VATT Working Papers 50

Are Voters Rational?

Teemu Lyytikäinen
Janne Tukiainen

VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

50

Are Voters Rational?

Teemu Lyytikäinen
Janne Tukiainen

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Helsinki 2016

VATT Institute for Economic Research and London School of Economics, Spatial Economics Research Centre. Email: teemu.lyytikainen@vatt.fi. Tel: +358295519431.

VATT Institute for Economic Research and Helsinki Center of Economic Research. Email: janne.tukiainen@vatt.fi. Tel: +358295519451.

We thank Manuel Bagues, Yosef Bhatti, Matz Dahlberg, Jon Fiva, Ronny Freier, Stefano Gagliarducci, Essi Eerola, Rafael Hortala-Vallve, Ari Hyytinen, Mika Kortelainen, Valentino Larcinese, David Myatt, Benjamin Lauderdale, Giacomo Ponzetto, Tuukka Saarimaa and Marko Terviö for useful comments. Seminar participants at FEAAM, EPSA, HECER, Jyväskylä, RES and VATT contributed to very useful discussions. Niklas Jahnsson provided outstanding research assistance. The paper has been awarded the Essex Social Science Data Analysis Award 2013 "for the best paper providing an empirical test of a theoretical argument presented at the 2012 European Political Science Association conference in Berlin". Janne Tukiainen thanks the Emil Aaltonen Foundation, the Jenny and Antti Wihuri Foundation and the Yrjö Jahnsson Foundation for generous funding and STICERD at the London School of Economics and Political Science for hospitality.

This version (January 19, 2016) replaces the September 27, 2013 version.

ISBN 978-952-274-092-2 (PDF)

ISSN 1798-0291 (PDF)

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, January 2016

Are Voters Rational?

VATT Institute for Economic Research
VATT Working Papers 50/2016

Teemu Lyytikäinen – Janne Tukiainen

Abstract

We test whether voters are rational in the sense that their decision to cast a vote depends on its expected impact on the election outcomes. We provide causal evidence on this rational voting hypothesis by using exogenous variation in pivotal probabilities that arise at population thresholds determining council sizes in Finnish municipal elections. First, we document statistically significant, economically relevant and robust effects of crossing the threshold on turnout. Second, we use simulations to measure the pivotal probabilities. Finally, we use a novel instrumental variables design to show that the changes in the pivotal probabilities rather than simultaneous changes in the number of available candidates or their quality explain the changes in turnout. Moreover, we document that turnout responds only to such pivotal probabilities that are salient to the voters, and that the effect of district magnitude on the election closeness in general or the proportionality of elections is not behind the response in turnout. Thus, the rational voter exists.

Key words: Local government elections, Instrumental variables, Rational voting, Regression discontinuity design

JEL classes: C26, D72

1. Introduction

The expected instrumental benefits of voting are often thought to be close to zero, because the probability that a vote has an effect on the election outcome is typically close to zero (e.g. Grofman 1993 and Blais 2000). Moreover, the act of voting incurs non-negligible costs such as the opportunity cost of time. Therefore, if voters are rational in the sense that they only vote when the benefits of voting exceed the costs, the turnout rates should be very low in most elections, and particularly so in large elections.[1] Yet, we observe relatively high turnout rates across countries and various elections, and it is not clear that turnout responds to the closeness of elections (e.g., Grofman 1993). This voting paradox (Downs 1957) has long puzzled political scientists and economists and continues to do so.

Various explanations for the paradox have been offered both within and outside the rational voting paradigm. A simple way to explain the empirical patterns of high turnout is to introduce an additional utility component by assuming that voters derive utility from the act of voting itself, irrespective of the election outcome. This expressive component may include elements such as identity, duty or social pressure. Indeed, there is convincing empirical evidence that such expressive motives matter (e.g., Gerber et al. 2008 and Funk 2010).

The purpose of our study is to analyze whether the instrumental components of the rational voting model also matter for the decision to vote or abstain. Empirical analysis of the relationship between turnout and how close the elections are is complicated due to standard endogeneity issues. In particular, reverse causality is an inherent issue in such studies, because pivotal probability depends on turnout. Moreover, numerous unobservable candidate,

[1] This hypothesis is, however, somewhat debated. Mulligan and Hunter (2003) show that the simple empirical frequency of pivotal vote is one out of 89,000 in U.S. Congressional elections. Gelman et al. (1998) discuss estimating these rare events and find that the pivotal probability is at best one out of 1 million for close U.S. national elections. Even though these are small numbers, the expected incremental benefit is not necessarily small if voters have social preferences (e.g. Edlin et al. 2007). Moreover, once aggregate uncertainty is introduced, also the pivotal probabilities are larger than usually thought (Myatt 2012).

voter and constituency characteristics may influence both turnout and how close the election is. Furthermore, it is not always easy to measure the pivotal probabilities, especially in proportional elections. In this paper, we solve both the issues of measurement and causal inference.

To overcome these challenges for causal inference, we utilize a natural experiment to analyze if turnout increases when the probability of one vote making a difference increases exogenously. In our quasi-experiment, we utilize population discontinuities in Finnish local government elections to construct a regression discontinuity design (RDD). In Finland, municipal council size is an increasing step function of population. This step function is determined by law. Therefore, we compare otherwise similar municipalities with different council sizes.

In a proportional open-list election system, pivotal events are relatively common. They can occur both between parties, i.e., situations where one abstaining voter could change the allocation of seats between parties by voting, and within parties, i.e., situations where one voter could change which of the party's candidates gets the last seat of the party. In this setting, an increase in the amount of seats leads to an increase in the probability of a vote being pivotal, ceteris paribus. The effect arises from various sources, but most importantly simply from the elections allocating the last seats with a smaller amount of votes and thus smaller margins both between and within parties. According to our bootstrap election simulations, crossing the threshold, where council size increases, increases the probability of pivotal occurrence between parties for the average party on average by about 18% (from 1.7 times out of a hundred to 2.0, see Table 2) and the probability of pivotal occurrence for the last seat within parties for the average candidate increases on average by about 17% (from 2.4 times out of a hundred to 2.8).

In the absence of precise manipulation of the municipal population, the treatment of increasing the council size is almost as good as random near the population thresholds. Therefore, this setup provides a valid causal test for the effect of council size on turnout. We find that turnout is higher just above the council size thresholds than just below. This result is robust to standard validity and robustness checks involved with RDD. One particular placebo test is especially convincing – here we use the same RDD to explain municipal level turnout in the national parliamentary elections where these thresholds play no role. We find that crossing the same thresholds has no effect on turnout there.

The main caveat of our natural experiment design in making conclusions about rational voting from the council size effect is that pivotal probabilities are not the only things that change at the thresholds. On the contrary, the relationship between turnout and district magnitude can arise from many sources. From the voter perspective, council size affects also proportionality (i.e. the minimum vote share required for getting elected) and hence potentially political efficacy, because voters may perceive that less votes are wasted (see e.g. Gallego et al. 2012, Karp and Banducci 2008). Also parties may respond to changes in political competition (i.e. closeness of elections in a more general sense that just one vote making a difference) and hence induce elite mobilization (Powell 1986), which may be reflected in campaigning and candidate placement. Indeed we observe that the number of candidates and their quality change at the discontinuities.[2] Besides a further impact of more and better candidates on the pivotal probability (vote distribution may be flatter), this may

[2] In Finland, the maximum allowed size of each party list is a deterministic function of a council's size (1.5 times the council size). Moreover, parties and potential candidates may respond to an increase in council size by presenting a longer list, even when the limit is not binding. Indeed, when our RDD is used on candidates, we observe that the number and type of candidates is different above the threshold than below it. Fortunately, this is the only confounding policy that uses any of our thresholds. For example, municipalities' responsibilities do not depend on any population thresholds. We discuss potential confounding responses by candidates and parties in Section 4 along with our empirical strategy.

also influence turnout due to some expressive components.[3] Since RDD identifies only the joint effect of all the events that take place at each threshold, we cannot be sure whether the voters react to the pivotal probabilities or to e.g. the number of candidates. Fortunately, we are able to analyze this question of real causal channel in detail using a novel design.

Our test of rational voting involves three steps. First, we document the overall effect of council size on turnout. Second, we simulate the pivotal probabilities for each municipality-party both within and between parties. There are many attempts in the literature that test for rational voting by correlating some measures of expected closeness of elections to turnout (see e.g. Geys 2006 for a survey). One fundamental issue with such studies is that the real pivotal probability is highly nonlinear in relation to typical closeness measures such as margin of victory and the relationship is context dependent, e.g. 1% vote share victory margin may be either very far or very near to one vote being decisive (Cox 1988). We solve these measurement issues by bootstrapping pivotal occurrences.

Finally, we use our measure of expected pivotal probability simultaneously with the other endogenous variables in an instrumental variables regression (IV), where the different thresholds are the instruments. We find that the number or quality of candidates has consistently no significant effect on turnout, but that the pivotal probability has a positive and significant effect. Moreover, we find that turnout responds only to the more salient within-party pivotality but not at all to the - harder to calculate - between-party pivotality. These results also largely rule out proportionality and elite mobilization as explaining the turnout response, because they are mainly between party concerns and also relate to political competition in general rather than to the influence of a single voter. Nonetheless, we also

[3] For example, in the spirit of spatial models of voting (e.g., Downs 1957), voters may find a better match for their preferences. Moreover, more candidates may mean more available information, which may decrease the costs of voting (e.g., Matsusaka 1995). More candidates could also imply more campaigning.

subject our measures of pivotality, general closeness of elections and proportionality to a horse race and show that only pivotality affects turnout.

This approach of using multiple thresholds to uncover and quantify the real causal channel at work at the thresholds in RDD is an interesting applied methodological contribution. Previous approaches to disentangle multiple treatments from each other in RDD with multiple thresholds have included focusing only on such ‘clean’ thresholds which affect only one of the treatments (see e.g. Ferraz and Finan 2009) or using variation in treatment timing by combining RDD with a difference-in-differences analysis (so called difference-in-discontinuity approach as first coined by Grembi et al. 2012).[4] We provide a new alternative that requires, first, in the case of, for example, two simultaneous treatments A and B, there to be some threshold(s) where treatment A has a large intensity (but treatment B only small) and some other threshold(s) were treatment B has a large intensity (but treatment A only small), and second, that there are at least as many thresholds as endogenous variables (simultaneous treatments). Therefore, we contribute to the emerging econometrics literature on the analysis of multiple cutoffs RDD (Eggers et al. 2015b, Cattaneo et al. 2015a).

Understanding whether voting is rational, in the sense that voters take into account the probability of their vote making a difference in elections, is important for various reasons. First, it is important to understand whether a central activity in a democratic society is outside the realm of rational behavior in this strict sense. Second, understanding the correct underlying model of voting is essential for implementing policies aimed at affecting turnout and thus political representation. Third, there are some views (e.g., Fiorina 1989, Green and Shapiro 1994 and Aldrich 1997) that turnout or the paradox of voting is proof against rational

[4] See also a recent survey and critical review of studies using multiple population thresholds in RDD by Eggers et al. (2015b) whose concerns we carefully address.

choice theory in general and thus challenges many models in social sciences. While there are convincing theoretical counterarguments (e.g., Myatt 2012), causal empirical evidence of rational voting has largely been missing.

2. Theoretical framework and prior empirical evidence

The classic rational voting model (e.g., Downs 1957 and Riker and Ordenshook 1968) can easily be derived from a choice tree, where the decision is between voting and not voting and the two outcomes under both alternatives are the preferred candidate (or party) getting elected or not. In this discrete choice model, standard utility comparison calculus leads to the following model:

$$(1) \qquad Y = 1(G + pB - C > 0),$$

where $Y = 1$ if voter votes and zero otherwise, 1 denotes an indicator function, G denotes the outcome independent utility from voting. This is the expressive component, which was not present in the original model by Downs (1957). C is the cost of voting. B is the benefit of preferred election outcome and $p=p_1-p_2$, where p_1 is the probability that the preferred candidate will get elected if the voter votes, and p_2 is the probability that the preferred candidate will get elected if the voter does not vote. In this model, p vanishes as the set of voters gets larger. Therefore, with positive C, only large G can explain large turnout in large elections.

Obviously, the existence or importance of G in the voting decision does not imply irrationality as such. The pure expressive (or irrational) utility model assumes that p does not matter even if it is not very small. Therefore, the p component is at the heart of the rationality in these models. At its simplest, the expressive utility voting model can be presented as

(2) $Y = 1(G > C)$.

This would imply that a statistical test based on the *pB* component would be able to separate between the rational and expressive voting hypothesis. On the contrary, showing that voters respond to the costs of voting does not allow differentiating between rational and expressive models.

However, the testing also has to account for the possible strategic behavior of voters. Since *p* depends on how many of the other potential voters actually vote, the decision to vote is strategic. Even if $G = 0$, one would expect some amount of turnout due to *p* getting larger the fewer of the other potential voters turn out (see e.g. Palfrey and Rosenthal 1983 and 1985). If voters differ in *G* or *C*, pure strategy equilibria exist that can be presented by threshold values in (G - C). In that scenario, only voters with high (G - C) participate.[5] Therefore, game theory models may result in very similar empirical predictions to the expressive models where all hinges on *G*. This makes testing for rationality difficult: High or low turnout alone is not sufficient to infer the correct voting model.

There are plenty of alternative explanations for voting behavior in the literature and we do not cover them all. Explanations include a behavioral explanation of voters inflating their individual *p*'s (Riker and Ordenshook 1968), and group voting where individuals are ethically obliged to a group and groups coordinate (first mentioned by Harsanyi 1955, empirical evidence by Coate and Conlin 2002). In addition, uncertain voter models (e.g., Matsusaka 1995 and Feddersen and Pesendorfer 1996, 1999), where the costs of voting have

[5] If all voters are identical, they play mixed strategy equilibrium that leads to some positive level of turnout that is decreasing in population. However, empirical regularities imply that population characteristics matter to turnout and that to a large extent the same people vote in sequential elections (e.g., Matsusaka and Palda 1999). This implies that mixed strategies are not typically used.

an endogenous relationship with the information available to voters on candidates, have been found to be empirically relevant in explaining turnout (Degan and Merlo 2011).

We use a natural experiment to estimate the causal effect of p on Y at an aggregate level. Thus, this analysis does not explicitly allow us to distinguish between alternative theories where pivotal probabilities play a role such as individually rational model or group voting model. However, we do show evidence that rational behavior with respect to the likelihood of changing the election outcome is a significant driver of the decision to vote, thus ruling out pure expressive models where none of the voters respond to the closeness of elections.

Some attempts to uncover the relation between Y and p have been made previously in both political science and economics. The evidence is mixed. For example, Geys (2006) reports in a survey article that 69% of the reviewed articles support the rational voting hypothesis, i.e., turnout responds to closeness of the elections. Foster (1984) also surveys many of these attempts and conducts such an analysis herself. Those results are also mixed. Unfortunately, all of these results are partial correlations and reveal no causal results. Among many others, Indridason (2008) deals with some endogeneity issues by focusing on run-off elections, but not to the extent of mimicking randomization. The main limitation of the run-off elections approach is that the first round election outcome is not exogenous either.

Coate et al. (2008) calibrate the parameters of a rational voting model using real data from small elections and then simulate the election outcomes. They find that although the patterns in the real data match the patterns in their simulation with respect to how turnout responds to the size of the elections, the margins of victory do not. This leads them to reject the rational voting model.

Natural experiments have been used previously to analyze other aspects of voting behavior. Andersen et al. (2014) show that by using a natural experiment in simultaneous elections for different offices, turnout responds to larger stakes (increasing *B*). While this can be seen as an indirect test of voters also responding to *p*, it is also possible that *B* is also partly determining *G*. Funk (2010) studies implementation of postal voting in Switzerland in a difference-in-differences estimation framework. She finds that the causal effect of removing social pressure from voting decisions on turnout (decreasing *G*) was larger than the effect of reduction in voting costs (decrease in *C*), especially in small communities. Fujiwara (2011), Lago (2012), Kawai and Watanabe (2013) and Saarimaa and Tukiainen (2016) provide causal evidence that some voters are strategic, i.e. given that voter participates, the decision who to vote for accounts for the pivotal probability concerning electing different candidates.

3. Institutional setting

In Finnish local government elections, voters elect the municipal councils. These elections are very important for several reasons. First, municipalities are responsible for the majority of public services in Finland, including health care and primary education. The GDP share of municipality spending is roughly 18 percent and they employ around 20 percent of the total workforce, whereas the central government share is only roughly five percent. Second, the municipalities are not heavily regulated in how they collect their income and they can also independently decide how to provide many of the public services. For example, they can freely set the income tax rate, which is the largest revenue source. Third, municipal councils are the main seat of power in municipal decision-making. Compared with many other countries, mayors or city managers do not wield much power in Finland, nor are they elected but are rather civil servants chosen by the council. Fourth, in most municipalities the probability of a vote being pivotal is fairly high, due to a relatively high number of council seats per capita and other details of the election rule, such as the open-list property and a multi-party system. Therefore, the pivotal occurrence rate is substantial and quite high on average, even in larger municipalities. Thus, we are able to observe significant responses in pivotality to the exogenous variation in council size.

Another important feature of the Finnish system is the large number of municipalities relative to the population, with a large variation in municipal population size. The largest municipality is the country's capital, which has over half a million inhabitants, whereas the smallest mainland municipality has roughly 800 inhabitants. The median municipal population is less than 6,000. Finland is also sparsely populated and population density varies substantially across municipalities, as do other population characteristics such as age structure and income. Thus, we have significant variation in our forcing variable (population) and the political preferences of the population. This influences the distribution of votes and thus we have a large variation in pivotal occurrences and turnout, even within municipalities of the same size.

Finland has a multi-party system. Currently, there are eight parties in the Finnish parliament and these parties also dominate municipal politics, but some local single-issue groups exist as well. For example, in the 2004 municipal elections the three largest parties (the Social Democrats, the Centre Party and the National Coalition) received around 68 percent of the votes with roughly similar overall shares but with large variation in shares between municipalities.

The municipal council is responsible for strategic and financial outlines and the main objectives of municipal activities. The council also chooses a municipal board which has a preparatory role. The composition of the board is based on party shares in the council, i.e., each party in the council gets seats in the municipal board according to its share of council seats. Thus, there is no real opposition in local politics. This is important, because the majority of votes for some party is not the only relevant dimension where a single vote could have a meaningful impact. For the implemented policies, every seat potentially counts.

The elections in our data were held in the October of 1996, 2000, 2004, 2008 and 2012. The Election Day was the same for all municipalities. No other issues are voted upon at the same time. The councils are elected using an open-list system. Each voter gives a single vote to a single candidate. Voters cannot vote only for a party without selecting a candidate. Individual votes rank the candidates within the parties. Thus, there are potentially pivotal events within parties. The total number of votes for all of the given party's candidates determines the number of votes for the party. The seats are allocated between parties based on the D'Hondt comparison method.[6] Therefore, there are also potentially pivotal events between parties. Each municipality has only one electoral district.

Council size is a step function of the municipality's population and it is determined by law.[7] Table 1 shows the amount of municipalities in different population groups with different council sizes. In Table 1, we show this distribution only up to a population of 45,000, because after that the data are not dense enough for RDD and are thus omitted from our analysis. As can be seen from Table 1, council size in our data varies from 13 for some municipalities with a population of 2,000 or less, up to 51 for municipalities with a population of between 30,000 and 45,000. This concave step function implies that there is large variation in the amount of council seats per capita.

---

[6] The candidates of each party in the municipality are ordered according to their votes and they are given comparison numbers calculated by dividing total votes for the party in the municipality first by 1, then by 2, then 3 etc. Council seats are allocated to the candidates with the highest comparison numbers.

[7] The council sizes for the different population groups are: population less than or equal to 2,000 (council size 13, 15 or 17), 2,001–4,000 (21), 4,001–8,000 (27), 8,001–15,000 (35), 15,001–30,000 (43), 30,001–60,000 (51), 60,001–120,000 (59), 120,001–250,000 (67), 250,001–400,000 (75) and over 400,000 (85).

Table 1. Population groups, council size and number of elections in data (population < 45,001).

Population	Council seats	N
<2000	17 (or 15 or 13)	274
2001-4000	21	465
4001-8000	27	478
8001-15000	35	307
15001-30000	43	168
30001-45000	51	55

4. Empirical strategy

The purpose of this section is to present our empirical approach. The analysis consists of three steps. First, we estimate the impact of council size on turnout in local elections using regression discontinuity design. The second step is to measure the pivotal probability, i.e., the probability that one vote changes the outcome of the election. The third step is the analysis of mechanisms behind the overall effect of council size on turnout, in particular, showing that change in the pivotal probability at the thresholds is behind the effects on turnout rather than alternative responses at the thresholds. To achieve this we outline an instrumental variables approach where the first stages consist of regression discontinuity designs and identification is based on the fact that different thresholds have different effects on the various endogenous variables. Therefore, we argue that our empirical analysis provides a test of whether the likelihood of a pivotal vote affects turnout in local elections.

4.1. Step 1: The effect of council size on turnout

The basic idea of our empirical strategy is to compare turnout in municipalities below and above the cut-off points. The main identifying assumption in such RDD is that individuals cannot precisely manipulate the forcing variable (see e.g. Lee and Lemieux 2010). This should be true in our case, because municipalities do not self-report their population. In this case, identification is based on a local randomization at the threshold.[8] Ideally, we would have many data points close to the threshold and then estimate the effect at the threshold. In our case, due to lack of observations and having multiple thresholds, we need to use specifications that use data points further away from the thresholds. More specifically, we estimate both parametric polynomial regressions and nonparametric local linear regressions. We prefer the parametric ones because they lend themselves more easily to the analysis with multiple thresholds and the instrumental variable approach, but show that in our particular case, the parametric specification produces similar results as the more flexible nonparametric approach.[9]

We estimate a regression model of turnout on a set of zero-one indicators for being above a cut-off point and include a flexible but smooth function of population as control variables. The population variables pick up the impact of all determinants of turnout correlated with population, apart from council size. Hence, we will obtain a reliable estimate of the effect of council size on turnout clean of confounding factors that might otherwise bias our estimates.

A further complication to our analysis is how to deal with multiple thresholds. One option would be to calculate the forcing variable as a population distance to the nearest threshold and simply define a single group for being above a threshold. We do not use the pooling option for two reasons. First, unless the share of identifying observations around each threshold would be the same (which would happen in large samples due to local randomization), we could be comparing, for example, a municipality of a population of 1,999 (just below) to a municipality of 30,001 population (just above). This is clearly not a valid

[8] Local randomization is not a requirement but rather one possible interpretation of RDD. What is required is that other covariates develop smoothly over the threshold. One difference between these two perceptions is that the latter allows there to be trends in the other covariates. See Cattaneo et al. (2015b) for further discussion.

[9] In general, it is not advised to use parametric specifications, see e.g. Imbens and Gelman (2014) for the many issues with the parametric approach.

comparison. Therefore, given a relatively small sample, this sort of normalization and pooling is not recommended. Second, we want to allow for different effects at different thresholds. This will be important in our instrumental variables method, where we have more than one endogenous variable. Therefore, we will estimate a model with the actual population as the forcing variable and then allow for a different effect at different thresholds. We will calculate the overall average effect as the weighted (by the number of observations around each threshold) average of these separate effects.

Our main specification is a parametric model where we simply fit a high order polynomial of population over the whole range of population and include dummy variables for groups defined based on council size. We estimate by OLS the following equation

(3) $Turnout_{it} = \beta_1 + \beta_2 Group2_{it} + \ldots + \beta_6 Group6_{it} + f(Pop_{it}) + u_{it}.$

The dependent variable is the turnout in a municipality i in an election year t. Function f is a polynomial of population. We use 1st – 7th order polynomials. The explanatory variables of interest are overlapping dummies *Group2*,…,*Group6,* indicating all municipalities above a certain threshold. For example, Group 2 includes all the municipalities with a population of more than 2,000.[10] Our estimating sample contains data from the first six groups, because we limit the analysis to municipalities with a population of less than 45,000 to keep the data somewhat dense. The respective group coefficients $\beta_2,\ldots,\beta_6$ give direct estimates of the effect on turnout of increasing council size by one step. The group dummies can be interpreted as individual treatment variables, with the previous group as the control group. Therefore, this specification allows for a different effect at each threshold.

[10] Non-overlapping dummies would produce the same fit of the model, but we prefer the overlapping dummies, because the related coefficients have directly our desired interpretation.

One drawback of Model 1 is that it uses data far from the cut-offs to estimate the value of the polynomial at the cut-off. Therefore, we introduce slightly more flexible models as a robustness checks. Our second approach relaxes the specification by allowing for different polynomials between different cut-offs. Our Model 2 is written as

$$(4) \quad Turnout_{it} = \beta_1 + \beta_2 Group2_{it} + \ \ldots + \beta_6 Group6_{it} + f_1(Pop_{it})$$

$$+Group2_{it} * f_2(Pop_{it} - 2000) + \cdots + Group6_{it} * f_6(Pop_{it} - 30000) + u_{it}.$$

Here the function f is a 1st to 3rd order polynomial of population and is allowed to vary between the cut-off groups. Note that already a linear specification maps turnout quite flexibly to the population. Normalizing population to zero at the cutoff when estimating functions $f_2,\ldots, f_6$ implies that, like in Model 1, the coefficients $\beta_{2,}\ldots, \beta_6$ on the dummies for groups above the cut-offs give direct treatment effect estimates of interest.

As a third approach, we estimate nonparametric local linear regressions using triangular kernel. We report these results for various bandwidths, including the Imbens and Kalyanaraman (2012) optimal ones. We conduct these estimations separately for the each threshold and verify that the parametric approaches give very similar results at each threshold. Therefore, we feel confident in using the inflexible parametric approach (3) as the main specification.

The identifying assumption of our models is that other determinants of turnout develop smoothly with respect to population and are therefore captured by the f function. Factors outside the model depending on the same population rule would violate this assumption. Ade and Freier (2011) and Eggers et al. (2015b) have raised this concern especially related to the case of analyzing population thresholds, since in many countries, municipalities' responsibilities, grants, politicians' salaries and regulation depend also on the same thresholds. In that case, there are simultaneous exogenous treatments and RDD is able to only identify their joint effect. None of these concerns is present in the Finnish system. However, in addition to the number of seats, the maximum number of candidates allowed for each party also changes at the cut-offs. More candidates could affect turnout if people find it easier (or harder) to find a suitable candidate from a larger pool. Moreover, also candidate quality can respond to a change in council size.[11] Finally, district magnitude also affects the closeness of elections more generally as well as the proportionality of the elections. In order to isolate the effect of seats from the other effects, we conduct an instrumental variables (IV) analysis that we describe in detail in Step 3.[12] Before we can proceed to IV, we need to be able to measure the pivotal probabilities.

4.2. Step 2:Measuring pivotal probabilities

We calculate the pivotal events both between parties, i.e., situations where one abstaining voter could change the allocation of seats between parties by voting, and within parties, i.e., situations where one voter could change which of the party's candidates gets the last seat of the party. A pivotal event takes place both in the case of actual draws (in which case a lottery would determine who gets elected) and also when giving one more vote creates a draw or a different seat allocation directly.

Kotakorpi et al. (2013) develop the bootstrap sampling method that we use as a basis for our simulation. They use this method to define close elections (lucky and unlucky candidates) at an individual level in the proportional election setting to study the causal effect of being elected on returns to office. We have an entirely different purpose. The basic idea of the

[11] See e.g., Cox (1997) or Fiva and Folke (2015) for analysis of strategic entry of candidates.

[12] Simply controlling for all the other things changing at the cutoffs is not satisfactory, because they are alternative response variables, and therefore 'bad controls'.

simulation is that the sampling procedure allows us to mimic the randomness involved in voters' decisions on whether to vote at all and who to vote for. For example, changes in the weather conditions may make a difference on which voters give their vote, and some voters may be sick one day and others another day, etc.

We sample votes with replacement for each candidate, so that the sampling probability of a vote for a particular candidate is the share of all votes that he or she received in the real elections. We sample as many votes as were originally given in the real elections. Each sampling produces a new distribution of votes among the candidates, while on average maintaining the original distribution of votes. After each sample, we take note of a pivotal event after each repetition. We repeat this exercise 1,000 times and count the share of times a pivotal event occurred for each party-election (between parties) and each candidate-election (within parties) observation. Thus, this is a bootstrap procedure for a particular non-standard statistic of the vote distribution, the pivotal vote probability. The simulation produces almost continuous variables between zero and one that are good measures of the probability of a pivotal event between and within parties.

We show in the Online Appendix A Tables A1 and A2 using counterfactual simulations that our measure works as intended and jumps at the cutoffs in the expected way, that is, it jumps much at the smaller thresholds and less at the larger. Besides showing that crossing the threshold has a significant impact on the pivotal probability, our simulations are essential in providing a metric that can be used to analyze whether the effect of council size on turnout is due to increases in pivotal probability at the thresholds or due to other changes at the threshold. To get the metric that we use in the IV, we conduct a counterfactual simulation, where we construct measures of pivotal probabilities that do not include the effects of crossing the threshold on turnout. This counterfactual is needed to correctly measure the effect of crossing the threshold on expected pivotality at the first stage of IV. In other words, the counterfactual corrects for the bias arising from simultaneity between turnout and pivotality. We accomplish this by manipulating the turnout. We use equation (3) and regress turnout on 1-7th order polynomials of population, election year dummies and the cut-off group dummies. Then we subtract from the resulting fit the effects of the group dummies on the fit. We use this adjusted fit as the counterfactual turnout that determines the number of votes given in our counterfactual simulations.

We stress that our measure of pivotality is not a measure of closeness of elections in any more general sense. We measure the incidence of cases were a single vote would make a difference. In contrast, elections where, for example, 3 votes would make a difference to the seat allocation could be seen a close, but would not count as pivotal in our metric. We can also construct a simple metric for the closeness of elections, or the victory margin. In the within party dimension were votes rank the candidates, this is simply the distance in personal vote share to the within party threshold of getting elected. For this metric we use the same threshold as Kotakorpi et al. (2013) defined as the mean votes of the elected candidate with the least votes and the non-elected candidate with the most votes. In our IV estimation we can test whether the response in turnout is explained by this within party margin of victory or by the within party pivotality. Between party margins of victory are harder to measure in PR systems (both to voters, parties, candidates and researchers), but it turns out that we do need study that confounder, because only the within party pivotality dimension matters for voters.

4.3. Step 3:Identifying the causal mechanisms

In our IV regression, we estimate the main RDD equation (3) in the first stage. The first endogenous variable (the first stage outcome) is the simulated counterfactual pivotal probability. As the second first stage outcome we use the number of candidates and their quality is the third. We assume that incumbency status is a sufficient metric for candidate quality. This assumption is based on Eggers et al. (2015a) who show that incumbency status captures all other observable candidate quality measures relevant for voters. The fourth variable is the proportionality measured as the minimum vote share among the elected candidates in the municipality. Other observables are shown not to jump at the threshold and therefore not considered endogenous (see Table B3, discussed later).

We conduct this estimation both at the election and election-party level. Therefore, we want to measure also pivotal probability at the election (or election-party) level. In order to achieve this aggregation, we take a weighted (by respective vote shares) average over both the within party pivotal probability (measured at the candidate level) and the between parties pivotal probability (measured at the party level) to calculate the probability that a randomly drawn voter would be pivotal in a given election. We use either a single measure for pivotality, which is simply the sum of these weighted between and within pivotalities, or both of them separately. These measures along with the other endogenous variables are described in Table 2 at both municipal and party level. For the pivotal probabilities we report the counterfactuals based on a 6^{th} order polynomial specification similar to Model 1. Differences to descriptions in Table A1 arise mainly from the weighting.

Table 2. Describing endogenous variables in IV.

Variable	N	Mean	Std. Dev.	Min	Max
Panel A: Election level					
Pivotality	1746	0.063	0.031	0.006	0.193
Number of candidates	1746	84	48	14	305
Share of incumbents	1746	0.26	0.07	0.12	0.71
Proportionality	1746	0.014	0.007	0.002	0.052
Panel B: Party-election level					
Pivotality, within	10171	0.006	0.009	0.000	0.164
Pivotality, between	10171	0.020	0.020	0.000	0.130
Number of candidates	10171	14	13	1	76
Share of incumbents	10171	0.21	0.17	0	1
Proportionality	8656	0.018	0.011	0.002	0.158

In the second stage, we use the fits of the first stages to explain turnout. In other words, we perform an IV regression of turnout on simulated pivotal probability, the number and type of candidates and proportionality using the threshold dummies as the excluded instruments. The estimation equations are written as

(5)

1st stage: $Pivotality_{it} = \beta_1 + \beta_2 \; Group2_{it} + \; ... + \beta_3 \; Group6_{it} \; + f(Pop_{it}) + u_{it}.$

1st stage: $Nro\ of\ candidates_{it} = \delta_1 + \delta_2 Group2_{it} + \; ... + \delta_6 Group6_{it} \; + f(Pop_{it}) + v_{it}$

1st stage: $Quality\ of\ candidates_{it}$

$$= \theta_1 + \theta_2 Group2_{it} + \; ... + \theta_6 Group6_{it} \; + f(Pop_{it}) + r_{it}$$

1st stage: $Proportionality_{it} = \kappa_1 + \kappa_2 Group2_{it} + \; ... + \kappa_6 Group6_{it} \; + f(Pop_{it}) + k_{it}$

2nd stage: $Turnout_{it}$

$$= \gamma_1 + \gamma_2 \widehat{Quality\ of\ candidates}_{it} + \gamma_3 \widehat{Nro\ of\ candidates}_{it}$$

$$+ \gamma_4 \widehat{Pivotality}_{it} + \gamma_5 \widehat{Proportionality} + f(Pop_{it}) + e_{it}$$

Despite substantial effort, we have not been able to find any policies, other than the number of seats and candidates, using the same thresholds in Finland. Neither did Pettersson-Lidbom (2012), who uses the same Finnish election thresholds as we do. However, he finds that government spending decreases after crossing the threshold. This behavioral effect on public spending could potentially be another confounding response to council size policy, but if anything this effect should make the elections less important because of smaller stakes, thus reducing the turnout rather than increasing it (e.g., Andersen et al. 2014). Moreover, using data from later years than him and different RDD specifications, we do not find any effect of council size on public spending (see Table B3, discussed later), and thus, we do not include public spending in our IV model.

Another factor that could make observing an effect at the threshold more difficult is that a larger council may decrease turnout because power per councilor may be lower, thus the benefit of getting one's own candidate elected may be lower. The identification and interpretation of the treatment effect would be further complicated if additional endogenous choices on other institutions were taken simultaneously, such as voting in multiple elections or ballots at the same time (see Ade and Freier 2011 and Eggers et al. 2015b). In our data, a vote is given only for the municipal council election and no other decisions are made simultaneously.

5. Data and descriptive analysis

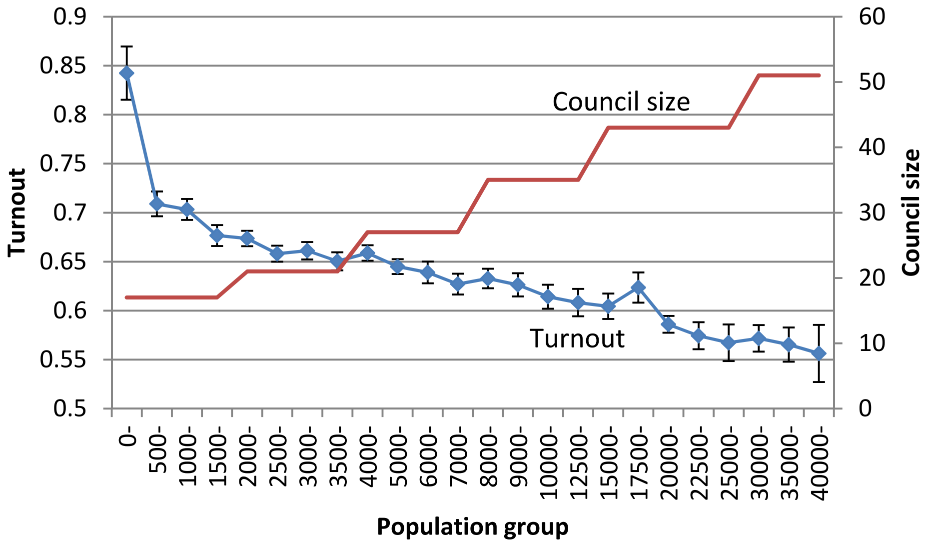
We have received the main data from Statistics Finland. The election data are publicly available from their website, but some control variables required access to proprietary databases. We limit the sample used in the analysis to five election years and to municipalities with a population below 45,000. This leaves us with 1,747 municipality-election year (i.e., election) observations. Besides the endogenous variables of interest described in Table 2, the other key variables for our analysis are council size, population and turnout. Table 5 reports summary statistics for these variables and other municipal characteristics that we will use for validity tests.

Table 3. Summary statistics for outcome and control variables (population < 45,000).

Variable	N	Mean	Std. Dev.	Min	Max
Turnout	1747	0.646	0.060	0.420	0.895
Population	1747	7506	7592	234	44804
Council size	1747	27	9	13	51
Political competition	1743	0.00033	0.00058	0	0.00831
Number of parties	1747	5.8	1.7	1	13
Tax revenue €1000/capita	1747	2.3	0.5	1.4	6.2
Municipality personnel/1000 people	1746	59	16	4	134
Unemployment rate %	1736	13.2	5.2	2.2	33.9
Grants €1000/capita	1747	1.5	0.8	-0.1	5.1
Share of 65+ year old	1747	0.196	0.049	0.049	0.386
Expenditure €1000/capita	1747	5.5	1.1	3.1	12.0

Graph 1 shows council size and turnout in population groups. The labels on the horizontal axis are the lower bounds of the group (note that the scale increases along the axis). Turnout clearly decreases with population but there appears to be an upward shift at the cut-off points where council size increases, suggesting that turnout is positively affected by the council size.

Graph 1. Council size and average voter turnout by population group.



6. Estimation results

6.1. The effect of council size on turnout

Table 4 reports the individual coefficients of the group dummies of Model 1 (individual treatment effects), as well as the weighted average treatment effect. Columns 1 – 7 use 1st – 7th order polynomials of population. The group dummies indicate that turnout increases with council size at every threshold for most specifications, but the individual effects are typically not significant. However, the average effect of a one-step increase in council size is positive across the board and statistically significant and quite robust in the specifications with a third or higher order polynomial of population.

The estimates e.g. in the 6th column implies that crossing a council size threshold increases turnout by 1.5 percentage points on average. It is not very easy to evaluate whether this effect is small or large. It seems to be relatively small compared to the cross sectional variation across elections of similar size (see Graph 2). On the other hand, relatively large changes in population size would be required to achieve 1.5 percentage point change in turnout. For example, based on the model fit of the 6th order polynomial model in Table 4, in our data, a municipality with population size 7528 has a predicted turnout of 63 % and a municipality with 5535 inhabitants has a fit of 64.5%. Thus, the population needs to decrease by about 2000 (27% decrease in population) to achieve 1.5 percentage points increase in turnout.

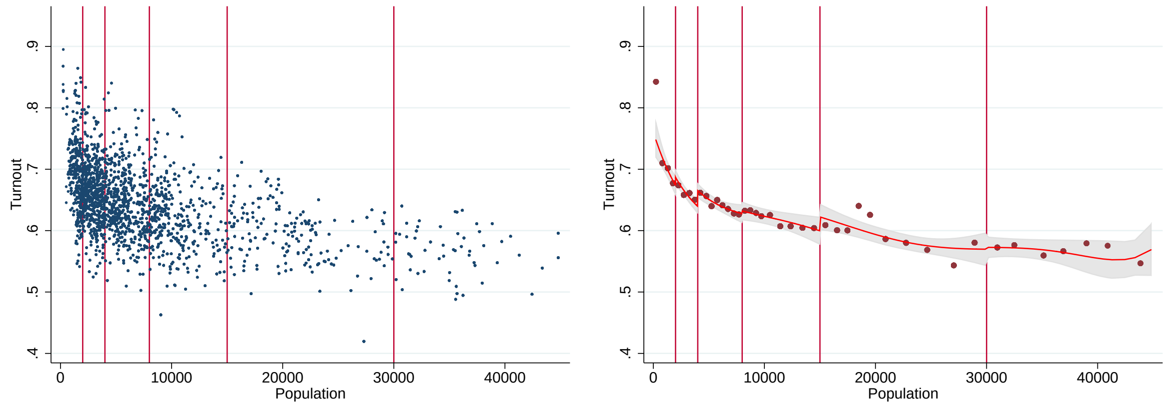
Table 4. Council size and voter turnout in municipal elections (Model 1 results).

Dep var: Turnout in municipal elections							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.025***	-0.016**	-0.01	0.001	0.008	0.013	0.015
	[0.007]	[0.008]	[0.008]	[0.009]	[0.010]	[0.011]	[0.012]
pop>4k	-0.002	0.011	0.019**	0.027***	0.028***	0.026***	0.023**
	[0.006]	[0.007]	[0.008]	[0.009]	[0.009]	[0.009]	[0.009]
pop>8k	0.001	0.018*	0.022**	0.016	0.009	0.005	0.005
	[0.008]	[0.010]	[0.010]	[0.010]	[0.011]	[0.012]	[0.012]
pop>15k	0.025**	0.036***	0.025*	0.007	0.010	0.023	0.023
	[0.012]	[0.014]	[0.014]	[0.016]	[0.016]	[0.018]	[0.018]
pop>30k	0.048***	0.005	0.001	0.033*	0.013	0.003	0.018
	[0.014]	[0.013]	[0.014]	[0.018]	[0.015]	[0.017]	[0.019]
Average effect	0.001	0.01*	0.013**	0.016***	0.015**	0.015**	0.016***
	[0.005]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]

Notes*:* Sample size is 1,747. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

In Graph 2, we illustrate these findings in two ways. In left hand graph we report a scatter plot of all the raw data and in the right hand graph we report bin averages. The population bin width varies somewhat based on the density of the data being 500 between 0 and 10000 population, 1000 between 10000 and 20000 and 2000 above that. The right hand graph also shows the fit of the equation (3) regression using on the 6th order polynomial of population. The band around the regression line shows the 95% confidence interval. The regression line jumps up at all the five cut-offs, indicating that council size increases turnout. The individual jumps are not statistically significant with the exception of the second jump, but the regression results in Table 4 show that the weighted average effect is significant and stable across the richer specifications. Furthermore, as correlation type evidence, we observe that turnout decreases with population, as we would expect if voters are rationally responding to pivotal probabilities.

Graph 2. Population and voter turnout (Model 1, 6th order polynomial of population).



The drawback of our main specification (Model 1) is that it uses data far from the cut-offs to estimate the function f at the cut-offs. The concern is that omitted election-level variables may confound the results if f does not adequately control for them. To analyze this issue, we look at the robustness of the results to adding control variables and adding flexibility to f. In Table B1 (Online appendix B), we repeat the analysis in Table 6 but add a set of municipality controls. The results do not change.

Table 5 shows the results of Model 2, where we allow for different polynomials between different cut-offs. The improvement in flexibility comes at the price of reduced efficiency and higher than 2nd order polynomials give estimates that are too imprecise to be informative, but the point estimate is robust. We report the estimated effect at each threshold and the average treatment effect estimates for specifications using 1st - 3rd order polynomials of population interacted with the group dummies.

The average treatment effect estimate is 1.4 percentage points and significant at the 5% level in the first column of Table 5 with the piecewise linear specification. Including the 2nd order term (column 2) reduces the estimate to 1.1 percentage points and it is now only weakly (10% level) significant. In the third column, the average effect increases but becomes insignificant. In the last three columns, we control for municipality attributes. The weighted estimated average effects are almost the same as without additional controls.

Table 5. Council size and voter turnout in municipal elections (Model 2 results).

Dep var: Turnout in municipal elections						
	Order of polynomial of pop					
Threshold	1st	2nd	3rd	1st	2nd	3rd
pop>2k	0.013	-0.002	0.013	0.015	0	0.014
	[0.011]	[0.012]	[0.014]	[0.011]	[0.012]	[0.013]
pop>4k	0.015*	0.019*	0.025**	0.013	0.018*	0.022*
	[0.009]	[0.010]	[0.012]	[0.008]	[0.010]	[0.012]
pop>8k	0.012	0.012	0.019	0.011	0.01	0.018
	[0.012]	[0.014]	[0.016]	[0.010]	[0.012]	[0.014]
pop>15k	0.018	0.014	-0.01	0.011	0.009	-0.012
	[0.016]	[0.022]	[0.026]	[0.015]	[0.021]	[0.025]
pop>30k	0.014	0.017	-0.017	0.019	0.013	-0.017
	[0.012]	[0.018]	[0.020]	[0.014]	[0.019]	[0.022]
Average effect	0.014**	0.011*	0.013	0.013**	0.010*	0.011
	[0.006]	[0.006]	[0.008]	[0.006]	[0.006]	[0.008]
Controls	NO	NO	NO	YES	YES	YES
N	1747	1747	1747	1736	1736	1736

Notes*:* Sample size is 1,747. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

In Table 6, we report the results from a nonparametric local linear estimation. We use triangular kernel. We report the results using the optimal bandwidth by Imbens and Kalyanaraman (2012) and also study the sensitivity of the results to half and double the optimal ones. We report the effect estimated at each threshold separately and also the weighted average effect. Standard errors for the average are not shown because the average is a combination of estimates from separate regressions. The results both at the individual thresholds and the average are very close to those obtained with the parametric specifications. Overall, we are comfortable in proceeding with the Model 1 in the IV.

Table 6. Nonparametric local linear RDD.

	Bandwidth		
Threshold	IK*0.5	IK	IK*2
pop>2k	-0.014	0.003	0.008
	[0.012]	[0.011]	[0.011]
Bandwidth	425	850	1700
N	218	411	507
pop>4k	0.034***	0.027***	0.023***
	[0.012]	[0.009]	[0.009]
Bandwidth	514	1028	2056
N	178	382	527
pop>8k	0.005	0.011	0.012
	[0.011]	[0.012]	[0.012]
Bandwidth	857	1713	3426
N	160	286	402
pop>15k	0.008	0.004	0.008
	[0.019]	[0.021]	[0.021]
Bandwidth	1952	3903	7806
N	77	143	199
pop>30k	-0.01	-0.001	0.005
	[0.022]	[0.015]	[0.014]
Bandwidth	3446	6891	13782
N	40	79	101
Average	0.006	0.012	0.013
	[N.A.]	[N.A.]	[N.A.]

Notes: Table shows nonparametric local linear estimation results for each threshold separately and the weighted average effect. Triangular kernel is used. IK refers to Imbens and Kalyanaraman (2012) optimal bandwidth. Standard errors are in brackets (clustered at the municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

6.2. Validity checks

We test the validity of the RDD through two placebo tests. First, we use data on turnout in national parliamentary elections to see if the overall propensity to vote is correlated with the treatment variables. These general elections take place one year prior to local elections and there pivotal probabilities are not affected by size of the municipality. Sample size is somewhat smaller than in the municipal election data because data on national elections in 1995 is not available to us. Graph B1 (Online appendix B) shows the fitted regression line of specification 6 in Table 4 with the general election turnout as the dependent variable. There are both upward and downward notches in the regression line at the cut-offs without a systematic pattern. Table B2 (Online appendix B) confirms that the average effect is close to zero and insignificant in all specifications (1^{st} to 7^{th} order polynomials). This suggests that the positive effect on turnout in local elections is indeed caused by the council size change and is not driven by other factors. Interestingly, the turnout in municipal elections is decreasing in population but in national elections there is no relationship between population and turnout. This is consistent with rational voting, since in national elections the pivotal probability is not influenced by municipal population whereas in municipal elections, local population has a large influence on the pivotal probability.

In our second placebo test, we estimate Model 1 with artificial cut-offs created by shifting the real cut-offs between -40% and 40%. We use the 6^{th} order polynomial of population. Graph B2 (Online appendix B) shows the results. The pattern is as it should be. Analysis with placebo thresholds results consistently in zero effect, unless the artificial location is very close the real one. When the location is shifted only between -1 % and +8 % the result is positive and statistically significant. These results reflect the somewhat inflexible specification rather than any threat to validity.

Manipulation and precise control over population measures would invalidate the research design. In our setup, the manipulation of population statistics would be very costly to municipalities, because this information is gathered independently by central government from the official population register. Furthermore, as is standard in the literature, we conduct a McCrary (2008) density test of manipulation separately for each threshold. The idea is to show that there are no discontinuities in the amount of observations at the thresholds, as there should be in the case of local randomization. We present these tests in Graph B3 (Online appendix B). We do not find evidence of manipulation at any of the five analyzed thresholds. Furthermore, the statistically insignificant jumps may go up or down depending on the threshold, implying that even a joint test would not (and does not) provide statistically significant evidence of a jump in any direction.

Finally, we test for the possibility that the results might be driven by confounding factors not adequately captured by the polynomial of the population by using background characteristics of municipalities as the dependent variable in Model 1. Table B3 (Online appendix B) reports these balancing tests for six municipality characteristics that are likely to correlate with turnout: number municipal employees per capita, unemployment rate, tax revenue per capita, share of over 65 year olds, central government grants per capita and municipal expenditure per capita. Three out of 30 estimates for individual threshold treatment effects are significant at the 5% level, but the average effect is insignificant for all of these covariates, supporting the validity of the RDD. Moreover, Table B3 reports two measures of political competition: the number of parties and the minimum within party margin of victory in the municipality. Neither of these measures jumps at the thresholds. This implies that we do not need to address political competition, i.e. closeness of elections, as an endogenous variable in the next section.

6.3. Instrumental variables regression

The purpose of the IV regression is to compare which channel, voters responding to pivotal probabilities, number of candidates or their quality, or proportionality explains the effect of council size on turnout. In Table 7, we report the results from the three first stages of our IV regression using municipal level analysis. We limit the analysis to 3rd to 6th order polynomials.[13] All the four variables seem to be relevant based on the significance of the average effects. Moreover, the individual thresholds seem to affect different outcomes somewhat differently, implying that the second stage regression may identify which of these mechanisms is behind the overall effect. Note also that the overall effect is strongest at the second threshold which is also significant for all the endogenous variables (at least in most specifications) implying that more than one mechanism may be behind the overall findings. The F tests tell that the first stage for proportionality is the strongest, for the pivotal probability the second strongest, but for the two candidate variables quite weak.

[13] We do not use the 1st - 2nd order, because the overall effect of council size on turnout was not present then. We do not use 7th or higher order, because the first stage F tests show that the first stages of all the endogenous variables become less powerful as the order of polynomial increase (see Online Appendix C). We also report the first stages graphically in the Appendix C.

Table 7. IV estimation, first stages, municipal level (Model 1).

	Dep var: Candidates				Dep var: Simulated pivotal probability			
	Order of polynomial of pop				Order of polynomial of pop			
Threshold	3rd	4th	5th	6th	3rd	4th	5th	6th
pop>2k	5.68***	2.47	3.8	1.31	-0.0086***	-0.0018	-0.0002	0.0007
	[1.99]	[2.12]	[2.34]	[2.06]	[0.0030]	[0.0034]	[0.0041]	[0.0049]
pop>4k	7.03***	4.72**	5.01**	6.30**	0.003	0.0079**	0.0077**	0.0068**
	[2.60]	[2.39]	[2.31]	[2.46]	[0.0023]	[0.0031]	[0.0034]	[0.0027]
pop>8k	3.83	5.68	4.37	6.61*	0.0179***	0.0142***	0.0107***	0.0095***
	[3.34]	[3.72]	[4.14]	[3.90]	[0.0026]	[0.0020]	[0.0018]	[0.0024]
pop>15k	8.03	13.38*	13.92*	8.07	0.0084***	-0.001	0.0002	0.0039
	[7.00]	[7.75]	[7.87]	[8.21]	[0.0022]	[0.0031]	[0.0027]	[0.0027]
pop>30k	31.49***	22.11**	18.17	22.70**	-0.0062**	0.0126***	0.0002	-0.0008
	[10.07]	[9.93]	[11.43]	[10.77]	[0.0029]	[0.0042]	[0.0027]	[0.0038]
Avg. effect	7.80***	6.96***	6.83***	6.63***	0.0042***	0.0062***	0.0050***	0.0051***
	[2.15]	[2.02]	[2.05]	[2.04]	[0.0015]	[0.0017]	[0.0017]	[0.0017]
1st stage F	3.9	2.8	2.8	2.8	21.5	15.0	8.7	7.4
	Dep var: Share of incumbents				Dep var: Proportionality			
	Order of polynomial of pop				Order of polynomial of pop			
Threshold	3rd	4th	5th	6th	3rd	4th	5th	6th
pop>2k	-0.013	0.005	0.016	0.033**	-0.00546***	-0.00400***	-0.00299***	-0.00187*
	[0.009]	[0.011]	[0.013]	[0.014]	[0.00048]	[0.00059]	[0.00079]	[0.00101]
pop>4k	0.019**	0.032***	0.034***	0.026**	-0.00222***	-0.00117	-0.00095	-0.00153**
	[0.010]	[0.012]	[0.012]	[0.010]	[0.00059]	[0.00081]	[0.00083]	[0.00061]
pop>8k	0.055***	0.044***	0.033***	0.019*	-0.0003	-0.00114***	-0.00214***	-0.00314***
	[0.011]	[0.010]	[0.010]	[0.010]	[0.00057]	[0.00038]	[0.00034]	[0.00052]
pop>15k	0.015	-0.015	-0.01	0.028**	-0.00059*	-0.00302***	-0.00261***	0.00001
	[0.010]	[0.014]	[0.013]	[0.014]	[0.00035]	[0.00072]	[0.00060]	[0.00055]
pop>30k	-0.018	0.034**	0.001	-0.029**	-0.00184**	0.00242**	-0.00058	-0.00261***
	[0.011]	[0.014]	[0.010]	[0.013]	[0.00071]	[0.00113]	[0.00079]	[0.00096]
Avg. effect	0.0173***	0.0219***	0.0208***	0.0222***	-0.0022***	-0.0018***	-0.0019***	-0.0019***
	[0.0061]	[0.0067]	[0.0064]	[0.0062]	0.0003	0.0004	0.0003	0.0003
1st stage F	7.9	6.6	3.9	3.3	47.2	34.4	20.8	20.6

Notes*:* Sample size is 1,746. 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

We report the municipality level second stage results in Table 8. The municipal level analysis is not able to identify separately which mechanism explains the increase in turnout. This is reflected both in the insignificant parameter estimates and in the underidentification test. The endogenous variables are too collinear and the instruments too weak to tell apart the mechanisms.

Table 8. The effects of pivotal probability, proportionality, number of candidates and share of incumbent candidates on turnout in municipal elections, IV estimates, second stage, municipal level.

Dep var: Turnout in municipal elections				
	Order of polynomial of pop			
Threshold	3rd	4th	5th	6th
Candidates	0.00076	0.002	0.00211	0.00158
	[0.00049]	[0.00178]	[0.00134]	[0.00099]
Share of incumbents	0.29313	1.42519	1.04383	0.67563
	[0.91511]	[2.28697]	[1.26416]	[0.46865]
Proportionality	1.0208	3.54528	5.8762	7.00966
	[3.32080]	[5.99473]	[6.17933]	[5.71161]
Pivotal probability	0.33077	-3.8488	-1.94513	0.63439
	[2.78577]	[7.39341]	[4.07795]	[1.97258]
N	1746	1746	1746	1746
Kleibergen-Paap underidentification test	0.998	0.349	0.719	2.47
p-value	0.607	0.84	0.698	0.29

Notes*:* Sample size is 1,746. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

In Table 9, we report the first stage results at the municipality-party level rather than the municipality level. This increases the number of observations from 1746 to 10171. For each party we can calculate the party level turnout (the share of eligible voters voting for a particular party), the number of candidates, the share of incumbent candidates and the pivotal probabilities both within the party and between the respective party and other parties, as well as the party proportionality as the seat share of the last candidate in the party. We also ask whether the political competition variable (the margin of victory within party) jumps at the threshold in the party-level data even though it did not in the municipal level analysis.

Both the between and within party pivotality jump at the thresholds and in roughly similar pattern. All the other four variables seem to be also relevant based on the significance of the average effects. Moreover, the individual thresholds seem to affect different outcomes somewhat differently, implying that the second stage regression may identify which of these mechanisms is behind the overall effect. We report the first stages graphically in Appendix C.

Table 9. IV estimation , first stages, municipality-party level (Model 1)

	Dep var: Between parties				Dep var: Within party pivotal probability			
	Order of polynomial of pop				Order of polynomial of pop			
Threshold	3rd	4th	5th	6th	3rd	4th	5th	6th
pop>2k	-0.005***	-0.001	-0.001	0.000	-0.004***	-0.002***	0.000	0.003
	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]
pop>4k	0.000	0.003**	0.003*	0.003*	0.001*	0.003***	0.004***	0.004***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
pop>8k	0.007***	0.007***	0.005***	0.004***	0.003***	0.002***	0.001**	-0.001*
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.000]	[0.001]
pop>15k	0.003***	-0.001	-0.001	0.001	0.001***	-0.002***	-0.003***	0.001*
	[0.001]	[0.001]	[0.002]	[0.002]	[0.000]	[0.001]	[0.001]	[0.001]
pop>30k	-0.002*	0.004**	0	-0.001	-0.002**	0.004***	0.000	-0.003**
	[0.001]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
Avg. effect	0.0012	0.0024***	0.0020**	0.0021**	0.0002	0.001**	0.0011**	0.0014**
	[0.0008]	[0.0009]	[0.0009]	[0.0009]	[0.0004]	[0.0005]	[0.0005]	[0.0006]
1st stage F	16.6	9.8	5.3	3.2	16.5	11.9	8.2	4.7
	Dep var: Number of candidates				Dep var: Proportionality			
	Order of polynomial of pop				Order of polynomial of pop			
Threshold	3rd	4th	5th	6th	3rd	4th	5th	6th
pop>2k	0.17	0.32	0.93*	0.92*	-0.0073***	-0.0061***	-0.0054***	-0.0045***
	[0.40]	[0.44]	[0.51]	[0.55]	[0.0007]	[0.0007]	[0.0008]	[0.0009]
pop>4k	1.10**	1.22**	1.48***	1.48***	-0.0041***	-0.0031***	-0.0028***	-0.0030***
	[0.54]	[0.54]	[0.53]	[0.53]	[0.0005]	[0.0006]	[0.0006]	[0.0006]
pop>8k	0.56	0.51	0.06	0.06	-0.0017***	-0.0021***	-0.0027***	-0.0033***
	[0.58]	[0.61]	[0.67]	[0.67]	[0.0004]	[0.0004]	[0.0004]	[0.0004]
pop>15k	0.67	0.44	0.42	0.41	-0.0017***	-0.0035***	-0.0035***	-0.0022***
	[1.12]	[1.26]	[1.25]	[1.33]	[0.0004]	[0.0005]	[0.0005]	[0.0004]
pop>30k	3.85***	4.20***	3.13**	3.14**	-0.0027***	0.0002	-0.0013**	-0.0025***
	[1.23]	[1.32]	[1.31]	[1.33]	[0.0005]	[0.0007]	[0.0005]	[0.0006]
Avg. effect	0.88**	0.93**	0.96***	0.96***	-0.0038***	-0.0034***	-0.0034***	-0.0033***
	[0.39]	[0.37]	[0.37]	[0.37]	0.0003	0.0003	0.0003	0.0003
1st stage F	3.0	3.3	2.8	2.8	55	38	34	32
	Dep var: Share of incumbents				Dep var: Political competition			
	Order of polynomial of pop				Order of polynomial of pop			
Threshold	3rd	4th	5th	6th	3rd	4th	5th	6th
pop>2k	-0.015	0.001	0.012	0.024*	-0.00085***	-0.00081***	-0.00084***	-0.00086**
	[0.009]	[0.010]	[0.012]	[0.013]	[0.00026]	[0.00028]	[0.00031]	[0.00034]
pop>4k	0.003	0.016	0.021**	0.019*	-0.00074***	-0.00071***	-0.00072***	-0.00072***
	[0.009]	[0.010]	[0.011]	[0.010]	[0.00019]	[0.00023]	[0.00024]	[0.00024]
pop>8k	0.048***	0.043***	0.035***	0.026**	-0.00038**	-0.00039**	-0.00037**	-0.00036*
	[0.010]	[0.010]	[0.010]	[0.011]	[0.00018]	[0.00017]	[0.00017]	[0.00020]
pop>15k	0.015	-0.009	-0.01	0.01	-0.00059***	-0.00064***	-0.00064***	-0.00067***
	[0.011]	[0.013]	[0.013]	[0.014]	[0.00017]	[0.00020]	[0.00020]	[0.00020]
pop>30k	-0.004	0.033*	0.013	-0.005	-0.00041**	-0.00033	-0.00028	-0.00026
	[0.014]	[0.017]	[0.015]	[0.016]	[0.00019]	[0.00025]	[0.00021]	[0.00026]
Avg. effect	0.0116*	0.0168***	0.0173***	0.0189***	-0.00063***	-0.00062***	-0.00062***	-0.00062***
	[0.0061]	[0.0062]	[0.0061]	[0.0061]	0.00012	0.00013	0.00014	0.00014
1st stage F	8.9	6.8	3.7	2.3	8.6	6.4	6.2	5.7

Notes*:* Unit of observation is party-election (N=8384 for proportionality and political competition, and 10,171 for the others). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

We begin the municipality-party level second stage analysis by asking how turnout responds to the between and within party pivotality in Table 10. The result is striking: Turnout responds only to the within party pivotal probabilities but not at all to the between party ones. The within party effect is large, robust and highly significant. The between party effect is negative but not significant. A likely explanation to this is that the within party dimension is much more salient. Within parties the election system is simply N past the post, where N refers to the number seats for the given party. The pivotal calculus is much simpler within party than between party, because it depends only on the votes given to members of that single party, and thus, voters can easily observe how many votes decided the elections at the margin of getting elected last time. On the contrary, between party pivotal calculus requires information on the votes to all the parties and understanding of the fairly complex election mathematics. According to the underidentification test in Table 10, the effects are separately identified despite the variables being correlated.

Table 10. The effects of between and within party pivotal probability on turnout, IV estimates, second stage, party level.

Dep var: Party turnout in municipal elections				
	Order of polynomial of pop			
Threshold	3rd	4th	5th	6th
Between party pivotality	-0.773	-0.971	-1.282	-1.282
	[0.897]	[0.833]	[0.898]	[0.898]
Within party pivotality	7.638***	8.045***	8.169***	8.169***
	[1.680]	[1.504]	[1.493]	[1.493]
N	10171	10171	10171	10171
Kleibergen-Paap underidentification test	8.84	11.8	12.1	12.9
p-value	0.065	0.019	0.016	0.012

Notes: Unit of observation is party-election (N=10,171). Only municipalities with a population below 45,000 are included. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Alternative explanation to the results in Table 10 is that within party pivotality is correlated with some other mechanism that is really behind the turnout response, whereas between party pivotality is not. We turn to this in Table 11 where the within party pivotality is analyzed jointly with the other endogenous variables. We find evidence that the most plausible channel of crossing the threshold on turnout is the pivotal probability, because it has a large, positive and significant effect in three out of four specifications and the other endogenous variables are not significant in any specification. Moreover, the underidentification tests suggest that for the 3rd order polynomial specification we are likely to be able to identify all the five effects separately. In Table C11 in the Appendix C, we also show that the result for pivotality are robust to including only one of the other endogenous variable at the time as well including these variables as squared to assure that nonlinearities in the effects of the other endogenous variables are not important.

Table 11. The effects of pivotal probability, number of candidates and share of incumbent candidates on turnout in municipal elections, IV estimates, second stage, party level.

Dep var: Party turnout in municipal elections				
	Order of polynomial of pop			
Threshold	3rd	4th	5th	6th
Candidates	0.000	0.004	0.003	-0.002
	[0.001]	[0.005]	[0.005]	[0.014]
Share of incumbents	-0.122	-0.189	-0.357	-1.025
	[0.152]	[0.202]	[0.711]	[3.515]
Proportionality	-0.106	-5.252	-8.217	-7.85
	[2.762]	[9.620]	[23.199]	[33.170]
Political competition	2.506	32.713	44.582	11.071
	[15.312]	[55.905]	[115.388]	[63.428]
Within party pivotality	6.017***	6.578***	7.098**	4.725
	[1.592]	[2.166]	[3.458]	[7.142]
N	8384	8384	8384	8384
Kleibergen-Paap underidentification test	4.42	0.737	0.274	0.128
p-value	0.035	0.39	0.60	0.72

Notes*:* Unit of observation is party-election (N=8384). Party list with none elected are excluded. Only municipalities with population below 45,000 are included. Standard errors are in brackets (clustered at the municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

It is hard to come up with possible omitted alternative mechanism than the ones already included in Table 11. For example any behavioral responses by the parties or the candidates, for example increased campaigning effort in the municipalities just above the threshold, should be captured by the included variables. Parties and candidates should respond to the political competition variable rather than the pivotality variable, because they can influence more than one voter simultaneously. Moreover, both this results and the fact that turnout only responds to the within party dimension largely rules out a role of elite mobilization.

7. Conclusions

We present quasi-experimental evidence that is consistent with the rational voting hypothesis. We use RDD to show that turnout increases when the number of available council seats in elections increases exogenously. We also use election simulations to show that the change in seats increases the probability of one single vote having an impact on the election outcome. This change is sharp and relatively large at the discontinuity. We also use a novel instrumental variables design utilizing the presence of multiple thresholds to show that the effect on turnout can be attributed to the increase in pivotal probability rather than to the simultaneous increase in the number of candidates, the candidates' quality, proportionality or political completion more generally. The voters seem to conduct calculus of voting and take this into account when making decisions on whether to vote or to abstain.

Our results do not imply that the expressive utility components do not matter. Moreover, our results cannot rule out some of the alternative explanations for pure rational voting such as group voting where larger groups consider their pivotality together. We can only state that the calculus of voting seems to also matter. It may also be the case that only some voters but not all conduct the calculus of voting. Moreover, we do not learn much concerning how exact or heuristic this calculus is, for example it may well be that voters learn about pivotal probabilities from the results of the past elections rather than actually calculating them. We can only say that whatever type of calculus is conducted, it is empirically consistent with the rational voting model in Finnish municipal elections.

The Finnish proportional open-list election system has features of both majority and closed-list elections, since the within party competition component is simply N-past–the-post and the between parties competition component is the same as in the closed-list proportional system. Therefore, the results can potentially apply to a wide range of other institutions and countries. Of particular interest here is that voters seem only to respond to the more salient within party pivotal probabilities but not to the between party ones. This result has implications on the generalizability of the results. One should expect to see rational voting under simple election rules such as first-past-the-post or majority elections, but less likely in more complex systems such as closed-list proportional elections.

Finally, the only interest of this study is not what mechanism voters respond to but also what mechanisms they do not respond to. It is interesting to observe that the possible increase in political efficacy due to proportionality effects are not likely to be behind the turnout results, nor are elite mobilization or other party or candidate responses.

References

Ade F, Freier R (2011). When Can We Trust Population Thresholds in Regression Discontinuity Designs? DIW Berlin Discussion Paper No. 1136. http://www.diw.de/documents/publikationen/73/diw_01.c.376204.de/dp1136.pdf.

Andersen J, Fiva J, Natvik G (2014). Voting When the Stakes Are High. Journal of Public Economics 110, 157-166.

Aldrich J (1997). When is it Rational to Vote? In Perspectives on Public Choice, A Handbook, pp.379-390. Mueller D (ed.), Cambridge University Press.

Blais A (2000). To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory. University of Pittsburg Press.

Cattaneo M, Keele L, Titiunik R and Vazquez-Bare G (2015). Interpreting Regression Discontinuity Designs with Multiple Cutoffs. Working paper, June 2015.

http://www-personal.umich.edu/~cattaneo/papers/Cattaneo-Keele-Titiunik-VazquezBare_2015_wp.pdf.

Cattaneo M, Frandsen B, Titiunik R. (2015b). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. Journal of Causal Inference 3, 1-24.

Coate S, Conlin M (2002). A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence. American Economic Review 94, 1476-1504.

Coate S, Conlin M, Moro A (2008). The Performance of Pivotal-voter Model in Small-scale Elections: Evidence from Texas Liquor Referenda. Journal of Public Economics 92, 582-596.

Cox G (1988). Closeness and Turnout: A Methodological Note. The Journal of Politics 50, 768-775.

Cox G (1997). Making Votes Count: Strategic Coordination in the World's Electoral Systems. Cambridge University Press.

Degan A, Merlo A (2011). A Structural Model of Turnout and Voting in Multiple Elections. Journal of the European Economic Association 9, 209-245.

Downs A (1957). An economic theory of democracy. Harper and Row, New York.

Edlin A, Gelman A, Kaplan N (2007). Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others. Rationality and Society 19, 293-314.

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

Eggers A, Freier R, Grembi V and Nannicini T (2015b). Regression Discontinuity Designs Based on Population Thresholds – Pitfalls and Solutions. DIW Berlin Discussion Papers 1503. https://diw.de/documents/publikationen/73/diw_01.c.513460.de/dp1503.pdf.

Feddersen T, Pesendorfer W (1996). The Swing Voter’s Curse. American Economic Review 86, 408-424.

Feddersen T, Pesendorfer W (1999). Abstention in Elections with Asymmetric Information and Diverse Preferences. American Political Science Review 93, 381-398.

Ferraz C, Finan F (2009). Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance. NBER Working Paper No. 14906. http://www.nber.org/papers/w14906.pdf.

Foster C (1984). The Performance of Rational Voter Models in Recent Presidential Elections. American political Science Review 78, 678-690.

Fiorina M (1989). Retrospective Voting in American National Elections. New Haven: Yale University Press.

Fiva J, Folke O (2015). Mechanical and Psychological Effects of Electoral Reform. British Journal of Political Science (Forthcoming).

Fujiwara T (2011). A Regression Discontinuity Test of Strategic Voting and Duverger's Law. Quarterly Journal of Political Science 6, 197–233.

Funk P (2010). Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System. Journal of the European Economic Association 8, 1077-1103.

Gallego A, Rico G, Anduiza E (2012). Disproportionality and Voter Turnout in New and Old Democracies. Electoral Studies 31, 159-169.

Gelman A, King G, Boscardin J (1998). Estimating the Probability of Events That Have Never Occurred: When Is Your Vote Decisive? Journal of the American Statistical Association 93, 1-9.

Gelman A, Imbens G (2014). Why Higher Order Polynomials Should not be Used in Regression Discontinuity Designs. NBER Working Paper 20405. http://www.nber.org/papers/w20405.

Gerber A, Green, D, Larimer, C (2008). Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment. American Political Science Review 102, 33-48.

Geys B (2006). Explaining Voter Turnout: A Review of Aggregate Level Research. Electoral Studies 25, 637-663.

Green D, Shapiro I (1994). Pathologies of Rational Choice Theory: A Critique of Applications in Political Science. Yale University Press.

Grembi V, Nannicini T, Troiano U (2012). Policy Responses to Fiscal Restraints: a Difference-in-Discontinuities Design. CESifo Working Paper 3999. http://www.econstor.eu/bitstream/10419/68212/1/733920993.pdf.

Grofman B (1993). Is Turnout the Paradox that Ate the Rational Choice Theory? In Information, Participation, and Choice: An Economic Theory of Democracy in Perspective, Grofman B (ed.), 93-103, University of Michigan Press.

Harsanyi J (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. Journal of Political Economy 63, 309-321.

Imbens G, Kalyanaraman K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. Review of Economic Studies 79, 933-59.

Indridason I (2008). Competition & Turnout: The Majority Run-Offs as a Natural Experiment. Electoral Studies 27, 699-710.

Karp J and Banducci S (2008). Political Efficacy and Participation in twenty-seven Democracies: How Electoral systems Shape Political Behaviour. British Journal of Political Science 38, 311-334.

Kawai K, Watanabe Y. (2013). Inferring Strategic Voting. American Economic Review 103, 624–662.

Kotakorpi K, Poutvaara P, Terviö M (2013). Returns to office in national and local politics. CESifo Working Paper 4542. http://www.cesifo-group.de/portal/page/portal/DocBase_Content/WP/WP-CESifo_Working_Papers/wp-cesifo-2013/wp-cesifo-2013-12/cesifo1_wp4542.pdf.

Lago I (2012). Strategic Voting in Proportional Representation Systems: Evidence from a Natural Experiment. Party Politics 18, 653–665.

Lee D, Lemieux T (2010). Regression Discontinuity Designs in Economics. Journal of Economic Literature 48, 281-355.

McCrary J (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. Journal of Econometrics 142, 968-714.

Myatt D (2012). A Rational Theory of Voter Turnout. London Business School mimeo. http://dpmyatt.org/uploads/turnout-2012-jan-wp.pdf.

Matsusaka J (1995). Explaining Voter Turnout Patterns: An Information Theory. Public Choice 84, 91-117.

Matsusaka J, Palda F (1999). Voter Turnout: How Much Can We Explain? Public Choice 98, 431-446.

Mulligan C, Hunter C (2003). The Empirical Frequency of a Pivotal Vote. Public Choice 116, 31-54.

Palfrey T, Rosenthal H (1983). The Strategic Calculus of Voting. Public Choice 41, 7-53.

Palfrey T, Rosenthal H (1985). Voter Participation and Strategic Uncertainty. American Political Science Review 79, 62-78.

Pettersson-Lidbom P (2012). Does the Size of the Legislature Affect the Size of Government: Evidence from Two Natural Experiments. Journal of Public Economics 96, 269-278.

Powell Jr. G (1986). American Voter Turnout in Comparative Perspective. The American Political Science Review 80, 17-43.

Riker W, Ordenshook P (1968). A Theory of the Calculus of Voting. American Political Science Review 62, 25-42.

Saarimaa T, Tukiainen J (2016). Local Representation and Strategic Voting: Evidence from Electoral Boundary Reforms. European Journal of Political Economy 41, 31-45.

Supporting information (Online)

ONLINE APPENDIX A: How pivotal probability changes at the thresholds based on simulations

We conduct two counterfactual experiments in the pivotal probability simulation to assess the effect of crossing the population threshold on the pivotal probabilities. In the first counterfactual, we assume that in each municipality, fewer seats than in the real elections were actually given while maintaining everything else the same. The number of seats is assumed to be what they would be below the next population threshold downwards. However, we keep the total number of votes given unchanged. Only the allocated amount of seats is different. We call this counterfactual “CF down” in Table 2. “CF up” is otherwise the same but the number of seats is as above the next threshold upwards.

In Table A1, we report a summary of our simulation results both for the simulated real council size elections and the simulated counterfactual elections. In the first three rows, we show the average between parties results for all the election-party observations and in the last three rows we show the average within party results for all the election-candidate observations. The counterfactual results work as expected, as pivotal probability is lower when fewer seats are allocated and higher with more seats.

Table A1. Descriptive statistics of the pivotal probability simulations.

Simulation	**N**	**Mean**	**Std. Dev.**	**Min**	**Max**
CF down between	10171	0.017	0.019	0	0.111
Real between	10171	0.020	0.020	0	0.123
CF up between	10171	0.024	0.021	0	0.231
CF down within	146234	0.024	0.047	0	0.408
Real within	146234	0.028	0.051	0	0.389
CF up within	146234	0.032	0.057	0	0.494

Notes*:* The unit of observation is election-party (first three rows) or election-candidate (last three rows). Only municipalities with a population below 45,000 are included. “between” and “within” refer to pivotalities between and within parties. “CF down” refers to counterfactual simulation where the council size is what it would be in the next population group below the real one and “CF up” refers to counterfactual simulation where the council size is what it would be in the next population group above the real one.

In Table A2, we report the simulated causal effect of crossing the nearest population threshold on the pivotal probability for each threshold separately. This effect is not yet the actual first stage regression of the IV estimation, but rather based on the counterfactual simulations. This effect is calculated as the difference between the “Real” and “CF down” results for those municipalities just above (10% bandwidth) the threshold and as the difference between “CF up” and “Real” for those municipalities just below the threshold. We find systematically larger effects for smaller thresholds as expected.

Table A2. The effect of crossing the threshold on the pivotal probabilities both between and within parties. Simulation results within 10% population bands around each threshold.

Threshold	**Band**	**Mean**	**Std. Dev.**	**N**	**Relative council size change**
2k between	10 %	0.0052	0.029	518	0.24
4k between	10 %	0.0045	0.024	717	0.29
8k between	10 %	0.0039	0.018	947	0.3
15k between	10 %	0.0019	0.015	430	0.23
30k between	10 %	0.0008	0.011	322	0.19
2k within	10 %	0.0070	0.056	4799	0.24
4k within	10 %	0.0066	0.052	8554	0.29
8k within	10 %	0.0055	0.043	14081	0.3
15k within	10 %	0.0030	0.031	8324	0.23
30k within	10 %	0.0024	0.023	7376	0.19

Notes: The unit of observation is election-party (first three rows) or election-candidate (last three rows). Only municipalities with a population below 45,000 are included. “Relative council size change” is the relative council size change at the given threshold. N is the number of observations (at party or candidate level) in the group around the threshold defined as being within the 10% population band of the population at the threshold. “Mean” is the average change in the pivotal probability of crossing the threshold defined as the average of the differences between both real and counterfactual down and counterfactual up and real.

ONLINE APPENDIX B: Robustness and validity

Table B1. Council size and voter turnout (municipality attributes controlled for).

Dep var: Turnout in municipal elections							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.0250***	-0.0164**	-0.0077	0.0028	0.0103	0.0153	0.0173
	[0.0071]	[0.0075]	[0.0081]	[0.0093]	[0.0101]	[0.0110]	[0.0113]
pop>4k	-0.0061	0.0072	0.0176**	0.0249***	0.0263***	0.0237***	0.0203**
	[0.0056]	[0.0071]	[0.0080]	[0.0086]	[0.0086]	[0.0083]	[0.0087]
pop>8k	-0.0006	0.0160*	0.0218**	0.0155	0.008	0.0035	0.0036
	[0.0075]	[0.0091]	[0.0095]	[0.0095]	[0.0100]	[0.0107]	[0.0106]
pop>15k	0.0234*	0.0329**	0.0179	0.0008	0.0046	0.0163	0.0167
	[0.0121]	[0.0132]	[0.0138]	[0.0158]	[0.0153]	[0.0169]	[0.0172]
pop>30k	0.0487***	0.0054	0.0046	0.0331*	0.0094	0.0014	0.0177
	[0.0140]	[0.0152]	[0.0148]	[0.0189]	[0.0159]	[0.0179]	[0.0196]
Average effect	-0.001	0.008	0.012**	0.014**	0.014**	0.014**	0.015**
	[0.005]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
N	1736	1736	1736	1736	1736	1736	1736

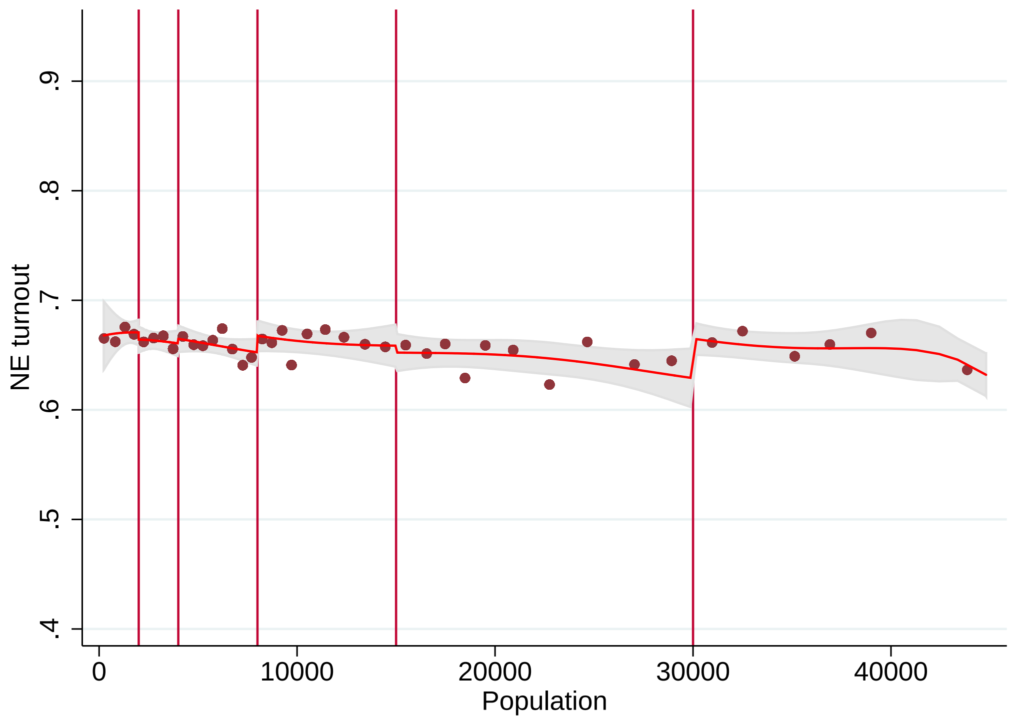
Notes*:* Sample size is 1,736. Controls: tax revenue/capita, municipality employees/capita, unemployment rate, central government grants/capita, share of over 65 year olds, municipal expenditure/capita. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Table B2. Placebo tests with national elections data.

Dep var: Turnout in national elections							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.0051	-0.0048	-0.0038	-0.0037	-0.0041	-0.0066	-0.0086
	[0.0063]	[0.0065]	[0.0071]	[0.0081]	[0.0090]	[0.0099]	[0.0104]
pop>4k	0.0011	0.0015	0.0028	0.0028	0.0028	0.0042	0.0079
	[0.0058]	[0.0074]	[0.0090]	[0.0100]	[0.0101]	[0.0101]	[0.0104]
pop>8k	0.0112	0.0117	0.0125	0.0125	0.0129	0.0151	0.0149
	[0.0071]	[0.0090]	[0.0096]	[0.0099]	[0.0106]	[0.0111]	[0.0111]
pop>15k	0.0013	0.0017	-0.0002	-0.0002	-0.0005	-0.0063	-0.006
	[0.0103]	[0.0115]	[0.0114]	[0.0137]	[0.0132]	[0.0153]	[0.0153]
pop>30k	0.0324**	0.0310***	0.0295**	0.0297*	0.0311**	0.0360**	0.0166
	[0.0138]	[0.0118]	[0.0134]	[0.0165]	[0.0134]	[0.0167]	[0.0168]
Average effect	0.004	0.005	0.005	0.005	0.005	0.005	0.004
	[0.004]	[0.005]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
N	1076	1076	1076	1076	1076	1076	1076

Notes: Sample size is 1,076. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph B1. Placebo test with national elections turnout (Model 1, 6th order polynomial).



Graph B2. Placebo tests with artificial cut-offs (Model 1, 6th order polynomial of pop).

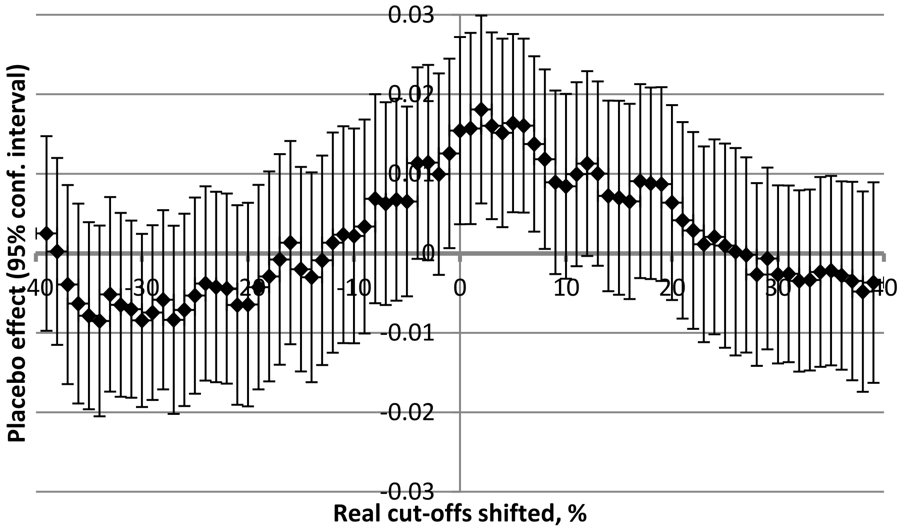
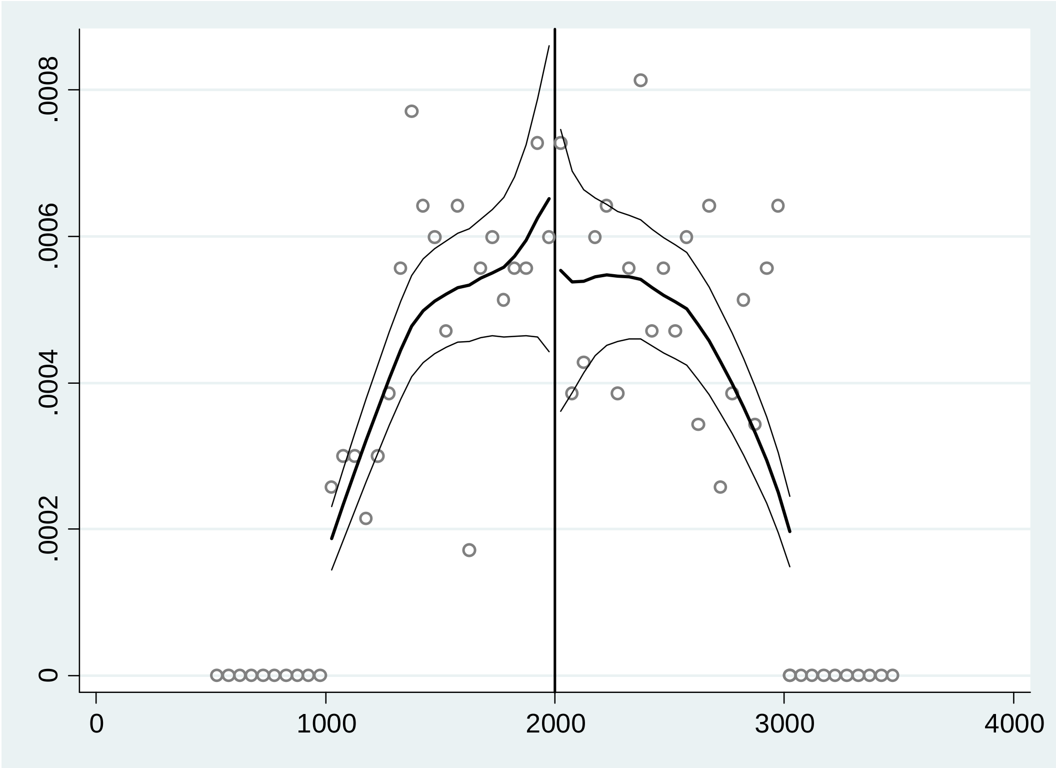


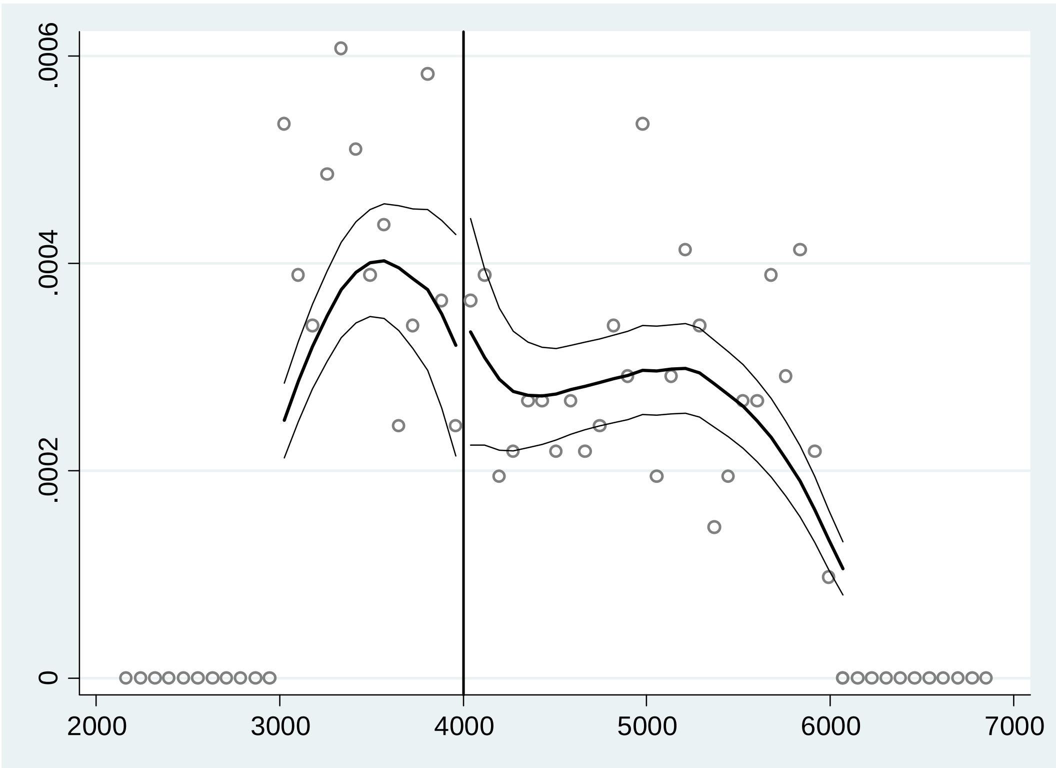
Table B3. Balancing tests (Model 1, 6th order polynomial of pop).

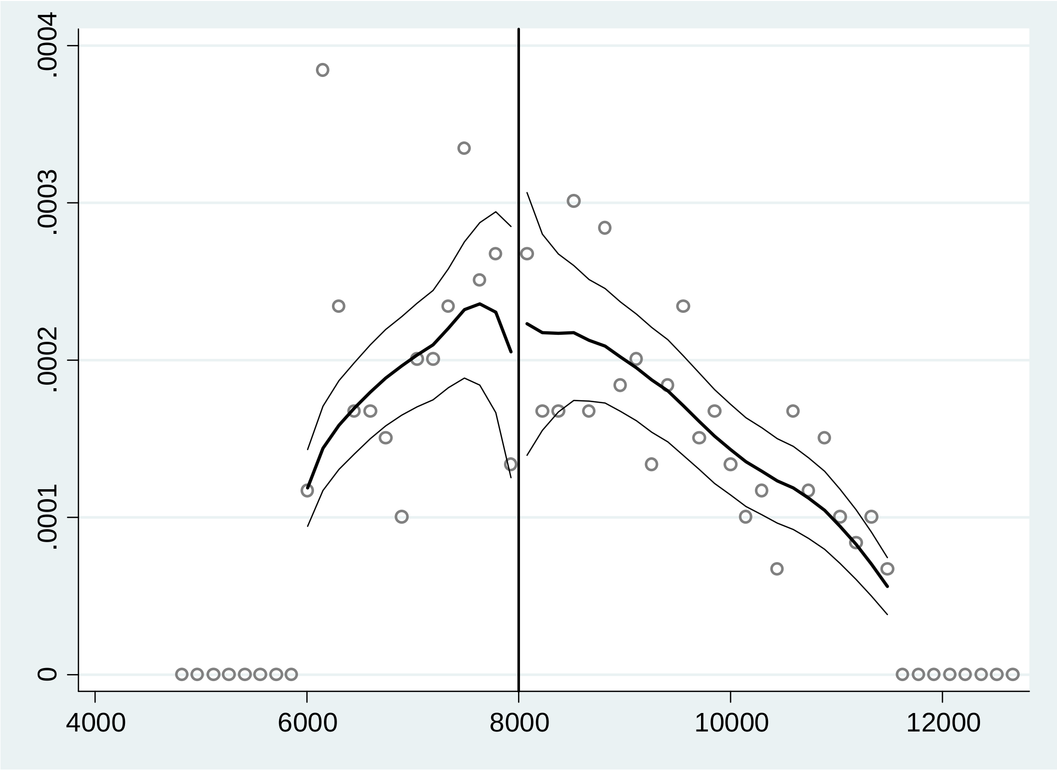
Threshold	Munic. Employees	Tax revenue	Share of over 65yo	Grants	Expenditures	Unemp. share	Number of parties	Political competition
pop>2k	111	-0.061	0.010	0.194	0.266	0.164	-0.0001	-0.0937
	[167]	[0.077]	[0.009]	[0.162]	[0.268]	[0.909]	[0.0001]	[0.2224]
pop>4k	111	0.074	-0.006	-0.053	0.091	0.051	0.0000	-0.1454
	[126]	[0.068]	[0.009]	[0.138]	[0.183]	[0.944]	[0.0001]	[0.1941]
pop>8k	35	-0.021	-0.010	0.045	-0.007	-0.403	-0.0002***	0.3644
	[134]	[0.141]	[0.009]	[0.149]	[0.218]	[1.060]	[0.0001]	[0.2298]
pop>15k	59	0.323**	0.013	-0.059	0.006	-2.029	0.0000	0.288
	[168]	[0.156]	[0.012]	[0.192]	[0.276]	[1.336]	[0.0001]	[0.3089]
pop>30k	-59	-0.331	-0.036**	-0.206	-0.660**	-0.53	-0.0002	-0.4268
	[180]	[0.286]	[0.018]	[0.183]	[0.326]	[2.355]	[0.0001]	[0.5287]
Avg. effect	72.6	0.0282	-0.00221	0.0179	0.0436	-0.386	-0.000065	0.0393
	79.2	0.0594	0.00495	0.0852	0.125	0.571	0.000040	0.1262
N	1736	1736	1736	1736	1736	1736	1733	1736

Notes*:* All models use the parametric RDD with 6th order polynomial. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

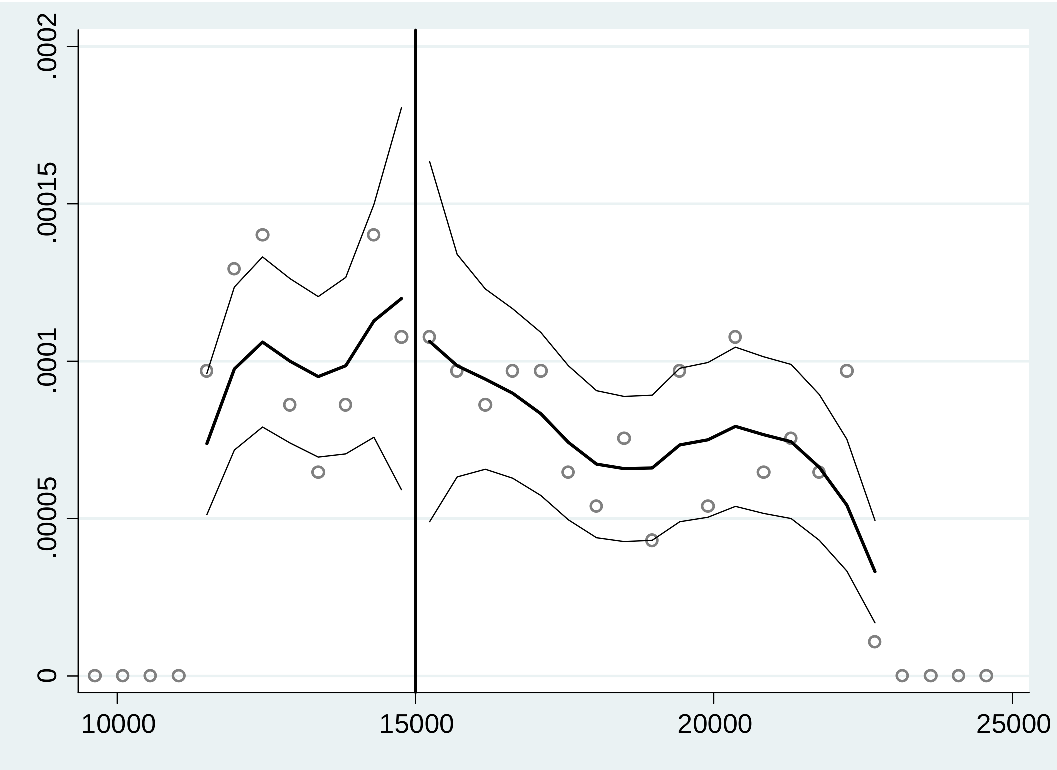
Graph B3. McCrary (2008) tests of manipulation of the forcing variable for each threshold.



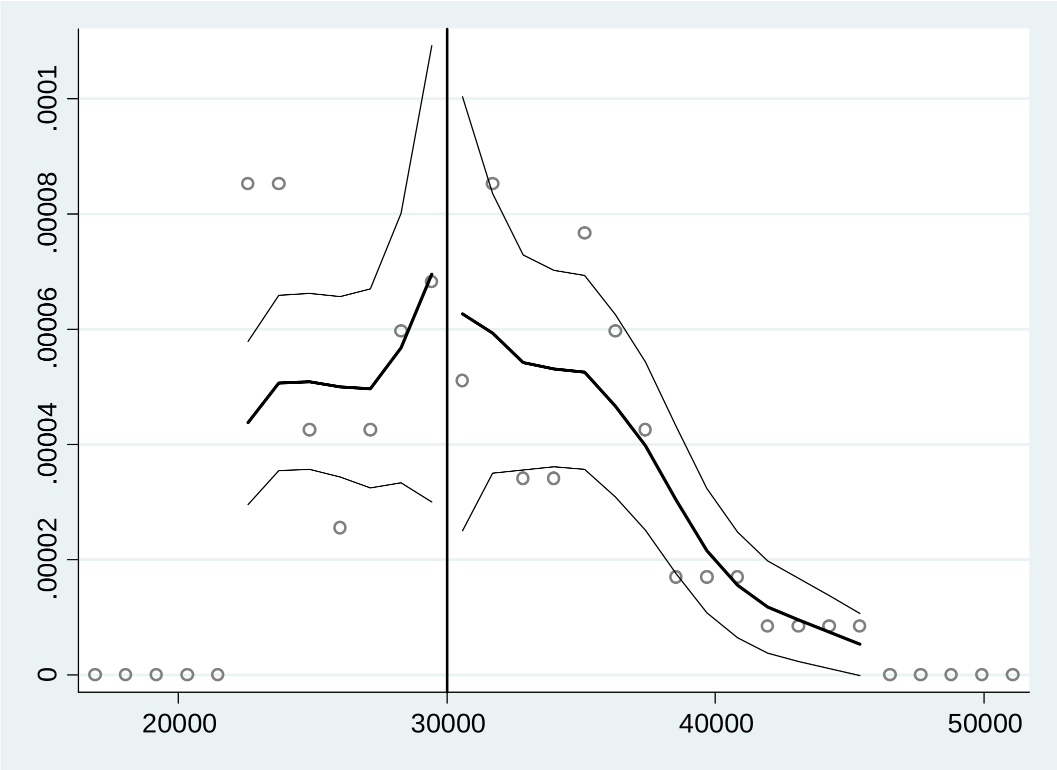




.0004
.0003
.0002
.0001
0
4000
6000
8000
10000
12000



.0002
.00015
.0001
.00005
0
10000
15000
20000
25000


0
.00002
.00004
.00006
.00008
.0001
20000
30000
40000
50000

ONLINE APPENDIX C: First stages of IV regression, municipality and municipality-party levels (Model 1).

Table C1. IV estimation, first stage for simulated pivotal probability, municipality level (Model 1).

Dep var: Simulated pivotal probability							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.0223***	-0.0159***	-0.0086***	-0.0018	-0.0002	0.0007	0.0035
	[0.0033]	[0.0032]	[0.0030]	[0.0034]	[0.0041]	[0.0049]	[0.0056]
pop>4k	-0.0163***	-0.0052***	0.003	0.0079**	0.0077**	0.0068**	0.0067***
	[0.0017]	[0.0020]	[0.0023]	[0.0031]	[0.0034]	[0.0027]	[0.0022]
pop>8k	-0.0006	0.0128***	0.0179***	0.0142***	0.0107***	0.0095***	0.0089***
	[0.0018]	[0.0024]	[0.0026]	[0.0020]	[0.0018]	[0.0024]	[0.0025]
pop>15k	0.0124***	0.0202***	0.0084***	-0.001	0.0002	0.0039	0.0048
	[0.0028]	[0.0030]	[0.0022]	[0.0031]	[0.0027]	[0.0027]	[0.0030]
pop>30k	0.0321***	-0.0015	-0.0062**	0.0126***	0.0002	-0.0008	-0.0013
	[0.0050]	[0.0055]	[0.0029]	[0.0042]	[0.0027]	[0.0038]	[0.0030]
Avg. effect	-0.0061***	0.0008	0.0042***	0.0062***	0.0050***	0.0051***	0.0056***
	[0.0015]	[0.0014]	[0.0015]	[0.0017]	[0.0017]	[0.0017]	[0.0019]
1st stage F	106.0	39.8	21.5	15.0	8.7	7.4	6.3

Notes: Sample size is 1,746. 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C1. Population and pivotality, municipality level (Model 1, 6th order polynomial of population).

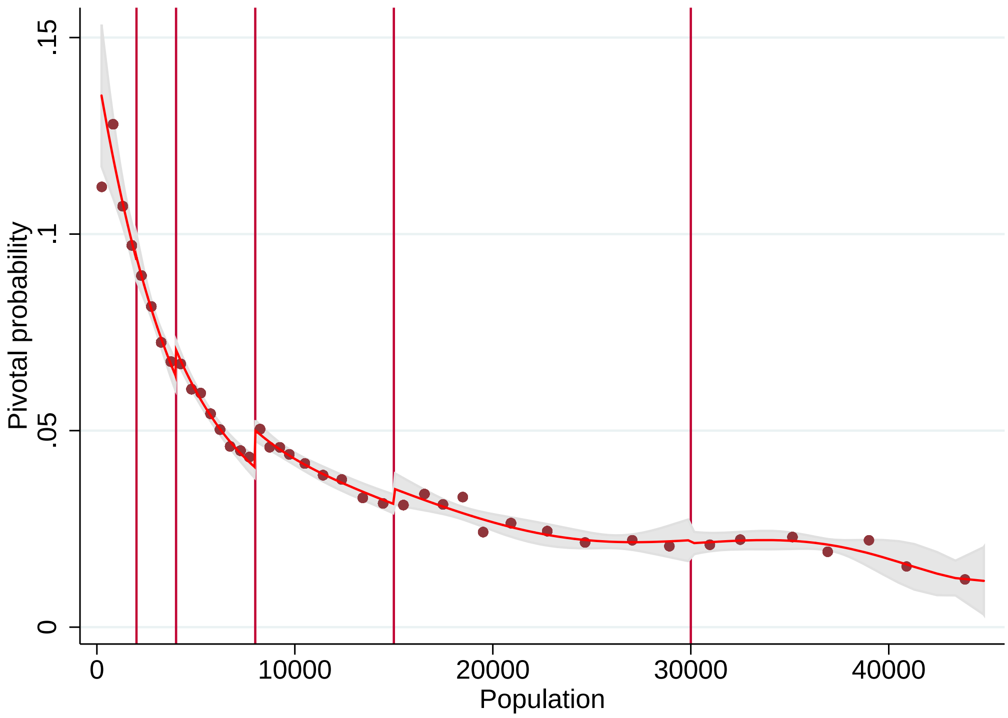


Table C2. IV estimation, first stage for the number of candidates, municipality level (Model 1).

Dep var: Candidates							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	11.08***	6.77***	5.68***	2.47	3.8	1.31	-0.31
	[1.48]	[1.77]	[1.99]	[2.12]	[2.34]	[2.06]	[1.85]
pop>4k	15.13***	8.37***	7.03***	4.72**	5.01**	6.30**	8.87***
	[1.98]	[2.54]	[2.60]	[2.39]	[2.31]	[2.46]	[2.68]
pop>8k	13.08***	4.63	3.83	5.68	4.37	6.61*	6.54*
	[3.23]	[3.57]	[3.34]	[3.72]	[4.14]	[3.90]	[3.86]
pop>15k	11.47*	6.22	8.03	13.38*	13.92*	8.07	7.58
	[6.62]	[6.03]	[7.00]	[7.75]	[7.87]	[8.21]	[8.16]
pop>30k	9.69	30.74***	31.49***	22.11**	18.17	22.70**	10.14
	[8.40]	[10.35]	[10.07]	[9.93]	[11.43]	[10.77]	[11.64]
Avg. effect	12.74***	8.32***	7.80***	6.96***	6.83***	6.63***	6.04***
	[2.27]	[2.20]	[2.15]	[2.02]	[2.05]	[2.04]	[2.05]
1st stage F	29.7	9.1	3.9	2.8	2.8	2.8	3.2

Notes: Sample size is 1,746. 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C2. Population and the number of candidates, municipality level (Model 1, 6th order polynomial of population).

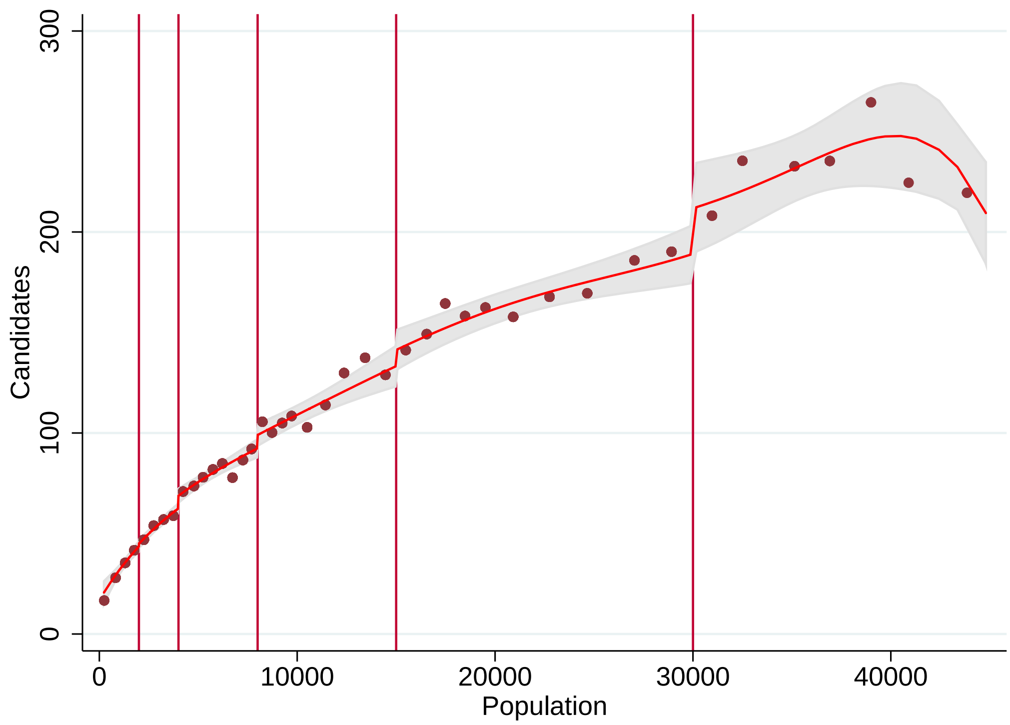


Table C3. IV estimation, first stage for the share of incumbents, municipality level (Model 1).

Dep var: Share of incumbents							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.039***	-0.025***	-0.013	0.005	0.016	0.033**	0.040***
	[0.009]	[0.009]	[0.009]	[0.011]	[0.013]	[0.014]	[0.014]
pop>4k	-0.017***	0.004	0.019**	0.032***	0.034***	0.026**	0.014
	[0.006]	[0.007]	[0.010]	[0.012]	[0.012]	[0.010]	[0.010]
pop>8k	0.019***	0.045***	0.055***	0.044***	0.033***	0.019*	0.019*
	[0.007]	[0.009]	[0.011]	[0.010]	[0.010]	[0.010]	[0.010]
pop>15k	0.020*	0.036***	0.015	-0.015	-0.01	0.028**	0.031**
	[0.011]	[0.011]	[0.010]	[0.014]	[0.013]	[0.014]	[0.014]
pop>30k	0.055***	-0.009	-0.018	0.034**	0.001	-0.029**	0.031***
	[0.012]	[0.010]	[0.011]	[0.014]	[0.010]	[0.013]	[0.011]
Avg. effect	-0.0024	0.0111**	0.0173***	0.0219***	0.0208***	0.0222***	0.025***
	[0.0043]	[0.0053]	[0.0061]	[0.0067]	[0.0064]	[0.0062]	[0.0061]
1st stage F	16.4	10.3	7.9	6.6	3.9	3.3	4.3

Notes: Sample size is 1,746. 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C3. Population and share of incumbents, municipality level (Model 1, 6th order polynomial of population).

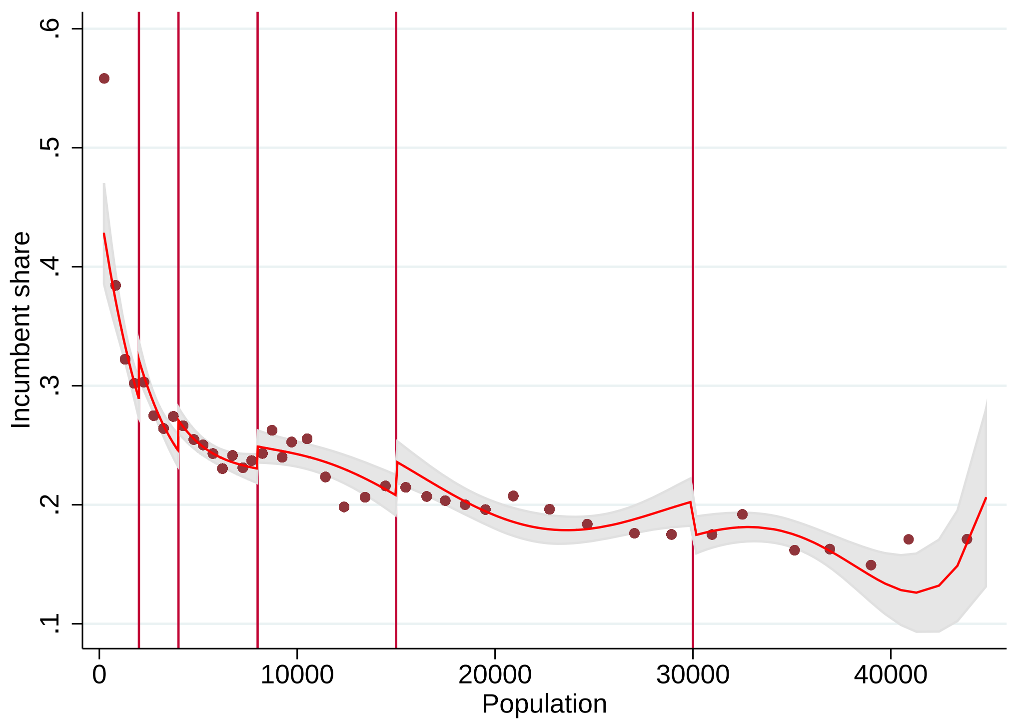


Table C4. IV estimation, first stage for proportionality, municipality level (Model 1).

Dep var: Proportionality							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.00693***	-0.00628***	-0.00546***	-0.00400***	-0.00299***	-0.00187*	-0.00147
	[0.00059]	[0.00053]	[0.00048]	[0.00059]	[0.00079]	[0.00101]	[0.00108]
pop>4k	-0.00425***	-0.00322***	-0.00222***	-0.00117	-0.00095	-0.00153**	-0.00217***
	[0.00027]	[0.00039]	[0.00059]	[0.00081]	[0.00083]	[0.00061]	[0.00045]
pop>8k	-0.00218***	-0.00091**	-0.0003	-0.00114***	-0.00214***	-0.00314***	-0.00312***
	[0.00025]	[0.00045]	[0.00057]	[0.00038]	[0.00034]	[0.00052]	[0.00050]
pop>15k	-0.00002	0.00077	-0.00059*	-0.00302***	-0.00261***	0.00001	0.00013
	[0.00040]	[0.00052]	[0.00035]	[0.00072]	[0.00060]	[0.00055]	[0.00055]
pop>30k	0.00191***	-0.00127*	-0.00184**	0.00242**	-0.00058	-0.00261***	0.00052
	[0.00064]	[0.00065]	[0.00071]	[0.00113]	[0.00079]	[0.00096]	[0.00081]
Avg. effect	-0.0033***	-0.0026***	-0.0022***	-0.0018***	-0.0019***	-0.0019***	-0.0017***
	0.0002	0.0002	0.0003	0.0004	0.0003	0.0003	0.0004
1st stage F	244.5	111.8	47.2	34.4	20.8	20.6	18.3

Notes*:* Sample size is 1,746. 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C4. Population and proportionality, municipality level (Model 1, 6th order polynomial of population).

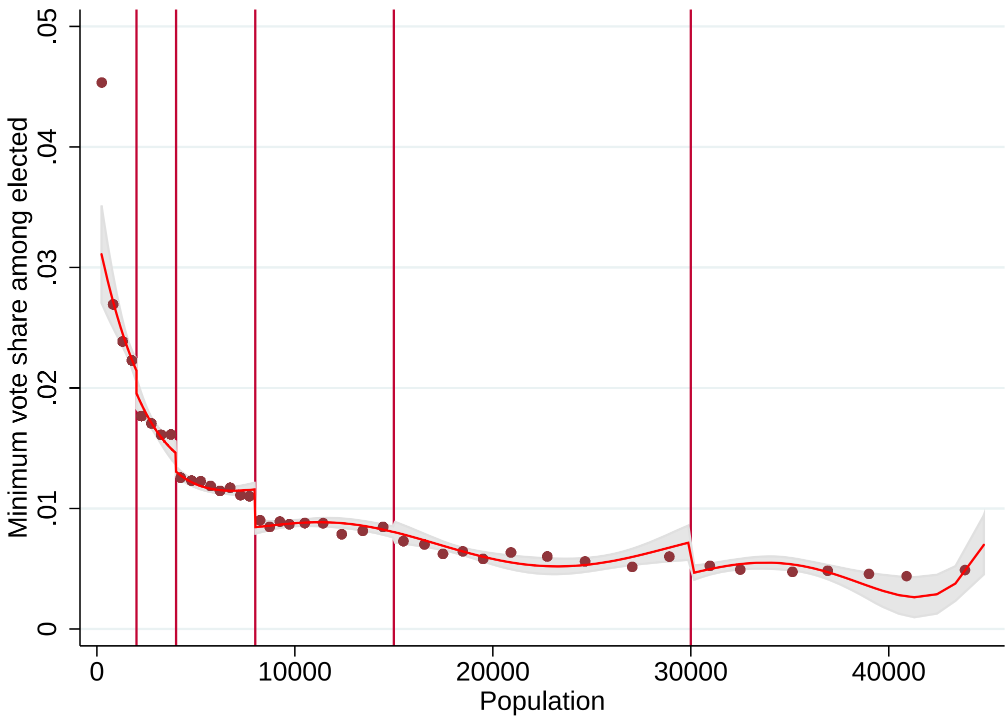


Table C5. IV estimation , first stage for between party pivotal probability, municipality-party level (Model 1).

Dep var: Simulated pivotal probability (between parties)							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.009***	-0.008***	-0.005***	-0.001	-0.001	0.000	0.002
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.003]
pop>4k	-0.008***	-0.003***	0.000	0.003**	0.003*	0.003*	0.003*
	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.001]
pop>8k	0.001	0.004***	0.007***	0.007***	0.005***	0.004***	0.004**
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
pop>15k	0.005***	0.007***	0.003***	-0.001	-0.001	0.001	0.002
	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]	[0.002]
pop>30k	0.012***	0.001	-0.002*	0.004**	0	-0.001	-0.001
	[0.003]	[0.002]	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]
Avg. effect	-0.0026***	-0.0006	0.0012	0.0024***	0.0020**	0.0021**	0.0026**
	[0.0009]	[0.0007]	[0.0008]	[0.0009]	[0.0009]	[0.0009]	[0.0011]
1st stage F	37.9	28.5	16.6	9.8	5.3	3.2	3.0

Notes: Unit of observation is party-election (N=10,171). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C5. Population and between party pivotal probability, municipality-party level (Model 1, 6th order polynomial of population).

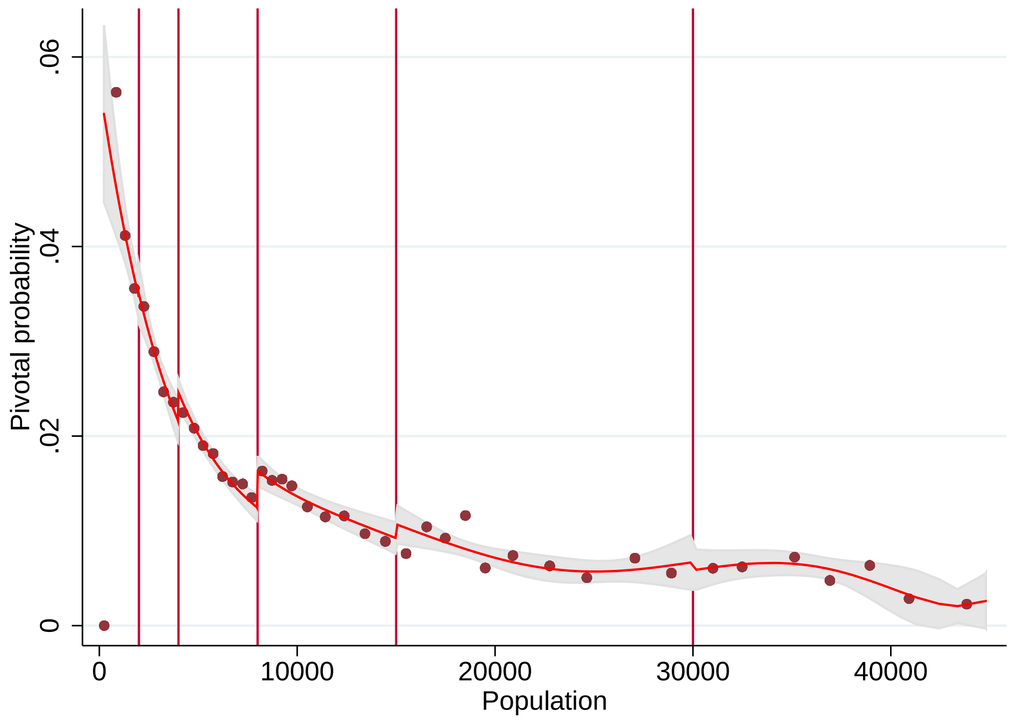


Table C6. IV estimation , first stage for within party pivotal probability, municipality-party level (Model 1).

Dep var: Simulated pivotal probability (within parties)							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.009***	-0.006***	-0.004***	-0.002***	0.000	0.003	0.004**
	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]
pop>4k	-0.002***	-0.001	0.001*	0.003***	0.004***	0.004***	0.003***
	[0.001]	[0.000]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
pop>8k	0.001	0.002***	0.003***	0.002***	0.001**	-0.001*	-0.002**
	[0.001]	[0.000]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
pop>15k	0.004***	0.003***	0.001***	-0.002***	-0.003***	0.001*	0.002**
	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]	[0.001]	[0.001]
pop>30k	0.007***	0.000	-0.002**	0.004***	0.000	-0.003**	0.002**
	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Avg. effect	-0.0015***	-0.0007**	0.0002	0.001**	0.0011**	0.0014**	0.0019***
	[0.0005]	[0.0003]	[0.0004]	[0.0005]	[0.0005]	[0.0006]	[0.0007]
1st stage F	21.1	34.7	16.5	11.9	8.2	4.7	4.5

Notes: Unit of observation is party-election (N=10,171). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C6. Population and within party pivotal probability, municipality-party level (Model 1, 6th order polynomial of population).

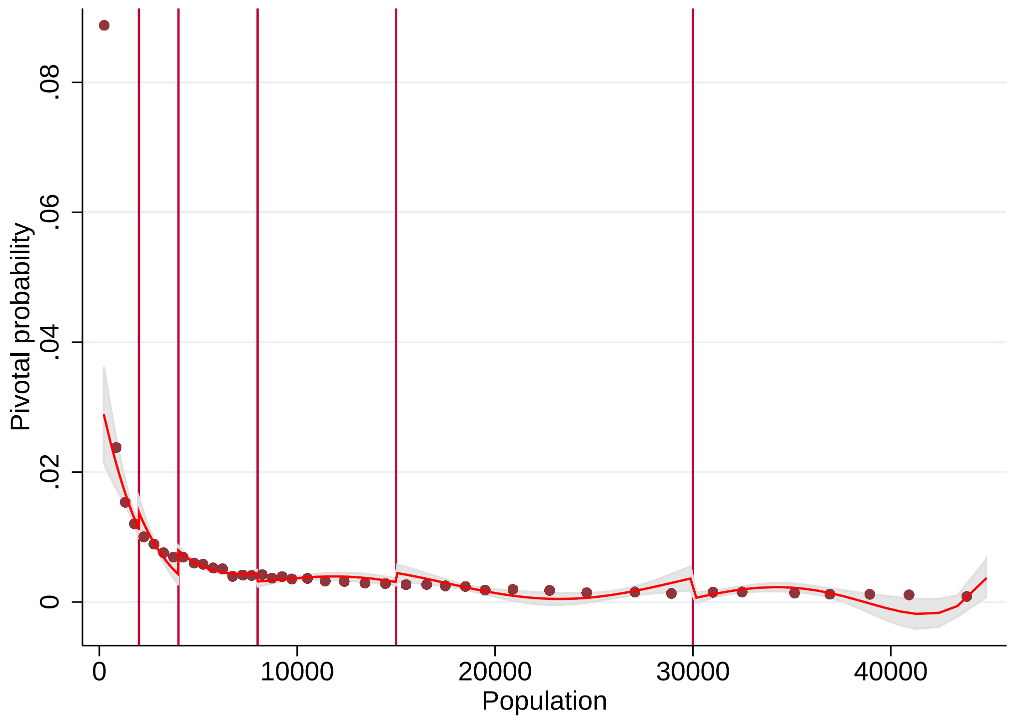


Table C7. IV estimation , first stage for the number of candidates, municipality-party level (Model 1).

Dep var: Candidates							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	1.04***	0.35	0.17	0.32	0.93*	0.92*	0.36
	[0.32]	[0.36]	[0.40]	[0.44]	[0.51]	[0.55]	[0.58]
pop>4k	2.46***	1.33***	1.10**	1.22**	1.48***	1.48***	1.89***
	[0.37]	[0.48]	[0.54]	[0.54]	[0.53]	[0.53]	[0.55]
pop>8k	2.15***	0.72	0.56	0.51	0.06	0.06	0.21
	[0.53]	[0.60]	[0.58]	[0.61]	[0.67]	[0.67]	[0.65]
pop>15k	1.64	0.46	0.67	0.44	0.42	0.41	0.08
	[1.06]	[1.00]	[1.12]	[1.26]	[1.25]	[1.33]	[1.34]
pop>30k	0.99	3.68***	3.85***	4.20***	3.13**	3.14**	1.26
	[1.10]	[1.23]	[1.23]	[1.32]	[1.31]	[1.33]	[1.47]
Avg. effect	1.82***	0.99**	0.88**	0.93**	0.96***	0.96***	0.80**
	[0.37]	[0.40]	[0.39]	[0.37]	[0.37]	[0.37]	[0.38]
1st stage F	19.1	5.0	3.0	3.3	2.8	2.8	2.6

Notes: Unit of observation is party-election (N=10,171). 1^{st} stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C7. Population and the number of candidates, municipality-party level (Model 1, 6^{th} order polynomial of population).

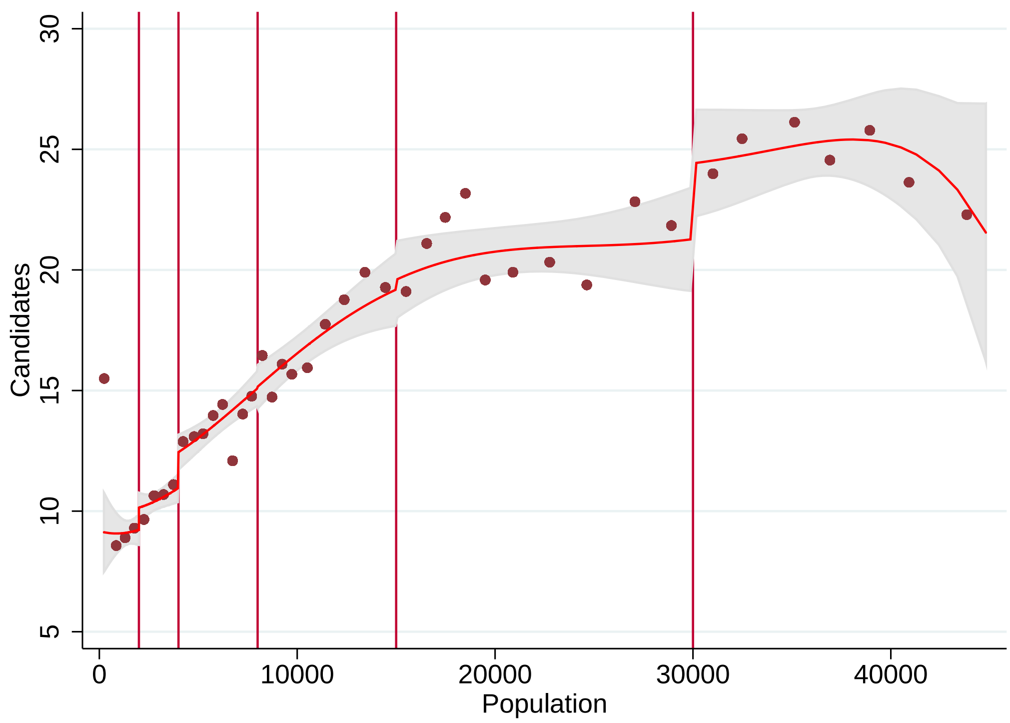


Table C8. IV estimation, first stage for the share of incumbents, municipality-party level (Model 1).

Dep var: Share of incumbents							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.025	-0.023***	-0.015	0.001	0.012	0.024*	0.034**
	[0.018]	[0.009]	[0.009]	[0.010]	[0.012]	[0.013]	[0.014]
pop>4k	-0.030**	-0.008	0.003	0.016	0.021**	0.019*	0.012
	[0.014]	[0.007]	[0.009]	[0.010]	[0.011]	[0.010]	[0.010]
pop>8k	0.023	0.040***	0.048***	0.043***	0.035***	0.026**	0.023**
	[0.016]	[0.009]	[0.010]	[0.010]	[0.010]	[0.011]	[0.011]
pop>15k	0.006	0.025**	0.015	-0.009	-0.01	0.01	0.015
	[0.024]	[0.011]	[0.011]	[0.013]	[0.013]	[0.014]	[0.016]
pop>30k	0.033	0.004	-0.004	0.033*	0.013	-0.005	0.027
	[0.033]	[0.012]	[0.014]	[0.017]	[0.015]	[0.016]	[0.018]
Avg. effect	-0.0059	0.0064	0.0116*	0.0168***	0.0173***	0.0189***	0.0216***
	[0.0101]	[0.0054]	[0.0061]	[0.0062]	[0.0061]	[0.0061]	[0.0064]
1st stage F	3.2	11.3	8.9	6.8	3.7	2.3	2.9

Notes*:* Unit of observation is party-election (N=10,171). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C8. Population and the share of incumbents, municipality-party level (Model 1, 6th order polynomial of population).

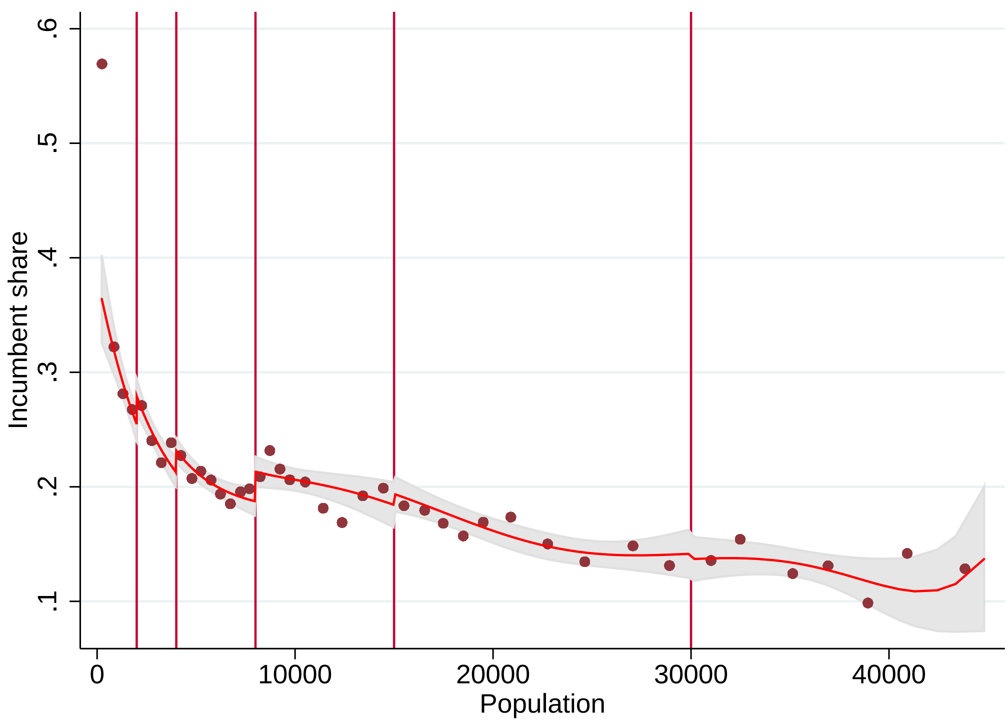


Table C9. IV estimation, first stage for proportionality, municipality-party level (Model 1).

Dep var: Proportionality							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.0086***	-0.0080***	-0.0073***	-0.0061***	-0.0054***	-0.0045***	-0.0042***
	[0.0007]	[0.0007]	[0.0007]	[0.0007]	[0.0008]	[0.0009]	[0.0009]
pop>4k	-0.0060***	-0.0050***	-0.0041***	-0.0031***	-0.0028***	-0.0030***	-0.0033***
	[0.0003]	[0.0004]	[0.0005]	[0.0006]	[0.0006]	[0.0006]	[0.0006]
pop>8k	-0.0036***	-0.0024***	-0.0017***	-0.0021***	-0.0027***	-0.0033***	-0.0034***
	[0.0003]	[0.0003]	[0.0004]	[0.0004]	[0.0004]	[0.0004]	[0.0004]
pop>15k	-0.0017***	-0.0008*	-0.0017***	-0.0035***	-0.0035***	-0.0022***	-0.0019***
	[0.0004]	[0.0005]	[0.0004]	[0.0005]	[0.0005]	[0.0004]	[0.0005]
pop>30k	0.0003	-0.0020***	-0.0027***	0.0002	-0.0013**	-0.0025***	-0.0010*
	[0.0005]	[0.0005]	[0.0005]	[0.0007]	[0.0005]	[0.0006]	[0.0006]
Avg. effect	-0.0049***	-0.0042***	-0.0038***	-0.0034***	-0.0034***	-0.0033***	-0.0031***
	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
1st stage F	284	123	55	38	34	32	27

Notes: Unit of observation is party-election (N=10,171). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C9. Population and proportionality, municipality-party level (Model 1, 6th order polynomial of population).

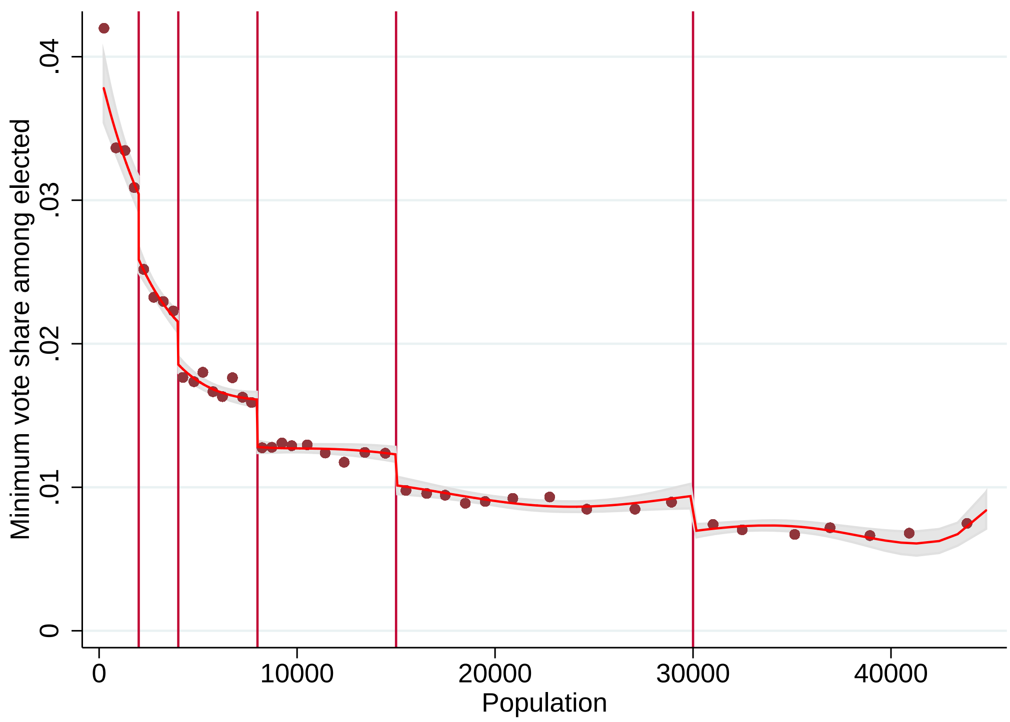


Table C10. IV estimation, first stage for political competition, municipality-party level (Model 1).

Dep var: Political competition							
	Order of polynomial of pop						
Threshold	1st	2nd	3rd	4th	5th	6th	7th
pop>2k	-0.00096***	-0.00092***	-0.00085***	-0.00081***	-0.00084***	-0.00086**	-0.00086**
	[0.00025]	[0.00025]	[0.00026]	[0.00028]	[0.00031]	[0.00034]	[0.00036]
pop>4k	-0.00090***	-0.00084***	-0.00074***	-0.00071***	-0.00072***	-0.00072***	-0.00072***
	[0.00014]	[0.00016]	[0.00019]	[0.00023]	[0.00024]	[0.00024]	[0.00024]
pop>8k	-0.00052***	-0.00045***	-0.00038**	-0.00039**	-0.00037**	-0.00036*	-0.00036*
	[0.00011]	[0.00015]	[0.00018]	[0.00017]	[0.00017]	[0.00020]	[0.00020]
pop>15k	-0.00055***	-0.00050***	-0.00059***	-0.00064***	-0.00064***	-0.00067***	-0.00067***
	[0.00015]	[0.00019]	[0.00017]	[0.00020]	[0.00020]	[0.00020]	[0.00021]
pop>30k	-0.0002	-0.00035**	-0.00041**	-0.00033	-0.00028	-0.00026	-0.00027
	[0.00018]	[0.00017]	[0.00019]	[0.00025]	[0.00021]	[0.00026]	[0.00025]
Avg. effect	-0.00072***	-0.00068***	-0.00063***	-0.00062***	-0.00062***	-0.00062***	-0.00062***
	0.00008	0.0001	0.00012	0.00013	0.00014	0.00014	0.00014
1st stage F	37.2	13.8	8.6	6.4	6.2	5.7	5.2

Notes: Unit of observation is party-election (N=10,171). 1st stage F is the F test of the excluded instruments. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.

Graph C10. Population and political competition, municipality-party level (Model 1, 6th order polynomial of population).

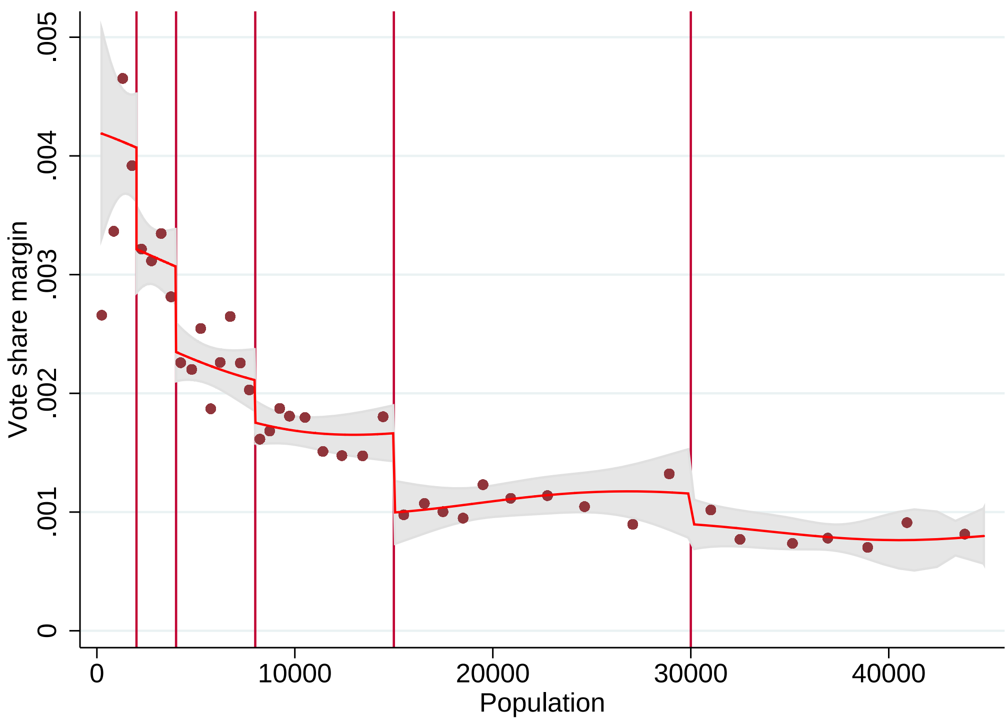


Table C11. IV estimation, second stage with nonlinear confounders, municipality-party level.

Dep var: Party turnout in municipal elections				
	Order of polynomial of pop			
	3rd	4th	5th	6th
Candidates	0.004	0.006	0.008	0.007
	[0.005]	[0.005]	[0.007]	[0.009]
Candidates^2	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Within party pivotality	6.606***	6.573***	6.118***	6.622***
	[0.861]	[1.036]	[1.314]	[2.380]
Share of incumbents	0.048	0.173	0.247	0.336
	[0.608]	[0.498]	[0.566]	[0.618]
(Share of incumbents)^2	-0.213	-0.414	-0.541	-0.651
	[0.963]	[0.783]	[0.857]	[0.875]
Within party pivotality	7.235***	7.645***	7.839***	8.259***
	[1.464]	[1.520]	[1.764]	[2.120]
Proportionality	-2.728	3.412	3.143	2.82
	[3.778]	[2.616]	[2.444]	[2.552]
Proportionality^2	57.903	-63.326	-51.231	-41.455
	[80.025]	[51.966]	[48.515]	[51.526]
Within party pivotality	3.79	7.935***	8.021***	8.173***
	[2.987]	[2.087]	[2.066]	[2.267]
Political competition	13.048	22.708	20.283	11.368
	[17.132]	[17.658]	[34.850]	[19.120]
(Political competition)^2	-825.825	-1698.355	-1535.205	-648.009
	[1323.135]	[1478.083]	[3090.007]	[1842.541]
Within party pivotality	5.702***	6.129***	6.183***	6.965***
	[1.284]	[1.274]	[1.417]	[1.638]

Notes*:* Unit of observation is party-election. Standard errors are in brackets (clustered at municipality level). Significance is denoted by asterisks: * p<0.1, ** p<0.05, *** p<0.01.