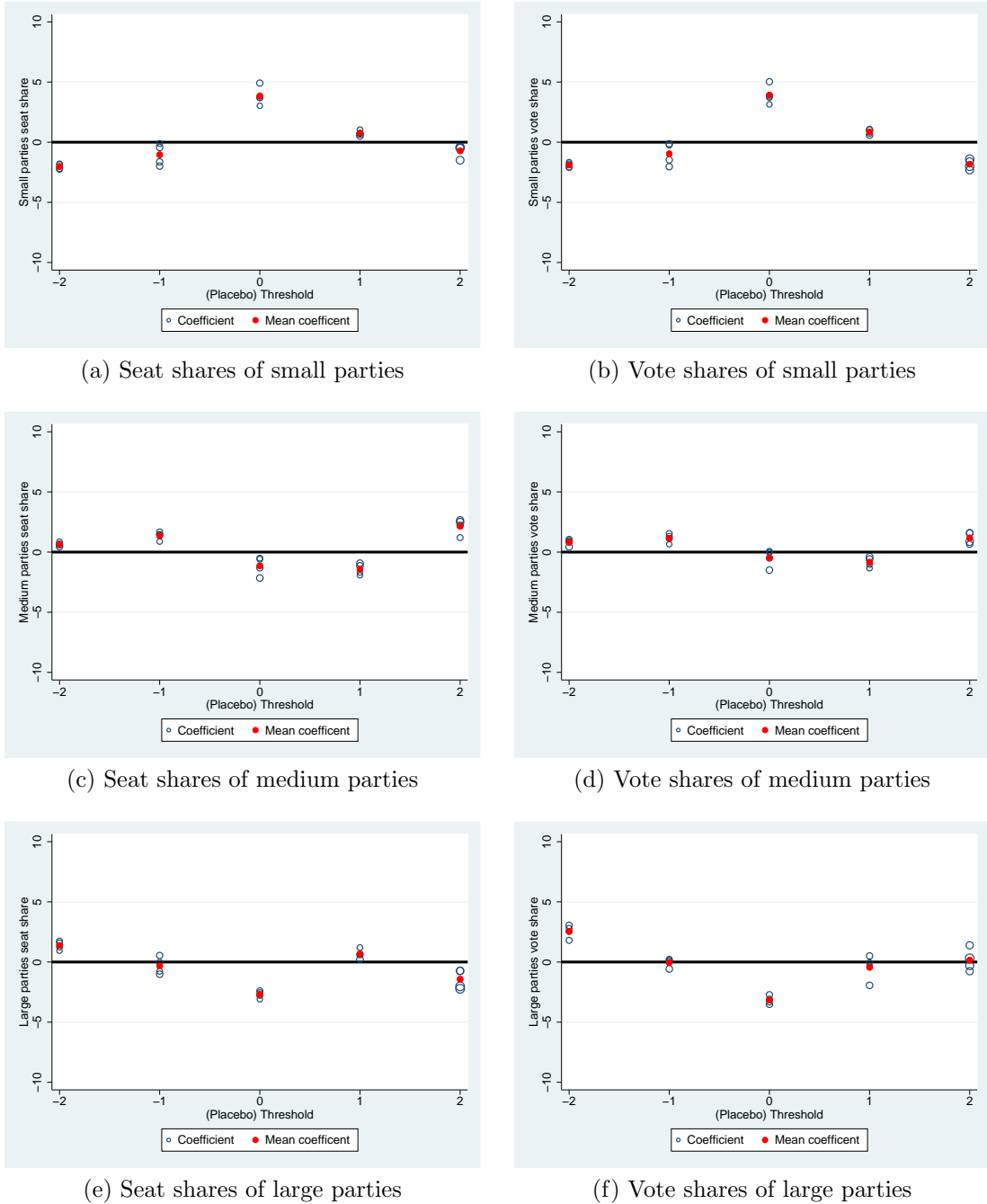


Fig A.3.2: Placebo Treatments with Fake Cutoffs. This figure shows coefficient estimates of the diff-in-disc model for the seat and vote shares of small, medium, and large parties with placebo treatments. The size of the dots indicates the standard error of each estimate. The cutoffs are redefined such that treatment sets in at $npop = -2, -1, 1, 2$. For comparison, the coefficient estimates at the true threshold of 0 are also indicated.

define D_i in Equation (3.1) such that it is 1 if $npop = -2, -1, 0, 1, 2$ ($npop = 0$ indicates the true cutoff). To save space, we summarize the results in graphs. The structure of the plots in Figure A.3.2 is as follows. For each fake cutoff, we plot the four coefficient estimates obtained by combining the bandwidths and polynomial degree used in the baseline regressions. We also indicate the mean value of the coefficient estimates with a red dot.

In Figure A.3.2, we present the placebo estimates for the seat and vote shares of the small, medium, and large parties. For the small parties, subfigures (a) and (b), the mean coefficient estimates revolve around zero at every fake cutoff. At the true cutoff, however, there is a large positive jump of the mean coefficient of the seat and vote shares of the small parties. Subfigures (c) and (d) in turn, show that the mean coefficient estimates for the medium parties revolve around zero for the fake and true cutoffs. Finally, both subfigures (e) and (f), for the large parties, show evidence of a noticeable negative jump in the mean coefficient estimate at the true cutoff, while estimates for the fake cutoffs are on average close to zero.

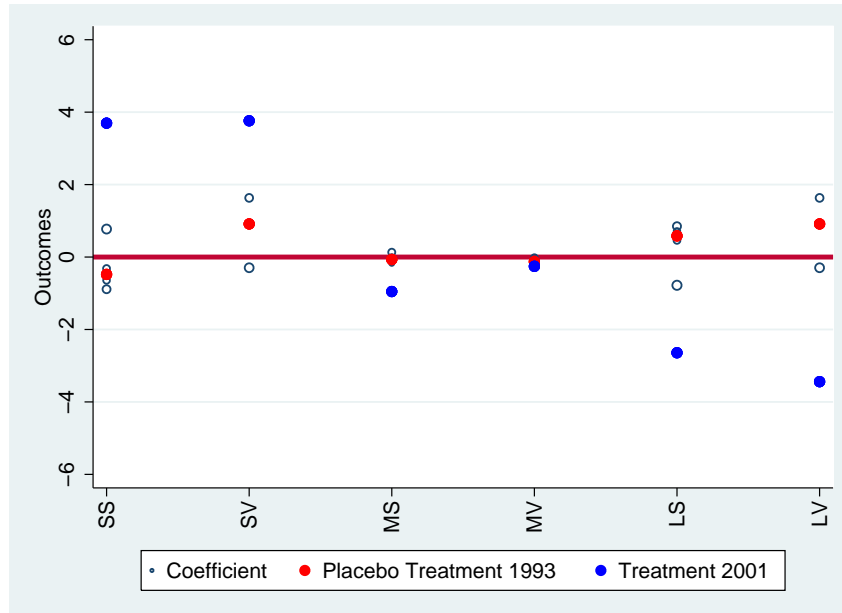


Fig A.3.3: Placebo Treatment for Fake Treatment Year. This figure shows coefficient estimates of the diff-in-disc model with a placebo treatment defined to begin in 1993. The sample covers the period 1989-1997. The size of the dots indicates the standard error of each estimate. Coefficient estimates are reported for small party seat share (SS), small party vote share (VS), medium party seat share (MS), medium party vote share (MV), large party seat share (LS) and large party vote share (LV). The mean estimate at the fake treatment year is indicated with a red dot. The mean estimate at the true treatment year is indicated with a blue dot.

As a second set of placebo tests, we let the treatment begin at a fake treatment year. That is, we limit the sample to the pre-treatment period (1989-1997) and let the treatment begin in 1993. Figure A.3.3 collects the results. The mean coefficient at the fake cutoffs for each of the outcome measures is indicated with a red dot. For comparison, we indicate the mean estimate at the true cutoff with a blue dot. Again, we find that the mean coefficient estimates for the seat and vote shares of the medium parties are around zero both at the

fake and the true treatment year. For the remaining variables, the graph shows that mean estimates for the fake treatment year are always close to zero, whereas estimates for the true treatment year are further away from zero. Overall, both the placebo test for the fake cutoffs and for the fake treatment year provide further robustness to our baseline estimates.

3.A.2.3 Individual Cutoffs

In order to assess whether the baseline results are driven by selected cutoffs, we report results for individual cutoffs. For compactness and since sample sizes are smaller in these regressions we only report results for specifications with a relatively large bandwidth of 0.5 and a quadratic control function. Even though treatment increases with council size, it is at the smallest population brackets that increase in treatment intensity is the largest. Also, the number of observations in the larger brackets is limited. Therefore, here we focus on the first four population cutoffs.

Table A.3.2: Individual Population Cutoffs.

	T=3001	T=5001	T=10,001	T=25,001
Small parties				
Seat share	6.285 (4.139)	3.805 (2.842)	0.798 (2.489)	0.645 (4.407)
Vote share	5.973 (4.060)	3.919 (2.868)	0.926 (2.409)	0.482 (4.300)
Medium parties				
Seat share	-0.498 (2.386)	-1.399 (1.810)	0.646 (2.024)	-1.374 (2.181)
Vote share	-0.243 (2.292)	-1.011 (1.680)	0.625 (1.858)	0.543 (2.058)
Large parties				
Seat share	-5.787 (4.173)	-2.405 (2.874)	-1.443 (2.321)	0.730 (3.543)
Vote share	-5.736 (4.155)	-2.907 (2.761)	-1.548 (2.202)	-1.023 (3.640)
N	625	1105	1122	430

Notes: This table presents diff-in-disc regressions at the individual cutoffs. Estimates for the average treatment effect of abolishing the election threshold are reported for a 0.5 bandwidth and a quadratic polynomial of normalized log population size. All models include municipality and legislative term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels at 10%(*), 5%(**) and 1%(***)

The results are collected in Table A.3.2. Results for the seat and vote shares of small, medium and large parties follow in general the baseline results. These results are never significant at conventional values, but this is presumably due to the smaller sample size in these regressions since the estimated coefficients are numerically large. According to

Baskaran and Lopes da Fonseca (2016) council sizes in municipalities falling into the first population cutoff did not decrease significantly (only larger municipalities reduced their council sizes). It is thus interesting to see that at that same cutoff the increase in the seat share of small parties is more than proportional to the increase in the vote share. The same holds for the decrease in the seat share of large parties which is higher than the decrease in vote share, though to a smaller degree. In the remaining cutoffs where council sizes did significantly decrease we see the contrary.

This analysis for individual cutoffs allows us to interpret the magnitude of the treatment effect. For example, crossing the population threshold of 3001 from below implies on average a reduction in the implicit threshold of about 3.33 to 2.17 percentage points, assuming all municipalities choose the highest possible council size. Hence, the intensity of treatment from abolishing the explicit threshold increases by around 1.16 percentage points at the 3001 threshold. Since the estimate for the treatment effect is around 6 percentage points, a percentage point decrease in the implicit threshold increases the vote share of smaller parties by about 5 percentage points. These values suggest that the psychological effects of electoral thresholds can be substantial.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table A.3.3: Summary Statistics.

Variable		Mean	SD	Min.	Max.	N
Turnout	overall	64.540	12.911	31	93.606	2554
	between		5.560	49.514	79.837	426
	within		11.654	42.579	87.954	5.995
Small party seat share	overall	18.336	15.348	0.000	100.000	2554
	between		13.746	0.000	100.000	426
	within		6.854	-16.447	57.659	5.995
Small party vote share	overall	18.428	15.212	-0.100	100.000	2554
	between		13.645	-0.017	100.000	426
	within		6.754	-16.904	56.178	5.995
Medium party seat share	overall	8.177	7.842	0.000	45.946	2554
	between		6.751	0.000	31.532	426
	within		4.009	-10.318	33.853	5.995
Medium party vote share	overall	8.582	7.771	0.000	46.400	2554
	between		6.823	0.000	30.766	426
	within		3.739	-10.819	33.727	5.995
Large party seat share	overall	73.487	14.434	0.000	100.000	2554
	between		12.594	0.000	100.000	426
	within		7.068	34.164	99.043	5.995
Large party vote share	overall	72.990	14.296	0.000	100.000	2554
	between		12.551	0.000	100.000	426
	within		6.862	35.240	98.373	5.995
Nr. parties	overall	3.782	0.990	1.000	10.000	2554
	between		0.834	1.000	6.833	426
	within		0.535	1.282	7.616	5.995
Max. seat share	overall	47.891	10.314	26.667	100.000	2554
	between		8.828	32.040	100.000	426
	within		5.347	29.612	72.529	5.995
Herfindahl index	overall	62.871	9.335	0.000	100.000	2555
	between		8.448	0.000	100.000	427
	within		4.294	29.601	82.531	5.984
Council size	overall	31.211	9.753	11.000	93.000	2554
	between		9.604	13.667	93.000	426
	within		1.736	24.211	38.211	5.995
Inhabitants	overall	13931.84	36795.72	638	669992	2554
	between		36801.83	727.5	648056.5	426
	within		1084.074	-12745.66	35867.34	5.995

Chapter 4

Candid Lame Ducks*

4.1 Introduction

Electoral accountability is often pointed as the foremost reason for holding periodical elections. The rationale is that the regular assessment of an incumbent's performance provides more than a chance to reward or punish an officeholder's conduct, it gives the possibility to constrain opportunistic behavior. Institutional barriers to re-election such as term limits may appear at odds with this theory of electoral accountability. Lame ducks, i.e. elected officials reaching the end of their tenure, are not electorally held accountable leaving voters unable to exert the appropriate control. It often follows that term-limited incumbents have less incentives to be responsive towards the electorate or responsible in performing their duties. A theory corroborated to some extent by empirical evidence inasmuch as Besley and Case (1995, p. 793) wonder "... why term limits exist at all".

However, and as recent literature suggests, by decreasing the value of holding office term limits may also provide the incentives for more truthful political behavior leading to a better selection effect of elections and a more qualified pool of politicians in the long-run (Smart and Sturm, 2013). The key distinction between the two theories is that one focuses on moral hazard whilst the other is concerned with the problem of adverse selection. By distinguishing politicians solely on the basis of their ability, the first looks at elections as a disciplining device and at term limits as the root of distortionary behavior while the second acknowledges a certain degree of heterogeneity in preferences and motivations and regards term limits as disciplinary.

The conflicting views on the role of elections and politicians' motivations make the quest of identifying the impact of term limits an empirical one. Still, there are only few empirical studies on the impact of gubernatorial term limits on fiscal policy and the existing evidence is shaped mainly by the U.S. experience. This paper attempts to broaden this literature by being one of the first micro-analysis on the impact of gubernatorial term limits on local policy choices in a context other than the U.S. institutional setting. The identification strategy relies on a natural experiment, a recent constitutional reform introducing mayoral

*This chapter originates from work previously circulated as "Lame but loyal ducks" (see Lopes da Fonseca (2015b)).

term limits at the Portuguese municipal level, and a difference-in-differences (diff-in-diff) quasi-experimental approach. The methodological approach is subject to a number of tests that validate its reliability in causally identifying the impact of term limits.

Regression results suggest a significant reduction in the property tax rate as well as a decrease in total current expenditures. Heterogeneous effects of the treatment in turn, indicate that results hinge on partisan affiliation suggesting that officeholders' motivations underlie the fiscal response to term limits. Furthermore, incumbents stop engaging in fiscal policy cycles once they are ineligible for re-election in contrast to eligible mayors who continue to pursue expansionary policies before the elections.

This analysis contributes to the literature by challenging previous results on the impact of term limits on fiscal policy. Results show that officeholders do behave systematically different in view of a binding term limit, but not in a manner indicating lack of effort or opportunism. By providing evidence consistent with distortionary elections and emerging policy preferences in face of ineligibility for re-election, this study suggests a new interpretation of previous evidence in light of Smart and Sturm's (2013) model of truth-inducing term limits.

4.2 Term Limits and Political Incentives

There are two distinct views on the role of competitive elections that produce disparate predictions in the term limits debate. The mainstream view is that elections are a disciplining mechanism that is disrupted by re-election constraints leading to situations of moral hazard (e.g. Barro (1973); Ferejohn (1986)). A different perspective deems elections as distortionary due to myopic or career concerned officeholders. This leads to adverse selection, warranting the introduction of term limits (Chari et al., 1997).

Political agency models capturing the key features of a representative democracy illustrate both views depending on the underlying assumptions on officeholders' characteristics or motivations.¹ Borrowing the principal-agent terminology, the political agent is compelled to act, or build a reputation of acting, in behalf of the principal, i.e. the electoral district, who since Ferejohn (1986) is generally modeled as a performance-oriented electorate, by the democratic institutional mechanism of periodic elections and thereby increase the probability of re-election.

In the context of disciplining competitive elections, fiscal policy outcomes are assumed to measure an incumbent's ability or competence. Attributing a finite and commonly known time horizon to officeholders, determined by the existence of term limits, creates a last-term effect where there is no control over the officeholder's behavior due to the lack of re-election incentives (Barro, 1973). This results in a departure from the optimal policy choice from the voters' perspective.

However, if politicians behave opportunistically in order to be re-elected, in line with reputation-building models predicting potentially negative effects of career concerned and

¹For a review of political agency models see e.g. Besley (2006).

myopic agents (Morris, 2001; Ely and Välimäki, 2003), elections are distortionary. In this context, a system of judicial power where no officeholder may be re-elected, i.e. a one-term limit, can dominate other forms of government (Maskin and Tirole, 2004).

The two theories boil down to one testable hypothesis: that ineligible and eligible incumbents behave differently. Moreover, being fiscal policy the instrument used by politicians to maximize re-election prospects implies that term limits are liable of having an impact on fiscal performance.

Empirical evidence, mostly based on the U.S. experience, appears to support the disciplining effect of elections and the existence of a term-limits-induced moral hazard (Besley and Case, 1995, 2003; Crain and Tollison, 1993; Crain and Oakley, 1995).² There is, nonetheless, one relevant exception: List and Sturm (2006) provide significant evidence of distorting policy choices for a sample of U.S. governors between 1970 and 2000 showing that environmental policy is substantially different in years where a governor may or may not run for re-election, with policy changes strategically hinging on the composition of the electorate.

More recently, the discussion has focused on determining the optimal length of term limits. Results support, in general, longer average tenures. Alt et al. (2011) show that from the perspective of disciplining elections, increasing average tenure matters as voters become able to weed out good from bad politicians. Smart and Sturm (2013) in turn, propose a model with a pool of public-spirited and biased politicians, where even the former are inclined to implement distorted policies in order to increase their re-election probability. In this context, term limits are welfare improving as they decrease the value of holding office, aligning officeholders' incentives and thereby inducing truthful behavior, again enabling voters to better select re-eligible candidates. They support a mix of electoral accountability and judicial power, i.e. two-terms or longer term limits.

This study fits the model of distortionary elections and public-spirited politicians by Smart and Sturm (2013). In addition, it tries to reconcile the early reputation-building models of moral hazard inducing term limits with the present evidence on ideologically driven politicians. Instead of a lack of effort, previous evidence could be interpreted in light of different preferences by politicians and voters that emerge in face of binding term limits. The problem at hands appears to be one of distortionary elections and adverse selection. Incumbents distort fiscal policy in pursue of pleasing the electorate and securing re-election, ambitions that become unimportant upon ineligibility. *Lame duck* terms thus show the incumbents true colors and distortions to the local economy cease. This reasoning is consistent with the electorally-induced political business cycles (PBC) literature.

²Recent evidence on the disciplining effect of elections relies on outcomes other than fiscal policy. See e.g. Ferraz and Finan (2011) and Janvry et al. (2012), that find less corruption and a more effective implementation of a decentralized conditional cash transfer program aimed at reducing school dropout, respectively, in Brazilian municipalities governed by re-eligible mayors.

4.3 Institutional Details

4.3.1 Local Politics and Finances

This paper relies on the Portuguese first sub-national level of government, i.e. municipalities, as a laboratory to study the impact of introducing mayoral term limits on local public accounts.³ In Portugal, there are in total 308 municipalities, the equivalent to a U.S. town/city. The focus is upon the 278 mainland municipalities for increased comparability and due to different institutional details regulating the autonomous regions of Azores and Madeira.

Municipalities are responsible for several aspects of the local public administration from the supply of public goods to territorial organization, promotion of local development and external cooperation.⁴ For the purpose, there are two political institutions at this level, the executive municipal council and the legislative municipal assembly.

The mayor is the top chief executive position in a municipality. He is the head of the municipal council, which is composed by an additional four to ten councilmen depending on the municipality's population size.⁵ The three-terms limitation introduced by the electoral reform concerns the mayoral position. In Portugal however, there are no explicit direct mayoral elections. Instead, the mayor is the first name on the winning list running for the municipal council elections.⁶ For this reason, these are the relevant election results used in the empirical analysis.

Elections for the municipal council are defined exogenously from the perspective of the local authorities. They take place simultaneously countrywide every four years with a proportional representation system of closed lists in place. Parties and independent lists of organized registered electors may contest the elections with seats being distributed according to the D'Hont method. Local politics are dominated by the local branches of the main parties represented in the national parliament. From left to right in the political spectrum these are the Communist Party (PCP), the center-left Socialist Party (PS), the center-right Social-Democrats (PSD) and the right-of-center Popular Party (CDS-PP).

On the financial side, the laws regulating local governments in Portugal are bound by the principle of decentralization. Municipalities have their own assets, finances and personnel. Still, and as is the case with several local governments across Europe, spending decentralization outweighs revenue decentralization. Conditional and unconditional transfers from higher levels of government still represent the main source of municipal funding. However, despite the reduced fiscal autonomy, in the past years there has been an increase in the relative importance of both local taxes and user charges.

³There is a second level of local administrative units in Portugal, the parishes. Currently, there are 3,902 parishes with each municipality consisting of at least one parish.

⁴Law no. 159/99 in *Diário da República* 215, Series I-A, 14th September 1999.

⁵Two exceptions are the Lisbon and Oporto municipalities with 16 and 12 councilmen, respectively.

⁶Even though votes are cast for the party, the electoral campaign is led by the mayoral candidates who are the faces of the different parties.

Municipalities publish their current and capital accounts yearly. The empirical analysis investigates the impact of term limits on the former as these are of an operative and year-by-year nature.⁷ On the revenue side the focus is on the few items entirely under the control of the municipality: user charges and the local property tax rate.⁸ User charges are set on a yearly basis by the municipality. However, the scope of services provided by the municipalities is very heterogeneous with no systematic records on the type of service and prices charged. Relying on revenues from fines and fees as a proxy gives the possibility to indirectly infer upon the impact of term limits on the level of user charges in the municipality. The local property tax rate for urban properties in turn, is since a reform in 2003 set within a range of 0.3% to 0.5% at the beginning of each year by the municipality.⁹ Additionally, local tax revenues as a whole are also analyzed in order to determine whether any change to the local property tax rate had an effect on the overall tax collection. Finally, on the expenditure side, the focus is on the impact of a binding term limit on current aggregate expenditures which primarily encompass expenditures with personnel and current transfers to parishes.

4.3.2 Electoral Reform

On July 25th 2005, the draft Law on the implementation of term limits for local office-holders was discussed and approved in Parliament leading up to the Law no. 49/2005 from August 29th 2005, which entered into force on January 1st 2006. The law sets a three consecutive terms limit for mayors after which they are not allowed to rerun for the mayoral position in the same municipality. However, upon entering into force it established that all incumbent mayors could rerun in the following 2009 elections. As a result, term limits were only first binding in the 2013 local elections for incumbent mayors serving their at least third consecutive term. In total, 150 mayors were in this situation in 2013 leaving the same number of municipalities to have exogenously determined open-seat elections for the municipal council.

The timeline of the reform creates an interesting natural experiment. First, the law was voted right before, and came into force right after, the 2005 local elections held in October. Second, upon entering into force in 2006 it allowed all incumbent mayors one last chance at re-election in the 2009 local elections, creating a stand-by period where the law was already in existence but not yet effective. Third, it was finally binding in the 2013 local elections in 150 municipalities. As only around half of the mayors were bound by a term limit it is possible to rely on both between- and within-municipality variation to identify the impact of gubernatorial term limits on local policy choices.

⁷Capital accounts often involve long-term commitments and are thus less susceptible to change.

⁸Mayors also set the local corporate income tax in a range of 0-1.5% since a reform in 2007. However, due to the small time span of data available this outcome is not considered in the empirical analysis. For the remaining municipal taxes the rate and/or base is set by higher tiers of government and thus exogenous to local authorities.

⁹The local property tax rate for rural properties is fixed at 0.8%.

4.4 Data

This analysis relies on a dataset combining fiscal variables and electoral results at the Portuguese municipal level for the past three complete electoral terms, i.e. 2002-2005, 2006-2009 and 2010-2013.¹⁰ The National Electoral Commission's (*Comissão Nacional de Eleições*) and the General Directorate for Internal Affairs's (*Direcção Geral da Administração Interna*) websites provide the data on electoral results. Data are provided at the party level per municipality and consist on the number of votes and seats allocated to each party. In order to ascertain the incumbency status of the mayor the data on electoral results was manually matched to the names of the members of the municipal council – in particular the mayor – provided in the Official Map by the National Electoral Commission published in *Diário da República*.

Data on local public finances are publicly available and can be retrieved from the General Directorate for Local Authority's (*Direcção Geral das Autarquias Locais*) website for the years from 2003 onward. For the previous years it is provided in the institution's annual publication entitled *Municipal Finances (Finanças Municipais)*. The variables measuring tax revenues, revenues from fines and fees and aggregate current expenditure are deflated to the year 2005 by the national consumer price index from the World Economic Outlook Database of the International Monetary Fund. Per capita values are obtained relying on resident population per municipality from the Portuguese National Statistics Institute (INE).

For robustness, the analysis relies on a number of control variables. The vector of control variables includes measures of municipal population size and municipal economic activity, as well as political dummies indicating whether there is a majority in the municipal assembly, whether the same party controls both the council and the assembly and if the mayor is left-leaning.¹¹ The political variables are constructed based on the local electoral results for both the municipal council and assembly from the sources mentioned above. Municipal population size coincides with the resident population per municipality series from INE and the measure for municipal economic activity is proxied by night light output over the years under study for each municipality. This variable is constructed by combining the following two databases in ArcGis: geodata on Portuguese municipal boundaries and location from The Global Administrative Areas database GADM, i.e. a spatial database collecting the location of the world's administrative boundaries, and data on night light output collected by the Defense and Meteorological Satellite Program satellites and treated

¹⁰The decision to analyze this period relates to the nature of the reform that determines a clear pre-treatment period, before 2005, a stand-by period between 2006 and 2009 and a post-treatment period, the electoral term 2010-2013 leading up to the term limited elections. Including only the three electoral terms from 2002 onward, is on the one hand enough to perform the intended analysis and on the other hand more robust to long term issues of co-treatment and past dynamics. Results relying on all electoral periods since 1994 support the evidence presented in the following sections. Before 1994 there is no systematic data on the composition of the municipal council.

¹¹The mayor is considered left-leaning if he is a representative of either the PS or PCP, or smaller parties known to be on the left side of the political spectrum.

and made available by the Earth Observation Group from the National Geophysical Data Center of the National Oceanic and Atmospheric Administration.

4.5 Methodology

4.5.1 Empirical Model

The variation in eligibility for re-election introduced by the law limiting the number of consecutive mandates is susceptible of causing an exogenous variation in short-term electoral incentives. Provided that treatment and control municipalities are comparable, it is possible to capture this variation and identify the causal impact of term limits on policy choices simply by assessing the difference in the change in fiscal outcomes for the two groups of municipalities from the pre- to the post-treatment period. Evidence of significant post-treatment group-specific fiscal policy choices points to a causal effect of treatment that resulted in a systematically different behavior between term-limited and re-eligible incumbents.

The nature of the constitutional reform under study establishes pre- and post-treatment periods that allow for a quasi-experimental diff-in-diff approach. As of now, there is only one wave of term limited elections. These took place in municipalities re-electing incumbents into office in the 2009 elections to serve their at least third electoral term. As the focus is on implementation effects, i.e. the impact of a binding term limit, treatment is implicitly assigned by the 2009 local elections results.

Let C_i be the dummy indicating treatment, equal to one for all municipalities that elected a lame duck in 2009 and zero otherwise.¹² Treatment assignment occurs at t_0 , corresponding to the 2009 local elections. Yet, 2009 is still part of the pre-treatment period as fiscal policy is decided in the yearly budget prepared at the beginning of each year. The post-treatment period starts in 2010, with the new local executive taking office and passing the first budget. Accordingly, let d_t be a time dummy that switches to one the year after treatment assignment, i.e. $d_t = 1[t > t_0]$. Meaning d_t indicates the four years of the 2010-2013 electoral term.

Inference on the average treatment effect of term limits on local policy choices is based on the following general diff-in-diff regression model:

$$Y_{it} = \gamma_i + \gamma_t + \delta (C_i \cdot d_t) + Z'_{it}\beta + \epsilon_{it} \quad (4.1)$$

where Y_{it} is any of the outcome variables under study, $B_{it} = (C_i \cdot d_t)$ indicates a binding term limit and Z'_{it} is the vector of socio-economic and political control variables described in the previous section. The parameter δ measures the average treatment effect of term limits on the different outcome variables for the entire electoral term.

The model is fully identified by including municipality and year fixed effects, γ_i and γ_t , respectively. For robustness, more conservative versions of the baseline model include

¹²In the following treatment is assumed to fall upon the mayor, council or municipality interchangeably.

district trends, $\gamma_s \cdot t$, and district-year fixed effects, γ_{st} .¹³ The former to control for district-specific trends and the latter to allow for unobservable district-specific variables to vary over time. Particular validity and robustness tests also include a linear time trend, t , and a treatment group specific time trend, $C_i \cdot t$.

In addition, the pattern of lagged effects is also of interest as it often provides further and more insightful information on the dynamics of the treatment effects. Therefore, average annual treatment effects are assessed through the following extension of the baseline regression model:

$$Y_{it} = \gamma_i + \gamma_t + \sum_{\tau=1}^m \delta_{-\tau} B_{i,t-\tau} + Z'_{it}\beta + \epsilon_{it} \quad (4.2)$$

where the sum allows for m lags or post-treatment effects. The number of lags is $m = 4$, one for each of the four years of the 2010-2013 electoral term. The remaining variables are defined as before.

Finally, for the purpose of studying heterogeneous treatment effects the regression models in Equations (4.1) and (4.2) are extended to encompass both a second dummy variable and its interaction with the binding term limit. Average term effects are obtained within the following regression framework:

$$Y_{it} = \gamma_i + \gamma_t + \delta B_{it} + H_{it}(\alpha + \rho B_{it}) + Z'_{it}\beta + \epsilon_{it} \quad (4.3)$$

where H_{it} is the additional dummy variable. Average yearly effects in turn, are estimated in the context of the regression model in Equation (4.4), with all variables defined as before.

$$Y_{it} = \gamma_i + \gamma_t + \sum_{\tau=1}^m \delta_{-\tau} B_{i,t-\tau} + H_{it}(\alpha + \sum_{\tau=1}^m \rho_{-\tau} B_{i,t-\tau}) + Z'_{it}\beta + \epsilon_{it} \quad (4.4)$$

This study focuses on one source of heterogeneous effects. In order to identify possible mechanisms behind any significant effect of the term limits treatment on local fiscal policy, partisan affiliation is introduced in the analysis with H_{it} indicating left-leaning mayors. Additionally, the heterogeneous effects regression models above are also estimated as an extension of the baseline results in Section 4.7.2 to analyze the resigning mayors phenomenon taking place in the electoral term 2010-2013.

4.5.2 The Common Trends Assumption

Internal validity of a diff-in-diff framework hinges on the common trends assumption. In short, the trend in each of the dependent variables under study must be the same for all municipalities in the absence of treatment. This assumption can be tested through different procedures. For the purpose, in the following, the sample is restricted to the pre-treatment period, i.e. 2002-2009.

¹³Municipalities are classically grouped into 18 districts created in 1835. Districts are not an official local administrative unit. Still, they group to a certain extent similar municipalities.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

One common approach is to compare the evolution of the different outcome variables in treated and control municipalities during the pre-treatment period (Angrist and Pischke, 2009, p. 231). Figure 4.1 provides mean plots for the four fiscal variables under study: revenue from fines and fees, the property tax rate, tax revenues and total current expenditure. With the exception of the initial distinct trends in the collection of revenue from fines and fees, the graphs do not provide substantive evidence of differential trends between treatment and control councils capable of undermining the empirical design.

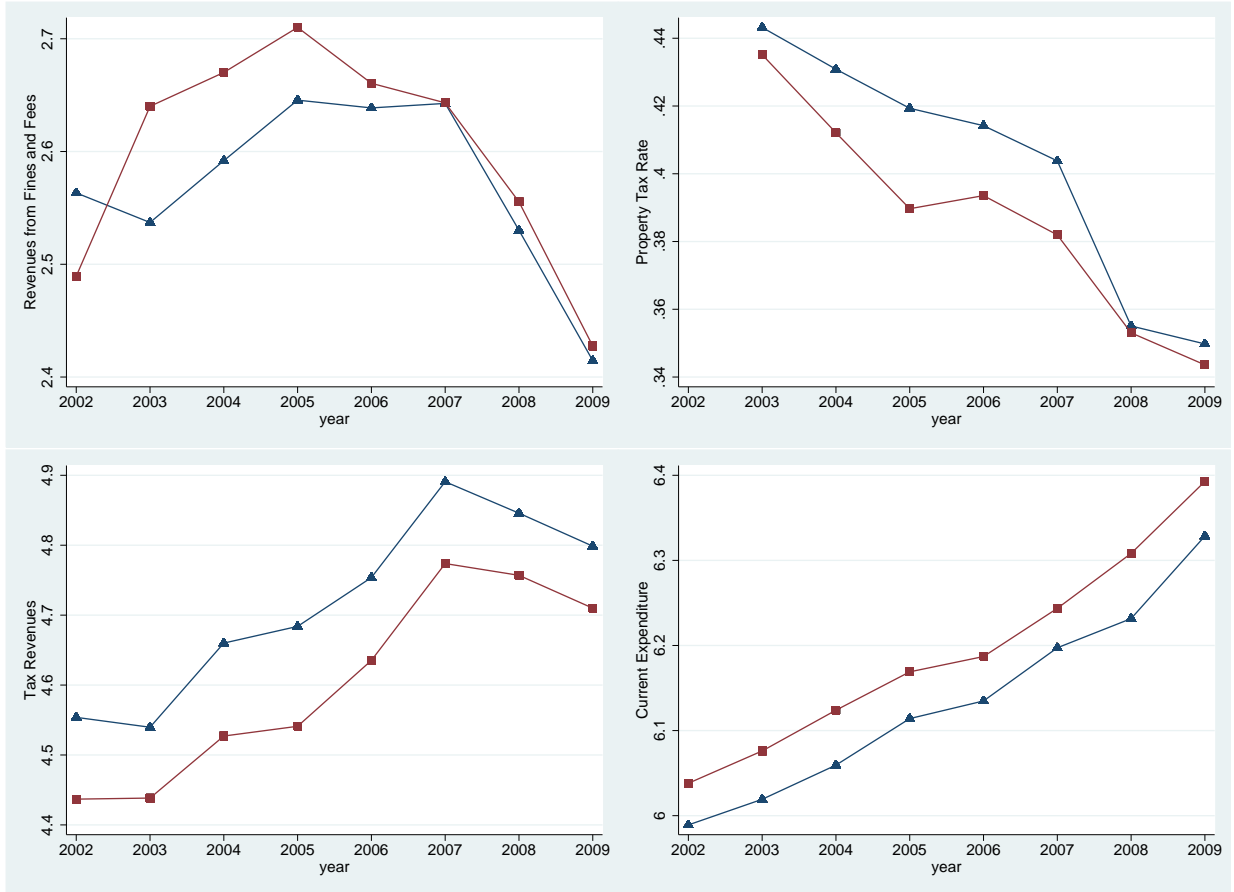


Fig 4.1: Mean Plots. This figure provides plots depicting the logarithm of the mean of each of the fiscal outcomes under study in treated (square) and control (triangle) municipalities for the pre-treatment period, 2002-2009. Data on the local property tax rate is only available from 2003 onward due to a fiscal reform.

A second approach consists in regressing the different outcome variables on yearly dummies indicating the treatment group (Moser and Voena, 2012). Similar to Equation (4.2) yet, instead of lags, this test estimates leads of the treatment variable, i.e. $\sum_{\tau=1}^q \delta_{+\tau} B_{i,t+\tau}$. The number of leads, $q = 8$, corresponds to each year of the pre-treatment period. The omitted category is 2004, the year before the law limiting mayoral terms came to a vote in Parliament, and the model specification includes municipality and year fixed effects. Coefficient estimates in Figure 4.2 measure how outcome variables differ between treatment and control municipalities. As suggested by the previous test, the only significant difference

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

relates to the revenues from fines and fees for the year 2002. Evidence on the remaining outcome variables shows that, on average, treatment and control councils did not execute significantly different fiscal decisions in the pre-treatment period.

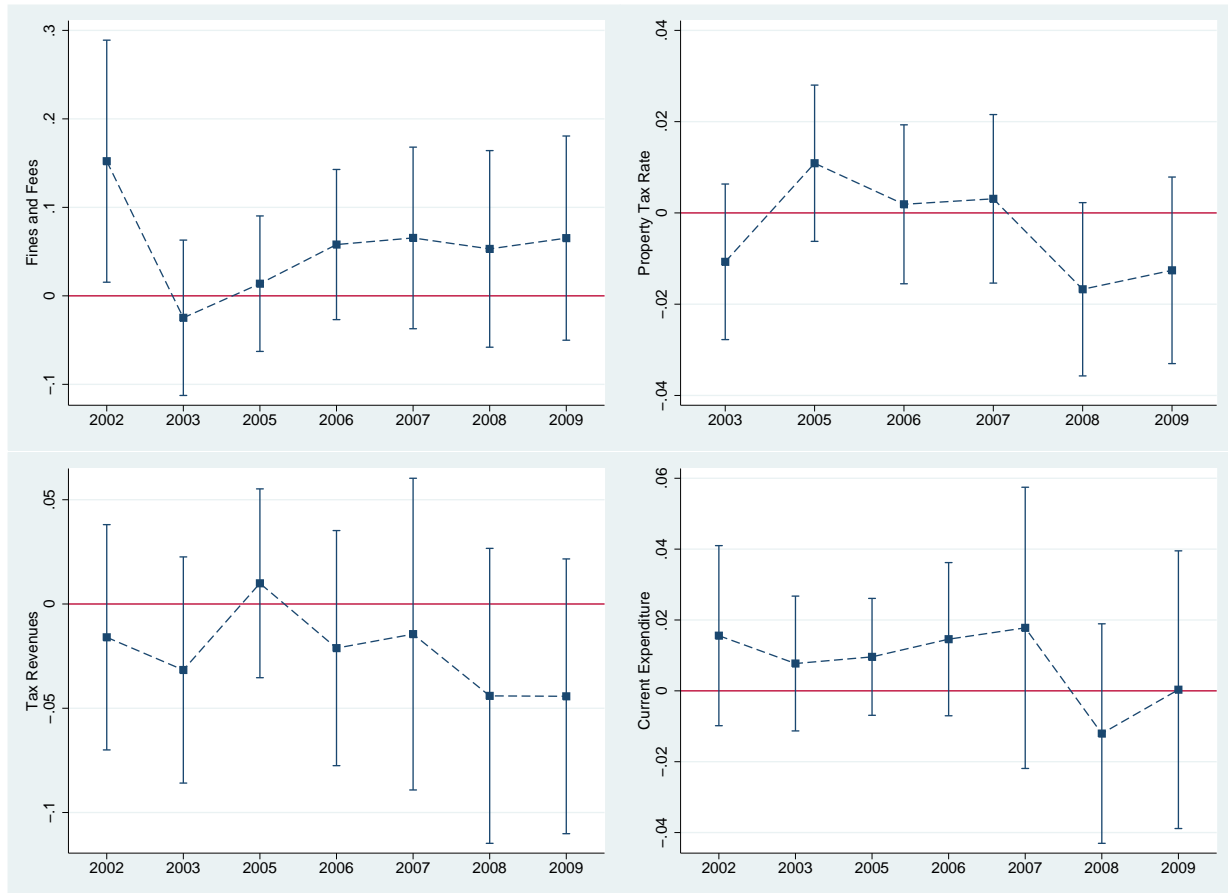


Fig 4.2: Yearly Dummies. This figure provides plots depicting the coefficient estimates for yearly dummy variables indicating the treatment group for the pre-treatment period, 2002-2009. Data on the local property tax rate is only available from 2003 onward due to a fiscal reform. Coefficients are obtained in a model controlling for municipality and year fixed effects. Caped lines indicate 95% confidence intervals.

Finally, Figure 4.3 provides a third and last standard test of the common trends assumption. Each outcome variable is regressed on a linear time trend and a group-specific trend (De Jong et al., 2011). As before, the model specifies both year and municipality fixed effects. Coefficient estimates assess the existence of a pre-existing differential trend in municipalities subject to treatment. All estimates are insignificant.

All in all, the three tests support the assumption of common trends in the fiscal outcomes under study for treatment and control municipalities. In light of these tests however, results on the impact of term limits on the revenues from fines and fees should be taken with a grain of salt whenever the model does not control for differential trends. For the remaining fiscal variables, diff-in-diff coefficient estimates are assumed to capture the causal effect of treatment.

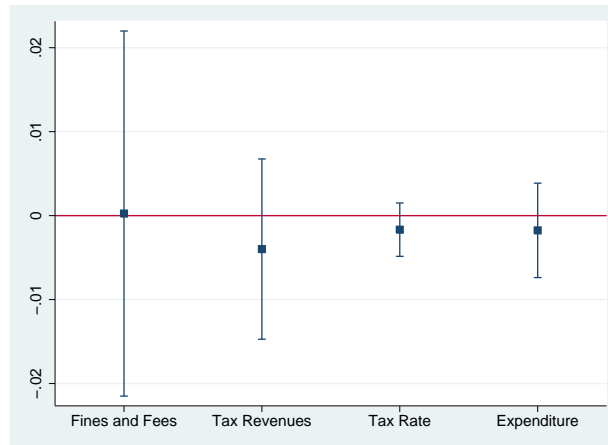


Fig 4.3: Group-specific Trends. This figure provides a plot depicting the coefficient estimates for a variable capturing group-specific trends in the pre-treatment period, 2002-2009. Coefficients are obtained in a model controlling for a linear time trend and municipality and year fixed effects. Caped lines indicate 95% confidence intervals.

4.6 Empirical Evidence

4.6.1 Electoral Term and Yearly Effects

Table 4.1 collects the diff-in-diff results for the average treatment effect of term limits on fiscal policy decisions for the whole 2010-2013 electoral term. Coefficient estimates are obtained from estimating different variations of the diff-in-diff model in Equation (4.1). Model (1) provides the baseline relying solely on municipality and year fixed effects, while models (2) and (3) test the robustness of the results to district-specific time trends and district-year fixed effects. Specifications (4), (5) and (6) replicate (1), (2) and (3) including the vector of control variables. Both the magnitude and significance of the coefficient estimates is consistent across all specifications indicating their robustness to omitted variables (Altonji et al., 2005).

Results suggest that lame ducks, on average, significantly reduce both the local property tax rate and aggregate current expenditures. Coefficient estimates point to a 0.013 percentage points reduction in the tax rate and a decrease of 3.6% (approx. \$ 780.000 for the average municipality) in total current spending.¹⁴

Accordingly, the average treatment effect on tax revenues is consistently negative. Estimates are however never significant. Tax revenues depend on a number of other municipal taxes and therefore do not mimic the movements of the property tax rate alone. Increases or decreases in this tax are thus susceptible of having no significant impact on the whole of tax revenue collection.

User charges also appear to decrease in the aftermath of treatment. Revenues from fines and fees, used as a proxy for the change in user charges, are negative though insignificant. It can be that lame ducks changed specific user charges with not enough economic impact

¹⁴See summary statistics in Table 4.11.

Table 4.1: Term Treatment Effects.

	(1)	(2)	(3)	(4)	(5)	(6)
Fines & fees	-0.079 (0.058)	-0.056 (0.054)	-0.054 (0.055)	-0.078 (0.058)	-0.060 (0.008)	-0.059 (0.055)
Tax rate	-0.014** (0.006)	-0.012** (0.006)	-0.012** (0.006)	-0.014** (0.006)	-0.012** (0.006)	-0.013** (0.006)
Tax revenues	-0.029 (0.030)	-0.013 (0.023)	-0.018 (0.023)	-0.026 (0.025)	-0.013 (0.021)	-0.016 (0.020)
C. Expenditure	-0.036** (0.016)	-0.034** (0.014)	-0.034** (0.014)	-0.036** (0.015)	-0.036*** (0.013)	-0.036*** (0.013)
District trends	No	Yes	Yes	No	Yes	Yes
District x year	No	No	Yes	No	No	Yes
Controls	No	No	No	Yes	Yes	Yes
Observations	3333	3333	3333	3297	3297	3297

Notes: This table presents the baseline results on the average treatment effect of a binding term limit on fiscal policy. The left column lists the dependent variables. Results are obtained from the estimation of Eq. (4.1). All estimates include municipality and year fixed effects. Model (2) adds district-specific time trends and model (3) district-year fixed effects. Models (4), (5) and (6) replicate (1), (2) and (3), respectively, adding the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

on the entire revenue item, similar to the relation between the property tax rate and the resulting collected tax revenues. On the other hand, the overall revenue from fines and fees is already small when compared to other sources of municipal funding, not leaving much leeway for incumbents to significantly decrease it further given the available instruments.

The annual average treatment effects in Table 4.2 in turn, obtained from estimating Equation (4.2) for the most conservative specification, i.e. model (6), are in line with the previous assessment and provide additional information as to the timing of the mayoral response to term limits.¹⁵ The yearly treatment effects show that the significant drop in the property tax rate occurs in the last two years of the electoral term, whilst the drop in current expenditure happens throughout the first three years. In fact, in the election year lame ducks no longer spend significantly less than re-eligible mayors. Either all mayors choose to spend less in the year leading up to the election or lame ducks use the cushion created in the preceding years to restore expenditure levels right before the election.

In addition, although the average electoral term treatment effect on tax revenues is insignificant, the study of annual effects provides evidence of a significant decrease in tax revenue collection in the election year. This comes presumably as a result of the significantly lower property tax rate in the later years of the electoral term.

Overall, the present results appear in contrast to previous evidence in the literature. Rather than having a positive and significant effect on spending (Besley and Case, 1995, 2003) and tax revenues (Alt et al., 2011), term limits lead to a general reduction in current

¹⁵Results are robust to the different specifications.

Table 4.2: Annual Treatment Effects.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010	-0.020 (0.051)	-0.004 (0.006)	0.004 (0.022)	-0.030* (0.016)
2011	-0.029 (0.061)	-0.007 (0.007)	0.007 (0.023)	-0.044*** (0.016)
2012	-0.069 (0.067)	-0.019** (0.007)	-0.009 (0.023)	-0.036** (0.015)
2013	-0.116 (0.073)	-0.020** (0.008)	-0.065** (0.030)	-0.033 (0.020)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the baseline results on the annual treatment effect of a binding term limit on fiscal policy. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.2). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

expenditure and the property tax rate, with lame ducks pursuing a more conservative fiscal policy. Still, evidence indicates a possible term-limits-induced fiscal policy cycle as argued in Besley and Case (1995). Section 4.7.1 studies whether that is the case in the present institutional setting. Nevertheless, average electoral term effects show that overall, both spending and tax rates are on average significantly lower in treated municipalities.

There are multiple possible explanations for the more conservative behavior exhibited by lame ducks. Even though certain hypothesis coming from the standard literature on term limits such as lack of effort or opportunistic behavior do not appear to fit the present evidence, theories base on the increased experience of long-term officials could motivate the assessed outcomes. It is however not clear what would be the incentives behind such commitment in the face of a binding term limit except for perhaps party loyalty or reputational concerns. In fact, 97 of the 150 termed-out mayors continued to pursue a political career closely related to public administration. The study of heterogeneous effects in the following section is meant to clarify the mechanisms behind these results.

4.6.2 Heterogeneous Effects

In order to identify the motivation behind the lame ducks' response to term limits, this section analyzes the presence of heterogeneous effects. For the purpose, inference is based on the estimation of both Equations (4.3) and (4.4) with H_{it} identifying different non-overlapping groups, in particular left- from right-leaning lame ducks.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Testing for ideologically motivated treatment effects provides a solid ground to infer upon a number of possible mechanisms beyond political preferences. If left- and right-leaning term-limited officeholders behave differently, encompassing explanations, as experience, party loyalty, or lack of effort, do not justify the treatment effects adequately. Indeed, in case any of these mechanisms is behind the assessed results, these should on average not significantly hinge on partisan affiliation. If there are significant heterogeneous effects however, the causal effect of term limits is presumably working through politicians' intrinsic preferences or motivations.

Results collected in Tables 4.3 and 4.4 show significant evidence of heterogeneous effects contingent on party affiliation. The fiscal policy choices of right-leaning ineligible mayors appear to be driving the baseline results, with coefficient estimates for left-leaning lame ducks being either insignificant, in the case of the property tax rate, or symmetric to the previously assessed results.

Table 4.3: Heterogeneous Term Treatment Effects.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010-13	-0.071 (0.061)	-0.0106 (0.00709)	-0.00522 (0.0244)	-0.0598*** (0.0149)
H_{it}	-0.006 (0.037)	0.0152** (0.00731)	0.0144 (0.0161)	-0.0117 (0.00994)
$H_{i,t>t_0}$	0.032 (0.075)	-0.00169 (0.00731)	-0.0190 (0.0281)	0.0510*** (0.0162)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on average heterogeneous treatment effects of a binding term limit on fiscal policy when a lame duck is left leaning. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.3). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

That right-leaning lame ducks are responsible for the fiscally conservative response to term limits identified in the baseline results is consistent with the findings in Besley and Case (1995) attributing the loose fiscal policy caused by binding term limits to the Democrats. Taken together, one possible explanation for both results is that when reputation becomes less important due to ineligibility, rather than less effort as suggest in Besley and Case (1995), officeholders act more truthfully as modeled in Smart and Sturm (2013). Ideology may show once mayors are no longer electorally held accountable.

This possible mechanism is in addition, in line with evidence suggesting that the U.S. electorate is fiscally conservative (Peltzman, 1992). The rationale is as follow, given a

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table 4.4: Heterogeneous Annual Treatment Effects.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010	-0.006 (0.054)	0.000 (0.006)	0.028 (0.025)	-0.054*** (0.019)
2011	-0.055 (0.069)	-0.002 (0.007)	0.014 (0.028)	-0.073*** (0.018)
2012	-0.079 (0.073)	-0.020** (0.009)	0.004 (0.027)	-0.059*** (0.018)
2013	-0.145* (0.082)	-0.021** (0.010)	-0.067* (0.035)	-0.054** (0.025)
H_{it}	-0.006 (0.037)	0.015** (0.007)	0.014 (0.016)	-0.012 (0.010)
$H_{i,2010}$	-0.021 (0.068)	-0.008 (0.007)	-0.048 (0.031)	0.051*** (0.018)
$H_{i,2011}$	0.063 (0.084)	-0.010 (0.007)	-0.011 (0.031)	0.064*** (0.021)
$H_{i,2012}$	0.028 (0.092)	0.004 (0.010)	-0.025 (0.031)	0.048** (0.020)
$H_{i,2013}$	0.061 (0.096)	0.007 (0.011)	0.009 (0.040)	0.041 (0.025)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on annual heterogeneous treatment effects of a binding term limit on fiscal policy when a lame duck is left leaning. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.4). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

conservative electoral district, re-eligible incumbents choose on average more conservative policies to boost their re-election probability. Once faced with binding term limits, Democrats choose unconstrained optimal fiscal policies that are on average looser than Republicans'. By the same token and relying on the evidence above, the Portuguese electorate appears to be fiscally liberal. Thus right-leaning ineligible incumbents feel free to pursue a more conservative fiscal policy. Furthermore, this theory would also be in line with the fact that the median voter in Portugal is likely poorer than in the U.S. and thus more prone to favor redistributive policies.

4.7 Extensions

4.7.1 Political Business Cycle

Evidence from the PBC literature usually identifies an electorally induced expansionary policy in the year(s) preceding the election financed by a recessionary beginning of the term. Even though the baseline results in the previous section provide evidence of a similar pattern for aggregate current expenditures, in order to assess whether term limits introduced a fiscal cycle in the current accounts of treated municipalities it is necessary to identify the general fiscal policy pattern in control municipalities. Hence, this section re-analyzes the impact of term limits on fiscal outcomes controlling for the specific year of the electoral term so that it is possible to identify any significant changes to the pattern of fiscal policy in treated municipalities. Inference is based on the following regression model:

$$Y_{itj} = \gamma_i + \gamma_t + \sum_{j=1}^4 ey_j(\gamma + C_i + \sum_{\tau=1}^m \delta_{-\tau} B_{i,t-\tau}) + Z'_{it}\beta + \epsilon_{it} \quad (4.5)$$

where $j = 1, \dots, 4$ indexes the four years in an electoral term. The dummy variables ey_j indicate each year of the electoral term and the resulting γ_j parameter estimates can be thought of as electoral year fixed effects. The second interaction controls for group-specific patterns in fiscal policy decisions for each specific year of the electoral term. Finally, the third interaction identifies the change in decision patterns from pre- to post-treatment period in treated municipalities controlling for the electoral year. In other words, the δ parameters estimate how policy decisions of lame ducks compare with the ones of re-eligible incumbents for the same year of the electoral term.

Regression results obtained from estimating Equation (4.5) indicate whether term limits introduced or simply altered a pre-existing cycle in fiscal policy. Table 4.5 collects the coefficient estimates for the most conservative specification.¹⁶ The first year of the electoral terms is the omitted category.

Altogether, results suggest that rather than introducing a fiscal cycle, term limits mainly induce a change in how the increase in current aggregate expenditure in election years is financed. Evidence points to the existence of a PBC in both treated and control municipalities, with mayors increasing expenditures in election years independently of eligibility for office. However, re-eligible officeholders finance this increase in expenditure through higher tax revenues whilst lame ducks appear to rely on the cushion created by lower spending on the first years of the electoral term.

4.7.2 Mayoral Resignation

In the 2010-2013 electoral term, as in previous terms, a number of mayors resigned from office. Yet, during this particular term 24 mayors stepped down in comparison to five and

¹⁶Results are robust to the different specifications.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table 4.5: Political Business Cycle.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
γ_2	-0.064 (0.125)	0.067* (0.036)	-0.057 (0.036)	0.009 (0.023)
γ_3	-0.050 (0.082)	-0.040 (0.034)	-0.070 (0.044)	-0.047 (0.029)
γ_4	-0.017 (0.064)	0.089* (0.050)	0.167*** (0.042)	0.082*** (0.028)
$\gamma_2 \cdot C_i$	-0.064* (0.037)	-0.001 (0.005)	-0.005 (0.020)	0.007 (0.010)
$\gamma_3 \cdot C_i$	-0.045 (0.036)	-0.005 (0.005)	0.006 (0.016)	-0.007 (0.009)
$\gamma_4 \cdot C_i$	-0.037 (0.036)	0.004 (0.003)	0.010 (0.015)	-0.001 (0.009)
$\gamma_2 \cdot B_{it}$	0.005 (0.070)	-0.006 (0.008)	0.014 (0.029)	-0.047*** (0.018)
$\gamma_3 \cdot B_{it}$	-0.055 (0.068)	-0.014* (0.008)	-0.013 (0.025)	-0.027 (0.016)
$\gamma_4 \cdot B_{it}$	-0.109 (0.072)	-0.024*** (0.008)	-0.073** (0.029)	-0.029 (0.019)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on the treatment effect of a binding term limit on the pattern of fiscal policy. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.5). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

eleven during the previous 2006-2009 and 2002-2005 electoral terms, respectively. Moreover, all resigning mayors were lame ducks.

It appears as though a number of officeholders, faced with the impossibility of running for re-election, strategically decided to resign and give their successor in the party list a chance to take the lead. In fact, when a mayor resigns, the vice-president, a fellow councilman often from the same party, steps in and assumes office until the next election. In 20 out of the 24 municipalities where a lame duck resigned, the substitute mayor ran for the top chief executive position in the 2013 local elections.

Including this strategic response by officeholders into the analysis as a second treatment shows whether results are somehow driven by this small sample of municipalities or if there are significant heterogeneous effects that can shed more light into the underlying mechanics

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

of term limitations. By letting H_{it} in Equations (4.3) and (4.4) indicate every municipality-year observation where a mayor resigned from office, it is possible to compare fiscal policy decisions by re-eligible incumbents, lame ducks, and eligible substitute officeholders.

Most resignations took place in the year preceding the election meaning that up to that point all term-limited municipalities were under the control of a lame duck, while in the election year substitute mayors were already in office. Thus in the following it is assumed that coefficient estimates for the first three years of the electoral term provide heterogeneous effects for resigning lame ducks, and election year coefficient estimates refer to policy choices by substitute eligible mayors.

Term and annual average treatment effects are collected in Tables 4.6 and 4.7, respectively. Eligible substitute mayors behave no different than the average re-eligible mayor. Their incentives should indeed be the same provided eligible officeholders on average run for the elections, as is most often the case. Also resigning and non-resigning term-limited incumbents appear to face similar incentives inasmuch as both significantly decrease aggregate current expenditure. Resigning mayors however, do it to a larger extent. This result is presumably due to the fact that the majority of resigning lame ducks belong to parties on the right side of the political spectrum.

Table 4.6: Resigning Mayors – Term Treatment Effects.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010-13	-0.058 (0.057)	-0.013** (0.006)	-0.006 (0.021)	-0.032** (0.013)
H_{it}	-0.063 (0.073)	0.001 (0.012)	-0.015 (0.038)	0.014 (0.019)
$H_{i,t>t_0}$	0.070 (0.135)	0.003 (0.015)	-0.051 (0.049)	-0.041 (0.029)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on average heterogeneous treatment effects of a binding term limit on fiscal policy when a lame duck resigns. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.3). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%** and 1%***).

Furthermore, evidence shows that lame ducks incur significantly lower current expenditures also in the election year. The insignificance of the baseline coefficient estimates is thus likely driven by the increased expenditure of substitute mayors. In fact, evidence suggests that term-limited incumbents on average pursue a consistent low spending policy combined with reduced tax rates throughout the whole electoral term whilst substitute mayors engage in a fiscal policy cycle as does the average re-eligible mayor.

Table 4.7: Resigning Mayors – Annual Treatment Effects

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010	-0.014 (0.051)	-0.004 (0.006)	0.009 (0.023)	-0.027 (0.017)
2011	-0.021 (0.064)	-0.007 (0.007)	0.019 (0.024)	-0.037** (0.016)
2012	-0.085 (0.070)	-0.020** (0.008)	-0.000 (0.024)	-0.029* (0.016)
2013	-0.110 (0.076)	-0.022*** (0.008)	-0.052* (0.031)	-0.035* (0.021)
H_{it}	-0.063 (0.073)	0.001 (0.012)	-0.015 (0.038)	0.014 (0.019)
$H_{i,2010}$	0.037 (0.170)	-0.005 (0.015)	-0.017 (0.046)	-0.035 (0.031)
$H_{i,2011}$	0.019 (0.135)	0.000 (0.015)	-0.070 (0.051)	-0.064* (0.033)
$H_{i,2012}$	0.188 (0.140)	0.002 (0.016)	-0.042 (0.055)	-0.068** (0.033)
$H_{i,2013}$	0.035 (0.158)	0.014 (0.019)	-0.076 (0.063)	0.004 (0.048)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on annual heterogeneous treatment effects of a binding term limit on fiscal policy when a lame duck resigns. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.4). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

4.8 Robustness Checks

4.8.1 Selection Bias

The first robustness test tackles a potential problem with diff-and-diff regressions that relates to issues of selection bias. In principle, municipalities are exogenously assigned to treatment and control group. Only then can a differential change in the outcome variables from pre- to post-treatment period be attributed to the change in treatment assignment. If that is not the case there may be intrinsic differences between the two groups that are affecting the results which can no longer be interpreted as causal effects of treatment.

It is possible to test for selection bias in a number of ways. The second test on pre-existing group-specific trends in section 4.5.2 for example, also serves the purpose of testing for selection. The fact that year dummies indicating treatment municipalities during the pre-treatment period are insignificant points to an exogenous treatment assignment. Moreover, the fact that coefficient estimates remain consistent both in terms of magnitude and significance after the introduction of a vector of variables controlling for a number of observables asserts their robustness not only to selection on observables but also to selection on unobservables as proposed in Altonji et al. (2005).

The most common approach in the context of a diff-in-diff framework however, is to devise placebo tests. Placebo tests for diff-in-diff methodologies usually consist on re-estimating the baseline results relying on a placebo treatment setting in at a fake treatment year (De Jong et al., 2011). Here, the sample is restricted to the pre-treatment period, i.e. 2002-2009, and the regression model in Equation (4.1) is re-estimated with d_t indicating a placebo treatment setting in after the 2005 election, for the 2006-2009 electoral term. As Table 4.8 shows, estimates are close to zero and always insignificant dismissing any remaining concerns of a possible selection bias.

Table 4.8: Placebo Treatment.

	(1)	(2)	(3)	(4)	(5)	(6)
Fines & fees	0.025 (0.046)	0.041 (0.044)	0.043 (0.044)	0.031 (0.046)	0.049 (0.044)	-0.051 (0.044)
Tax rate	-0.006 (0.006)	-0.004 (0.006)	-0.004 (0.006)	-0.007 (0.006)	-0.005 (0.006)	-0.005 (0.006)
Tax revenues	-0.022 (0.025)	-0.006 (0.021)	-0.005 (0.021)	-0.021 (0.023)	-0.005 (0.020)	-0.006 (0.020)
C. Expenditure	-0.003 (0.012)	-0.003 (0.012)	-0.003 (0.012)	-0.005 (0.012)	-0.006 (0.012)	-0.007 (0.012)
District trends	No	Yes	Yes	No	Yes	Yes
District x year	No	No	Yes	No	No	Yes
Controls	No	No	No	Yes	Yes	Yes
Observations	2221	2221	2221	2197	2197	2197

Notes: This table presents the results of a placebo treatment on fiscal policy. The left column lists the dependent variables. Results are obtained from the estimation of Eq. (4.1) with d_t indicating a placebo treatment for the electoral year 2006-2009. All estimates include municipality and year fixed effects. Model (2) adds district-specific time trends and model (3) district-year fixed effects. Models (4), (5) and (6) replicate (1), (2) and (3), respectively, adding the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

4.8.2 Controlling for Anticipation Effects

As suggested in Malani and Reif (2015), often pre-trends can and should be interpreted as anticipation effects. In the context of the reform under study, controlling for anticipation

effects is of particular interest as the law limiting consecutive mayoral mandates established a stand-by period of one electoral term before becoming effective.

It is possible that forward-looking officeholders reacted to the now limited perspectives of long-term mayorships already in the 2006-2009 electoral term. If that was the case, credibly identifying the treatment effect involves controlling for anticipation effects. Following Malani and Reif (2015) and taking into account that the stand-by period during which anticipation effects may occur is known and finite, inference is based on the following quasi-myopic model:

$$Y_{it} = \delta B_{it} + \sum_{\tau=1}^q \delta_{+\tau} B_{i,t+\tau} + Z'_{it}\beta + \gamma_i + \gamma_t + \gamma_s \cdot t + \gamma_{st} + \epsilon_{it} \quad (4.6)$$

where the sum allows for q leads or anticipation effects with the dummies $B_{i,t+\tau}$ equal to one if the reform takes place in year $t + \tau$. The number of leading indicators that address anticipation is $q = 4$, the four years of the electoral term 2006-2009.

Table 4.9: Quasi-myopic model.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
Term limit	-0.031 (0.066)	-0.016* (0.009)	-0.018 (0.026)	-0.038** (0.016)
Ex ante effect ($t + 1$)	0.060 (0.061)	-0.009 (0.008)	-0.021 (0.025)	-0.017 (0.018)
Ex ante effect ($t + 2$)	-0.059 (0.058)	-0.015** (0.008)	-0.012 (0.027)	-0.019 (0.014)
Ex ante effect ($t + 3$)	0.057 (0.053)	0.002 (0.007)	0.012 (0.032)	0.005 (0.019)
Ex ante effect ($t - 4$)	0.042 (0.043)	-0.001 (0.006)	0.003 (0.022)	0.008 (0.009)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the quasi-myopic results on the average treatment effect of a binding term limit on fiscal policy. The first row lists the dependent variables. Results are obtained from the estimation of Eq. (4.6). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

Results in Table 4.9 are in line with the baseline average treatment effects. There is a significant anticipation effect for the property tax rate series in the year 2008 in line with the actual treatment effect for the year 2012. This effect suggests that strategic decisions already took place in anticipation of treatment. Still, the magnitude of the

average treatment effect on the property tax rate even after controlling for anticipation effects remains basically unchanged. So does the magnitude of the average treatment effect on current expenditures.

4.8.3 Controlling for Pre-existing Trends



Fig 4.4: Yearly Effects. This figure provides plots depicting the coefficient estimates for yearly dummy variables indicating the treatment group for the whole sample period 2002-2013. Data on the local property tax rate is only available from 2003 onwards due to a fiscal reform. Results are obtained from the estimation of Eq. (4.7). All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Dashed lines indicate 90% confidence intervals.

Finally, although the pre-trends analysis in section 4.5.2 does not provide evidence of differential trends – except for the revenues from fines and fees item that is insignificant throughout the regression results – and the previous section shows that results are robust to anticipation effects, this section re-estimates the treatment effects controlling for pre-existing time trends. Group-specific trends are controlled for in two ways based on the second and third tests in section 4.5.2 and the focus here is on the annual treatment effects.

The first approach combines Equations (4.2) and (4.6) in a regression model allowing for leads, or group-specific differential outcomes, and lags, or annual treatment effects

(Angrist and Pischke, 2009; Autor, 2003; Reber, 2005). In particular, coefficient estimates are obtained for the following extended regression model:

$$Y_{it} = \sum_{\tau=1}^m \delta_{-\tau} B_{i,t-\tau} + \sum_{\tau=1}^q \delta_{+\tau} B_{i,t+\tau} + Z'_{it}\beta + \gamma_i + \gamma_t + \gamma_s \cdot t + \gamma_{st} + \epsilon_{it} \quad (4.7)$$

where the sums specify m lags and q leads. Whilst the number of lags is again $m = 4$, one for each year of the 2010-2013 electoral term, the number of leads is $q = 8$, as in section 4.5.2 instead of $q = 4$ as before. Each lead corresponds to one year of the pre-treatment period, from 2002 to 2009. Again, 2004 is the omitted category.

Table 4.10: Group-specific Trends.

	Fines & Fees	Property Tax Rate	Tax Revenues	Current Expenditure
2010	0.009 (0.086)	-0.010 (0.013)	0.006 (0.035)	-0.041* (0.024)
2011	0.000 (0.092)	-0.013 (0.014)	0.009 (0.036)	-0.055** (0.021)
2012	-0.040 (0.095)	-0.025* (0.014)	-0.007 (0.036)	-0.047** (0.022)
2013	-0.087 (0.099)	0.026* (0.014)	-0.064 (0.042)	-0.043 (0.027)
District trends	Yes	Yes	Yes	Yes
District x year	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	3297	3025	3297	3297

Notes: This table presents the results on the annual treatment effect of a binding term limit on fiscal policy controlling for group-specific pre-existing trends. The first row lists the dependent variables. Results are obtained from the estimation of an extended version of Eq. (4.2) including a linear time trend and a treatment group specific time trend. All estimates include municipality and year fixed effects, district-specific time trends, district-year fixed effects and the vector of control variables. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%** and 1%(***)

For simplicity and compactness evidence is presented in the form of a graphical analysis. Figure 4.4 shows a graph for each outcome variable under study. The first plot shows that even when controlling for pre-existing group-specific trends, revenues from fines and fees are not affected by the introduction of term limits.

Graphs for the remaining outcome variables show insignificant coefficient estimates for the whole duration of the pre-treatment period and depict, for the post-treatment years, a path in line with the results obtained for the baseline annual treatment effects. The average property tax rate is significantly lower in treatment municipalities for the last two years of the 2010-2013 electoral term, tax revenues are significantly lower in 2013 and the meaningful drop in current expenditure occurs in the early years of the term, even when controlling for pre-existing differential trends.

In the second approach, Equation (4.2) is re-estimated for a specification including, in addition, both a linear time trend and a treatment group specific pre-trend (Moser and Voena, 2012). Table 4.10 collects the coefficient estimates. These too, are in line with the previously assessed results lending increased robustness to the baseline estimates. Altogether, the various tests validate the reliance on the diff-in-diff approach as a suitable identification strategy.

4.9 Conclusion

The introduction of term limits at the Portuguese local elections level provides an opportunity to analyze the economic consequences of an exogenous variation in eligibility for re-election. Using electoral and fiscal data on the 278 municipalities in continental Portugal for the past 12 years encompassing the last three complete electoral terms, inference is based on a quasi-experimental diff-in-diff approach that allows for a causal interpretation of the estimated treatment effects.

In contrast to the U.S. centered literature, results provide robust evidence of a decrease in the local property tax rate and aggregate current expenditure in treated municipalities with lame ducks engaging in a more conservative fiscal policy driven by a lack of re-election incentives. Significant heterogeneous effects indicate that political preferences are behind the change in fiscal policy choices of ineligible mayors. Results fit a model of distortionary elections predicting a more truthful behavior of politicians bound by term limits as in Smart and Sturm (2013). Not only do fiscal cycles pre-date the introduction of term limits, but also only eligible officeholders, rather than lame ducks, appear to engage in distortionary policies. Furthermore, despite the long tenures of the present lame ducks there is no clear evidence of their relatively higher experience.

This study challenges the standard approach of disciplining elections, opportunistic lame ducks and term limits induced fiscal cycles established by Besley and Case (1995). However, even the results therein could fit a model predicting a more candid behavior by term limited incumbents as Democrats are responsible for the identified loose fiscal policy of lame ducks. This theory is in turn consistent with the evidence of a conservative electorate in the U.S. (Peltzman, 1992) and is consistent with distortionary elections and the existence of electorally driven PBC instead of disciplining elections and undesirable term limitations.

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table 4.11: Summary Statistics.

Variable		Mean	SD	Min.	Max.	N
Fines and Fees	overall	16.572	18.088	.741	589.077	3333
	between		13.183	2.526	113.342	278
	within		12.404	-69.329	492.308	11.989
Property Tax	overall	.380	.085	.2	.5	3058
	between		.064	.218	.464	278
	within		.056	.143	.607	11
Tax revenues	overall	140.068	124.081	6.582	1312.731	3333
	between		117.361	36.677	811.033	278
	within		40.725	-134.751	660.493	11.989
Current Expenditures	overall	554.899	272.574	85.451	2007.682	3333
	between		254.888	211.185	1650.912	278
	within		97.580	110.862	1311.723	11.989
Resignation rate	overall	.050	.219	0	1	3336
	between		.126	0	.667	278
	within		.179	-.616	.717	12
Aligment	overall	.976	.152	0	1	3336
	between		.087	.417	1	278
	within		.125	.0597	1.560	12
Left	overall	.511	.500	0	1	3336
	between		.434	0	1	278
	within		.249	-.405	1.428	12
Local Economy	overall	283.503	214.436	12.073	3792.383	3300
	between		193.111	12.286	2498.046	275
	within		93.890	-677.495	1577.84	12
Inhabitants	overall	36165.72	57587	1634	549766	3336
	between		57607	1768.083	514263	278
	within		2939.726	-8588.284	83074.13	12

Chapter 5

Identifying the Source of Incumbency Advantage through a Constitutional Reform*

5.1 Introduction

Incumbency advantage, generally defined as the returns from holding office, has long been regarded exclusively as a personal affair. The early literature focuses on the role of the officeholder as the driver of the incumbency effect, offering several reasons for its existence, such as name recognition and media exposure (Erikson, 1971), the performance of constituency services (Fiorina, 1989), the control over relevant resources and institutional privileges (Mayhew, 1974), the conviction that incumbency works rather as a voting cue (Kostroski, 1973), the ability to scare-off high quality challengers (Jacobson, 1978) or the officeholder's quality and experience (Cox and Katz, 1996).

This “direct officeholder benefit” (Levitt and Wolfram, 1997) or “personal vote” (Ansolabehere et al., 2000) is referred to in the following as the personal incumbency advantage. A quantity scholars repeatedly tried to credibly identify often running into issues of selection on observables and endogeneity. While early estimates of the effect rely on variants of two popular measures, the sophomore surge (Erikson, 1971) and the retirement slump (Alford and Brady, 1989), simply opposing vote margins in open versus incumbent-contested seats, testing and controlling for each of the different explanations above eventually led to the use of increasingly more sophisticated methodologies.¹ Still, only recently was the non-randomness of incumbency advantage systematically addressed with Lee (2008) applying a Regression Discontinuity Design (RDD) to the analysis. This methodological breakthrough came, however, at the expense of an abstraction from the standard concept of incumbency advantage. Rather than analyzing the extent to which a candidate benefits by virtue of being the incumbent, the new framework tries to assess the benefits that

*See Lopes da Fonseca (2015a).

¹ Some examples include explicitly or implicitly addressing challenger quality following the paper by Jacobson (1978). Later, Gelman and King (1990) and Levitt and Wolfram (1997) also control for the normal vote and partisan tides. In Ansolabehere et al. (2000) and Ansolabehere and Snyder (2004), endogeneity concerns are first tackled using decennial redistricting as a natural experiment and an IV approach aimed at correcting for strategic retirement, respectively.

accrue to a candidate from running for the incumbent party, i.e. the partisan incumbency advantage.

Despite the breach with the previous literature, exploring the existence of a partisan incumbency advantage is of relevance.² Just as parties benefit from running an incumbent officeholder, it is also reasonable to assume that a candidate benefits from running for an incumbent party. In fact, many of the explanations for the existence of a personal incumbency advantage may also justify a partisan one. Parties are likewise in control of institutional resources and may have access to privileges, as well as benefit from media exposure and attested quality. Moreover, studies show that parties introduce strategic changes to the political system in order to maximize their electoral success (Trebbi et al., 2008). In addition, it is well-established that parties enjoy loyalty from voters (Aldrich and Rohde, 2001) and their relevance as political representatives is latent in the main role they play in Downs's (1957) theory of democracy. Still, only putting the emphasis on the candidate can explain certain phenomena such as the successful re-election of candidates who "cross the floor".³ These can be found even in the setting at the genesis of the incumbency advantage literature: the U.S. House of Representatives.⁴

Therefore, instead of focusing on either the candidate or the party, credibly estimating the incumbency effect implies that both sources of incumbency advantage are taken into account. By overlooking this aspect, Lee's (2008) framework fails to provide an unbiased estimate of either. Hence, the objective of this paper is to causally estimate the personal and the partisan effects that contribute to an incumbent's advantage at the ballot box. Put differently, this paper measures to what extent being the incumbent officeholder, or alternatively being the incumbent party, has a causal impact on electoral success.

The identification strategy relies on two quasi-experimental methods, the RDD and the difference-in-discontinuities (diff-in-disc) design, and on a natural experiment, the introduction of mayoral term limits in Portuguese municipalities. In Portugal, mayoral elections correspond to the proportional elections for the local council, the municipality's executive organ, with the first name on the winning list becoming the mayor. As each mayoral candidate represents a party, there is a need to dissociate the two in order to assess their independent returns to incumbency. This is only possible due to the recent introduction of term limits. The diff-in-disc design captures the second source of variation resulting from the exogenous heterogeneity in the implementation of term limits and combines it with the exogenous variation in incumbency status captured by the RDD focus on close elections.

² The partisan incumbency advantage as defined in Lee (2008) is different from the partisan strength present in the previous literature on incumbency where it usually refers to the normal vote and/or national tides (see e.g. Kostroski (1973)).

³ This term originates from the British House of Commons where an MP would have to literally cross the floor in order to join the opposition. The term has two different interpretations. Firstly it can refer to voting against party lines. Secondly and of relevance here, it refers to switching parties altogether.

⁴ Recent examples include: Philip Gramm, a Texas Congressman, who resigned his seat in the House as a Democrat in 1983 and was re-elected as a Republican a month later; Rodney Alexander, who first won his seat as a Democrat in 2002 and later as a Republican in the following five elections; Virgil Goode, a member of the House from 1979 to 2009, serving as a Democrat until 2000 and as a Republican thereafter.

Finally, a system of equations relating partisan to personal incumbency advantage in the spirit of Fowler and Hall (2014) provides credible estimates of the two quantities.

This paper primarily contributes to the incumbency advantage literature, by providing one of the first causal estimates of both the personal and partisan effects, and to the literature on voting behavior to the extent that it provides evidence on the relative importance of party versus candidate for electoral success. To my knowledge the only study offering causal evidence on the disentangled partisan and personal incumbency effects is Fowler and Hall (2014). Their findings, based on a sample of U.S. state legislative elections, show that incumbency is personal.

The main differences of this study consist of its reliance on a recent natural experiment that introduces a second source of exogenous variation into the analysis and on a novel identification strategy that uses the diff-in-disc design to disentangle the personal and partisan effects. In addition, this paper contributes to the discussion by extending Fowler and Hall's (2014) result in three ways. First, this study looks at a different country in what is a very American-centric topic. Second, it investigates an executive rather than legislative position. Despite the greater visibility of an executive position, most explanations for a positive personal incumbency advantage are usually linked to the role of a legislator. Third, it deals with a proportional representation (PR) system. Even though a significant partisan effect is more likely in a PR system, as the focus is most often directed to the parties and the cost of updated political information increases with increasing electoral candidates, results still support the early focus on candidates.

5.2 Incumbency Advantage and the Electoral Process

Whilst it may simply convey an electorate's satisfaction with the elected representative, the existence of an incumbency advantage can also pose a threat to equality of opportunities jeopardizing the competitiveness of electoral races and political accountability (Carson et al., 2007). Identifying the source of the incumbency effect not only unveils who is held responsible in the eyes of the voters but also provides a better general understanding of voting behavior, eventually leading to better policy choices.

The literature on voting behavior offers four standard explanations on how people vote: retrospectively, according to partisan affiliation, and based on candidates' issue positions or likeability. Evidence on the first is largely inconclusive. Peltzman (1987), for example, shows that voters in U.S. gubernatorial elections respond to economic conditions while Oliver and Ha (2007) find no relation between voters' ratings of government performance and actual votes for a sample of U.S. municipalities. It is often argued that, at the local level, retrospective voting is made difficult by the wide range of responsibilities accruing to the officeholders. In general, evidence suggests that retrospective voting hinges on the extent to which information is available to the electorate (Brender, 2003; Berry and Howell, 2007).

Also scant evidence shows voters basing their decisions on candidates' issue positions (Oliver and Ha, 2007). Indeed, information on candidates' stances is arguably hard to ob-

tain (Conover and Feldman, 1989) and assuming an at large informed electorate is unlikely to hold in most political contexts (Downs, 1957). Therefore, information may well lie at the core of voting behavior, with better informed voters being less likely to be swayed by party loyalty or incumbency status (Bartels, 1996), and the average uninformed electorate relying heavily on partisan or incumbency cues.

Accordingly, several studies across different levels of government demonstrate the importance of party affiliation and partisan endorsements on voting behavior (Lee et al., 2004; Snyder and Groseclose, 2000) as well as the relevance of candidates' characteristics and likeability (Lieske, 1989). The latter to the extent that inferences of competence from candidates' facial appearance correctly predict actual election outcomes of French parliamentary and U.S. congressional elections (Antonakis and Dalgas, 2009; Todorov et al., 2005). This is in line with evidence in Oliver and Ha (2007) that shows that, contrary to the standard view that incumbency is an advantage particularly in smaller settings due to increased visibility and familiarity, incumbency rather becomes more important with increasing population size as access to detailed information diminishes.

Yet, there is no conclusive evidence on whether the partisan or the candidate cue holds a higher electoral strength. Inasmuch as incumbency shapes voting behavior, identifying the driver of incumbency advantage can provide the means to infer upon the relative importance of each. A significant personal (partisan) effect means the candidate (party) is guiding voter's decisions by being ultimately held accountable at the ballot. The effect can be either positive or negative, and any but a similar personal and partisan effect implies an uneven balance of powers between the officeholder and the party.

Whether the incumbency advantage is personal, partisan or a combination of the two has different implications for politics at large. If the personal effect overrides the partisan one, parties have an incentive to renominate officeholders as well as to acquiesce to their demands. This may result in political dynasties and the capture of the political system by individual politicians. Electoral mechanisms such as term limits may be useful in this case to stimulate competition.

If, on the other hand, parties benefit solely by virtue of their incumbency status, this means they retain a political advantage even when officeholders retire allowing for more room to nominate candidates conditional on their alignment with the party and impose the partisan ideals leading to a centralization of political decisions. Just as before, electoral competition is threatened, in particular taking into account the evidence on the endogeneity of political institutions. The introduction of term limits does not imply an increase in political competition in this case but electoral rules benefiting minority parties could be a solution.

5.3 Institutional Background

5.3.1 Local Politics

The institutional setting encompasses all 278 municipalities in continental Portugal for increased comparability.⁵ Municipalities have two local political institutions: the municipal council and the municipal assembly. The former and subject of the analysis is the collegial executive organ accountable to the assembly, the legislative organ with decision-making power. It is composed by the mayor and other four to ten councilmen depending on the municipality's population size.⁶ With regards to electoral proceedings and term responsibilities, the mayor may be regarded as a local Prime-Minister and the municipal council as a small parliament. Elections occur simultaneously countrywide every four years. A PR system of closed lists is in place with seats being distributed according to the D'Hont method. The first name on the winning list becomes the mayor.

Local politics are dominated by the four largest national parties that generally run a list in every municipality. From left to right in the political spectrum, these are the Communist Party (PCP), the center-left Socialist Party (PS), the center-right Social-Democratic Party (PSD) and the right-of-center Popular Party (CDS-PP). Winning on average over 100 municipalities each, the two major parties, PS and PSD, are responsible for over two thirds of the municipal mayorships. Despite the high level of centralization and party discipline, municipal electoral campaigns are locally-oriented featuring the regional branches of the different political parties and are led by the mayoral candidates themselves.

5.3.2 Constitutional Reform

For decades, there were no limits to the number of times local politicians could run for re-election. This changed on January 1st, 2006, upon the entry into force of Law no. 46/2005.⁷ This law establishes a three consecutive terms limit for incumbent mayors after which they cannot rerun for the municipal council.⁸ It was introduced as means to void unconstrained overstay in office, not unusual at the Portuguese local level, and was first implemented in the 2013 local elections and binding for 150 municipalities.

Table 5.1 shows that re-election rates are very high at the Portuguese local level rendering a particularly interesting setting to investigate the magnitude and source of incumbency advantage. Before 2013, parties were re-elected on average in 81% of the municipalities. On the other hand, the absence of term limits enabled a rerunning rate of mayoral candidates of 83%, of which over 86% were on average re-elected. In 2013, while only 75% of the municipalities re-elected the incumbent party, 83% out of 91% of the mayors who

⁵The 30 municipalities comprising the autonomous regions of Azores and Madeira are not included in the analysis due to different institutional details.

⁶Two exceptions are the Oporto and Lisbon municipalities with 12 and 16 councilmen, respectively.

⁷Law no. 46/2005 in *Diário da República* 165, Series I-A, 29th August 2005.

⁸Upon entering into force in 2006, the law stipulated a stand-by period of one election, the 2009 elections, where no mayor was termed-out.

Table 5.1: Rerunning & Success Rates.

		1997	2001	2005	2009	2013
Panel A: Officeholder						
Rerun	N	218	227	233	238	117
	%	79	83	84	86	42
Win	N	186	188	211	205	97
	%	85	83	91	86	83
Panel B: Party						
Win	N	219	211	234	230	208
	%	80	76	84	83	75
Obs.		275	278	278	278	278

Notes: Rerunning and success rates of incumbent officeholders and parties.

stood for re-election won.⁹ In addition, there is considerable inter-party mobility, with candidates changing parties or running as independents, with relatively few costs of shifting allegiances. In fact, 24 municipalities have consecutively re-elected mayors upon their crossing the floor.

5.3.3 Data

The analysis is based on the electoral results of all 278 municipal councils in mainland Portugal in the past 20 years. Data on local elections are available at the National Electoral Commission (Comissão Nacional de Eleições) and the General Directorate for Internal Affairs (Direcção Geral da Administração Interna) websites.¹⁰ The identity of mayoral candidates, obtained in the Official Map from the National Electoral Commission published in *Diário da República*, is only available from the 1993 local elections onwards. As such, the underlying dataset collects complete information on electoral results at the party and mayoral candidates level for the past six elections taking place in 1993, 1997, 2001, 2005, 2009 and 2013.

5.4 Methodology

5.4.1 Incumbency Advantage and the RDD

The two effects that this study quantifies can be formally defined as follows. The personal incumbency advantage is the difference in vote share received by a candidate running for

⁹ The incumbent mayor stood for re-election in 117 out of the 128 municipalities without a binding term limit, being successful in 97.

¹⁰<http://www.cne.pt> and <http://www.dgai.mai.gov.pt>

the incumbent party as an incumbent and as a non-incumbent in an open-seat election in the same municipality i at the same time t . The partisan incumbency advantage, in turn, is the difference in vote share received by a non-incumbent candidate running in an open-seat election for the incumbent party and for the same party but with a non-incumbent status in the same election, i.e. same municipality i at time t .

As can be understood from the above definitions, assessing incumbency advantage is a classical example of an attempt of causal estimation with missing data as in the context of a Rubin Causal Model with potential outcomes (Rubin, 1974; Holland, 1986). In an ideal setting one would be able to compare the above outcomes, in reality only part of these are observable. Lee (2008) proposes the implementation of an RDD due to its focus on close elections in order to overcome this difficulty. The rationale runs as follows. Characterized by a small vote margin, i.e. difference in vote share between the winner and runner-up, close elections can be thought of as a mechanism simulating a randomized experiment. Parties with a positive vote margin run as incumbents in the next election whilst parties with a negative vote margin are non-incumbents. The zero threshold splits the treatment and control group. In a neighborhood of this threshold, comprising bare winners and bare losers, all variables determined prior to the election are independent of the resulting incumbency status. Consequently, any discontinuous jump in the vote share in the neighborhood of the threshold is entirely attributable to the change in treatment assignment and interpreted as a causal effect of the treatment, i.e. incumbency.

However, the focus on party-level variables does not by itself imply an unbiased estimation of the partisan incumbency effect as suggested in Lee (2008). Let the treatment dummy d_{it} denote whether a specific party is the incumbent. Underlying this dummy variable is the party's vote margin in the preceding election, the running variable $v_{i,t-1}$, with treatment deterministically depending on the value of the running variable as follows $d_{it} = 1 [v_{i,t-1} > 0]$. According to Fowler and Hall (2014), estimating the following general RDD model

$$y_{it} = \gamma_0 + \gamma_1 v_{i,t-1} + \rho_0 d_{it} + \rho_1 d_{it} \cdot v_{i,t-1} + \epsilon_{it} \quad (5.1)$$

where y_{it} is the outcome variable measuring electoral success and ρ_0 the average treatment effect, provides a combination of personal and partisan incumbency advantage as in Equation (5.2).

$$\rho_0 = 2 \cdot \text{Partisan Adv.} + 2 \cdot \text{Prob(Incumbent Reruns)} \cdot \text{Personal Adv.} \quad (5.2)$$

Since the incumbency status is always working for one party while harming the other the RDD estimate double counts both sources of incumbency advantage (Erikson and Titiunik, 2015).¹¹ Additionally, the personal incumbency advantage is multiplied by the probability that the incumbent is rerunning as this is often, but not always, the case. Parties, on the contrary, always field a candidate and therefore their probability of rerunning equals one.

¹¹In other words, the incumbent enjoys not only the partisan and personal advantage coming from the incumbency status, but also the foregone partisan and personal advantage of the competing party had it won the election. Intuitively, as the vote shares always sum up to 100%, whatever gain an incumbent collects must be lost by a non-incumbent.

As a starting point, it is useful to assess this combined effect. For this purpose, inference is based on two estimation methods. First, a non-parametric local linear regression restricts the sample in intervals $v_{i,t-1} \in [-h, h]$ to estimate Equation (5.1) (Hahn et al., 2001). For transparency, in addition to the use of the optimal bandwidth h , selected as in Imbens and Kalyanaraman (2012), two other intervals are considered: $[-5, 5]$ and $[-10, 10]$.¹² Second, a spline polynomial approximation of order p in $v_{i,t-1}$ as in Equation (5.3) estimates the effects using the full sample (Van der Klaauw, 2008).

$$y_{it} = \sum_{k=0}^p \gamma_k v_{i,t-1}^k + d_{it} \sum_{k=0}^p \rho_k v_{i,t-1}^k + \epsilon_{it} \quad (5.3)$$

The reliance on both estimation methods attempts to reinforce the robustness of the results and balance the usual trade-off between precision and bias.¹³ Nevertheless, results should be interpreted with caution as observations move away from the zero threshold.

Unlike the U.S., there are a myriad of parties in Portugal, which poses a challenge in defining the variables at large.¹⁴ The strategy employed attempts to take into account regional heterogeneity and relevant political developments. As such, the 18 Portuguese districts are treated as 18 independent regions. Upon assessing the strongest party within each district over the entire sample period, the outcome and the running variables are constructed accordingly.¹⁵ This approach not only maximizes the number of observations around the zero threshold but also enables a similar interpretation of the coefficient estimates as in a bi-partisan system. Even though the identity of the two major parties differs significantly by region, it is quite homogeneous within a district and particularly within a municipality. In fact, it is rare that a municipality has more than two different parties in control of the municipal council during the sample period and elections are usually effectively fought between these two. Following the literature on incumbency advantage, the main outcome variable is the baseline party vote share and a second outcome variable measuring the probability of winning conditional on the incumbency status – a dummy variable indicating whether the party wins the election at time t – serves as a robustness test.

In a second step, it is necessary to disentangle the partisan from the personal effect given by the RDD estimates. Partisan and personal incumbency are, often and non-randomly, jointly assigned but term limits are a reliable instrument in disentangling the two (Fowler and Hall, 2014). The following approach relies on the implementation of term limits as a natural experiment that keeps the incumbency status of the party constant but creates an exogenous variation in the incumbency status of the candidate. Using a diff-in-disc design to capture this variation provides for a new method to separate the personal from the partisan incumbency effect on electoral success.

¹² Results are robust to the use of alternative bandwidth selectors proposed in the literature, namely in Ludwig and Miller (2007) and Calonico et al. (2014). Results are also robust to the interval $[-2.5, 2.5]$.

¹³For a discussion on the advantages and drawbacks of the different estimation methods see e.g. Imbens and Lemieux (2007) or Lee and Lemieux (2010).

¹⁴ Over 40 different parties and coalitions ran for local elections during the sample period.

¹⁵ In practice, different regions have a different baseline party against which incumbency is measured.

5.4.2 Identification Strategy

The reform introduces a difference-in-differences (diff-in-diff) aspect to the analysis, with term limits as a second treatment applied in 2013 to municipalities whose incumbent officeholder has served as a mayor for the past three consecutive terms, loosely defining a pre- and post-reform period.¹⁶ Given the heterogeneity in treatment assignment, it is possible to rely on both between- and within-municipality variation in estimating the treatment effect of introducing term limits on electoral success.

Combining this diff-in-diff setting with the RDD methodology in the previous section results in an extended version of the RDD known as the diff-in-disc design (Grembi et al., 2015). As the name suggests, the diff-in-disc design estimates the difference between the two discontinuities in the observed outcome, the one after and the one before the reform. In its essence, it applies a diff-in-diff to two RDD estimations.

Formally, let $Y_{it}(d, b)$ and $\tilde{Y}_{it}(d, b)$, with $d = 0, 1$ and $b = 0, 1$, define the potential outcomes before and in 2013, respectively. The parameter b_{it} indicates a binding term limit and d_{it} indicates incumbency as before. The RDD estimator in the previous section is defined as $\hat{\rho}_0 = Y(1, 0) - Y(0, 0)$ and identifies the average treatment effect of incumbency at the threshold as in Equation (5.2). The diff-in-disc estimator in turn, explores the two discontinuities in the outcome variable as follows:

$$\hat{\beta}_0 = (\tilde{Y}(1, 1) - \tilde{Y}(0, 0)) - (Y(1, 0) - Y(0, 0)) \quad (5.4)$$

and identifies the difference in the average incumbency effect between term-limited and re-eligible incumbents at the threshold. Moreover, in the context of the diff-in-disc regression $\tilde{Y}(0, 0) = Y(0, 0)$ as the control group is the same for the two assessed discontinuities. Hence, $\hat{\beta}_0 = \tilde{Y}(1, 1) - Y(1, 0)$, clearly defining the diff-in-disc estimator as the change in an incumbents' electoral success when facing a term limit. Following the same rationale as in obtaining Equation (5.2), the first bracket on the right-hand-side of Equation (5.4) provides the following combination of partisan and personal advantage:

$$\tilde{Y}(1, 1) - \tilde{Y}(0, 0) = 2 \cdot \text{Partisan Adv.} + \text{Prob(Incumbent Reruns)} \cdot \text{Personal Adv.} \quad (5.5)$$

since the term-limited incumbent enjoys only a partisan incumbency advantage but in addition benefits from the foregone partisan and personal advantages of the counterfactual winner of the election at time $t - 1$. The personal incumbency advantage coming from the incumbency status is eliminated as the rerunning probability of an ineligible incumbent is zero.

Empirically, the treatment effect is obtained simply by adding the second treatment variable, b_{it} , and the appropriate interactions to Equations (5.1) and (5.3) above. Likewise, it is estimated in the context of a local linear regression model as the following:

$$y_{it} = \gamma_0 + \gamma_1 v_{i,t-1} + d_{it}(\rho_0 + \rho_1 v_{i,t-1}) + b_{it}[\alpha_0 + \alpha_1 v_{i,t-1} + d_{it}(\beta_0 + \beta_1 v_{i,t-1})] + \epsilon_{it} \quad (5.6)$$

¹⁶As not every incumbent is exposed to a term limit, $t = 2013$ does not clearly divide the sample into pre- and post-reform periods. Effectively, in 2013 there are three different groups of parties: the incumbents (all parties that won the election in 2009), the double-treated (the parties that in addition to holding the incumbency status also face a term limit), and the control.

restricting the sample to the same bandwidths h around the zero threshold as before, and resorting to a spline polynomial regression as in Equation (5.7).

$$y_{it} = \sum_{k=0}^p \gamma_k v_{i,t-1}^k + d_{it} \sum_{k=0}^p \rho_k v_{i,t-1}^k + b_{it} \left[\sum_{k=0}^p \alpha_k v_{i,t-1}^k + d_{it} \sum_{k=0}^p \beta_k v_{i,t-1}^k \right] + \epsilon_{it} \quad (5.7)$$

The parameter β_0 combines both treatments giving the average treatment effect as in Equation (5.4). Substituting Equations (5.2) and (5.5) into (5.4) defines the diff-in-disc estimator as follows

$$\beta_0 = -\text{Prob}(\text{Incumbent Reruns}) \cdot \text{Personal Advantage} \quad (5.8)$$

assuming that partisan and personal incumbency advantage remain on average identical across all election years. Solving the system of two equations and two unknowns given by Equations (5.2) and (5.8) gives the following expressions:

$$\text{Personal Advantage} = -\frac{\beta_0}{\text{Prob}(\text{Incumbent Reruns})} \quad (5.9)$$

$$\text{Partisan Advantage} = \frac{\rho_0}{2} + \beta_0 \quad (5.10)$$

that allow for identifying the partisan and personal independent returns to incumbency. In the last step of the analysis, the RDD and diff-in-disc estimates are substituted into this system of equations to solve for the two quantities of interest. The regressions and system of equations are re-estimated simultaneously 100,000 times in a non-parametric bootstrap with replacement clustered at the municipal level, in order to obtain a credible uncertainty measure for the resulting estimates.

5.5 Empirical Results

5.5.1 RDD and Diff-in-disc Estimates

As is standard in the literature, a first assessment of the results is based on graphical evidence on the treatment effects. All graphs are constructed by first dividing the running variable into bins of size one within a bandwidth of $[-40, 40]$. The outcome variables – vote share and probability of winning – are then averaged within each bin for the pre- and post-reform period, i.e. \bar{y}_{np} with the subscripts n and p indexing bin and period. The RDD graphs display these observations fit by a local polynomial plot of quadratic degree on each side of the threshold, relying on a triangular kernel. To obtain the diff-in-disc graphs it is necessary to first calculate the difference within each bin between the post- and pre-reform period, $\Delta y_n = (\bar{y}_{n,1} - \bar{y}_{n,0})$, to then plot these differences against the running variable and fit the observations with a polynomial plot in the same fashion as before.

Figure 5.1 provides the RDD and the diff-in-disc graphs depicting the vote share and probability of winning in period t against the vote margin in $t - 1$. All graphs show a

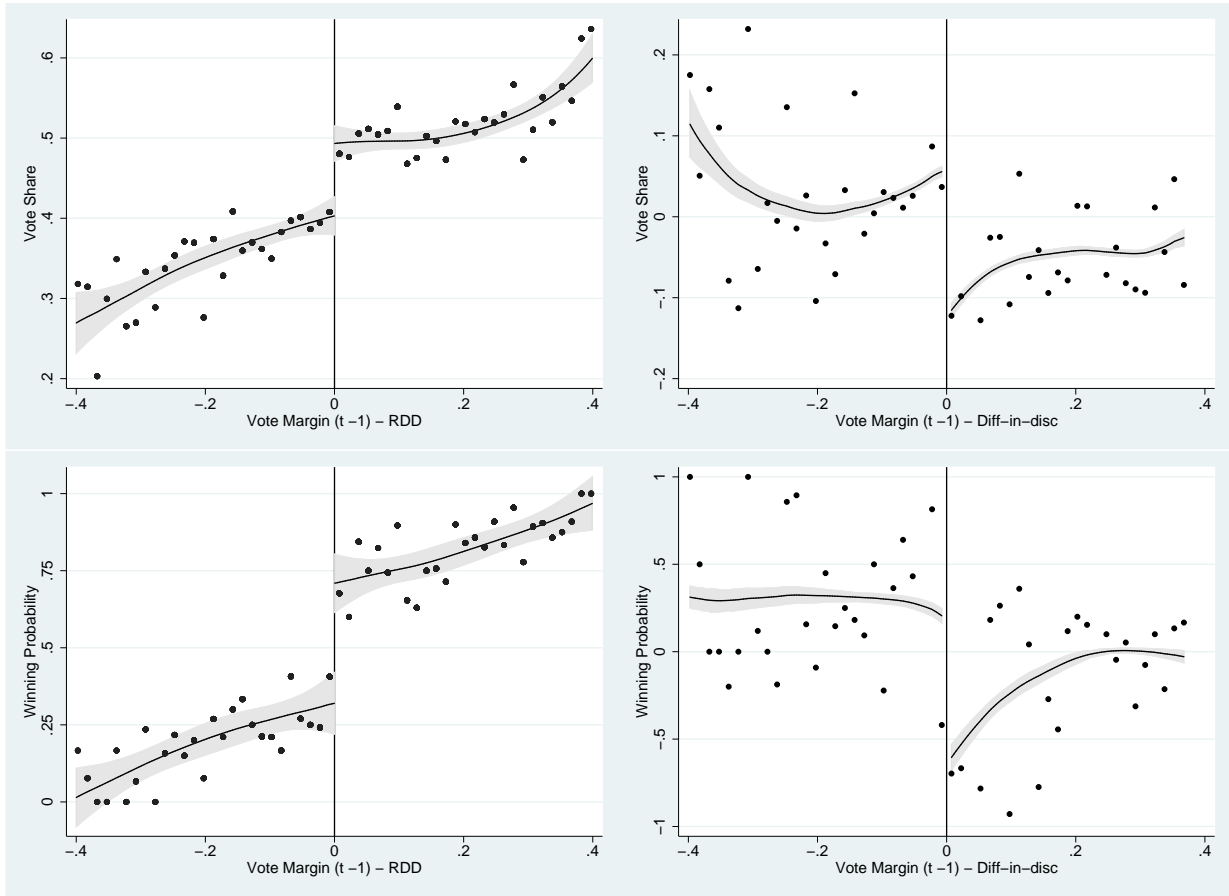


Fig 5.1: Vote Share and Winning Probability. Graphs plot the vote share at time t against the vote margin at $t - 1$. Observations in a bandwidth of $[-40, 40]$ are averaged within bins of size 1% and second degree polynomial plots are constructed using a triangular kernel. 95% confidence intervals are indicated in gray.

sizable jump around the zero threshold providing evidence of a significant incumbency effect. For the main outcome variable, the RDD graph suggests a combined personal and partisan incumbency advantage of around 8 to 10 percentage points, while the diff-in-disc graph shows a negative discontinuity amounting to 15 percentage points for municipalities affected by the term limit. The latter plot suggests that in case there exists a partisan effect, it is not as large as the personal one. The graphs depicting the probability of winning at time t against the vote margin at $t - 1$ are in line with this assessment.

Coefficient estimates in Table 5.2 support the graphical evidence. Estimates follow a similar pattern across specifications with the ones for the spline polynomial approximation, which were obtained relying on a cubic specification of the running variable in the models in Equations (5.3) and (5.7), suggesting in general larger treatment effects.¹⁷ The choice of which estimates to substitute into the expressions quantifying the personal and partisan incumbency advantages is therefore inconsequential. In the following, results are discussed

¹⁷Results are robust to different order polynomials.

Table 5.2: RDD and Diff-in-disc Estimates.

Design	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: Vote Share				
RDD	0.090*** (0.014)	0.067** (0.028)	0.071*** (0.021)	0.097*** (0.016)
Diff-in-disc	-0.154*** (0.050)	-0.207** (0.089)	-0.193*** (0.065)	-0.205*** (0.063)
Obs.	728	207	397	1384
Panel B: Winning Probability				
RDD	0.386*** (0.059)	0.224 (0.138)	0.322*** (0.098)	0.393*** (0.072)
Diff-in-disc	-1.155*** (0.253)	-0.772* (0.460)	-1.159*** (0.399)	-1.160*** (0.314)
Obs.	791	207	397	1384

Notes: IK stands for the bandwidth selected as in Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Robust standard errors are clustered at the municipal level. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

for the specification relying on the bandwidth selected as in Imbens and Kalyanaraman (2012) due its conventional use in the RDD literature.

5.5.2 Partisan and Personal Incumbency Advantages

In addition to the RDD and diff-in-disc estimates, a third input is necessary in order to obtain the personal and partisan incumbency advantages: the probability that an incumbent reruns. This is obtained by simply regressing a dummy variable indicating whether the incumbent officeholder is running for re-election on a linear function of the running variable.

The first three rows of Table 5.3 show the estimates of the three inputs required to compute the incumbency effects. These are substituted into Equations (5.9) and (5.10) and 95% bias-corrected confidence intervals are obtained relying on a non-parametric bootstrap with replacement clustered at the municipal level. This involves repeatedly re-estimating the three inputs and substituting them into the equations in a simultaneous fashion using different random samples from the original dataset.

The last two rows of the table provide the estimates for the personal and partisan incumbency advantage with and without fixed effects. The personal effect has a positive causal impact on electoral success. A party benefits by up to 17 percentage points from running an incumbent officer. The incumbency status of the party, however, is not sta-

Table 5.3: Personal and Partisan Incumbency Advantage.

	Vote Share		Winning Probability	
RDD	0.090 [0.063, 0.118]	0.111 [0.084, 0.139]	0.386 [0.262, 0.503]	0.227 [0.087, 0.362]
Diff-in-Disc	-0.154 [-0.259, -0.043]	-0.184 [-0.295, -0.084]	-1.154 [-1.655, -0.592]	-1.557 [-2.153, -0.888]
Prob.	0.922 [0.627, 1.213]	0.979 [0.589, 1.381]	0.830 [0.569, 1.091]	0.886 [0.515, 1.277]
	Incumbency Advantage			
Personal	0.167 [0.048, 0.322]	0.188 [0.076, 0.372]	1.391 [0.693, 2.348]	1.758 [0.893, 3.446]
Partisan	-0.109 [-0.021, 0.003]	-0.129 [-0.239, 0.026]	-0.961 [-1.456, -0.390]	-1.444 [-2.063, -0.797]
FE	No	Yes	No	Yes

Notes: The first three rows provide the inputs for the system of equations relating personal to partisan incumbency advantage. The last two rows provide the estimates of personal and partisan incumbency advantage. Bias-corrected confidence intervals in brackets. FE indicates whether municipality fixed effects. All models include election term fixed effects.

tistically different from zero thus having no sizable causal impact on the resulting vote share.

The results for the probability of winning, in turn, identify statistically significant personal and partisan effects but of opposite signs. These results are presumably rooted in the high magnitude and significance of the personal incumbency effect on the resulting vote share. Arguably, based on the descriptive evidence in Table 5.1, being the incumbent is a sure win. By ascribing this effect to the candidate rather than the party, it follows that running a new candidate undermines a party's electoral prospects. The high magnitude of the effect is likely due to the significant loss in vote share that led to a considerable reduction in the party re-election rate in 2013 in term-limited elections.¹⁸ Still, the significantly negative partisan effect should not be interpreted as a partisan incumbency disadvantage but rather as the disadvantage of not running an incumbent candidate.

5.5.3 Explanatory Hypotheses

This section tests two possible explanations for the baseline results. First, whether there exist ruling costs, as suggested by a strand of the literature on the effects of political representation on electoral outcomes (Paldam, 1986). Second, whether the fact that a challenger is rerunning for the election matters. Results are based on re-estimations of the

¹⁸Parties were re-elected in around two thirds of the 150 term limited elections compared to an average of above 80% in the pre-reform period.

diff-in-disc models with the dummy variable b_{it} indicating mayors that did not run for re-election after at least three consecutive terms in office in the first case, and municipalities with a binding term limit where no challenger is rerunning for the election in the second.

Table 5.4: Hypotheses Test.

Design	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: Ruling Costs				
Diff-in-disc	-0.087** (0.038)	-0.091 (0.062)	-0.073 (0.050)	-0.092** (0.043)
Obs.	506	170	331	1106
Panel B: Challenger				
Diff-in-disc	-0.090*** (0.025)	-0.128*** (0.043)	-0.118*** (0.025)	-0.110*** (0.029)
Obs.	728	207	397	1384

Notes: The outcome variable is the vote share of the baseline party. In Panel A the sample is restricted to the pre-treatment period. IK stands for the bandwidth selected as in Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

Coefficient estimates collected in Table 5.4 suggest no consistently significant ruling costs despite the almost professionalization of the mayor-career that characterized the pre-reform period. On the other hand, incumbent challengers appear to be driving part of the effect. The diff-in-disc estimates remain significantly negative but their magnitude is down by half of the baseline estimates. This result is presumably due to the fact that the major opposition party nominee running as an incumbent challenger is not unknown to the electoral district. In fact, he was most likely part of the municipal council and the closest to the mayor figure in terms of public recognition. Accordingly, in more than half of the 2013 term limited elections where the incumbent challenger or a former mayor reran, they won the mayorship. The relevance of the candidates thus transpires even through the success of rerunning challengers once the incumbent mayor is excluded from the electoral race. This result reinforces the baseline findings and lends increased support to the thesis that incumbency is personal and parties play only a secondary role.

5.6 Internal Validity

The validity of an RDD hinges on fulfilling the continuity assumption that implies no sorting around the zero threshold. This assumption is satisfied as long as individuals are not able to precisely manipulate their location relative to the threshold. This is plausible in the context of competitive elections as parties cannot precisely manipulate their vote

shares and even to a lesser extent the resulting vote margin between winner and runner-up. Two methods can be used to validate this assumption: testing for discontinuities in the density of the running variable through a McCrary (2008) plot, and checking for any significant differences in baseline characteristics between the treatment and control groups, i.e. incumbents and non-incumbents. The McCrary (2008) plot on the left side of Figure 5.2 and the graphs collected in Figure 5.3 testing for differences on six pre-determined characteristics – the parties’ vote share and winning probability in $t - 1$, whether the party is one of the two major center-left or -right parties, if it is a coalition, whether the candidate is male or female and his or her previous experience – show no significant discontinuities around the zero threshold implying no manipulation and no selection bias.

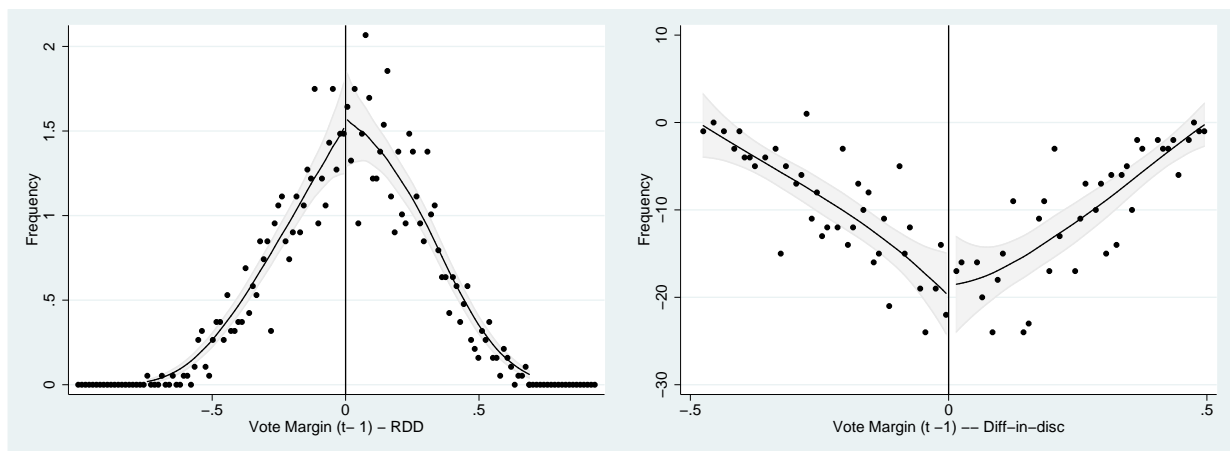


Fig 5.2: Manipulation of the Running Variable. The McCrary plot on the left graphs the density of the running variable. The McCrary style plot on the right graphs the change in the density of the running variable from the pre- to the post-treatment period.

The diff-in-disc design, in turn, is superior to the RDD as it captures a second source of variation combining diff-in-diff and RDD methodologies. The diff-in-diff aspect of the approach addresses any concerns with the continuity assumption described above. The RDD aspect controls for differential trends and unobservable municipal characteristics. Still, one assumption must be fulfilled in order to validate a diff-in-disc design: that the ability or incentives to manipulate the vote margin did not change from the pre- to the post-treatment period. This is tested relying on a variant of the McCrary (2008) plot that depicts the change in the number of observations around the threshold from one period to the other. There is no reason to suspect that parties are now able to manipulate their vote shares as the right hand side graph in Figure 5.2 shows no discontinuity in the change in the number of observations close to the cutoff from pre- to post-treatment period.

In addition, in order to be able to identify the partisan and personal effects relying on the proposed approach, these two effects are assumed unchanged from pre- to post-treatment period. There is no obvious reason why this should not be the case. Still, this assumption can be verified by restricting the analysis to the 2005 and 2009 elections, when the law establishing term limits already exists but is not applied, to compare expiring to

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

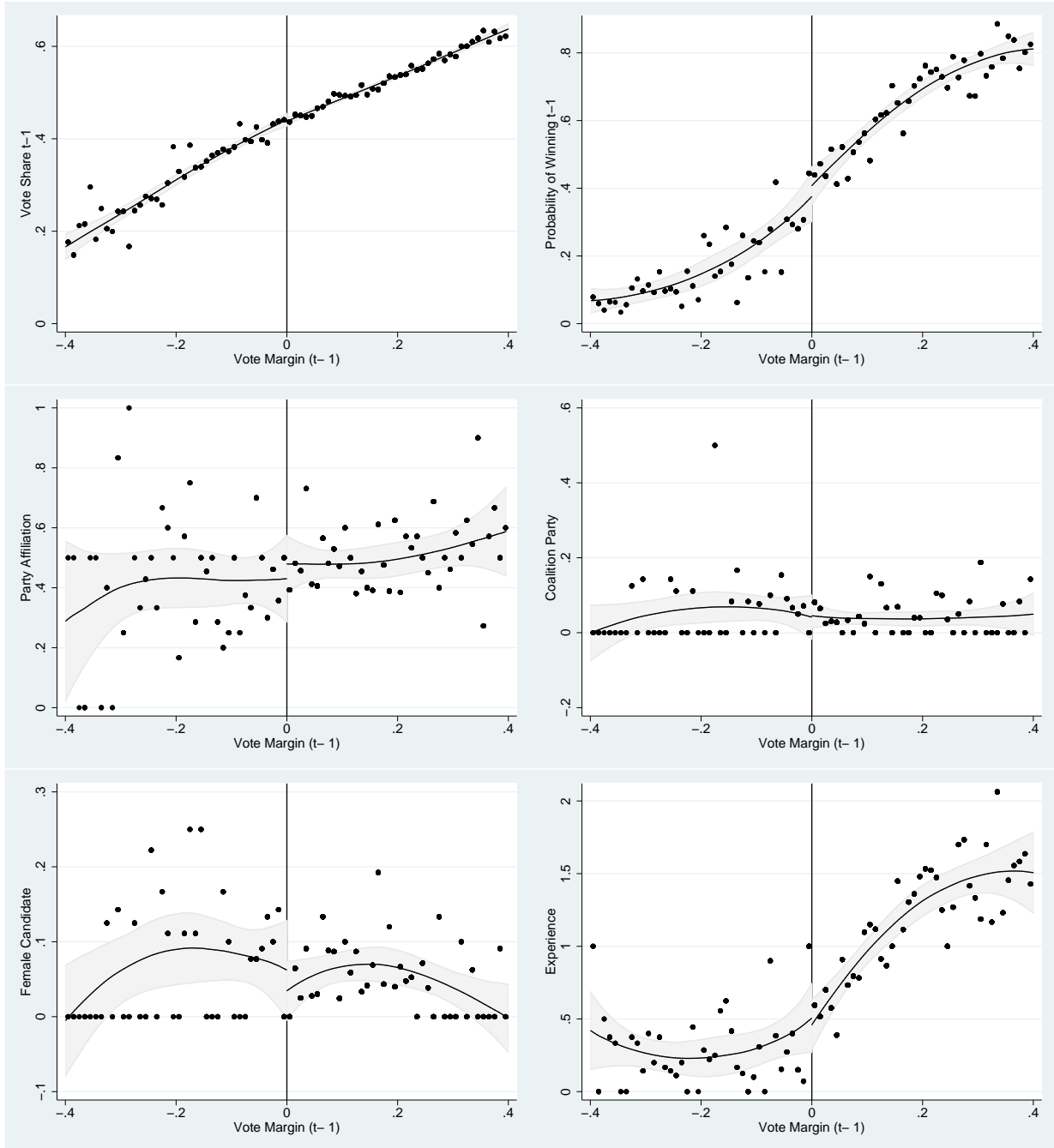


Fig 5.3: Selection Bias. Graphs test for selection into treatment for six different pre-determined characteristics indicated on the y-axis. Observations in a bandwidth of $[-40, 40]$ are averaged within bins of size 1% and second degree polynomial plots are constructed using a triangular kernel. 95% confidence intervals are indicated in gray.

non-expiring mayors.^{19,20} This test involves restricting the sample size, therefore inference relies on the three data driven bandwidth selectors suggested in the literature: Ludwig and Miller (2007), Imbens and Kalyanaraman (2012) and Calonico et al. (2014).

Results are collected in Table 5.5 and are obtained relying on a local linear regression as in Equation (5.1) for all bandwidths above including half and two times the bandwidth selected as in Imbens and Kalyanaraman (2012). The RDD estimates suggest a positive and significant combined partisan and personal effect for all incumbents in line with the baseline estimates. The diff-in-disc estimates, obtained by defining b_{it} in Equation (5.6) as to indicate expiring officeholders, provide no evidence of a significantly different incumbency effect between the two groups.

Table 5.5: Diff-in-disc Assumption.

Design	Local Linear Regression				
	CCT	1/2 IK	IK	2 IK	CV
RDD	0.123*** (0.024)	0.116*** (0.036)	0.114*** (0.022)	0.076*** (0.016)	0.118*** (0.021)
Diff-in-disc	0.039 (0.066)	0.088 (0.126)	0.051 (0.057)	0.029 (0.035)	0.066 (0.051)
Obs.	270	149	299	494	332

Notes: The outcome variable is the vote share of the baseline party. CCT, IK and CV stand for the bandwidths selected as in Calonico et al. (2014), Imbens and Kalyanaraman (2012) and Ludwig and Miller (2007), respectively. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

Overall, this section validates each of the assumptions underlying the proposed approach. There exists no manipulation at the zero threshold neither in the pre- nor in the post-treatment period, there is no evidence of selection bias and both the personal and partisan incumbency advantage are unchanged by the electoral reform introducing term limits. Baseline results are, therefore, susceptible of causal interpretation.

5.7 Conclusion

The aftermath of a constitutional reform introducing term limits at the local level elections is the perfect laboratory to identify the personal and partisan returns to incumbency. The institutional setting, comprising 278 homogenous municipalities, an exogenous constitutional reform, heterogeneity in treatment resulting from the law ruling the application of term limits and the verified comparability between bare winners and bare losers, allows for the implementation of both RDD and diff-in-disc design providing coefficient estimates suitable of causal interpretation.

¹⁹The law was voted and approved in Parliament before the 2005 local elections.

²⁰ This terminology is borrowed from Fowler and Hall (2014). Expiring mayors are the ones running for their last possible term in office. Non-expiring mayors have no short-term perspectives of being termed-out.

Results show that the original definition of incumbency advantage purposely focuses on the candidate rather than the party. Positive returns accrue to the rerunning officeholder by virtue of his incumbent status independently of the incumbency status of the party. Even though it would be reasonable to assume that parties play a significant role in attracting and retaining votes, particularly in a multi-party context with a PR system of closed lists (Shugart et al., 2005), fact is that the assessed partisan effect is indistinguishable from zero.

The magnitude of the effect conditional on the institutional setting in which it emerges, together with the evidence on the importance of incumbency and personal characteristics in explaining voting behavior at various levels of government motivate the external validity of the results. These reinforce the relevance of candidates vis-a-vis parties in politics at large.

Conceivably both voters and politicians will over time adapt to the new institutional setting. Preliminary assessments show e.g. an increase in candidates' mobility between municipalities whereas traditionally it followed a bottom-up pattern. This is in line with the theory of political ambition and its interaction with the political structure in shaping incentives and provides an interesting avenue for further research (Schlesinger, 1966; Black, 1972).

Appendix to Chapter 5

5.A.1 Robustness tests

This section offers a variety of tests checking for the robustness of the baseline estimates to the inclusion of control variables and municipality fixed effects, other bandwidths and polynomial functions of the running variables and different specifications of the main variables. First, using the same specifications as in the baseline results, all regression models are re-estimated including the following nine municipality-specific control variables: municipal population size and municipal population size squared, a measure of municipal economic activity proxied by night time light satellite observations, the level of current and capital expenditures in the election year and three political dummy variables indicating whether the mayor's party is also in control of the central government, whether the incumbent party is on the left side of the political spectrum and whether the Prime Minister is from the center-right PSD. Results are collected in Table A.5.1. Second, still using the same baseline specifications, instead of control variables, results are re-estimated including municipality fixed effects in addition to election year fixed effects. This implies the resulting estimates capture only within- but no longer between-variation, possibly compromising the significance of the estimates for smaller samples. Results are collected in Table A.5.2. Overall, coefficient estimates in both tables reinforce the robustness of the baseline estimates.

Table A.5.1: Control Variables.

Design	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: Vote Share				
RDD	0.089*** (0.014)	0.066** (0.028)	0.072*** (0.021)	0.096*** (0.016)
Diff-in-disc	-0.159*** (0.051)	-0.261*** (0.096)	-0.178*** (0.064)	-0.226*** (0.067)
Obs.	720	204	393	1369
Panel B: Winning Probability				
RDD	0.379*** (0.059)	0.190 (0.140)	0.309*** (0.097)	0.389*** (0.072)
Diff-in-disc	-1.081*** (0.250)	-0.578 (0.507)	-1.140*** (0.397)	-1.115*** (0.313)
Obs.	781	204	393	1369

Notes: All models include the vector of control variables. IK stands for the bandwidth selector proposed by Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

Table A.5.2: Municipality Fixed Effects.

Design	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: Vote Share				
RDD	0.111*** (0.014)	0.059* (0.034)	0.094*** (0.024)	0.103*** (0.016)
Diff-in-disc	-0.184*** (0.047)	-0.133 (0.103)	-0.179*** (0.068)	-0.250*** (0.054)
Obs.	689	127	327	1384
Panel B: Winning Probability				
RDD	0.227*** (0.069)	-0.202 (0.191)	0.097 (0.136)	0.233*** (0.081)
Diff-in-disc	-1.557*** (0.298)	-0.690 (0.494)	-1.769*** (0.376)	-1.650*** (0.335)
Obs.	759	127	327	1384

Notes: IK stands for the bandwidth selector proposed by Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

The third robustness check involves re-estimating the baseline results for different bandwidths and polynomials of the running variable. Using a local linear regression, estimates are obtained for three additional bandwidths: two data driven bandwidth selectors in the literature, the Calonico et al. (2014) and the Ludwig and Miller (2007) optimal bandwidth selectors, and a more conservative bandwidth restricting the sample to the interval $[-2.5, 2.5]$. The spline polynomial regression model in turn, is re-estimated relying on second- and fourth-order approximations. Results collected in Table A.5.3 are also in line with the baseline estimates.

Finally, the last robustness test makes use of different specifications of the main variables to ascertain that the baseline results are not contingent on how variables are measured nor on the particular treatment year. Following the approach in Uppal (2009) the RDD is re-estimated relying on variables structured at the candidate level. Since rerunning decisions are systematically different between incumbents and non-incumbents the analysis is conditional on incumbents rerunning the election to prevent overestimating the effect. This should provide a similar combination of personal and partisan incumbency advantage as the baseline results. However, in contrast to Equation (5.2), here the probability of an incumbent rerunning equals one and the partisan incumbency advantage is multiplied by the probability of a candidate being loyal, i.e. not crossing the floor. But first, conditioning on re-running may introduce a selection bias if rerunning decisions are systematically different between rerunning and non-rerunning losers. Testing for differences in the four

THE POLITICAL ECONOMY OF ELECTORAL REFORMS

Table A.5.3: Different Bandwidths and Polynomials.

Design	Local Linear Regression			Spline Polynomial Regression	
	CCT	CV	$h = 2.5\%$	Quadratic	Quartic
Panel A: Vote Share					
RDD	0.077*** (0.019)	0.087*** (0.013)	0.076** (0.036)	0.093*** (0.013)	0.084*** (0.020)
Diff-in-disc	-0.155** (0.061)	-0.125*** (0.046)	-0.238*** (0.063)	-0.124** (0.050)	-0.193*** (0.071)
Obs.	582	1054	100	1384	1384
Panel B: Winning Probability					
RDD	0.388*** (0.066)	0.370*** (0.043)	0.159 (0.189)	0.356*** (0.050)	0.386*** (0.094)
Diff-in-disc	-1.148*** (0.252)	-0.669*** (0.159)	0.056 (0.203)	-0.940*** (0.245)	-1.220*** (0.376)
Obs.	835	1381	100	1384	1384

Notes: CCT and CV stand for the optimal bandwidth selectors suggested in Calonico et al. (2014) and Ludwig and Miller (2007), respectively. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

pre-determined baseline characteristics in Table A.5.4 shows no evidence that this is the case. This approach is therefore a valid robustness test and coefficient estimates on Panel A of Table A.5.5 are in all likelihood similar to baseline estimates, reinforcing the results.

Table A.5.4: Internal Validity: Losers Self Selection.

	$h = 1\%$	$h = 5\%$	$h = 10\%$	Full Sample
Vote Share $t - 1$	-0.014 (0.014)	0.002 (0.007)	0.004 (0.005)	0.001 (0.003)
Affiliation	-0.018 (0.104)	-0.006 (0.060)	-0.002 (0.043)	-0.011 (0.027)
Coalition	0.065 (0.055)	0.001 (0.030)	0.030 (0.024)	0.012 (0.015)
Female	0.032 (0.039)	0.042 (0.032)	0.031 (0.024)	0.021 (0.017)
Obs.	52	238	449	1471

Notes: This table depicts differences between rerunning and non-rerunning losers for the variables on the left column. The first row indicates the intervals which the sample is restricted to. The second column provides only a difference in means. The remaining columns show the regression of each characteristic on a dummy that takes value one if the candidate reruns in the next election and zero otherwise, the running variable and an interaction between these two for the intervals defined on the first row. The last column includes in addition a third-order polynomial of the running variable and interaction terms. Standard errors are robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

To test the diff-in-disc results, the sample is reduced to the pre-treatment period and the dummy variable $b_{i,t}$ is constructed to indicate first, any election where the incumbent mayor is not rerunning, i.e. open seat elections, and second, elections the incumbent mayor is not rerunning due to death, promotion or suspension. The first specification tests how incumbent parties that nominate a new candidate fare against the ones that run with the incumbent mayor for elections from 1993 through to 2009. And the second focuses on a subset of these, where the incumbent mayor did not complete the whole term. This may be the case for four different reasons: death, promotion, suspension or personal motives. Whilst the latter may be closely related to strategic retiring the remaining reasons are assumed exogenous. Given that the number of mayors in these situations is reduced, all are taken together to compare incumbent parties whose mayor stepped down for justified reasons against incumbent parties whose mayor deliberately chose whether to rerun or not. Contrary to the baseline diff-in-disc estimates here there is no systematically credible second source of exogenous variation. Results from these two tests should thus be taken with a grain of salt.

Table A.5.5: Different Specifications.

	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: RDD				
Candidate	0.090*** (0.016)	0.045* (0.026)	0.080*** (0.019)	0.087*** (0.015)
Obs.	618	257	471	1394
Panel B: Diff-in-disc				
Open Seats	-0.112*** (0.024)	-0.134 (0.088)	-0.197*** (0.063)	-0.169*** (0.052)
Subset	-0.026 (0.050)	-0.040 (0.102)	-0.021 (0.048)	-0.173* (0.089)
Obs.	506	170	331	1106

Notes: Outcome variable is the vote share of the baseline party. IK stands for the bandwidth selector proposed by Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***)

Results are collected in Panel B of Table A.5.5. Coefficient estimates for all open seat elections are in line with baseline estimates, negative, mostly significant but of a smaller magnitude. As in Ansolabehere and Snyder (2004), when strategic retirement is controlled for as in the baseline estimates, the effect is of a higher rather than of a smaller magnitude. As discussed in the same paper, the importance of strategic retirement appears to be overestimated in the literature. Results for the sample of suspended, deceased and promoted mayors are consistently negative but only twice significant, for the largest

bandwidths. This could be due to the small amount of mayors in these situations. For the whole sample, 24 mayors were promoted mid-term, six passed away and two were suspended. Still, it is remarkable to see that results are still in line with baseline estimates even though in the majority of cases parties can publicize the promotion, or voters are mourning the death, of a previous mayor. All in all, voters show a coherent response to a new mayoral candidate running for the incumbent party reinforcing the robustness and validity of the baseline estimates.

5.A.2 Heterogeneous Effects

Table A.5.6: Heterogeneous Effects.

	Local Linear Regression			Spline Polynomial Regression
	IK	$h = 5\%$	$h = 10\%$	
Panel A: Central Government				
Aligned	-0.032* (0.019)	0.002 (0.033)	-0.023 (0.024)	-0.029 (0.021)
Panel B: Party				
PS	-0.013 (0.020)	-0.024 (0.033)	-0.010 (0.026)	-0.014 (0.023)
PSD	0.018 (0.020)	0.016 (0.034)	0.003 (0.025)	0.011 (0.023)
PCP	-0.015 (0.020)	0.021 (0.034)	0.003 (0.027)	-0.023 (0.022)
CDS	0.017 (0.028)	0.015 (0.030)	0.013 (0.027)	0.011 (0.028)
Panel C: Region				
North	0.022 (0.021)	-0.011 (0.034)	0.032 (0.027)	0.011 (0.024)
Center	0.006 (0.019)	0.061* (0.033)	0.011 (0.024)	0.017 (0.022)
Alentejo	-0.023 (0.019)	-0.014 (0.032)	-0.025 (0.025)	-0.018 (0.022)
Algarve	-0.032 (0.042)	-0.181** (0.086)	-0.051 (0.049)	-0.035 (0.060)
Obs.	728	207	397	1384

Notes: IK stands for the bandwidth selector proposed by Imbens and Kalyanaraman (2012). The spline polynomial regression estimates are obtained relying on a third-order approximation. All models include election term fixed effects. Standard errors are clustered at the municipal level and robust to heteroscedasticity. Stars indicate significance levels of 10%(*), 5%(**) and 1%(***).

This section analyzes three possible channels of heterogeneous effects. First, the focus is upon political alignment and thus whether incumbency advantage hinges on the mayor being from the same party as the Prime Minister. Second, whether candidates of any of the four major national parties enjoy a distinguishing incumbency advantage. Third, the concern is geographical and whether incumbency advantage is specific of any given region in Portugal.

Results are obtained relying on the diff-in-disc models with $b_{i,t}$ indicating every of the two different and non-overlapping groups for the bandwidths and polynomial used in the baseline estimates. Coefficient estimates are collected in Table A.5.6. Overall there is no consistent pattern of statistically significant results. Coefficient estimates suggest that incumbency effects are stronger than alignment effects, and that incumbency advantage is not particular of neither a party nor a region. This test for heterogeneous effects can instead be taken as another robustness test of the consistency of the incumbency advantage effect across parties and regions.

Bibliography

- Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a fundamental cause of long-run economic growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1, pp. 385–472. Elsevier.
- Acemoglu, D. and J. Robinson (2006). *Economic origins of dictatorship and democracy*. Cambridge, UK: Cambridge University Press.
- Acemoglu, D. and J. A. Robinson (2000). Why did the West extend the franchise? Growth, inequality and democracy in historical perspective. *Quarterly Journal of Economics* 115, 1167–1119.
- Aghion, P., A. Alesina, and F. Trebbi (2004). Endogenous political institutions. *Quarterly Journal of Economics* 119, 565–611.
- Aldrich, J. H. and D. W. Rohde (2001). The logic of conditional party government: Revisiting the electoral connection. In L. Dodd and B. Oppenheimer (Eds.), *Congress Reconsidered* (7th ed.), pp. 269–292. CQ Press.
- Alford, J. and D. Brady (1989). Personal and partisan advantage in U.S. congressional elections. In L. Dodd and B. Oppenheimer (Eds.), *Congress Reconsidered* (4th ed.). New York, NY: Praeger Press.
- Alt, J., E. Bueno de Mesquita, and S. Rose (2011). Disentangling accountability and competence in elections: Evidence from U.S. term limits. *The Journal of Politics* 73(1), 171–186.
- Altonji, J., T. Elder, and C. Taber (2005). Selection on observed and unobserved: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* 113(1), 151–184.
- Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless economics: An empiricist's companion*. Princeton University Press.
- Ansolabehere, S. and J. Snyder (2004). Using term limits to estimate incumbency advantage when officeholders retire strategically. *Legislative Studies Quarterly* 29(4), 487–515.
- Ansolabehere, S., J. 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.

- Antonakis, J. and O. Dalgas (2009). Predicting elections: Child’s play! *Science* 323(5918), 1183.
- Arnold, F. and R. Freier (2015). Signature requirements and citizen initiatives: Quasi-experimental evidence from Germany. *Public Choice* 162, 43–56.
- Asatryan, Z., T. Baskaran, T. Grigoriadis, and F. Heinemann (2015). Direct democracy and local public finances under cooperative federalism. *Scandinavian Journal of Economics* (forthcoming).
- Autor, D. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics* 21, 1–42.
- Baqir, R. (2001). Government spending, legislature size, and the executive veto. IMF Working Paper 01/208.
- Barro, R. (1973). The control of politicians: An economic model. *Public Choice* 14(1), 19–42.
- Bartels, L. M. (1996). Uninformed voters: Information effects in presidential elections. *American Journal of Political Science* 40(1), 194–230.
- Baskaran, T. (2015). Intergovernmental transfers, local fiscal policy, and the flypaper effect: Evidence from a German State. *Finanzarchiv / Public Finance Analysis* (forthcoming).
- Baskaran, T. and M. Lopes da Fonseca (2014). Electoral thresholds and the success of minor parties. CEPR Discussion Papers 177.
- Baskaran, T. and M. Lopes da Fonseca (2016). Electoral competition and endogenous political institutions: quasi-experimental evidence from Germany. *Journal of Economic Behavior and Organization* 122, 43–61.
- Bellettini, G., C. B. Ceroni, and G. Prarolo (2014). Knowing the right person in the right place: Political connections and resistance to change. *Journal of the European Economic Association* 12(3), 641–671.
- Berry, C. R. and W. G. Howell (2007). Accountability and local elections: Rethinking retrospective voting. *The Journal of Politics* 69(3), 844–858.
- Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford University Press.
- Besley, T. and A. Case (1995). Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits. *The Quarterly Journal of Economics* 110(3), 769–798.
- Besley, T. and A. Case (2003). Political institutions and policy choices: Evidence from the United States. *Journal of Economic Literature* 41(1), 7–73.

- Black, G. S. (1972). A theory of political ambition: Career choice and the role of structural incentives. *The American Political Science Review* 66(1), 144–159.
- Borge, L.-E. (2005). Strong politicians, small deficits: Evidence from Norwegian local governments. *European Journal of Political Economy* 21, 325–344.
- Brender, A. (2003). The effect of fiscal performance on local government election results in Israel: 1989-1998. *Journal of Public Economics* 87, 2187–2205.
- Buchanan, J. M. (2008). Same players, different game: How better rules make better politics. *Constitutional Political Economy* 19, 171–179.
- Calónico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for Regression-Discontinuity designs. *Econometrica* 82(6), 2295–2326.
- Calvo, E. and J. P. Micozzi (2005). The governor’s backyard: A seat-vote model of electoral reform for subnational multiparty races. *The Journal of Politics* 67 (4), 1050–1074.
- Carey, J. M. and S. Hix (2011). The electoral sweet spot: Low-magnitude proportional electoral systems. *American Journal of Political Science* 55 (2), 383–397.
- Carson, J., E. Engstrom, and J. Roberts (2007). Candidate quality, the personal vote , and the incumbency advantage in congress. *The American Political Science Review* 101(2), 289–301.
- Chari, V. V., L. E. Jones, and R. Marimon (1997). The economics of split-ticket voting in representative democracies. *American Economic Review* 87(5), 957–976.
- Conover, P. J. and S. Feldman (1989). Candidate perception in an ambiguous world: Campaigns, cues, and inference processes. *American Journal of Political Science* 33(4), 912–940.
- Cox, G. (1997). *Making votes count*. Cambridge: Cambridge University Press.
- Cox, G. and J. Katz (1996). Why did the incumbency advantage in U.S. elections grow? *American Journal of Political Science* 40(2), 478–497.
- Crain, W. M. and L. K. Oakley (1995). The politics of infrastructure. *Journal of Law and Economics* 38(1), 1–17.
- Crain, W. M. and R. D. Tollison (1993). Time inconsistency and fiscal policy: Empirical analysis of U.S. states, 1969-89. *Journal of Public Economics* 51(2), 153–159.
- De Jong, P., M. Lindeboom, and B. van der Klaauw (2011). Screening disability insurance applications. *Journal of the European Economic Association* 9(1), 106–129.
- Doron, G. and M. Maor (1991). Barriers to entry into a political system. *Journal of Theoretical Politics* 3, 175–188.

- Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper and Row.
- Drometer, M. and J. Rincke (2014). Electoral competition and endogenous barriers to entry. *European Journal of Political Economy* 34, 253–262.
- Duverger, M. (1954). *Political parties: their organization and activity in the modern state*. New York: Wiley.
- Egger, P. and M. Koethenbueger (2010). Government spending and legislative organization: Quasi-experimental evidence from Germany. *American Economic Journal: Applied Economics* 2, 200–212.
- Eggers, A., R. Freier, V. Grembi, and T. Nannicini (2015). Regression discontinuity designs based on population thresholds: Pitfalls and solutions. DIW Discussion Papers 1503.
- Ely, J. C. and J. Välimäki (2003). Bad reputation. *The Quarterly Journal of Economics* 118(3), 785–814.
- Erikson, R. (1971). The advantage of incumbency in congressional elections. *Polity* 3(3), 395–405.
- Erikson, R. and R. Titunik (2015). Using regression discontinuity to uncover the personal incumbency advantage. *Quarterly Journal of Political Science* 10(1), 101–119.
- Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice* 50(1), 5–25.
- Ferraz, C. and F. Finan (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review* 101(4), 1274–1311.
- Fiorina, M. (1989). *Congress: Keystone of the Washington establishment*. Yale University Press, New Haven.
- Fiorina, M. P. (1976). The voting decision: Instrumental and expressive aspects. *Journal of Politics* 38, 390–415.
- Fiva, J. H. and O. Folke (2014). Mechanical and psychological effects of electoral reform. *British Journal of Political Science* (forthcoming).
- Fowler, A. and A. Hall (2014). Disentangling the personal and partisan incumbency advantage: Evidence from close elections and term limits. *Quarterly Journal of Political Science* 9(4), 501–531.
- Gebethner, S. (1997). Free elections and political parties in transition to democracy in Central and Southeastern Europe. *International Political Science Review* 18 (4), 381–399.

- Gelman, A. and G. King (1990). Estimating incumbency advantage without bias. *American Journal of Political Science* 34(4), 1142–64.
- 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.
- Grembi, V., T. Nannicini, and U. Troiano (2015). Do fiscal rules matter? *American Economic Journal: Applied Economics* (forthcoming).
- Gulino, G. (2014). Do electoral systems affect the incumbent probability of re-election? Evidence from Italian municipalities. Mimeo (University of Bologna).
- Hahn, J., P. Todd, and W. Ven der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1), 201–209.
- Hayo, B. and S. Voigt (2010). Determinants of constitutional change: Why do countries change their form of government? *Journal of Comparative Economics* 27, 283–305.
- Hayo, B. and S. Voigt (2013). Endogenous constitutions: Politics and politicians matter, economic outcomes don't. *Journal of Economic Behavior & Organization* 88, 47–61.
- Hessami, Z. (2014). Appointed versus elected mayors and incentives to pork-barrel: Quasi-experimental evidence from Germany. University of Konstanz Working Paper Series 2014-23.
- Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81(396), 945–960.
- Hoxby, C. (2000). The effects of class size on student achievements: New evidence from population variation. *Quarterly Journal of Economics* 115, 1239–1985.
- Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies* 79(3), 933–959.
- Imbens, G. and T. Lemieux (2007). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2), 615–635.
- Jacobson, G. (1978). *Money in Congressional Elections*. New Heaven, CT: Yale University Press.
- Janvry, A., F. Finan, and E. Sadoulet (2012). Local electoral incentives and decentralized program performance. *The Review of Economics and Statistics* 94(3), 672–685.
- Kostroski, W. L. (1973). Party and incumbency in post senate elections: Trends, patterns and models. *The American Political Science Review* 67(4), 1213–1234.

- Lee, D., E. Moretti, and M. Butler (2004). Do voters affect or elect politicians? Evidence from the U.S. House. *Quarterly Journal of Economics* 119, 807–859.
- Lee, D. J. (2013). Third-party threat and the dimensionality of major-party roll call voting. *Public Choice* 159, 515–531.
- Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics* 142(2), 675–97.
- Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2), 281–355.
- Levitt, S. and C. Wolfram (1997). Decomposing the sources of incumbency advantage in the U.S. House. *Legislative Studies Quarterly* 22(1), 45–60.
- Lieske, J. (1989). The political dynamics of urban voting behavior. *American Journal of Political Science* 33(1), 150–174.
- List, J. A. and D. M. Sturm (2006). How elections matter: Theory and evidence from environmental policy. *The Quarterly Journal of Economics* 121(4), 1249–1281.
- Lizzeri, A. and N. Persico (2005). The drawbacks of political fragmentation. *Journal of the European Economic Association* 3, 809–827.
- Lopes da Fonseca, M. (2015a). Identifying the source of incumbency advantage through a constitutional reform. CEGE Discussion Papers 239.
- Lopes da Fonseca, M. (2015b). Lame but loyal ducks. CEGE Discussion Papers 254.
- Ludwig, J. and D. L. Miller (2007). Does head start improve children’s life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* 122(1), 159–208.
- Malani, A. and J. Reif (2015). Interpreting pre-trends as anticipation: Impact on estimate treatment effects from tort reform. *Journal of Public Economics* 124, 1–17.
- Maskin, E. and J. Tirole (2004). The politician and the judge: Accountability in government. *American Economic Review* 94(4), 1034–1054.
- Mayhew, D. R. (1974). *Congress: The electoral connection*. Yale University Press, New Haven.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142, 698–714.
- Morris, S. (2001). Political correctness. *Journal of Political Economy* 109(2), 231–265.

- Moser, P. and A. Voena (2012). Compulsory licensing: Evidence from the trading with the enemy act. *American Economic Review* 102(1), 396–427.
- Moser, R. G. and E. Scheiner (2004). Mixed electoral systems and electoral system effects: Controlled comparison and cross-national analysis. *Electoral studies* 23, 575–599.
- Moser, R. G. and E. Scheiner (2012). *Electoral systems and political context*. Cambridge, UK: Cambridge University Press.
- Oliver, J. E. and S. E. Ha (2007). Vote choice in suburban elections. *The American Political Science Review* 101(3), 393–408.
- Paldam, M. (1986). The distribution of election results and the two explanations of the cost of ruling. *European Journal of Political Economy* 2(1), 5–24.
- Pellicer, M. and E. Wegner (2014). The mechanical and psychological effects of legal thresholds. *Electoral Studies* 33, 258–266.
- Peltzman, S. (1987). Economic conditions and gubernatorial elections. *American Economic Review* 77(2), 293–297.
- Peltzman, S. (1992). Voters as fiscal conservatives. *Quarterly Journal of Economics* 107(2), 327–361.
- Perea, E. A. (2002). Individual characteristics, institutional incentives and electoral abstention in Western Europe. *European Journal of Political Research* 41, 643–673.
- Pettersson-Lidbom, P. (2012). Does the size of the legislature affect the size of the government: Evidence from two natural experiments. *Journal of Public Economics* 96, 269–278.
- Rae, D. W. (1971). *The political consequences of electoral laws*. New Haven: Yale University Press.
- Reber, S. J. (2005). Court-ordered desegregation: Successes and failures integrating american schools since Brown versus Board of Education. *The Journal of Human Resources* 40(3), 559–590.
- Rokkan, S. (1968). Electoral systems. In *International Encyclopedia of the Social Sciences*. Macmillan.
- Rubin, D. (1974). Estimating causal effects of treatment in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5), 688–701.
- Schlesinger, J. A. (1966). *Ambition and politics: Political careers in the United States*. Rand McNally, Chicago.

- Shepsle, K. and B. Weingast (1981). Political preferences for the pork barrel: A generalization. *American Journal of Political Science* 25, 96–111.
- Shugart, M., M. Valdini, and K. Suominen (2005). Looking for locals: Voter information demands and personal vote-earning attributes of legislators under proportional representation. *American Journal of Political Science* 49(2), 437–449.
- Smart, M. and D. M. Sturm (2013). Term limits and electoral accountability. *Journal of Public Economics* 107, 93–102.
- Snyder, J. and T. Groseclose (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science* 44, 187–205.
- Taagepera, R. (2007). *Predicting party sizes: the logic of simple electoral systems*. Oxford, UK: Oxford University Press.
- Taagepera, R. (2008). *Making social sciences more scientific: The need for predictive models*. Oxford, UK: Oxford University Press.
- Ticchi, D. and A. Vindigni (2010). Endogenous constitutions. *Economic Journal* 120, 1–39.
- Todorov, A., A. N. Mandisodza, A. Goren, and C. C. Hall (2005). Inference of competence from faces predict election outcomes. *Science* 308(5728), 1623–1626.
- Trebbi, F., P. Aghion, and A. Alesina (2008). Electoral rules and minority representation in U.S. cities. *Quarterly Journal of Economics* 123(1), 325–357.
- Uppal, Y. (2009). The disadvantaged incumbents: Estimating incumbency effects in Indian state legislatures. *Public Choice* 138(1/2), 9–27.
- Van der Klaauw, W. (2008). Regression-discontinuity analysis: A survey of recent developments in economics. *Labour* 22(2), 219–245.
- Vatter, A. (2003). Legislative party fragmentation in Swiss cantons: A function of cleavage structures or electoral institutions? *Party Politics* 9 (4), 445–461.