Is There an Incumbency Advantage or Cost of Ruling in Proportional Election Systems?*

Che-Yuan Liang#

July 21, 2009

Abstract: This paper empirically investigates the effects of political representation on the electoral outcome at the party and coalition levels in proportional election systems. There are two notions of representation, namely, to hold seats and to belong to the ruling coalition. I refer to the effect of the former as the incumbency effect and the effect of the latter as the ruling effect. I find that incumbency determines the distribution of 12 percent of the total vote, which is similar to the advantage found in majoritarian systems. Further, I find no ruling effect, contrary to the commonly-found cost of ruling in proportional systems.

Keywords: incumbency advantage, cost of ruling, proportional election systems, regression discontinuity, local governments

JEL classification: D72, D73

* I thank Niklas Bengtsson, Sören Blomquist, Mikael Elinder, Jon Fiva, Olle Folke, Per Johansson, Henrik Jordahl, Mikael Lindahl, Matti Liski, Natalia Pakhotina, Panu Poutvaara, Jeffrey Smith, Hannelore Weck-Hannemann, participants at the IEB summer school in public economics in Barcelona, the IIPF annual congress in Maastricht, the HECER political economy conference in Helsinki, and seminar participants at Uppsala University for valuable comments and suggestions, Henrik Jordahl for providing me some of the data used, as well as the Jan Wallander and Tom Hedelius Foundation for financial support.

Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, P.O. Box 513, SE-75120, Uppsala, Sweden; Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE-10215, Stockholm, Sweden; Phone: +46-(0)18-4711633; Fax: +46-(0)18-4711478; E-mail: che-yuan.liang@nek.uu.se.

1. Introduction

There are two main types of election systems in the world's democratic countries, namely, majoritarian and proportional systems. Majoritarian systems are based on the "winner take all" principle in which one party receives the (usually one and only) seat in an election district, which often results in two dominating parties (Duverger, 1954; Lijphart, 1994). In contrast, proportional systems typically distribute several mandates to parties roughly proportional to their votes in an election district, which encourages a greater number of large parties. Most English-speaking democracies have majoritarian systems, whereas most European democracies have proportional systems. Two different literatures, namely, the incumbency literature and the vote function literature, reach two apparently conflicting results regarding whether holding an office is beneficial for office holders in future elections. The incumbency literature finds an incumbency advantage in majoritarian systems (Gelman and King, 1990; Alford and Brady, 1993; Jacobson, 1997), whereas the vote function literature finds a cost of ruling in proportional systems (Paldam and Nannestrand, 1994).

Although striking, the results are not directly comparable, since the incumbency advantage usually relates to legislators and parties, and the cost of ruling usually relates to coalitions of parties. In addition, the incumbency literature has developed sophisticated identification strategies, whereas the vote function literature has identified the cost of ruling as a by-product of analyses that focus on economic or other political determinants of the electoral outcome. Nevertheless, it is surprising how distinctly the two literatures have evolved and that no previous study has tried to reconcile these results, as the differences are interesting from a comparative-politics perspective. This paper empirically investigates the effects of political representation on the electoral outcome for representatives in proportional election systems more carefully than previously. I use regression discontinuity (RD) designs for causal inference, which has previously only been used to analyze majoritarian systems. This puts the research on proportional systems in conversation with that on majoritarian systems.

Two different kinds of office-holding effects can be distinguished. First, parties can hold seats; this effect is henceforth referred to as the "incumbency effect". Second, a ruling coalition of parties can hold a majority of the seats; this effect is henceforth referred to as the "ruling effect". In majoritarian systems, these two effects coincide in an election district, as holding the only seat implies holding a majority of the seats there. In proportional systems, several parties can hold different shares of seats in a district and thus experience varying degrees of incumbency effects. However, only the ruling coalition of parties can, additionally,

experience a ruling effect. With this distinction, there can be an incumbency advantage *and* a cost of ruling in proportional systems. This paper investigates both effects of office holding in proportional systems. To my knowledge, the incumbency effect as defined here has never been investigated before in proportional systems.

The most commonly-suggested explanation for the incumbency advantage in majoritarian systems is that incumbents have more resources at their disposal than do challengers. Holding office makes available greater opportunities to promote the incumbent's popularity through media exposure, which improves fund-raising capabilities for electoral campaigns. Funds can in turn be spent on image building. Another resource is the extra experience that incumbents obtain by being in office an extra term as compared to their competitors. An indirect effect of resources may be that some challengers are intimidated from running for office in the face of this asymmetry. Cox and Katz (1996) find that all these components contribute significantly to the incumbency advantage. There are also other differences between incumbents and challengers; Bernhardt and Ingberman (1985) show that voters can perceive incumbents to be less risky alternatives due to reputational mechanisms, while Anderson and Glomm (1992) point out that incumbents are first-movers in selecting electoral platforms, though this is a dynamic that may sometimes become disadvantageous for the incumbents.

Paldam (1986) suggests back-swinging as an explanation for the cost of ruling in proportional systems. A coalition's popularity is assumed to have a permanent long-run and a temporary short-run component. Ruling coalitions are more likely temporarily successful than not. Ruling coalitions are then also more likely to perform worse in future elections as they tend to back-swing to their long-run level. Thus, the cost of ruling is a spurious relationship that disappears by accounting for the temporary component. Mueller (1970) suggests asymmetric voter reactions as another explanation. According to this explanation, the voters reward ruling coalitions less for good events than they punish them for bad events. Paldam and Skott (1995) instead propose a median-gap model. They use a two-party median-voter framework to analyze a divergent equilibrium in which coalitions are located on each side of the median voter. The median voter is expected to alternatingly vote for the two coalitions to moderate the average policy back toward its position. Since this voter's vote is assumed to be decisive, power shifts forth and back.

From a theoretical point of view, it is unclear whether political representation is beneficial for the representatives, as arguments in both literatures are to some degree applicable in both majoritarian and proportional systems. Resources may, for example, also depend on representation in proportional systems, and voters could, for instance, also respond asymmetri-

cally in majoritarian systems. Furthermore, the arguments in both literatures are applicable to both the incumbency effect ruling effects. Some resources may, for example, be connected to the share of seats controlled by a party, while other resources may be determined by whether a party belongs to the ruling coalition. The direction of the net effects is therefore an empirical matter that this paper aims to analyze for proportional systems. To separate the different mechanism is an interesting task that, however, is beyond this paper.

The main econometric difficulty in identifying the causal effects of political representation on the electoral outcome is omitted variables. A correlation between the current seat share and the future vote share may, for instance, reflect that a third variable, such as current voter preferences, directly affects both variables rather than a causal relationship between the first two variables. In other words, parties may have been selected into representation based on a large number of factors. The estimated effect is biased if we fail to adjust for all these factors, and it is usually difficult to account for all of them with a simple selection-on-observables approach, since some factors are unobservable, at least to the analyst.

Traditionally, three sophisticated approaches have been used to estimate the incumbent legislator effect in majoritarian systems. The sophomore surge approach, used by, for example, Erikson (1971) and Levitt and Wolfram (1997), uses the difference in vote shares between the first and second terms for winning challengers. The retirement slump approach, used by, for example, Payne (1980) and Ansolabehere and Snyder (2002), uses the difference between last-period winning incumbents and their freshmen successors. The Gelman-King approach, used by, for example, Gelman and King (1990) and Cox and Katz (1996) uses the difference between all incumbents and freshmen successors. Estimates of the incumbent legislator advantage vary between two and ten percent of the total vote. The incumbent party effect includes the legislator advantage, but it may have other components as well. Recently, the RD approach has gained popularity due to its credibility with regard to causal inference. Using this approach, Lee (2008) finds an incumbent party advantage of ten percent of the total vote for the U.S. House of Representatives.

In contrast, most studies in the vote function literature use a selection-on-observables approach to estimate the ruling effect in proportional systems. The earlier literature investigates country-level macro data, usually time series. The papers by Goodhart and Bhansali (1970), Mueller (1970), and Kramer (1971) are the seminal contributions. Newer literature, starting with the papers by Fiorina (1978) and Kinder and Kiewiet (1979), instead turns its attention to individual micro panels. Estimates of the cost of ruling are usually of the order of one to two percent of the total vote. However, this effect cannot be considered causal, but ra-

ther a correlation found as a byproduct of analyses in which the ruling effect is not the focus. Paldam (1986) provides the most comprehensive empirical study of the ruling effect. However, he confines the study to statistical regularities in the distribution of election results rather than attempting to identify a causal effect.

I use data on Swedish municipality elections. Local government data provide a larger number of observations than country-level data, which is essential in the RD designs that I use. The attention in the vote function literature has traditionally been on central government elections. Recently, the cost of ruling result has, however, also been confirmed in local government elections by, for example, Tellier (2005) and Akarca and Tansel (2006). There is therefore no reason to believe that the results for local governments would be irrelevant at other levels. The local governments are also major actors in the Swedish economy.

I use the discontinuous variation in the seat share as the vote share varies for parties in order to isolate exogenous variation in incumbency in a RD-like design. Folke (2007, 2008) uses a similar design to study the incumbency effect of parties on policy outcomes in Swedish local governments.[1] I find a statistically and economically significant incumbency advantage of 0.12 percent of the total vote for each percent of seats. The current seat distribution therefore determines the distribution of 12 percent of the total vote in the next election. This advantage is of the same magnitude as in majoritarian systems.

I use the discontinuous variation in ruling at the 50 percent seat share cutoff for coalitions in order to isolate exogenous variation in ruling. Pettersson-Lidbom (2008) uses such a RD design to investigate the ruling effect of parties on policy outcomes in Swedish local governments. Lee et. al. (2004), Lee (2008), and Ferreira and Gyourko (2009) use this design to investigate the incumbency effect of legislators or parties on policy and electoral outcomes in majoritarian systems. I find no cost of ruling contrary to the previous vote function literature. However, I can reproduce the cost of ruling result with a conventional specification like those typically used in the vote function literature. This indicates that previous studies do not satisfactorily adjust for omitted variables and that the cost of ruling is not a causal relationship.

The paper proceeds as follows. The next section presents Sweden's institutional background and the data. Section three reports some summary statistics. Section four investigates the incumbency effect. Section five investigates the ruling effect. The final section concludes.

---

[1] The idea of exploiting these discontinuities for causal inference was presented by Folke (2007) in a slide presentation held at the Institute for Labor Market Policy Evaluation in May 2007. Folke (2008) is a preliminary manuscript on this topic. The presentation and manuscript investigate party effects on tax rates, environmental policy, and refugee placement. The idea that incumbency could be investigated with a similar design was mentioned to me by Henrik Jordahl. As far as I know, no other studies use this kind of design.

2. Institutional Background and Data

Sweden has three levels of governments, namely, the central government, the counties and the municipalities. As of 2006, there were 21 counties (24 before 1998) and 290 municipalities. The local governments are major actors in the Swedish economy. The counties are responsible for health care, while the municipalities are responsible for day care, education, and care of the elderly. Together, their share of national GDP spending is around 20 percent, with one third spent by the counties and two thirds spent by the municipalities. The counties and municipalities also employ around 25 percent of the total Swedish workforce. The municipalities that existed in 2006 are the local governments of interest in this study.

Each municipality has a council, which is its highest decision-making assembly; the law prescribes a minimum size depending on the number of eligible voters. The size varies from 31 to 101 members. Elections are held simultaneously as parliamentary elections every third year before 1994 and every fourth year after that. The election date is the third Sunday in September. I use a complete data set obtained from Statistics Sweden for the period 1982–2006, which spans eight elections. Pre-eighties data are not easily accessible and somewhat problematic to merge with the current sample, as there were several extensive municipal amalgamation reforms in the fifties, sixties, and seventies. During the sample period, there was some redistricting of municipal borders, as 10 municipalities were split into 21. These municipalities are included as precision increases a bit. The results are, however, insensitive to their exclusion.

The election system is a multi-member-district proportional system in which seats are assigned to parties according to the modified Saint-Laguës highest averages method. First, all votes are divided by 1.4. The party with the highest quotient receives the first seat. A new quotient is then calculated for that party by dividing its quotient with another divisor. A second comparison is then made among the parties, and the party with the highest quotient receives the second seat. The procedure is repeated until all seats have been distributed. The divisor series is 3, 5, 7, 9, and so on. A real-life example that illustrates the procedure is given later in this section. The seat distribution method somewhat favors larger parties.

There is an explicit small-party barrier of four percent that excludes small parties from receiving seats in parliamentary elections, though not in municipal elections. Before 1998, the parties had full power as to the ordering of candidates, that is, a closed-list system was used. In 1998, an open-list system was introduced, although this has had little practical impact with regard to voter input on candidate ordering. In the parliament, for example, only around 10

out of 349 seats are usually allocated differently due to the open-list system as opposed to the closed-list system. Election turnout is generally high at over 80 percent.

During the sample period, 73 percent of local governments consisted of a single election district. As recognized in the vote function literature, the interesting level of analysis in proportional systems is the council level rather than the election district level, since political decision making is conducted mainly by parties and coalitions of parties in the council. This contrasts the incumbency literature on majoritarian systems, where the interesting level of analysis is the election district level since a single legislator from one party represents an election district. The results are, however, insensitive to using only local governments that consist of a single election district, in which case the two levels coincide.

Politics in the municipal council are usually characterized by stable coalitions during election periods. A coalition holding a majority of the seats can rule by uniform voting in legislative issues. The municipal council elects the municipal executive board. The board's main tasks are preparatory work for and the execution of the council's decisions. It consists of a legally prescribed minimum number of members. The exact size and composition is decided by the council, but there are usually 15 members. A chairman is also elected to lead the board. Representation in the council is the main object of analysis in this paper, since the council always has final decision-making power on all issues. It may also dismiss members of the executive board. In practice, there is also a high congruence between the composition of the council and of the executive board.

Since 1970, the Local Government Act has given municipalities the legal right to dedicate resources from the local budget to the parties represented in the council, although this has to be done in a way that does not disproportionally favor any parties (Svensk Författningssamling, 1991). The municipal public support for the parties is now an important source of revenue for the parties. The exact formula differs between municipalities, but it usually involves basic support to all parties with representation and additional support that depends on each party's share of seats. The support makes up a third of the parties' revenues. The rest of the party revenues come from the county council, the national parliament, and membership fees. None of these other sources are tied to representation in the municipal councils.

The variables of interest in the empirical analysis are described in Table 1. *Votes* is the municipal vote share of a party or coalition and expressed in terms of the percentage of the total vote. *Seats* is the municipal seat share of a party or coalition and is determined by the vote share and the seat assignment rule; it is expressed in terms of the percentage of seats. It is the measure of incumbency. As the number of seats is an integer, the seat share can only in-

crease discretely. Since the vote share is continuous[2], full proportionality cannot be achieved. *Difference* is the degree of overrepresentation (or underrepresentation) calculated as *Seats* – *Votes* and is a measure of the discrepancy from proportionality.

Table 1. *Description of notation*

Notation	Description
Votes	The vote share in the municipal election in terms of the percentage of the total vote
Seats	The seat share in the municipal council in terms of the percentage of seats
Difference	The discrepancy from proportionality in percentage points, Seats – Votes
Majority	The theoretical majority dummy; 1 if Seats > 50 and 0, otherwise
Coalition	The actual ruling dummy; 1 if in the ruling coalition and 0, otherwise
Chair	The executive board chairman dummy; 1 if chairing and 0, otherwise
M	The moderate party (conservative)
KD	The Christian democratic party
FP	The liberal party
C	The center party (agrarian)
S	The social democratic party (labor)
V	The left party
MP	The green party
Rest	Other parties
Right	The right bloc consisting of M, KD, FP, and C
Left	The left bloc consisting of S and V

Majority is a dummy that takes the value of one if the coalition receives more than 50 percent of seats in the municipal council and zero, otherwise. *Coalition* is a dummy that takes the value of one if the coalition is recognized as the ruling coalition and zero, otherwise. *Chair* is a dummy that takes the value of one if the coalition holds the chairman position in the municipal executive board and zero, otherwise.

An index *i* is used to index the unit of observation. A party in a municipality forms a unit in the case of incumbency and a coalition of parties in a municipality in the case of ruling. An index *t* is used to index election periods. An observation is therefore a unit in an election period.

The seven main parties in Swedish politics, abbreviated by *M*, *KD*, *FP*, *C*, *S*, *V*, and *MP*, are also listed in Table 1. Only these parties have consistently held seats in the parliament during the sample period, and they are also the most important parties at the municipal level. Except for *KD* and *MP*, these parties have historical roots that go back to the introduction of universal suffrage. *Rest* groups together all other parties. Except for *MP*, the parties can be

[2] Although this is not true, technically the vote share is much more continuous than the seat share.

ordered along a right-left scale roughly as presented in the table, with *M* being the rightmost party and *V* the leftmost party.

Although Sweden technically is a genuine multiparty system, politics are rather polarized with two blocs, namely, a right-liberal and a left-socialist bloc. This has lead Alesina et al. (1997) and Laver and Shofield (1990) to classify Sweden as a bipartisan system. *Right* is used to denote the right bloc consisting of *M*, *KD*, *FP*, and *C*, and *Left* is used to denote the left bloc consisting of *S* and *V*. This dividing line is commonly accepted within Swedish politics. However, *MP* is sometimes also classified as a left-bloc party, as they sometimes are included in left-bloc coalitions. At the municipal level, *MP* has ruled together with other left-bloc parties four times more often than with right bloc parties. However, it claims itself to be bloc-independent. The results are, however, insensitive to including *MP* in the left bloc.

To illustrate how the municipal election system works, the seat distribution is derived given the vote distribution in the Torsby municipality after the 1994 election in Table 2 as an example. Torsby had 14,981 inhabitants, among which 11,856 were eligible voters, 49 council members, and a turnout rate of 85 percent, which is quite close to the municipal median in these respects. The number and percentage of votes received are reported in the *# votes* and *% votes* rows, respectively. The next rows report the first and last five quotients according to which the marginal seats were distributed. The party with the highest quotient marked in italics was the one that won the marginal seat, while a new quotient is calculated for that party with a higher divisor as described previously. We see, for example, that *S* received the first seat and *V* the last seat. The distribution of the number and percentage of seats are shown in the *# seats* and *% seats* rows, respectively. In this example, all parties became incumbent. The left bloc gained a majority of the seats, and they also formed a ruling coalition and elected a left bloc chairman (S) to the executive board.

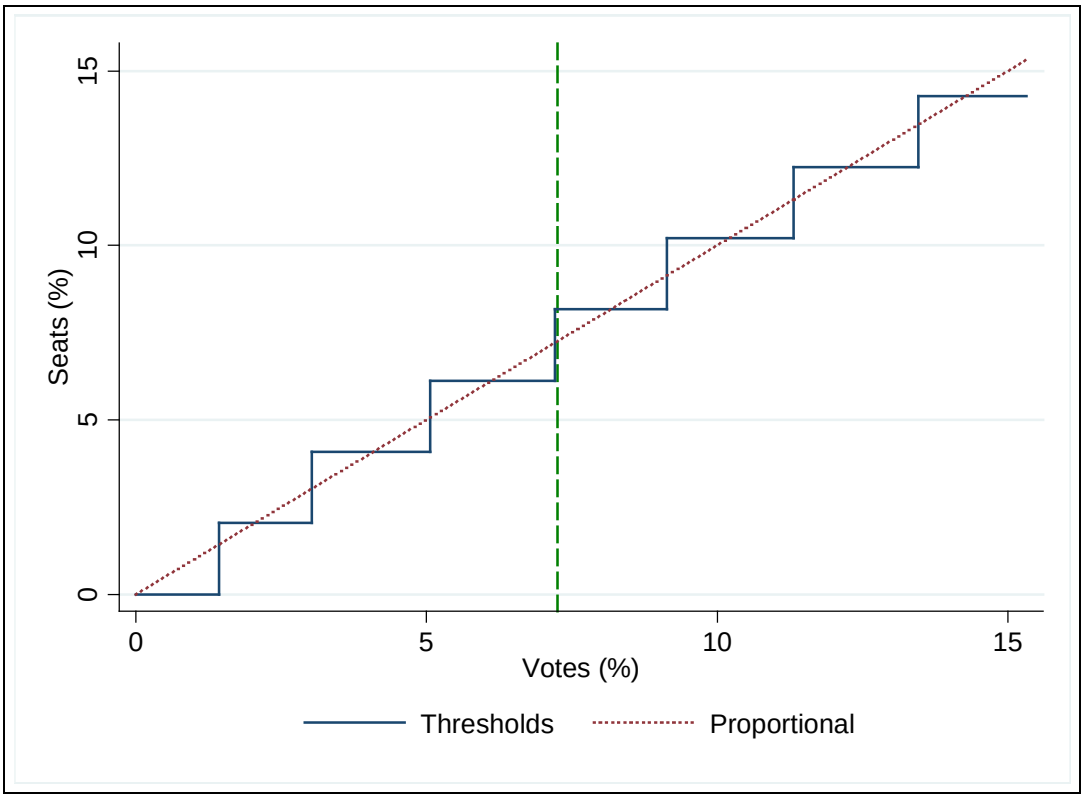
To illustrate when additional seats are received by *V*, the seat share is plotted against the vote share for *V*, given the council size and the number of votes received by other parties, in Figure 1. The discontinuous variations in the seat share create a staircase-like threshold structure, as Folke (2007) recognizes. A vertical line marks the actual vote share received by *V*, that is, four seats. A diagonal dashed line with unit slope marks when representation becomes fully proportional. We see that *V* was overrepresented, although it would have been underrepresented had it received just below the number of votes necessary to make the previous threshold.

Table 2. *The seat distribution in Torsby after the 1994 election*

Party	M	KD	FP	C	S	V	MP
# votes	2807	156	170	1059	4861	726	233
% votes	27.964	1.554	1.694	10.550	48.426	7.233	2.321
1st	2,005	111	121	756	*3,472*	519	166
2nd	*2,005*	111	121	756	1,620	519	166
3rd	936	111	121	756	*1,620*	519	166
4th	936	111	121	756	*972*	519	166
5th	*936*	111	121	756	694	519	166
…	…	…	…	…	…	…	…
45th	104	111	57	96	*113*	104	78
46th	104	*111*	57	96	108	104	78
47th	104	52	57	96	*108*	104	78
48th	*104*	52	57	96	103	104	78
49th	97	52	57	96	103	*104*	78
# seats	14	1	1	5	23	4	1
% seats	28.571	2.041	2.041	10.204	46.939	8.163	2.041

Notes: There were also 26 votes for other parties, an amount that did not qualify for any seats in the council. x^{th} refers to the quotients on which the distribution of the x^{th} marginal seat is based.

Figure 1. *The threshold structure for V in Torsby after the 2002 election*

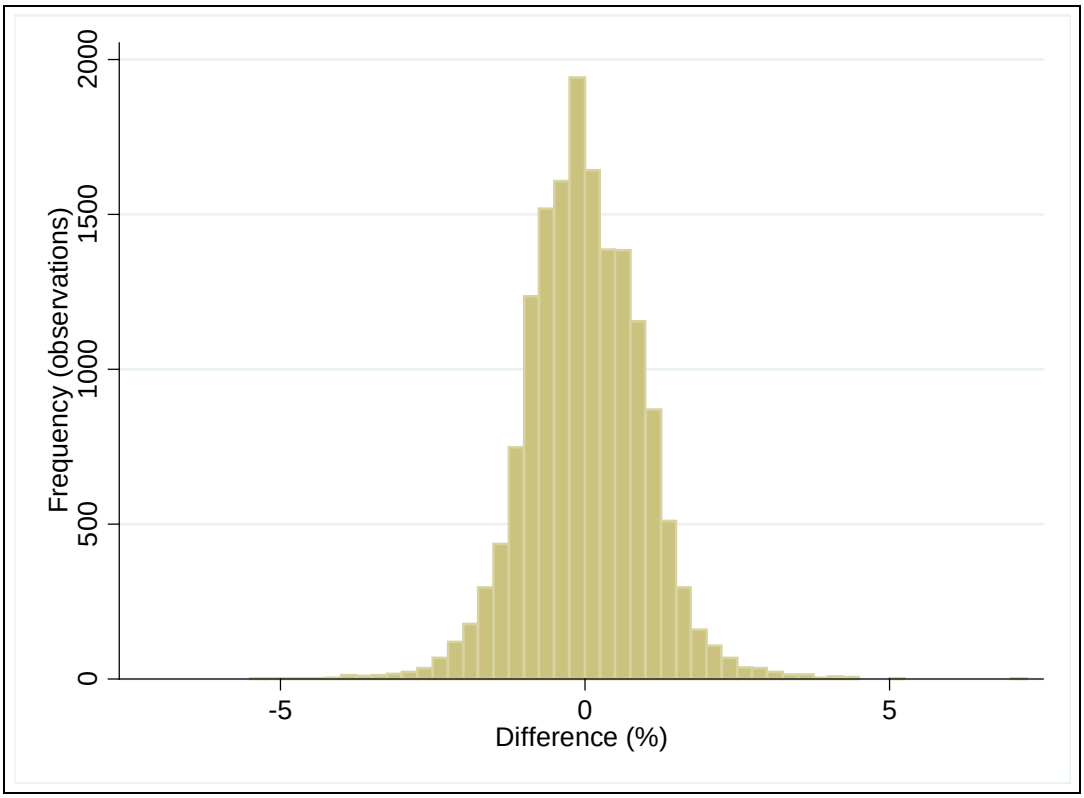


3. Summary Statistics

In the analysis of the incumbency effect, there are 2,320 units (each party in each local government is a unit) from eight parties (counting *Rest* as a party) and 290 local governments. Due to the use of first-lagged variables, a cross-section of observations is lost. Observations of units without candidates in an election period (which is the case for 20 observations) are dropped, although the results are insensitive to their inclusion. The panel is unbalanced due to redistricting, resulting in a loss of 200 observations. This leaves 16,020 observations from 2,320 units and 7 elections.

Like Folke (2007), I use the variation in the discrepancy from proportionality for identification. To get a sense of the scope of this variation, the frequency distribution for *Difference* is plotted in the histogram in Figure 2. Although the discrepancy is small in general, there are many observations with substantial discrepancy from proportionality in absolute terms. There are, for example, more than 500 observations with over two percentage point of overrepresentation or underrepresentation, and the extrema span twelve percentage points.

Figure 2. *Histogram of the discrepancy from proportionality*



The incumbency effect is also separately investigated for each party and bloc. Summary statistics for the vote share (no parentheses) and the discrepancy from proportionality (parentheses) are reported in Table 3. We see that the right bloc consists of two larger and two smaller parties, whereas the left bloc consists of one larger and one smaller party. Standard deviations are high and extremum values show a large span. The average deviation from proportionality is close to zero, and standard deviations are around one percentage point. We also see that there is a positive correlation between the two variables, a pattern produced by the seat distribution method. At the bloc level, the two blocs are quite equal-sized and show a similar distribution. This contrasts with what is observed at the parliamentary level, where there is clearer left-wing dominance.

Table 3. *Descriptive statistics of the share of votes and the discrepancy from proportionality*

Party	Mean	Std. Dev.	Min	Max
M	17.348	8.656	0.205	68.070
	(0.266)	(0.858)	(-2.211)	(4.118)
KD	5.251	3.976	0.012	44.313
	(-0.120)	(0.848)	(-4.634)	(2.343)
FP	8.083	4.117	0.121	30.117
	(0.077)	(0.778)	(-2.442)	(2.956)
C	14.628	8.630	0.532	51.900
	(0.137)	(0.841)	(-2.886)	(2.778)
S	39.974	9.461	8.101	66.952
	(0.686)	(1.173)	(-2.677)	(7.017)
V	6.160	5.012	0.019	58.297
	(-0.062)	(0.802)	(-3.749)	(3.093)
MP	3.665	2.360	0.082	43.679
	(-0.200)	(0.810)	(-3.577)	(2.315)
Rest	4.942	6.268	0.008	44.378
	(-0.790)	(1.024)	(-5.431)	(2.300)
Right	45.310	11.635	10.081	87.883
	(0.359)	(1.284)	(-4.141)	(5.719)
Left	46.128	11.571	8.675	82.248
	(0.624)	(1.303)	(-3.102)	(6.509)

Notes: *Votes* in percentage is reported without parentheses, and *Difference* in percentage points is reported within parentheses. The data come from 290 local governments and 7 elections.

In the analysis of the ruling effect, I use the variation in the theoretical majority of a bloc for identification. Since the ruling of one bloc implies that the other bloc is in opposition in a two-bloc system, the blocs are almost mirrors of each other. I choose the left bloc for examination, though the results are insensitive to using the right bloc. There are 290 units (left-bloc coalitions), which is equal to the number of local governments. Due to the need to use first-lagged variables and redistricting, the seven elections provide 2,005 observations.

Since there are bloc-independent parties in Sweden, there are cases in which none of the blocs has a majority of the seats. The number of right-wing, left-wing, and unclear majorities are reported in Table 4. Although one bloc usually has the majority, there are many unclear majorities; moreover, their number increases over time. The results are, however, insensitive to different strategies used to address these unclear majorities, as discussed in section 5.

Table 4. *The number of right-wing, left-wing, and unclear majorities*

Majority	1982	1985	1988	1991	1994	1998	2002	Sum
Right	115	126	94	143	64	92	99	733
Left	148	127	124	72	145	113	108	837
Unclear	21	31	66	71	79	84	83	435
Sum	284	284	284	286	288	289	290	2,005

Another complication that arises due to concentrating on within-bloc coalitions is the existence of unconventional coalitions with bloc-independent or local parties or with parties across the bloc lines. Such coalitions are more likely when majorities are unclear at the bloc level. It is commonly believed that unconventional coalitions have gained in importance over time. Coalition data are available for the sub-period 1994–2002 and used in some specifications to address this complication. The number and different types of conventional and unconventional coalitions during this sub-period are reported in Table 5. We see that a majority of the coalitions consist of conventional coalitions, that 23 percent of the coalitions involve the bloc-neutral *MP* or a local party, and that only 16 percent are cross-bloc coalitions. The number of unstable coalitions is reported in the *Other* column. 6.7 percent of the coalitions are minority coalitions in which majority support is gathered on an issue-by-issue basis.

Table 5. *The number and different types of coalitions 1994–2002*

	Conventional	+MP/Local	Cross-bloc	Other	Sum
Frequency	504	197	138	28	867
Percentage	58	23	16	3	100

Despite these complications, empirical studies using Swedish election data, such as the papers by Johansson (2003), Jordahl (2006), Svaleryd and Vlachos (2007), and Pettersson-Lidbom (2008) frequently concentrate on bloc effects, implicitly assuming that the blocs form permanent coalitions. Several of them also find that the theoretical majority of the blocs affects or is affected by other variables. For example, Pettersson-Lidbom (2008) finds that left-bloc majorities cause higher taxes and spending.

4. The Incumbency Effect

4.1 Empirical Strategy

The incumbency effect of representation is the effect of the current vote share of a party in a municipal council on the vote share in the next election, that is, the effect of $Seats_{i,t}$ on $Votes_{i,t+1}$, where i indexes the parties in the local governments, and t indexes the election periods. There is an omitted-variables endogeneity problem when simply running a bivariate regression of $Votes_{i,t+1}$ on $Seats_{i,t}$ since $Seats_{i,t}$ is determined by several other observable and unobservable variables that may have direct effects on $Votes_{i,t+1}$. Voter preferences are such omitted variables. The resulting correlation is not a causal relationship. It is unlikely that a selection-on-observables strategy, such as adding control variables, could account for all omitted variables. What is needed is exogenous variation in the seat share.

I use an identification strategy that makes use of the variation in the seat share around each vote share cutoff. Recall that the cutoff determines when additional seats are received. An example of such cutoffs was illustrated in Figure 1 above. Locally, if $Votes_{i,t}$ is close enough to the cutoffs, the variation in $Seats_{i,t}$ is "as good as random", since a negligible variation in $Votes_{i,t}$ causes $Seats_{i,t}$ to vary, ceteris paribus. This is a RD-like design. See Imbens and Lemieux (2008) and Lee (2008) for an introduction to such designs. The difference between the design used here as compared to the standard design is that several thresholds that are located at different cutoffs for different units are used.

I estimate the following regression with the within estimator:

$$Votes_{i,t+1} = \beta_S Seats_{i,t} + f\left(Votes_{i,t}\right) + \mu_i + \lambda_{Party,t+1} + \varepsilon_{i,t+1}. \qquad (1)$$

$f(Votes_{i,t})$ is a polynomial in $Votes_{i,t}$; μ_i is a party-specific local government fixed effect; and $\lambda_{Party,t+1}$ is a party-specific election period dummy. β_S is an estimate of the incumbency effect. $\varepsilon_{i,t+1}$ is an error term, and I assume that it is uncorrelated with $Seats_{i,t}$ conditional on the other regressors, that is, $E(\varepsilon_{i,t+1}|Seats_{i,t}Votes_{i,t},\mu_i,\lambda_{Party,t+1}) = E(\varepsilon_{i,t+1}|Votes_{i,t},\mu_i,\lambda_{Party,t+1})$. Equation (1) is a dynamical model as the first-lagged dependent variable enters the regressors. A consistent estimation of the parameters of $Votes_{i,t}$ requires that the number of time periods is large (Nickell 1981). These are, however, only nuisance parameters without intrinsic interest in this setting, as we are only interested in the consistent estimation of β_S.

The threshold structure of a unit is characterized by the position of the cutoffs and is determined by the distribution of votes across parties, the number of parties, the council size, and the seat distribution method. Consider first the case in which the threshold structure is

fixed and can be treated as exogenously given for a unit. $Votes_{i,t}$ would then completely determine $Seats_{i,t}$. $Votes_{i,t}$ is therefore equivalent to the forcing variable in the standard RD design in which it would be the only variable that needs to be accounted for. Since the vote share is persistent between elections, $Votes_{i,t}$ does have an own effect on the dependent variable $Votes_{i,t+1}$. It is important to account for the forcing variable flexibly, as we do not know how it affects the outcome; I do so with a higher-order polynomial.

Identification of the incumbency effect is possible because the seat and vote shares are not perfectly correlated, which can be seen from Figure 1 above. In fact, between the thresholds at which additional seats are received, there is no correlation between the two variables, as varying the vote share slightly does not affect the seat share. At the thresholds, a negligible continuous variation in the vote share changes the seat share discontinuously. The seat share received is just the sum of these local discontinuous changes and thus has a random component. In fact, we do not even need to know where the thresholds are exactly in each case to use the discontinuities they create.

Since the first term in $f(Votes_{i,t})$ is $Votes_{i,t}$, the two first right-hand-side components in equation (1) can be rearranged as $\beta_S Seats_{i,t} + f(Votes_{i,t}) = \beta_S Difference_{i,t} + \beta_S Votes_{i,t} + f(Votes_{i,t}) = \beta_S Difference_{i,t} + g(Votes_{i,t})$. This reformulation shows that I try to isolate the random component in the variation of *Difference* for identification. This does not mean that the resulting estimate is confined to the effect of discrepancies rather than incumbency but only that the discrepancies provide the variation in incumbency used for identification. *Difference* is correlated with *Votes* for two reasons. Between the thresholds, holding the seat share constant, there is a perfect negative correlation, as *Difference* = *Seats* – *Votes*. Across threshold, there is, however, a positive correlation, as the seat assignment method favors larger parties. The vote share is therefore the main variable to control for.

Including $Votes_{i,t}$ removes most omitted variables concerns. We may, for instance, worry that demographics or the charisma of individual politicians affect $Seats_{i,t}$ as well as $Votes_{i,t+1}$. However, the effect on $Seats_{i,t}$ has to work through the variables determining it. In the present example, demographics or the charisma of individual politicians likely affect $Seats_{i,t}$ through party popularity via $Votes_{i,t}$. But since we control for $Votes_{i,t}$, we have taken care of this channel of causation.

A complication that arises in the present design is that the threshold structure varies between units and is hence not exogenous. Together with the vote share, the threshold structure determines the seat share, and the determinants of this structure may have their own effects on the outcome. The discrepancy from proportionality is, for example, smaller when

there are few large parties and/or when the council size is large. The influence of these factors on the dependent variable is unlikely to have a simple functional form. Party-specific local government fixed effects and party-specific election period dummies are included to address threshold structure effects in a flexible way.

The unit fixed effects address party-specific local-government invariant threshold effects. They remove any remaining endogeneity in local governments in which the council size and the vote distribution are fairly constant over time, and they eliminate concerns, such as exceptional discrepancies from proportionality due to the presence of a very large party in a local government. The time dummies address party-time invariant effects. They remove any remaining endogeneity caused by time trends for parties, and eliminate concerns such as exceptional discrepancies from proportionality due to temporary party popularity at the national level. Only the variation in seat shares across units and over time is used for identification. Fixed effects and time dummies also reduce error variance and can increase precision.

The high correlation between the seat share and the vote share as well as the inclusion of fixed effects and time dummies reduce the amount of variation in the seat share drastically. Is there sufficient variation left for identification? I run regressions of $Seats_{i,t}$ and $Difference_{i,t}$ on the other explanatory variables using a linear term in $Votes_{i,t}$ and obtain:

$$\hat{Seats}_{i,t} = \underset{(0.003)}{1.074} * Votes_{i,t} - \underset{(0.047)}{0.166} + \mu_i + \lambda_{Party,t} \quad R^2 = 0.996, \tag{2}$$

$$\hat{Difference}_{i,t} = \underset{(0.003)}{0.074} * Votes_{i,t} - \underset{(0.047)}{0.166} + \mu_i + \lambda_{Party,t} \quad R^2 = 0.167. \tag{3}$$

Heteroskedasticity-robust standard errors allowing for clustering at the local government level for each party are reported in parentheses. The seat share and the vote share co-vary almost one-to-one, and the explanatory variables explain almost all variation in the seat share as the coefficient of determination is close to one. However, most of the variation in the deviation from proportionality that is useful for identification and depicted in Figure 2 above remains. Using a fourth-order polynomial, which is the maximal order that is used later, increases the coefficient of determination indistinguishably in equation (2) and to 0.174 in equation (3).

4.2 Results

The incumbency effect results, estimated with equation (1), are reported in Table 6. The polynomial order in $Votes_{i,t}$ is varied horizontally. Including further orders has small effects. The specifications in Table 6 include party-specific local government fixed effects and party-

specific election period dummies. Heteroskedasticity-robust standard errors allowing for clustering at the local government level for each party are reported in parentheses.

Table 6. *The incumbency effect on the election result*

Dep: $Votes_{i,t+1}$	None	1st	2nd	3rd	4th
$Seats_{i,t}$	0.416***	0.118***	0.117***	0.113***	0.115***
	(0.019)	(0.029)	(0.029)	(0.029)	(0.029)
$Votes_{i,t}$		0.333***	0.349***	0.415***	0.320***
		(0.038)	(0.040)	(0.047)	(0.056)
$Votes_{i,t}^2/10^2$			-0.035	-0.368*	0.524
			(0.050)	(0.203)	(0.376)
$Votes_{i,t}^3/10^4$				0.397*	-2.155**
				(0.219)	(0.961)
$Votes_{i,t}^4/10^6$					2.172***
					(0.796)
FE (local party)	Yes	Yes	Yes	Yes	Yes
Year*Party	Yes	Yes	Yes	Yes	Yes
R^2	0.48	0.49	0.49	0.49	0.49

Notes: 16,020 observations are used in each regression. Heteroskedasticity-robust standard errors allowing for clustering within local governments for each party are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

The estimate of $Seats_{i,t}$ is an estimate of the incumbency effect. The *None* column reports a positive correlation. When including a polynomial in $Votes_{i,t}$, removing the positive vote share persistency, the point estimate decreases. The estimated effect does not vary much with polynomial order. This effect is of the order 0.12 percent of the total vote for each percent of seats and is statistically significant at the one percent level. That is, a party that moves from having 35 to 36 percent of the seats would increase its incumbency advantage by 0.12 percentage points in votes from (0.12*35=) 4.2 to (0.12*36=) 4.32 percent of the votes. If a party holds all (that is, 100 percent of) seats, it would enjoy an advantage of 12 percent of the total vote (with some extrapolation).

Another way to summarize the results is that the present seat distribution determines the distribution of 12 percent of the total vote in the next election. This advantage is close to the incumbency advantage found in majoritarian systems. It is large and economically significant. However, the relative advantage of one party versus another is much smaller in proportional systems, as the distribution of seats is more even across parties, while only one party holds all seats in majoritarian systems. Even a party that holds a small majority of the seats only has an edge of six percentage points in votes compared to a party without seats. The variation in discrepancy from proportionality, which spans 12 percentage points in seats, only induces a relative incumbency effect of (0.12*12=) 1.44 percentage points in votes.

Despite the small relative incumbency effects and the small, curious effects of discrepancy from proportionality, it should be emphasized that the predetermination of 12 percent of the total vote in an election makes incumbency an important determinant of the electoral outcome that is not merely of curious interest. Another remark is that an incumbency advantage does not imply any kind of growth without a limit pattern of support for incumbent parties; it does not even imply that an incumbent party on average can expect more votes in the next election than the previous one.

Normally, the "treatment effect" obtained when exploiting a discontinuity is a local effect that cannot be generalized to elsewhere. In the present case, several discontinuities scattered over a large domain of vote shares are used. The results are therefore very general, although not to the point at which one party holds all seats, as there are no such observations in the data.

The estimates of the polynomial show that the persistency between elections is statistically significant and large. Based on the first-order polynomial specification, we see that a third of the total vote in an election have the same distribution as in the previous election when the effect of the seat share has been partialled out. The election result therefore has a roughly three times larger effect on the next election result as compared to incumbency, although the estimates of the polynomial do not have any causal interpretation. The coefficient of determination is around 0.5.

The results should be insensitive to adding additional control variables if there is no remaining bias from omitted variables. I have added a large number of party-specific controls that may affect either the vote share or the threshold structure, and as such, they are potential sources of endogeneity. These estimates are not reported here, but they are available upon request. The first group of controls consists of council size and the number of eligible voters, which may affect the threshold structure. The second group involves population, population changes, population density, per capita tax base, the percentage of the population under 18, and the percentage of the population over 65. These demographic variables may affect the vote share. The third group is a set of dummies describing whether there was a left bloc, right bloc, or unclear majority before the election. This is to control for eventual effects of ruling. Note that the estimates of the controls do not have any causal interpretation.

Although most of these controls are normally considered to be predetermined, I choose to include them from the same election as $Seats_{i,t}$, that is, from period t, which is one period prior to the dependent variable from period $t+1$. This avoids potential bias from including post-treatment controls, which may remove some channels of causation from the main inde-

pendent variable (Rosenbaum, 1994; Rubin, 2004). The results are, however, similar when including controls from period *t+1*. Many of these controls are statistically significant. However, including them does not change the point estimates or standard errors of the incumbency effect much. Another sensitivity test that has been performed is to leave out local governments involved in local government redistricting. The results stay the same. A third test is to leave out local governments with multiple election districts for which there are multiple threshold structures. This also does not change the results.

By pooling data across parties in the main analysis, the incumbency effect and the persistency of the vote share are assumed to be constant across parties. To relax these assumptions, party-specific incumbency effects allowing for party-specific vote share persistency are also estimated. Equation (1) is then separately estimated for each party. These party-wise incumbency effect results are reported in Table 7. The polynomial order in $Votes_{i,t}$ is varied horizontally, and party is varied vertically. The polynomial estimates are left out. Each estimate is hence an estimate of the incumbency effect for one party or bloc from one separate regression. All specifications include local government fixed effects and election period dummies. Heteroskedasticity-robust standard errors allowing for clustering at the local government level are reported in parentheses.

Table 7. *The incumbency effect on the election result by party*

Dep: $Votes_{i,t+1}$	None	1st	2nd	3rd	4th
M	0.360***	0.139*	0.149**	0.142**	0.140**
	(0.045)	(0.072)	(0.072)	(0.071)	(0.071)
KD	0.478***	0.075*	0.096**	0.074*	0.080*
	(0.080)	(0.042)	(0.046)	(0.043)	(0.044)
FP	0.270***	0.108*	0.103	0.083	0.080
	(0.028)	(0.063)	(0.063)	(0.064)	(0.064)
C	0.486***	0.134*	0.129	0.141*	0.149*
	(0.040)	(0.081)	(0.084)	(0.084)	(0.084)
S	0.362***	0.058	0.055	0.053	0.053
	(0.030)	(0.076)	(0.076)	(0.076)	(0.076)
V	0.529***	0.150*	0.175**	0.133*	0.137*
	(0.100)	(0.076)	(0.073)	(0.072)	(0.074)
MP	0.231***	0.153**	0.088**	0.088**	0.091**
	(0.019)	(0.061)	(0.040)	(0.042)	(0.043)
Rest	0.457***	0.233*	0.233*	0.234*	0.180
	(0.033)	(0.139)	(0.139)	(0.141)	(0.136)
FE	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes

Notes: Each estimate is an estimate of $Seats_{i,t}$ for one party from one separate regression. The polynomial order in $Votes_{i,t}$ is varied horizontally. Heteroskedasticity-robust standard errors allowing for clustering within local governments are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

The estimates vary somewhat with polynomial order, but they are quite stable across the higher-order specifications. There is an incumbency advantage for all parties, even if the effects vary substantially between 0.05 and 0.19 percent of the total vote for each percent of seats, with the lowest value for *S* and the highest value for *Rest*. Statistical confidence levels vary, but the estimates are statistically significant for a majority of the parties. For the right-wing parties, the incumbency advantage is larger for the two larger parties (*M* and *C*) and smaller for the two smaller parties (*KD* and *FP*). For the left-wing parties, the incumbency advantage is smaller for the larger party (*S*) and larger for the smaller party (*V*). The differences between some parties are statistically significant at the ten percent level, both for parties within and between blocs. Some statistically significant differences are, however, expected even when there are no real differences. Precision is in fact too low to rule out that the differences are not due to noise.

I have also estimated the effect of the first seat received separately by using only observations in which the number of seats equals one or zero. The advantage of the first seat turns out to be 0.296 percent of the total vote for each percent of seats, which is 2.5 times the advantage of other seats. Since this first seat on average corresponds to 2.35 percent of seats on average, it provides an advantage of 0.696 percent of the total vote for parties with one seat versus parties without seats, which is a quite large advantage for small parties. This result suggests that there could be a discrete incumbency effect in addition to the continuous effect allowed for in the main specification. Including a binary dummy for incumbency in addition to the seat share variable in the previous analysis does, however, generally lead to imprecise estimates of the discrete effect and does not change the estimate of the continuous effect much.

I have also conducted tests regarding whether the advantage is heterogeneous along other dimensions. These estimates are not reported in detail here, but they are available upon request. The effect has been separately investigated for sub-samples within different seat share intervals, for sub-samples with only overrepresented or underrepresented units, for sub-samples in the first and second half of the time period available, and for parties in ruling and non-ruling coalitions. All these tests reveal homogenous effects.

5. The Ruling Effect

5.1 Empirical Strategy

The ruling effect of representation is the effect of currently being in the ruling coalition in a municipal council on the vote share in the next election. There is an omitted-variables endogeneity problem when simply running a bivariate regression of the vote share on a previous ruling, since the coalition-forming process is determined by several other observable and unobservable variables that may have direct effects on the vote share. The relationship between the parties in the municipality is such an omitted variable. It is unlikely that a selection on observables strategy, such as adding control variables, could account for all omitted variables. What is needed is exogenous variation in ruling.

I confine the analysis to within-bloc coalitions and choose arbitrarily to work with the left bloc, as the right bloc should in principle provide a mirror image of the same results. Hence, I am interested in the effect of $CoalitionLeft_{i,t}$ on $VotesLeft_{i,t+1}$, where i indexes the local governments and t indexes the election periods. There are two reasons for only looking at within-bloc coalitions. First, this makes the results comparable to the previous literature. This is important, since I use conventional selection-on-observables during part of the analysis to replicate the cost of ruling result found in previous studies. Secondly, I use an identification strategy that is only applicable to within-bloc coalitions.

The identification strategy is to make use of the variation in the actual ruling of a bloc close to the 50 percent seat cutoff in which the theoretical majority shifts between blocs. Locally, if $SeatsLeft_{i,t}$ is close enough to the cutoff, the variation in $MajorityLeft_{i,t}$ is "as good as random", since a negligible variation in $SeatsLeft_{i,t}$ causes $MajorityLeft_{i,t}$ to vary, ceteris paribus. This is a RD design (Imbens and Lemieux, 2008; Lee, 2008) that has been used in majoritarian systems (Lee et. al., 2004; Lee, 2008; Ferreira and Gyourko, 2009) with the modification that the 50 percent cutoff in the *vote share* completely determines the *actual ruling*. Pettersson-Lidbom (2008) uses this design to investigate party effects on policy outcomes in Sweden. To interpret the effect of the theoretical majority as a ruling effect requires the assumption that the theoretical majority in most cases determines the actual ruling. Since the two variables co-vary strongly in Sweden, it is common to treat Sweden as a two-bloc system and to use the theoretical majority as a proxy for the actual ruling (Johansson, 2003; Jordahl, 2006; Svaleryd and Vlachos, 2007). This is also the starting point in this paper.

In my basic specification, I implement the RD design by estimating the following regression with ordinary least squares (OLS):

$$VotesLeft_{i,t+1} = \beta_0 + \beta_M MajorityLeft_{i,t} + f\left(SeatsLeft_{i,t}\right) + \varepsilon_{i,t+1}. \quad (4)$$

$f(SeatsLeft_{i,t})$ is a polynomial in $SeatsLeft_{i,t}$. A piecewise polynomial with two independent pieces is used, one on the non-ruling and one on the ruling side, divided by the 50 percent seat share cutoff, that is, when $MajorityLeft_{i,t}$ switches sign. The parameters of $SeatsLeft_{i,t}$ are nuisance parameters and should not be interpreted causally. β_M is an estimate of the effect of holding the theoretical majority for coalitions at the bloc level. $\varepsilon_{i,t+1}$ is an error term, and I assume that it is uncorrelated with $MajorityLeft_{i,t}$ conditional on the other regressors; that is, $E(\varepsilon_{i,t+1}|MajorityLeft_{i,t},SeatsLeft_{i,t}) = E(\varepsilon_{i,t+1}|SeatsLeft_{i,t})$. β_0 is a constant.

Since the seat share completely determines the majority, it is the forcing variable and thus the only control variable that needs to be accounted for. Since it is persistent between elections (as the vote share is persistent), it may have its own effect on the dependent variable that need to be accounted for. It is important to flexibly account for the forcing variable, as we do not know how it affects the outcome. I do this with a higher-order polynomial on each side of the threshold.

The identification condition requires that $SeatsLeft_{i,t}$ affects $VotesLeft_{i,t+1}$ continuously[3], whereas $MajorityLeft_{i,t}$ is allowed to affect $VotesLeft_{i,t+1}$ discontinuously when $SeatsLeft_{i,t}$ = 50 percent. In other words, we make use of the local random assignment in the majority at the threshold at which a negligible continuous variation in the seat share changes the majority discontinuously. This requires that locally at the cutoff, blocs cannot precisely manipulate the results to place themselves on a certain side of the cutoff. This assumption is reasonably fulfilled, as there is always some uncertainty in election results due to factors outside the control of the politicians. Additional controls are not needed for consistency in the RD design. The majority is conditionally exogenous, as any variable that affects it must do so through the seat share, which is partialled out. The continuity requirement is very mild compared to the identification conditions needed in most other identification strategies.

Although the theoretical majority is exogenous in the RD framework, the direct interpretation of the estimate as a ruling effect relies on it being a perfect proxy for the actual ruling. As a non-perfect proxy, it measures ruling with a measurement error. Since the proxy is conditionally exogenous, complications do not arise due to it being correlated with the error

[3] The seat share is technically not strictly continuous, but it is much more continuous than the majority.

term. The measurement error does, however, induce an attenuation bias and hence constitute a downward-biased estimate of the true ruling effect. This can be remedied by using the actual ruling variable, $CoalitionLeft_{i,t}$. This variable is, however, only available for a shorter sub-period.

The coalition variable is a complex endogenous variable. The theoretical majority of within-bloc coalitions does, however, provide exogenous variation. The correlation between the variables can be used to let the former variable isolate exogenous variation in the latter variable. At the threshold, a negligible change in the seat share alters the theoretical majority, which in turn alters the probability of the actual ruling, as the council decides on the actual ruling. We can investigate whether this locally random probability of the actual ruling affects the vote share in the next election at the bloc level. When the forcing variable fully determines the main independent variable, as in the basic specification, the design is a sharp RD design. Here, the forcing variable only determines the main independent variable probabilistically, and the design is then a fuzzy RD design.

The fuzzy RD design can be conceptualized as an instrumental-variables (IV) design in which the discontinuity is the instrumental variable. I implement the design by estimating the following regression with two-stage least squares (2-SLS):

$$VotesLeft_{i,t+1} = \beta_0 + \beta_X XLeft_{i,t} + f\left(SeatsLeft_{i,t}\right) + \varepsilon_{i,t+1}, \quad (5)$$

$$XLeft_{i,t} = \alpha_0 + \alpha_M MajorityLeft_{i,t} + g\left(SeatsLeft_{i,t}\right) + e_{i,t}, \quad (6)$$

where $X = Coalition$. Equation (5) is the second-stage structural equation, and equation (6) is the first-stage reduced-form equation. $XLeft_{i,t}$ is the instrumented variable, and $MajorityLeft_{i,t}$ is the instrument. Again, a piecewise polynomial with two independent pieces is used for the forcing variable $SeatsLeft_{i,t}$, one on each side of the 50 percent seat share cutoff, that is, when $MajorityLeft_{i,t}$ switches sign. The parameters of $SeatsLeft_{i,t}$ are nuisance parameters and should not be interpreted causally. β_X is an estimate of the effect of X, which is thus an estimate of the ruling effect. $\varepsilon_{i,t+1}$ and $e_{i,t}$ are error terms, and β_0 and α_0 are constants.

The first IV assumption is instrument exogeneity and requires that the instrument is validly excluded from the first-stage equation. Formally, I assume that $MajorityLeft_{i,t}$ is uncorrelated with $\varepsilon_{i,t+1}$ conditional on the other regressors, that is, $E(\varepsilon_{i,t+1}|MajorityLeft_{i,t}, SeatsLeft_{i,t}) = E(\varepsilon_{i,t+1}|SeatsLeft_{i,t})$ as in the sharp RD design in equation (4). The second instrumental variables assumption is instrument relevance, which require that $MajorityLeft_{i,t}$ is strongly enough correlated with $XLeft_{i,t}$. This assumption is formally tested later.

The exogeneity assumption in the fuzzy RD context requires the same kind of continuity assumption required for the sharp RD case, that is, that $SeatsLeft_{i,t}$ affects $VotesLeft_{i,t+1}$ continuously, whereas $MajorityLeft_{i,t}$ is allowed to affect $VotesLeft_{i,t+1}$ discontinuously when $SeatsLeft_{i,t} = 50$ percent. In addition, the exogeneity assumption in the instrumental variables context also entails an exclusion restriction, which requires that the instrument does not have any own effect on the outcome and that any effect of the instrument on the outcome goes through the instrumented variable, that is, that the theoretical majority only matters insofar as it sometimes leads to actual ruling.

Note that equation (4) in the sharp RD design is a reduced-form equation from the perspective of the fuzzy RD design implemented with the IV specification in equations (5) and (6). β_X in equation (5) can be obtained as β_M/α_M in equations (4) and (6). Hence, the ruling effect in the fuzzy design is just the effect of the theoretical majority on the vote share, inflated with the inverse of the change in the probability of the actual ruling that accompanies the change in the theoretical majority. This inflation is to correct for the attenuation bias that arises when using the theoretical majority as an imperfect proxy for the actual ruling.

I also investigate if chairing the executive board is of any importance. Again, this variable is complex and endogenous. I use the same identification strategy as for the actual ruling and use $MajorityLeft_{i,t}$ to isolate exogenous variation in $ChairLeft_{i,t}$. At the threshold, a negligible change in the seat share alters the theoretical majority status, which in turn alters the probability of chairing the executive board. The same IV specification described in equations (5) and (6) can be used with X now being *Chair*. The assumptions required are of the same type as previously described.

In RD designs, observations close to the threshold help most in identification, as we essentially compare close winners and losers assuming that these two groups become similar when the difference in the forcing variable becomes small. An alternative to extrapolate a control function in the forcing variable as in the specification in equations (4) to (6) is to directly compare observations close enough to the thresholds. That method is less efficient, as many observations are discarded. However, it does not rely on fitting the forcing variable properly by using observation far from the threshold; it can therefore be used to indicate whether the extrapolation method is consistent. Sensitivity tests are performed that combine both methods.

The results should be insensitive to adding additional control variables if there is no remaining bias from omitted variables. Local government fixed effects, election period dummies, as well as a large number of other control variables are included in sensitivity tests.

These are the same controls used in the sensitivity tests of the incumbency effect, although the controls for ruling have been dropped, as ruling is included as the main independent variable here, and the vote share is not included, even though it was included as the main covariate in the incumbency effect analysis. To ensure that these controls are predetermined, they are from the same period as the other independent variables, that is, from period *t*, although the inclusion of *t+1* variables from the same period as the dependent variable yields similar results.

5.2 Results

The basic reduced-form estimates of the majority, estimated with equation (4), are reported in Table 8. Seen as a proxy for ruling, they are conservative estimates of the ruling effect.. The polynomial order in $Seats_{i,t}$ is varied horizontally, and specification along other dimensions is varied vertically. The polynomial estimates are left out. Each estimate is hence an estimate from one regression. Sensitivity test results are reported in the last three specifications. The *5%* rows limit the sample to five percentage points on each side of the majority cutoff; the *FE+Year* rows add local government fixed effects and election period dummies to the basic full sample analysis; and the *FE+Year+Con* rows add additional controls. Heteroskedasticity-robust standard errors allowing for clustering at the local government level are reported in parentheses.

Table 8. *The basic reduced-form majority effect on the election result for the left bloc*

Dep: $VotesLeft_{i,t+1}$	None	1st	2nd	3rd	4th
Full	16.630***	-0.537	0.049	-0.228	0.192
	(0.811)	(0.485)	(0.627)	(0.749)	(0.884)
5%	4.323***	-0.221	0.162	-1.291	2.363
	(0.470)	(0.854)	(1.819)	(4.098)	(7.687)
FE+Year	1.250***	-1.110***	-0.136	-0.716	-0.494
	(0.386)	(0.341)	(0.416)	(0.468)	(0.561)
FE+Year+Con	-1.133***	-1.285***	-0.303	-0.408	-0.179
	(0.378)	(0.378)	(0.446)	(0.534)	(0.642)

Notes: The polynomial order in $Seats_{i,t}$ is varied horizontally. The polynomials are independently estimated for the $MajorityLeft_{i,t} = 0$ and the $MajorityLeft_{i,t} = 1$ sides in all regressions. Heteroskedasticity-robust standard errors allowing for clustering within local governments are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

First, focus on the *None* column. When using the full sample, the estimated effect is large and positive, indicating a positive raw correlation. When confined to five percentage points on each side of the threshold, the estimated effect decreases. As fixed effects and year dummies are added, the estimated effect decreases further, but it is still statistically significant and pos-

itive. When additional controls are added, the estimated effect remains statistically significant but becomes negative.

The last two estimated effects are conventional selection-on-observables estimates with different sets of controls, and they are comparable to those in the vote function literature. These estimates show that the sign depends on the exact set of controls that is included. Using the most extensive set of controls, I can replicate a statistically and economically significant cost of ruling result. Experimenting with other combinations of controls, I can produce both a larger or smaller benefit as well as cost of ruling.[4] However, serious doubts can be raised regarding interpreting such estimates as causal effects.

The forcing variable is accounted for by using different polynomials in the 1^{st} to 4^{th} columns. Partialling out the effect of this variable is necessary and sufficient for consistency, as discussed above. We see that the estimated effect disappears as the polynomial order is increased. For the higher-order specifications, the estimated effect is not statistically significant. The results stay robust when using the five percentage point sample, except for fluctuations at the highest orders, which is due to difficulties in fitting high-order polynomials well with only observations close to the threshold. Adding different sets of controls does not change the results, as expected with this design. The size of the estimates is also always economically small at less than a cost of ruling of 0.54 percent of the total vote.

Standard errors increase when the analysis is confined to the five-percentage point sample as expected. They decrease as fixed effects and year dummies are added and increase as additional controls are added. The *FE+Year* specifications with the highest precision have standard errors small enough to rule out a one percent ruling effect. My results therefore indicate a certain absence of a ruling effect.

Although the RD design very convincingly identifies causal effects, one should usually interpret the results cautiously. What is strictly obtained is the local treatment effect at the majority cutoff. In many cases, this may not be generalized to other domains of interest. In the present case, we are, however, exclusively interested in the effect at the cutoff, as majorities are always gained there.

I have conducted additional sensitivity tests. These estimates are not reported here, but they are available upon request. First, *MP* has been included as a left bloc party, with the same results. Second, the majority has been redefined as holding more than 50 percent of the

[4] The two set of controls are chosen to serve as sensitivity tests in the polynomial specifications rather than exactly replicating the two percent cost of ruling found in previous studies. Fixed effects and year dummies are standard controls in panel regressions and the other controls are simply the most extensive set at my disposal.

sum of the two blocs' seats, again with the same results. Third, observations with unclear majorities have been excluded, without changing the estimated ruling effect. Fourth, election periods have been removed, one at a time, starting with the last election, to address the large number of unclear majorities in recent years (as can be seen from Table 4 above), with similar results after each elimination. Fifth, a dummy has been included to allow unclear majorities to have an independent effect on the election result. The majority effect is still unaffected, and the effect of an unclear majority is close to zero as well. In sum, these tests show that the results are insensitive to different ways of addressing unclear majorities.

Sixth, corresponding regressions have been run for the right bloc, with the same absence of a majority effect. Seventh, the bloc majority effect on the vote share has been examined for each of the parties individually, without revealing any party-specific effects. Seventh, local governments involved in redistricting have been removed from the sample, and this does not change the results. Finally, estimations have been carried out with a binary variable that indicates whether a party has received a majority of the votes as the dependent variable in linear probability, Probit, or Logit models with similar results.

The IV estimates of the ruling effect and the effect of chairing the executive board, estimated with equations (5) and (6), are reported in Table 9. The ruling results are reported in the first rows, and the chairing results are reported in the latter rows. The polynomial order in $Seats_{i,t}$ is varied horizontally. The polynomial estimates are left out. Each estimate is hence an estimate from one regression. The *1st stage* rows report the first-stage estimates, and the *F-statistics* rows report the partial first-stage heteroskedasticity and autocorrelation robust F-statistics of the instrument $MajorityLeft_{i,t}$. The *2nd stage* rows report the second-stage estimates of $CoalitionLeft_{i,t}$ and $ChairLeft_{i,t}$ respectively. Heteroskedasticity-robust standard errors allowing for clustering at the local government level are reported in parentheses.

Gaining the theoretical majority in the council increases the probability of actual ruling for the left bloc according to the first-stage estimates of the coalition variable. The increase is around 60 percent in the higher-order estimates. The partial first-stage F-statistics of the theoretical majority instrument is well beyond the critical value of ten, which indicates no weak instrument concerns. The actual ruling does, however, not change the vote share in the next election according to the second-stage estimates. The higher-order estimates are statistically insignificant. This is expected, as the basic reduced-form estimates of the effect of the theoretical majority reported previously were statistically insignificant. The second-stage structural estimates just inflate the previous estimates.

Table 9. *The ruling effect and the effect of chairing the executive board for the left bloc*

Dep: $VotesLeft_{i,t+1}$		None	1st	2nd	3rd	4th
$CoalitionLeft_{i,t}$	1st stage	0.784***	0.666***	0.517***	0.588***	0.598***
instrumented		(0.023)	(0.048)	(0.065)	(0.080)	(0.099)
$MajorityLeft_{i,t}$	F-statistics	1143.03	195.46	62.61	54.27	36.24
instrument	2nd stage	21.438***	-0.321	1.706	1.180	1.961
98–06 sample		(1.159)	(0.899)	(1.543)	(1.631)	(1.894)
$ChairLeft_{i,t}$	1st stage	0.749***	0.502***	0.356***	0.294***	0.285***
instrumented		(0.019)	(0.031)	(0.041)	(0.051)	(0.061)
$MajorityLeft_{i,t}$	F-statistics	1551.49	261.97	77.71	33.07	21.89
instrument	2nd stage	22.197***	-1.065	0.106	-0.816	0.699
Full sample		(1.008)	(0.959)	(1.743)	(2.531)	(3.106)

Notes: 2-SLS is used. Each estimate is an estimate of $MajorityLeft_{i,t}$ from one separate regression in the *1st stage* rows, and an estimate of $CoalitionLeft_{i,t}$ or $ChairLeft_{i,t}$ in the *2nd stage* rows. The polynomial order in $Seats_{i,t}$ is varied horizontally. The polynomials are independently estimated for $MajorityLeft_{i,t} = 0$ and $MajorityLeft_{i,t} = 1$ in all regressions. Heteroskedasticity-robust standard errors and F-statistics allowing for clustering within local governments are reported in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Gaining the theoretical majority in the council also increases the probability of receiving the chairman position in the executive board for the left bloc according to the first-stage estimates of the chairman variable. The increase is around 30 percent in the higher-order estimates. The partial first-stage F-statistics of the theoretical majority instrument is well beyond the critical value of ten, which indicates no weak instrument concerns. Chairing the executive board does not, however, change the vote share in the next election. The higher-order estimates are statistically insignificant. Again, this is expected, as the second-stage structural estimates just inflate the previous basic reduced-form estimates of the theoretical majority.

The standard errors are larger than in the previous basic estimates. This is often the case with IV. For the ruling effect, the use of a small subsample contributes to the higher standard errors. For the chairing effect, the weaker first-stage relationship contributes to the higher standard errors.

Strictly, we obtain the local treatment effect at the majority cutoff also in the fuzzy RD design. Moreover, it is the compliers that are identifying, that is, those for which actual ruling or chairing of the executive board depend on the theoretical majority in the council. This restricted generalizability of the results is not severe, since actual ruling or chairing the executive board almost always is gained at or around the theoretical majority cutoff.

Let us take a step back and look at the basic reduced-form estimates in Table 8, and interpret them together with the first-stage reduced-form estimates in Table 9. The insignificant effects and small standard errors of former estimates and the significant effects of the latter indicate a reliable conclusion that there is no effect of the theoretical majority when any chan-

nels of causation are included, that is, when including the effect of the theoretical majority *per se* and its significant effects on the actual ruling and the chairing. With the exclusion restriction that the effect of the theoretical majority only may work by causing an actual ruling, we can additionally claim that there is no ruling effect. Similarly, we can claim that there is no chairing effect. However, the two exclusion restrictions cannot be tested and cannot hold simultaneously.

5.4 Illustration of Results

To illustrate the estimated results graphically, local averages of $VotesLeft_{i,t+1}$ are plotted against $SeatsLeft_{i,t}$ in Figure 3. The x-axis is divided into 250 non-overlapping bins, each spanning 0.4 percent of the total vote, and a dot represents the average of the observations in a bin along the y-dimension and the midpoint of a bin along the x-dimension. The fitted line for the basic reduced-form regression with a piecewise fourth-order polynomial allowing for a jump as ruling switches is also plotted. A vertical dashed line indicates when the switch takes place. We see from the local averages that the relationship is smooth. Nothing happens when ruling switches as compared to elsewhere. The jump allowed for in the specification cannot be distinguished visually. There is no visual effect of holding the theoretical majority.

Figure 3 also reveals that holding a small majority of the seats generally leads to less than 50 percent of the total vote. A larger majority is required before the majority bloc becomes the favorite. A dotted diagonal line with a unit slope is also added in Figure 3. We see that most local averages above 40 percent of the seats lie below this line. Performing well in an election and holding many seats are therefore associated with performing less well in the next election. There is a smaller opposite behavior for the local averages below 40 percent. This is a pattern of reversion toward 40 percent, which is consistent with the back-swinging explanation of the cost of ruling. However, this is a non-causal effect of holding seats on the election result and not a causal ruling effect.

The first-stage relationships between $CoalitionLeft_{i,t}$ and $ChairLeft_{i,t}$, and $SeatsLeft_{i,t}$ are plotted in Figures 4 and 5. The local averages and fitted lines are constructed as previously. Clearly, gaining the theoretical majority in the council has a significant impact on the chance of forming an actual ruling coalition and on the chance of chairing the executive board. The impact of the theoretical majority is discontinuous in comparison with the impact of the seat share elsewhere, which is continuous. However, the rise in these probabilities does not entail any changes in the election result according to Figure 3.

Figure 3. *Plot of local averages of election results against representation*

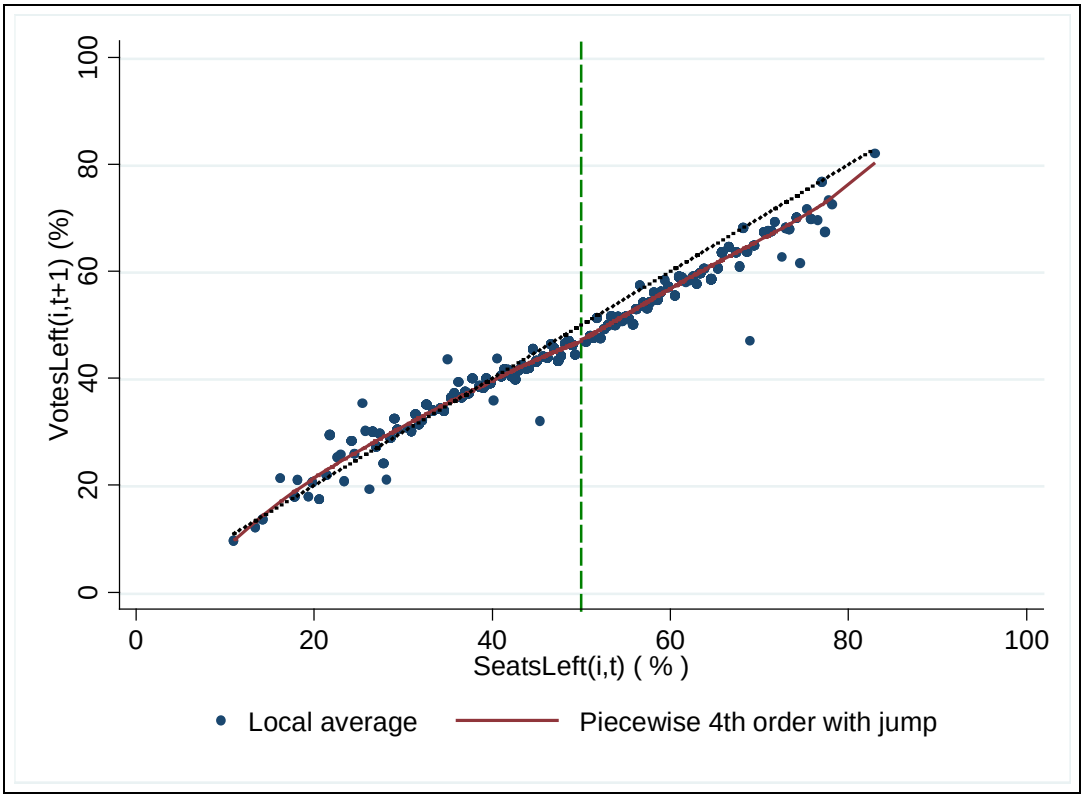


Figure 4. *Plot of local averages of ruling against representation*

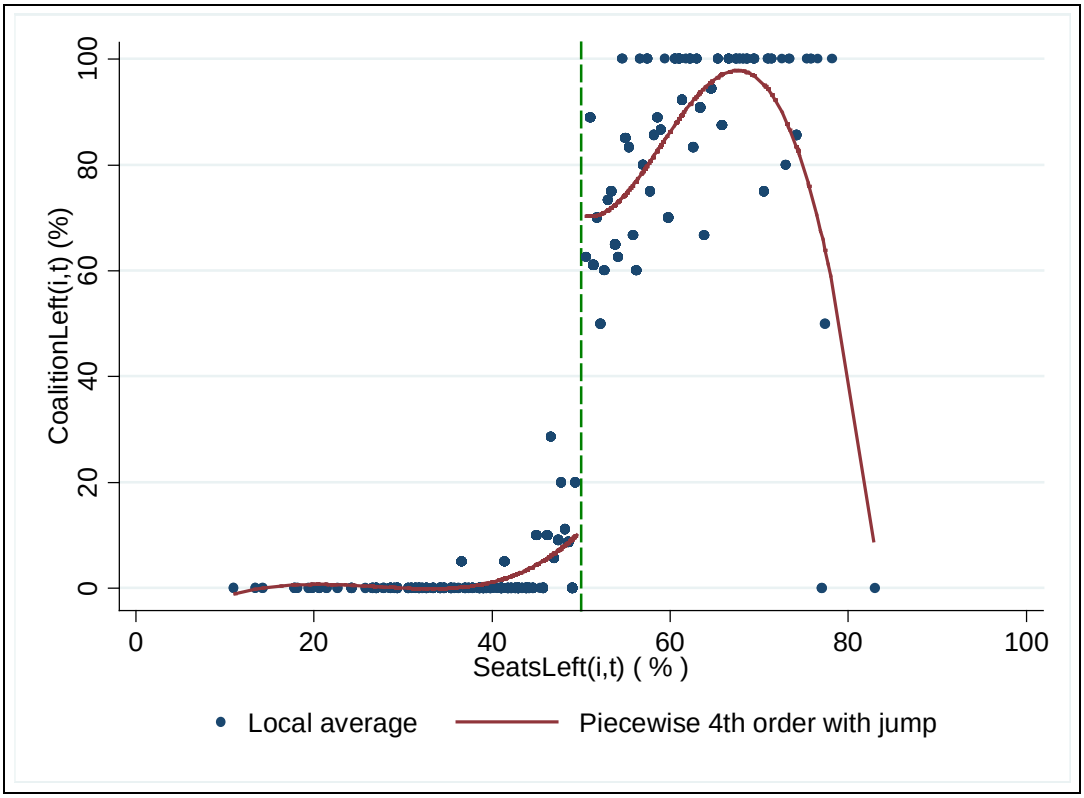
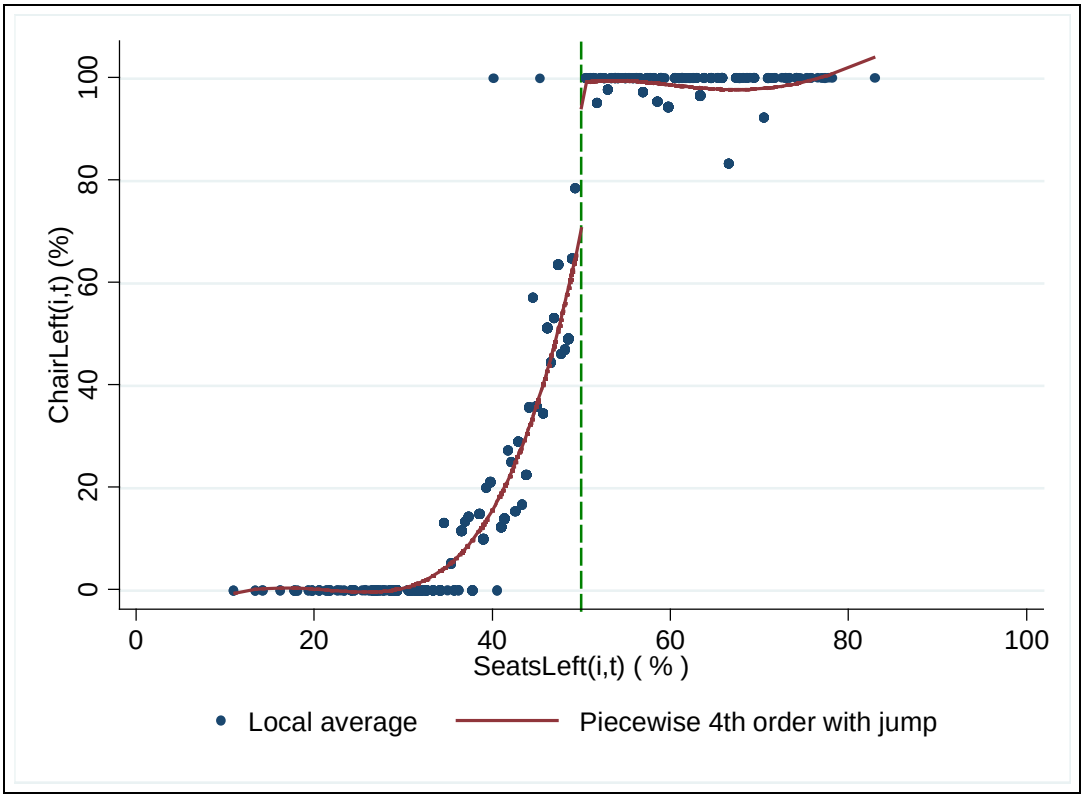


Figure 5. *Plot of local averages of chairing the executive board against representation*



6. Concluding Discussion

This paper estimates two effects of political representation on the electoral outcome at the local government level in Sweden, which has a proportional election system. I find an incumbency advantage of 0.12 percent of the total vote for each percent of seats, which implies that the present seat distribution determines the distribution of 12 percent of the total vote in the next election. However, I find no ruling effect.

The extra direct monetary and indirect resources, such as potential larger media exposure, that parties with larger incumbency receive therefore provide them with an electoral advantage in proportional election systems as well. This advantage is similar in size to the advantage in majoritarian systems. This is not surprising, since most theoretical arguments that predict an advantage are applicable to both majoritarian and proportional systems. One difference between the systems is that the focus lies on parties and coalitions in proportional systems and on legislators in majoritarian systems. The incentives for incumbents and rulers to promote their party's or their coalition's popularity in proportional systems may be fewer than the incentives to promote their own popularity in majoritarian systems, as there may be a free-

rider problem among incumbents from the same party or rulers from the same coalition. The results here do not indicate the existence of such a free-rider problem.[5]

The absence of a ruling effect rejects the asymmetric voter evaluation hypothesis as well as the median-gap model. Voters do not punish ruling coalitions more for bad events than reward them for good event, nor does the median-voter vote alternatingly. However, the absence of a cost of ruling is not incompatible with the back-swinging hypothesis. Performing better than usual in an election may be associated with performing less well in the following election as performance returns to its long-run level.

The pattern found here, namely, an incumbency advantage without a cost of ruling, can be understood in light of how the municipal public support for the parties is distributed; this monetary resource is directly linked to the incumbency of parties but not to the ruling of coalition of parties in the municipal councils in Sweden. If these locally-received funds are partly spent by the parties in the local campaign and this campaigning is effective, we would expect this to contribute to an incumbency advantage but not any cost of ruling. There are, however, many other reasons than monetary resources to expect incumbency and ruling to be either beneficial or harmful, as discussed previously. Investigating the different potential mechanisms at work in detail is an interesting question left for future research.

The phenomena investigated in this paper raise fundamental questions concerning political opportunism and the functioning of election systems in achieving representative governance. These broader and deeper topics are also left for future research.

[5] Even when the incumbency effect is investigated at the party level in majoritarian systems as in Lee (2008), an incumbent party's interest in a district usually does concur with that party's legislator's interest in that district as incumbent legislators usually campaign for re-election.

References

Akarca, A. T., Tansel, A., 2006, ”Economic Performance and Political Outcomes: An Analysis of the Turkish Parliamentary and Local Election Results Between 1950 and 2004”, *Public Choice* 129, 77–105.

Alesina, A., Roubini, N., and Cohen, G. D., 1997, *Political Cycles and the Macro Economy*, MIT Press.

Alford, J. R., Brady, D. W., 1993, ”Personal and Partisan Advantage in U.S. Congressional Elections”, Dodd, L. C., and Oppenheimer, B. I. (eds.), *Congress Reconsidered*, 5th Edition, CQ Press.

Anderson, S. P., and Glomm, G., 1992, “Incumbency Effects in Political Campaigns”, *Public Choice* 74, 207–219.

Ansolabehere, S., and Snyder, J., 2002, ”Using Term Limits to Estimate Incumbency Advantages When Officeholders Retire Strategically.”, *Legislative Studies Quarterly* 39, 487-515.

Bernhardt, M. D., and Ingberman, D. E., 1985, “Candidate Reputations and the Incumbency Effect”, *Journal of Public Economics* 27, 47–67.

Cox, G. W., and Katz, J. N., 1996, “Why Did the Incumbency Advantage Grow”, *American Journal of Political Science* 40, 478–497.

Duverger, M., 1954, *Political Parties*, Meuthuen.

Erikson, R. S., 1971, “The Advantage of Incumbency in Congressional Elections”, *Polity* 3, 395–405.

Ferreira, F., and Gyourko, J., 2009, “Do Political Parties Matter? Evidence from U.S. Cities.” *Quarterly Journal of Economics* 124, 399- 422.

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

Folke, O., 2007, Party Effects in Proportional Election Systems, Slide Presentation at the Institute for Labour Market Policy Evaluation.

Folke, O., 2008, “The Effect of Political Parties on the Placement of Refugee Immigrants in Swedish Municipalities”, *Politics and Preferences: Explanations to Policy Outcomes in Swedish Municipalities*, Licentiate Thesis, Swedish University of Agricultural Sciences.

Gelman, A., and King, G., 1990, “Estimating Incumbency Advantage without Bias”, *American Journal of Political Science* 34, 1142–64.

Goodhart, C. A. E., and Bhansali, R. J., 1970, "Political Economy", *Political Studies* 18, 43–106.

Imbens, G. W., and Lemieux, T., 2008, "Regression Discontinuity Designs: A Guide to Practice", *Journal of Econometrics* 142, 615–635.

Jacobson, G. C., 1997, *The Politics of Congressional Elections*, Longman.

Johansson, E., 2003, "Intergovernmental Grants as a Tactical Instrument: Empirical Evidence from Swedish Municipalities", *Journal of Public Economics* 87, 883–915.

Jordahl, H., 2006, "An Economic Analysis of Voting in Sweden", *Public Choice* 127, 251–273.

Kinder, D. R., and Kiewiet, D. R., 1979, "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgements in Congressional Voting", *American Journal of Political Science* 23, 495–527.

Kramer, G. H., 1971, "Short-Term Fluctuations in U.S. Voting Behavior", *American Political Science Review* 77, 131–143.

Laver, M., and Schofield, N., 1990, *Multiparty Government: The Politics of Coalition in Europe*, Oxford University Press.

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

Lee, D. S., 2008, "Randomized Experiments from Non-Random Selection in U.S. House Elections", *Journal of Econometrics* 142, 675–697.

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

Lijphart, A., 1994, *Electoral Systems and Party Systems*, Yale University Press.

Mueller, J. E., 1970, "Presidential Popularity from Truman to Johnson", *American Political Science Review* 64, 18–34.

Nannestad, P., and Paldam, M., 1994, "The VP-function: A Survey of the Literature on Vote and Popularity Functions after 25 Years", *Public Choice* 79, 213–245.

Nickell, S., 1981, "Biases in Dynamic Models with Fixed Effects", *Econometrica* 49, 1417–1426.

Paldam, M., 1986, "The Distribution of Election Results and the Two Explanations of the Cost of Ruling", *European Journal of Political Economy* 2, 5–24.

Paldam, M., and Skott, P., 1995, "A Rational Voter Explanation of the Cost of Ruling", *Public Choice* 83, 159–172.

Payne, J. L., 1980, “The Personal Electoral Advantage of House Incumbents”, *American Politics Quarterly* 8, 375–398.

Pettersson-Lidbom, P., 2008, “Do Parties Matter for Economic Outcomes? A Regression-Discontinuity Approach”, *Journal of the European Economic Association* 6, 1037–1056.

Rosenbaum, P. R., 1984, “The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment”, *Journal of the Royal Statistical Society Series A*, 147, 656–666.

Rubin, D. B., 2004, “Direct and Indirect Causal Effects via Potential Outcomes”, *Scandinavian Journal of Statistics*, 31, 161–170.

Svensk Författningssamling, 1991, *Svensk Författningssamling*, 1991:900, Chapter 2, Paragraph 9 and 10.

Svaleryd, H., and Vlachos, J., 2007, Political Rents in a Non-Corrupt Democracy, IFN Working Paper 698, Research Institute of Industrial Economics, Stockholm.

Tellier, G., 2005, “Explaining Subnational Government Popularity in Federal Systems: The Canadian Provincial Case”, Manuscript, University of Ottawa.