VATT Working Papers 59

Does Regression Discontinuity Design Work? Evidence from Random Election Outcomes

Ari Hyytinen
Jaakko Meriläinen
Tuukka Saarimaa
Otto Toivanen
Janne Tukiainen

VATT INSTITUTE FOR ECONOMIC RESEARCH

VATT WORKING PAPERS

59

Does Regression Discontinuity Design Work? Evidence from Random Election Outcomes

Ari Hyytinen
Jaakko Meriläinen
Tuukka Saarimaa
Otto Toivanen
Janne Tukiainen

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Helsinki 2015

This version (September 21, 2015) replaces the November 20, 2014 version.

ISBN 978-952-274-131-8 (PDF)

ISSN 1798-0291 (PDF)

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

Helsinki, September 2015

Does Regression Discontinuity Design Work?

Evidence from Random Election Outcomes

Ari Hyytinen, Jaakko Meriläinen, Tuukka Saarimaa, Otto Toivanen and Janne Tukiainen*

21 September 2015

Abstract

We use elections data in which a large number of ties in vote counts between candidates are resolved via lottery. We benchmark non-experimental RDD estimates of personal incumbency advantage against the estimate produced by this experiment that takes place exactly at the cutoff. The experimental estimate suggests that there is no personal incumbency advantage. In contrast, standard local linear RDD estimates suggest a moderate and statistically significant effect. Bias-correction and under-smoothing procedures however bring the RDD estimate(s) in line with the experimental estimate. Therefore, careful implementation of RDD can meet the replication standard in the context of close elections.

JEL Codes: C21, C52, D72.

Keywords: Close elections, experiment, incumbency advantage, regression discontinuity design

---

* Hyytinen: University of Jyväskylä, P.O. Box 35, 40014 University of Jyväskylä, Finland, ari.hyytinen@jyu.fi, Tel +358400777086; Meriläinen: Institute for International Economic Studies, Stockholm University, SE-10691, Stockholm, Sweden, jaakko.merilainen@iies.su.se, Tel +468163174; Saarimaa: VATT Institute for Economic research, Arkadiankatu 7, Helsinki, FI-00101, tuukka.saarimaa@vatt.fi, Tel +358295519444; Toivanen: Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium, otto.toivanen@econ.kuleuven.be, Tel +3216326753; Correspondence should be addressed to Tukiainen: VATT Institute for Economic research, Arkadiankatu 7, Helsinki, FI-00101, janne.tukiainen@vatt.fi, Tel +358295519451. We thank Manuel Bagues, Christine Benesch, Jon Fiva, Dominik Hangartner, Kaisa Kotakorpi, Benoit Le Maux, Yao Pan, Torsten Persson, Miikka Rokkanen, Riikka Savolainen and conference and seminar participants at EPCS, EPSA, LSE, Rennes 1, Stockholm University, Warwick and VATT for useful discussions and comments.

1 Introduction

A non-experimental empirical tool meets a very important quality standard if it can reproduce the results from a randomized experiment (LaLonde 1986, see also Fraker and Maynard 1987). In a regression discontinuity design (RDD), individuals are assigned dichotomously to a treatment if they cross a given cutoff of an observable and continuous forcing variable, whereas those who fail to cross the cutoff form the control group (Thistlethwaite and Campbell 1960, Lee 2008, Imbens and Lemieux 2008). Provided that the conditional expectation of the potential outcome is continuous in the forcing variable at the cutoff, correctly approximating the regression function above and below the cutoff and then comparing the values of the regression function for the treated and control groups at the cutoff gives the RDD treatment effect. We study whether standard implementation of RDD can reproduce an experimental estimate that we obtain by utilizing data from electoral ties between two or more candidates in recent Finnish municipal elections.[1]

The unique feature of our data is that the ties were resolved via a lottery and that the random assignment occurs *right at the cutoff*. This means that if RDD works, it should deliver a treatment effect estimate that exactly matches our experimental estimate. To explore whether this is the case, we utilize a data set that includes nearly 200 000 candidates that run for a seat in municipal councils in local Finnish elections every fourth year during 1996–2012. The elections were organized in a shared institutional environment and allow us to study whether there is a personal

[1] Investigating the performance of RDD in an electoral setting seems particularly important, as numerous applications of RDD in economics and political science have used close elections to estimate the effects of electoral results on a variety of economic and political outcomes (see, e.g., Lee et al. 2004, DiNardo and Lee 2004, Ferreira and Gyourko 2009, Gerber and Hopkins 2011 and Folke and Snyder 2012).

incumbency advantage, i.e., extra electoral support that an incumbent politician of a given party enjoys when she runs for re-election, relative to her being a non-incumbent candidate from the same party and constituency (see, e.g., Erikson and Titiunik 2015). Our experimental estimate of the personal incumbency advantage is estimated from data on 1351 candidates for whom the (previous) electoral outcome was determined via the random seat assignment.[2] The experimental estimate provides no evidence for a personal incumbency advantage: It is close to zero and quite precisely estimated. As we explain later, this null finding is neither surprising nor in conflict with the prior evidence when interpreted in the context of local proportional representation (PR) elections.

Since the seminal paper on RDD by Hahn, Todd and van der Klaauw (2001), non-parametric local linear regression has been used widely in applied work to approximate the regression function near the cutoff (see also Porter 2003). The standard implementation of the local linear regression calls for estimating a linear regression function separately above and below the cutoff in a neighborhood that is determined by a choice of a bandwidth. The bandwidth defines how close to the cutoff the estimation is implemented and various methods have been proposed for selecting it (e.g., Ludwig and Miller 2007, Imbens and Kalyanaraman 2012, Calonico et al. 2014a). For example, a mean-squared-error (MSE) optimal bandwidth trades off the bias due to not getting the functional form completely right for wider bandwidths with the increased variance of the estimate for narrower bandwidths. We find that when

[2] Use of lotteries to solve electoral ties is not unique to Finland. For example, some US state elections and many US local elections have used lottery-based rules to break ties in elections (see, e.g., UPI 14.7.2014, The Atlantic 19.11.2012, and Stone 2011). Lotteries have been used to determine the winner in case of ties also in the Philippines (Time 17.5.2013), in India (The Telegraph India 7.2.2014), in Norway as well as in Canada and the UK (http://en.wikipedia.org/wiki/Coin_flipping#Politics).

RDD is applied to our elections data and implemented in the standard fashion using the local linear approximation and the often used (cross-validated or MSE optimal) bandwidths, the estimates suggest a statistically significant positive personal incumbency advantage.

The disparity between the experimental and RDD estimates may at first glance be seen as a piece of bad news for RDD. A major finding of this paper is, however, that the disparity is actually in line with the recent theoretical econometric work on RDD (see, e.g., Calonico et al. 2014a, Card et al. 2014). There are two primary reasons why the experimental estimate and the estimate that our standard implementation of a close election RDD generates might not match. First, the key RDD assumption that the conditional expectation of the potential outcome is continuous at the cutoff may be violated. A typical close election RDD analysis relies on the presumption that the political candidates, who get elected at the margin, win because of random factors (chance) and are therefore, on average, comparable to the barely losing candidates. Second, it is possible that the standard implementation of RDD, using the local linear approximation and the optimal bandwidths, is deficient. We find no clear signs of the key RDD assumption being violated using covariate balance checks. However, we can provide novel evidence that the standard implementation of RDD is deficient, precisely in the way(s) that the recent econometric work on RDD suggest.

Our evidence is the following: First, all nonparametric methods may produce biased estimates if the parametric specification is not a good approximation of the true regression function within the bandwidth. If the bias is relatively large when for example the local linear regression is used, the MSE optimal bandwidth does not provide a reliable basis for inference, as it then produces confidence intervals that have incorrect asymptotic coverage (see, e.g., Imbens and Lemieux 2008, Calonico et al. 2014a, Card et al. 2014). We find that the often used solution to this problem (see, e.g., Imbens and Lemieux 2008), the under-smoothing procedure of using smaller (than MSE-optimal) bandwidths, works as predicted. It allows us to recover the experimental estimate in the sense that with under-smoothing, the null hypothesis of no personal incumbency advantage is no longer rejected. Second, we find that the curvature of the regression function matters. Using richer local polynomial specifications within the bandwidth *optimized for the linear specification* can eliminate the bias. However, when higher order local polynomials are used and the bandwidths are accordingly optimized, the bias typically remains. This implies that in our case, MSE optimal bandwidths may be problematic more generally (somewhat in contrast with the findings of Card et al. 2014). Third, the bias-correction and robust inference method introduced recently by Calonico et al. (2014a) works well, too, provided that one does not allow for a too wide pilot bandwidth.

In sum, our findings provide a word of caution to practitioners, since the local linear regression with optimal bandwidths, which is often used in applied work, appears to lead to an incorrect conclusion. However, standard bias-correction and under-smoothing procedures bring the RDD estimate(s) in line with the experimental estimate. Thus, our results show that careful implementation of RDD can meet the replication standard in the context of close elections.

Our findings also bear on three other strands of the literature. First, it has been argued that in close elections, the conditions for randomization (and covariate balance) around the cutoff do not necessarily hold, especially in post-World War II U.S. House elections (Snyder 2005, Caughey and Sekhon 2011, and Grimmer et al. 2012).

Eggers et al. (2015) convincingly challenge this conclusion.[3] We contribute to this ongoing debate by showing whether and when the close election RDD, as it is currently often implemented, is capable of replicating the experimental estimate. Second, there is an emerging literature on within-study comparisons of RDDs (see e.g., Black et al. 2007, Cook and Wong 2008, Green et al. 2009, and Shadish et al. 2011, Wing and Cook 2013) that explores how the performance of RDD depends on the context in which it is used.[4] We contribute to this literature, because no within-study comparison on how a close election RDD performs appears to exist and because typically the randomized experiment does not take place exactly at the cutoff (meaning that it may identify a treatment effect different from the one that RDD targets).[5] Finally, as we discuss later in the paper, our findings add to the accumulating evidence on limited personal incumbency advantage in PR systems (see, e.g., Dahlgaard 2013, Golden and Picci 2015, Lundqvist 2011, Kotakorpi et al. 2013 and Redmond and Regan 2015).

The rest of this paper is organized as follows: In Section 2, we describe the institutional environment and our data. The experimental and non-experimental results are reported and compared in Section 3. We discuss the validity and robustness of our findings in Section 4. Section 5 concludes. A large number of additional analyses are reported in an online appendix that supplements this paper.

---

[3] The criticism on the close election RDD builds on the observation that outright fraud, legal and political manipulation and/or sorting of higher quality or better positioned candidates may naturally characterize close elections. However, Eggers et al. (2015) show that post-World War II U.S. House elections are a special case and that there is no imbalance in any of the other elections that their dataset on 40 000 close political races cover. Eggers et al. do not find evidence for imbalance in any other elections either in the U.S. or in other countries, nor in the U.S. House at times other than the post-World War II era. On the other hand, Vogl (2014) shows that in U.S. city elections in the South, black candidates seem to have an advantage in close races compared to white candidates.

[4] See also Cook et al. (2008). The terminology that is used to refer to these types of comparative studies is not entirely settled across the various disciplines. The current view of this literature appears to be that RDD is able to reproduce - or at least to approximate - experimental results in most, but not in all, settings (see in particular Cook et al. 2008 and Shadish et al. 2011).

[5] Black et al. (2007) seem to come closest, because their experiment targets a population within a small bandwidth around the cutoff. See also Cook and Wong (2008).

2 Institutional context and data

2.1 Institutional environment

Finland has a two-tier system of government, consisting of a central government and a large number of municipalities at the local level.[6] Finnish municipalities have extensive tasks and considerable fiscal autonomy. In addition to the usual local public goods and services, municipalities are responsible for providing most of social and health care services and primary and secondary education. Municipalities are therefore of considerable importance to the whole economy.[7]

Municipalities are governed by municipality councils. The council is by far the most important political actor in municipal decision making. For example, mayors are public officials chosen by the councils and can exercise only partial executive power. Moreover, municipal boards (i.e., cabinets) have only a preparatory role. The presentation in the boards follows the same proportional political distribution as the presentation in the council.

Municipal elections are held simultaneously in all municipalities. All municipalities have one electoral district. The elections in our data were held on the fourth Sunday of October in 1996, 2000, 2004, 2008 and 2012. The four year council term starts at the beginning of the following year. The seat allocation is based on proportional representation, using the open-list D'Hondt election rule. There are three (1996-2004 elections) or four (2008-2012 elections) major parties, which dominate the political landscape of both the municipal and national elections, as well as four other parties

[6] In 1996, Finland had 436 municipalities and in 2012, 304.

[7] Municipalities employ around 20 percent of the total workforce. The most important revenue sources of the Finnish municipalities are local income taxes, operating revenues, such as fees, and funding from the central government.

that are active both locally and nationally. Moreover, some purely local independent political groups exist. In the elections, each voter casts a single vote to a single candidate. One cannot vote for a party without specifying a candidate. In this setting, voters (as opposed to parties) decide which candidates are eventually elected from a given list, because the number of votes that a candidate gets determines the candidate's rank on her party's list.

The total number of votes over the candidates of a given party list determines the votes for each party. The parties' votes determine how many seats each party gets. The procedure is as follows: First, a comparison index, which equals the total number of votes cast to a party list divided by the order (number) of a candidate on the list, is calculated for all the candidates of all the parties. The candidates are then ranked according to the index and all those who rank higher than $(S+1)^{\text{th}}$ (S being the number of council seats) get a seat.

An important feature of this election system is that in many cases, there is an exact tie in the number of votes at the margin where the last available seat for a given party list is allocated. This means that within a party, the rank of two or more candidates has to be randomly decided. For example, it is possible that a party gets k seats in the council and that the k^{th} and $(k+1)^{\text{th}}$ ranked candidates of the party receive exactly the same number of votes. For them, the comparison index is the same. The applicable Finnish law dictates that in this case, the winner of the marginal (k^{th}) seat has to be decided using a randomization device. Typically, the seat is literally allocated by drawing a ticket (name) from a hat. The procedure appears to be very elementary: One of the (typically female) members of the municipal election committee wears a blindfold and draws the ticket in the presence of the entire committee.[8] While we have not run an experiment nor implemented a randomized controlled trial, we can use the outcomes from these lotteries to generate an experimental treatment effect estimate for the effect of incumbency status on electoral support.

It is also possible that two (or more) candidates from *different parties* face a tie for a marginal seat. However, within party ties are much more common in practice. Therefore, we do not analyze ties between candidates from different parties. Besides resulting in a larger sample in which the candidates that had a tie, there are three additional reasons to focus on the within party ties. First, using the within party ties allows for a simpler implementation of RDD, as we do not have to worry about discontinuities and possible party-level incumbency effects that are related to party lines.[9] Second, focusing on the within party dimension also allows a cleaner identification of the personal incumbency effect, net of the party incumbency effect. Third, the use of only within party ties increases the comparability of our RDD analysis, which uses multi-party PR elections data, with the prior studies that use data from two-party (majoritarian) systems. This is so as within a party list, the Finnish elections follow the N-past-the-post rule.

2.2 Data

Our data originate from several sources. The first source is election data which consist of candidate-level information on the candidates' age, gender, party affiliation, the number of votes they received, their election outcomes (elected status) and the

[8] See e.g. an article in one of the major Finnish tabloids, Iltasanomat, on 12.4.2011.

[9] See Folke (2014) for the complications that multi-party-systems generate and Snyder et al. (2015) on issues with partisan imbalance in RDD studies.

possible incumbency status.[10] These data were linked to data from KEVA (formerly known as the Local Government Pensions Institution) to identify municipal workers, and to Statistics Finland's data on the candidates' education, occupation and socio-economic status. We further added income data come from the Finnish tax authority. Finally, we matched the candidate-level data with Statistics Finland's data on municipal characteristics.[11]

We have data on 198 121 candidates from elections held in years 1996, 2000, 2004, 2008 and 2012.[12] Summary statistics (reported in Appendix A) show that the elected candidates differ substantially from those who are not elected: They have higher income and more often a university degree and are less often unemployed. The difference is particularly striking when we look at the incumbency status: 58% of the elected candidates were incumbents, whereas only 6% of those who were not elected were incumbents.

3 Main results

3.1 Experimental estimates

In this section, we estimate the magnitude of the personal incumbency advantage using the data from the random election outcomes. We define this added electoral support as the treatment effect of getting elected today on the probability of getting

[10] These election data are publicly available from the Statistics Finland. Our dataset was provided by the Ministry of Justice, as we need the social security numbers of the candidates in order to be able to link the election data to other data sources.
[11] The candidate-level demographic and occupation data usually refers to the election year, with the exception that occupation data from 1995 (2011) is matched to 1996 (2012) elections data.
[12] Two further observations on the data are in order: First, to be careful, we omit all data (about 150 candidates) from one election year (2004) in one municipality, because of a mistake in the elected status of one candidate. Mistake is apparently due to one elected candidate being disqualified later. Second, the data on the candidates running in 2012 are only used to calculate the outcome variables.

elected in the next election. We measure the event of getting elected today by a binary indicator, E_{it}, which takes value of one if candidate i was elected in election year t and is zero otherwise. Our main outcome is a binary variable, $Y_{i,t+1}$, which equals one if candidate i is elected in the next election year t+1 and is zero otherwise.

In elections between 1996 and 2008, 1351 candidates had a tie within their party lists for the last seat(s), i.e. at the margin which determines whether or not the candidates get a seat.[13] In these cases, a lottery was used to determine who got elected. This implies that E_{it} was randomly assigned in our *lottery sample*, i.e. among the candidates that had a tie.

Covariance balance tests for the lottery sample

Was the randomization required by the law conducted correctly and fairly? To address this question, we study whether candidates' characteristics balance between the treatment (randomly elected) and the control group (randomly not elected) within the lottery sample. The results are reported in Table 1. The differences are statistically insignificant and small in magnitude. These findings support the view that E_{it} is randomly determined in the lottery sample.[14]

[13] In addition, there were 202 ties in 2012. We do not include them in the lottery sample, because we cannot yet observe subsequent election outcomes for these candidates. When we include these ties in the balancing tests, the results do not change. Notice also that a tie may involve more than two candidates and more than one seat. For example, three candidates can tie for two seats.

[14] The candidates' party affiliations and municipal characteristics should be balanced by design, because we analyze lotteries within the party lists. The corresponding balancing tests (reported in Appendix B) confirm this.

Table 1. Covariate balance tests for the lottery sample

Individual characteristics

Variable	Elected (N = 671) N	Mean	Std. Dev.	Not elected (N = 680) N	Mean	Std. Dev.	Difference
Vote share (from all votes)	671	1.54	0.69	680	1.53	0.67	0.00
Vote share (from party votes)	671	6.49	6.34	680	6.49	6.31	0.00
Number of votes	671	41	39	680	41	38	0
Female	671	0.39	0.49	680	0.38	0.49	0.01
Age	671	45.42	11.87	680	45.69	11.54	-0.27
Incumbent	671	0.29	0.45	680	0.31	0.46	-0.02
Municipal employee	671	0.24	0.43	680	0.25	0.44	-0.01
Wage income	429	19190	12098	431	20207	12940	-1017
Capital income	429	2641	11420	431	2022	16683	618
High professional	671	0.18	0.38	680	0.18	0.38	0.00
Entrepreneur	671	0.24	0.43	680	0.24	0.43	0.00
Student	671	0.02	0.15	680	0.03	0.16	0.00
Unemployed	671	0.06	0.24	680	0.05	0.22	0.01
University degree	537	0.13	0.34	545	0.13	0.34	0.00

Notes: Difference in means has been tested using t test adjusted for clustering at municipality level. Sample includes only candidates running in 1996-2008 elections. Income data are not available for 2012 elections, and in 1996 elections they are available only for candidates who run also in 2000, 2004 and 2008 elections. Income and income per capita are expressed in euros.

Experimental estimate for the personal incumbency effect

Is there a personal incumbency effect? Before we can answer this question, we have to point out that a subsequent electoral outcome is observed for 820 out of the 1351 candidates who participated in the lottery between 1996 and 2008, because they reran in a subsequent election. We do not know what happened to those who decided not to run again. This attrition is a possible source of concern for us, because the decision not to rerun may mirror for example the candidates' expected performance. If it does, analyses based on the selected sample, from which those who did not rerun are excluded, would not provide as us with the correct treatment effect. Rerunning is an (endogenous) outcome variable and we therefore cannot condition on it, unless the treatment has no effect on the likelihood of rerunning. Relying on such an assumption would be neither harmless nor conservative.[15] Our baseline results therefore refer to the entire lottery sample. This means that we code our main outcome variable so that it is equal to one if the candidate is elected in the next election, and is set to zero if the candidate is not elected or does not rerun.

The raw data suggest that there is no personal incumbency advantage; the fraction of candidates who get elected in election year t+1 conditional on not winning the lottery in election year t is 0.325, whereas the same fraction conditional on winning the tie lottery is 0.329. The difference between the two fractions is small (≈ 0.004). To quantify the difference more formally, we regress $Y_{i,t+1}$ on E_{it} using OLS and the sample of candidates who faced within-party ties. Because E_{it} is randomly assigned in the lottery sample, its coefficient is the average treatment effect (ATE). Note that due to the way the lottery sample is constructed, this ATE is estimated *precisely at the cutoff point* of political support which determines whether or not a candidate gets elected. It is therefore an ideal benchmark for the non-experimental RDD estimate, because the sharp RDD targets exactly the same treatment effect.

Table 2 reports our experimental estimates of the personal incumbency effect. In the leftmost column, $Y_{i,t+1}$ is regressed on E_{it} and a constant, using OLS. In the remaining columns we report the OLS results from a set of specifications that include control variables and fixed effects. Three main findings emerge. First, there is no evidence of a personal incumbency advantage: the estimated effect is close to zero. Second, the coefficient of E_{it} is relatively stable across the columns and is thus not correlated with the added controls or fixed effects. This further supports the view that

[15] Uppal (2010) and Klašnja and Titiunik (2015) report the results for a sample that includes all candidates and for a sample that only includes those who rerun. De Magalhaes (2014) argues in favor of including all the candidates.

E_{it} is random. Third, the estimates are relatively precise, as the standard errors are around 0.03.

Table 2. Experimental estimates of the personal incumbency advantage (Lottery sample)

Outcome: Elected next election

	(1)	(2)	(3)	(4)
Elected	0.0044	0.0009	-0.0101	-0.0099
	(0.0247)	(0.0249)	(0.0291)	(0.0332)
N	1351	1351	1351	1351
R^2	0.00	0.03	0.28	0.44
Controls	No	Yes	Yes	Yes
Municipality fixed effects	No	No	Yes	No
Municipality-year fixed effects	No	No	No	Yes

Notes: Only actual lotteries are included in the regressions. Set of controls includes age, gender, party affiliation, socio-economic status and incumbency status of a candidate, and total number of votes. Some specifications include also municipality or municipality-year fixed effects. Standard errors shown in parentheses are clustered at municipality level. Unit of observation is a candidate *i* at year *t*.

We have considered the robustness of the experimental estimate(s) in various ways. First, 0.9% of the candidates run in another municipality in the next elections. For Table 2, they were coded as not rerunning. The results (not reported) are robust to accounting for their reelection in municipalities other than the one where they faced the tie in the previous election. Second, 118 of the candidates that lost in the lottery became council members during the term of the council because some council member stepped down. This may lead to a contamination bias. We have therefore recalculated our experimental estimates after excluding these candidates from the lottery sample. The experimental estimates do not practically change. Third, we have considered the vote share in the next election as an alternative outcome. While more problematic, we follow the same practice with this alternative outcome as above and set it to zero if the candidate did not rerun in the next election. The results (reported in

Appendix B) show that E_{it} has no impact on the alternative outcome.[16] Fourth, we considered small and large elections separately (see Appendix B), but found no evidence of a personal incumbency advantage. Finally, we get an experimental estimate close to zero (for both the elected next election and vote share next election outcomes) if we use a trimmed lottery sample that only includes the rerunners (reported in Appendix B).

Discussion of the experimental estimate

The personal incumbency advantage refers to the added electoral support that an incumbent politician of a given party enjoys when she runs for re-election, relative to her being a non-incumbent candidate from the same party and constituency.[17] Such advantage could stem from various sources, such as from having been able to serve the constituency well, having enjoyed greater public visibility while holding the office, improved candidate quality (through learning while in power), reduced competitor quality (due to a "scare-off" effect; see Cox and Katz 1996, Erikson and Titiunik 2015), and the desire of voters to disproportionately support politicians with past electoral success ("winners"). The earlier (mostly U.S.) evidence suggests that the existence of an incumbent personal advantage in two-party systems is largely beyond question (see, e.g., Erikson and Titiunik 2015, and the references therein). It is clear that the size of the advantage may nevertheless vary and be context specific; see e.g. Desposato and Petrocik (2003), Grimmer et al. (2012), Uppal (2009) and Klašnja and Titiunik

[16] We have also checked that if the event of rerunning in the next election is used as the dependent variable, the experimental estimate is small and statistically not significant.

[17] This politician-level electoral gain is not the same as the advantage that a party enjoys from being the incumbent party in an election (Gelman and King 1990, Lee 2008, Erikson and Titiunik 2015). The party incumbency advantage measures the electoral gain that a candidate enjoys when she is from the incumbent party, independently of whether she is an incumbent politician or not. Following Lee (2008), most of the earlier RDD analyses refer to the party advantage (e.g., Broockman 2009, Butler 2009, Uppal 2010, Caughey and Sekhon 2011, Trounstine 2011), even though it seems that what is being actually estimated is a combination of personal and party advantage (e.g., Fowler and Hall 2014).

(2015), who find evidence of party-level disadvantage in systems characterized by weak parties.

In our view, the null finding of no personal incumbency advantage is neither surprising nor in conflict with the prior evidence, for three reasons: First, we are looking at personal incumbency advantage in the rather special context of small local PR elections. It is possible that in this context, the randomized political victories take place at a relatively unimportant margin. For example, such a political win does not, per se, typically lead to a visible position in media or a prominent position in the wider political landscape. Perhaps being the last elected candidate of a party in the Finnish municipal elections conveys limited opportunities to serve one's constituency or to improve one's quality as a candidate through learning-by-doing.[18] What's more, it is certainly plausible that getting the last seat by a lottery does not work to scare off good competitors in the subsequent elections. Such a political victory provides the voters with a limited opportunity to picture and support the candidate as a political winner. It is thus not surprising if there is no personal incumbency *advantage* at the margin that we study.

Second, it is important to recall that most of the recent evidence on the positive and large incumbency effects refers to the advantage that *parties* get from being incumbent. For example, in Lee (2008), the estimated incumbency effect refers to the margin at which all the political power shifts from one party to another. In contrast, the random election outcomes in our data allow recovering a treatment effect estimate for the personal incumbency advantage that specifically *excludes* the party

[18] Similarly, being the first non-elected candidate of a party may convey some opportunities to participate in the municipal decision making, e.g., by serving as a deputy councilor or as a member in municipal committees.

effect, because it is estimated from within-party variation in the incumbency status. Moreover, the existing studies that look at *personal* incumbency advantage in the PR systems of developed countries find typically only modest or no incumbency effects (Dahlgaard 2013, Golden and Picci 2015, Lundqvist 2011 and Kotakorpi et al. 2013).

Third, our null finding is not as much in conflict even with the prior U.S. between party evidence as it first appears, because – as we will show in the next section – we also would claim to have found a positive, moderate and statistically significant incumbency effect if we had just relied on the standard implementation of RDD, using local linear regression and optimal bandwidths. As we illustrate below, we also would have reported significant and large effects if we had used the parametric global polynomial RDD specifications (like the earlier RDD literature often did; see e.g. Lee 2008).

3.2 Non-experimental estimates

Implementing RDD for PR elections

The identification of the treatment parameter using RDD relies on the assumption that both the expected electoral support for a candidate who is not an incumbent, given her electoral support in the previous election, and the expected electoral support for a candidate who is an incumbent, given her electoral support in the previous election, are continuous in the share/number of votes at the cutoff (Hahn et al. 2001, see also Lee 2008 and Imbens and Lemieux 2008). A special feature of a PR election system is that it is much harder than in a two-party majoritarian system for a candidate or a party to accurately predict the precise location of the cutoff that determines who gets elected from a given party-list. The reason for this is that the number of seats allocated to the party also depends on the election outcome of the other parties.[19] This makes it more likely that the forcing variable cannot be perfectly manipulated and thus that the key RDD assumption is satisfied.

Our forcing variable is constructed as follows. We measure closeness *within a party list* in order to focus on within-party variation in the incumbency status and, as we explained earlier, to abstract from multi-party issues and potential party effects in PR systems (see Folke 2014). To this end, we calculate for each ordered party list the pivotal number of votes as the average of the number of votes among the first non-elected candidate(s) and the number of votes among the last elected candidate(s). A candidate's distance from getting elected is then the number of votes she received minus the pivotal number of votes for her list (party). We normalize this distance by dividing it by the number of votes that the party list got in total and then multiply it by 100.[20] This normalized distance is our forcing variable v_{it}.

Four observations about our forcing variable are in order: First, it measures closeness within a party list in vote shares. It is thus in line with the existing measures for majoritarian systems. As usual, all candidates with $v_{it} > 0$ get elected, whereas those with $v_{it} < 0$ are not elected. All those candidates for whom $v_{it} = 0$ face a tie and get a seat if they win in the lottery. Second, the histogram of the forcing variable

[19] In PR multi-party election systems, such as those used in the Nordic countries, the location of the RDD cutoff(s) and the nature of the associated forcing variable are much more elusive than in majority systems. As Folke (2014) stresses in the seminal implementation of RDD in such a system, the main issue is that the seats allocated to a given party are not only a function of their own vote share, but rather depend on the entire vote vector and often in a nonlinear way. Freier and Odendahl (2012), Fiva et al. (2013) and Kotakorpi et al. (2013) provide subsequent methodological contributions. Dahlgaard (2013), Golden and Picci (2015), Lundqvist (2011) and Kotakorpi et al. (2013) study quasi-randomization that takes place within parties in a PR system using a very similar approach as we do.

[20] This way of defining the forcing variable means that all those party lists from which no candidates or all candidates got elected are dropped out from the analysis. In total, this means omitting about 6000 candidate-election observations. This corresponds to roughly 3% of the observations in the elections organized between 1996 and 2012.

close to the cutoff (reported in Appendix C) shows that there are observations close to the cutoff and thus that some, but not extensive, extrapolation is being done in the estimation of the RDD treatment effect. Third, the assumption of having a continuous forcing variable is not at odds with our forcing variable. For example, among the 100 closest observations to the cutoff, 92 observations obtain a unique value of v_{it} and there are 4 pairs for which the value is the same within each pair. Finally, our normalized forcing variable and the (potential alternative) forcing variable based on the absolute number of votes operate on a very different scale, but they are correlated (their pairwise correlation is in our data 0.34, p-value < 0.001; see also Appendix C).[21] Moreover, as we discuss later in connection with robustness tests, our RDD results are robust to using alternative definitions of the forcing variable.

The function of the forcing variable is estimated separately for both sides of the cutoff. Choice of the bandwidth determines the subsample near the cutoff to which the function of the forcing variable is fitted and from which the treatment effect is effectively estimated (Imbens and Lemieux 2008, Lee 2008, Lee and Lemieux 2010). For our baseline RDD, we use a triangular kernel and the optimal bandwidths of Imbens and Kalyanaraman (2012, IK), Calonico et al. (2014a, CCT) and Ludwig and Miller (2007, LM).[22]

RDD estimations: Baseline results

Table 3 reports our baseline RDD estimation results. We report results from a sharp RDD for the subsample of candidates that excludes the randomized candidates,

[21] In large elections, it is more likely that small vote share differences are observed (rather than small differences in the number of votes). The opposite holds for small elections.

[22] We have also calculated the bandwidths proposed by Fan and Gijbels (1996). As those were always broader than the IK bandwidths, we do not report them.

because a typical close election RDD would not have such lotteries in the data. The table consists of four panels (A-D). We describe each of them in turn.

In Panel A of Table 3, the bandwidth is selected optimally for the local linear specification using either the IK, CCT or LM method. The panel reports for these bandwidth choices the local linear (specifications (1)-(3)), quadratic (specifications (4)-(6)) and cubic (specifications (7)-(9)) RDD estimates of the personal incumbency advantage. As specifications (1)-(3) show, all local linear RDD specifications with bandwidths that are optimally chosen for the linear specification indicate a positive and statistically significant incumbency advantage. The local linear RDD with optimal bandwidth is thus not able to replicate the experimental estimate. This finding is consistent with the view that when the MSE-optimal bandwidths are used in the local linear regression (and when the bias is large), there is a risk of over-rejection because of the distributional approximation being poor. This is likely to happen when the regression function has curvature within the optimal bandwidth that the linear approximation cannot capture. The next specifications (specifications (4)-(9)) in the panel show that the curvature of the regression function indeed matters. Using the richer quadratic and cubic local polynomials aligns the RDD estimates with the experimental results for the bandwidths that are MSE-optimal, as determined by IK and CCT *for the linear specification*. For the much wider LM bandwidth even the quadratic and cubic local polynomials produce a positive and significant effect.[23]

[23] In our case, IK appears to result in the narrowest bandwidth among the IK, CCT and LM methods. The IK and CCT bandwidths are nevertheless pretty close to each other and they give similar results. The optimal IK and CCT bandwidths correspond to around 0.60% of the total votes of a list (that is 6 votes out of 1000). This typically translates into a small number of votes. Notice, however, that the bandwidths are not that small when compared to the vote shares that the candidates at the cutoff get in our data (6.5 % vote share, see Table 1). We use here only the CCT bandwidth selection criteria but not yet the bias-correction or robust inference method that they also propose, i.e. CCT-correction.

In Panel B of the table, we report the results using bandwidths that are half the optimal bandwidth of the local linear specification. This under-smoothing ought to reduce the (asymptotic) bias, which it indeed appears to do. All the estimates decrease in size. When bandwidths half the size of the optimal IK or CCT bandwidths are used, the results are in line with the experimental benchmark (specifications (10)-(11)). These results continue to hold when the quadratic and cubic polynomials are used. The wider LM bandwidth produces larger estimates that are not in line with our experimental finding in all but the cubic specification (18).

In Panel C of Table 3, we report the results using bandwidths that are twice the optimal ones. We do this to see how the richer polynomials work when the neighborhood around the cutoff is widened. Compared to Panel A and B, the estimates increase in size. Now only the local cubic polynomial regressions with IK and CCT bandwidths do not reject the null hypothesis of no effect. This further illustrates how the curvature of the regression function matters.

Finally, in Panel D, we report the results for the quadratic and cubic specifications, with IK and CCT bandwidths that have been re-optimized for these more flexible polynomial specifications. As the panel shows, we again get positive and statistically significant effects. In the cubic specification that uses the IK bandwidth, the effect is significant only at 10% level.

The above findings are in line with the well-known result that the MSE-optimal bandwidth is too large for inference. As e.g. Imbens and Lemieux (2008) and CCT have noted, the use of the MSE-optimal bandwidth leads to over-rejection of the null hypothesis if the bias caused by the linear approximation is non-negligible. What also is in line with the recent econometric work is that holding the order of the polynomial constant, smaller bandwidths align our RDD results with the experimental benchmark (see CCT for a discussion of under-smoothing). We find that holding the bandwidth constant, richer polynomials align our RDD results with the experimental benchmark, too. This result is in line with those reported by Card et al. (2014) in the sense of advocating the use of higher order local polynomials. The difference is that in our data, the MSE-based rule proposed by Card et al. appears not to reproduce the experimental estimate.[24]

[24] Card et al. (2014) propose selecting the order of the local polynomial by minimizing the asymptotic MSE. We have used polynomials of orders 0 – 5 with the IK optimal bandwidth (which in our data is narrower than the CCT optimal bandwidth). We failed to reproduce the experimental estimate.

Table 3. Local polynomial RDD estimates

Outcome: Elected next election									
Panel A: Bandwidth optimized for local linear specification									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Linear			Quadratic			Cubic		
Elected (conventional)	0.0387*	0.0521**	0.1562**	0.0077	0.0217	0.0683**	-0.0220	-0.0042	0.0364**
	(0.0156)	(0.0129)	(0.0090)	(0.0239)	(0.0203)	(0.0099)	(0.0338)	(0.0272)	(0.0132)
N	19407	26999	89094	19407	26999	89094	19407	26999	89094
Bandwidth	0.53	0.74	2.51	0.53	0.74	2.51	0.53	0.74	2.51
Bandwidth selection method	IK	CCT	LM	IK	CCT	LM	IK	CCT	LM
Panel B: Bandwidth optimized for local linear specification * 0.5									
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Linear			Quadratic			Cubic		
Elected (conventional)	0.0068	0.0240	0.0831**	-0.0224	-0.0148	0.0359**	-0.0180	-0.0250	0.0243
	(0.0232)	(0.0190)	(0.0096)	(0.0364)	(0.0303)	(0.0146)	(0.0549)	(0.0421)	(0.0207)
N	9808	13496	47898	9808	13496	47898	9808	13496	47898
Bandwidth	0.27	0.37	1.25	0.27	0.37	1.25	0.27	0.37	1.25
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * LM	0.5 * IK	0.5 * CCT	0.5 * LM	0.5 * IK	0.5 * CCT	0.5 * LM
Panel C: Bandwidth optimized for local linear specification * 2									
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
	Linear			Quadratic			Cubic		
Elected (conventional)	0.0708**	0.0975**	0.2418**	0.0334*	0.0412**	0.1344**	0.0159	0.0259	0.0737**
	(0.0104)	(0.0092)	(0.0083)	(0.0162)	(0.0131)	(0.0094)	(0.0228)	(0.0187)	(0.0099)
N	40111	57227	121501	40111	57227	121501	40111	57227	121501
Bandwidth	1.07	1.48	2.51	1.07	1.48	2.51	1.07	1.48	2.51
Bandwidth selection method	2 * IK	2* CCT	2* LM	2 * IK	2 * CCT	2 * LM	2 * IK	2 * CCT	2 * LM
Panel D: Bandwidths optimized for each specification separately									
	(28)	(29)		(30)	(31)		(32)	(33)	
	Linear			Quadratic			Cubic		
Elected (conventional)	0.0387*	0.0521**		0.0393**	0.0569**		0.0495**	0.1105**	
	(0.0156)	(0.0129)		(0.0136)	(0.0106)		(0.0114)	(0.0095)	
N	19407	26999		54465	78470		70577	112399	
Bandwidth	0.53	0.74		1.41	2.09		1.84	3.98	
Bandwidth selection method	IK	CCT		IK	CCT		IK	CCT	

Notes: Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. All estimations use a triangular kernel. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t*.

Even though a typical applied researcher does not have access to an experimental estimate and hence cannot benchmark her RDD estimate to the experimental one, it is of some interest to ask whether the experimental estimate (Table 2, specification (1)) is statistically different from the non-experimental estimates that the local linear RDD with optimal bandwidths produce (see Table 3, specification (1)-(3)). The reason is that an alternative interpretation for our findings is that our experimental estimate is imprecise and, in fact, consistent with a small and positive incumbency effect. The experimental estimate

(0.0044) is 88.6% smaller than the RDD estimate (0.0386) produced by the local linear RDD with the IK optimal bandwidth, but we cannot reject the null hypothesis that the two estimates are equal (p-value = 0.24). However, the difference is statistically significant at 10% level when the estimates based on the CCT bandwidths are used (p-value = 0.087) and highly significant when the estimates based on the LM bandwidths are used (p-value < 0.0001). It is important to stress that this comparison is *not* what a typical applied researcher using RDD absent the experiment could do and would rely on, and thus these tests are not a good benchmark for evaluating the RDD estimate(s).

RDD estimations: Curvature analysis

The raw data support the view that there is substantial curvature in the relation between the forcing variable and the outcome variable close to the cutoff. We demonstrate this in Panel A of Figure 1. It plots the data using 100 bin averages around the optimal IK bandwidth of the local linear specification and the fits of linear (on the left), quadratic (in the middle) and cubic (on the right) regressions. The figure on the left clearly shows that that there is curvature in the data near the cutoff, making the linear approximation inaccurate. The quadratic local polynomial in the middle seems to capture the curvature quite well. This finding suggests that a polynomial specification of order 2 is flexible enough for the bandwidth that has been optimized for a polynomial of order 1, possibly explaining the different performance of these estimators in Table 3.

The same observation can be made from Panels B and C of Figure 1, where the bandwidths are optimal for the quadratic (Panel B) and cubic (Panel C) specifications. Like in Panel A, the graphs on the left side of these panels display the fits that are based on the same order of the local polynomial specification, p, for which the optimal bandwidth is calculated. In the middle graph, the fit uses a p+1 local polynomial, but the bandwidth is the same as on the left. In the graphs on the right side, the displayed fits are based on a p+2 local polynomial. A visual inspection of these graphs again suggests that a polynomial of order p+1 is flexible enough for the bandwidth that has been optimized for a polynomial of order p.[25] The approximation is better especially near the cutoff when the richer p+1 polynomial is used. Moreover, the experimental estimate indicates that there should not be a jump at the cutoff. The graphs on the left are therefore consistent with a poor local approximation, because there a jump can be detected. The jumps are nearly invisible or completely non-existent in the graphs displayed in the middle (p+1) or on the right (p+2).

We have checked that these findings are not specific to the way we define the forcing variable. The same patterns can be observed also if we use the absolute number of votes as the forcing variable (reported in Appendix C).

[25] We checked that this holds in our case from p=0 to p=5, but did not check p>5.

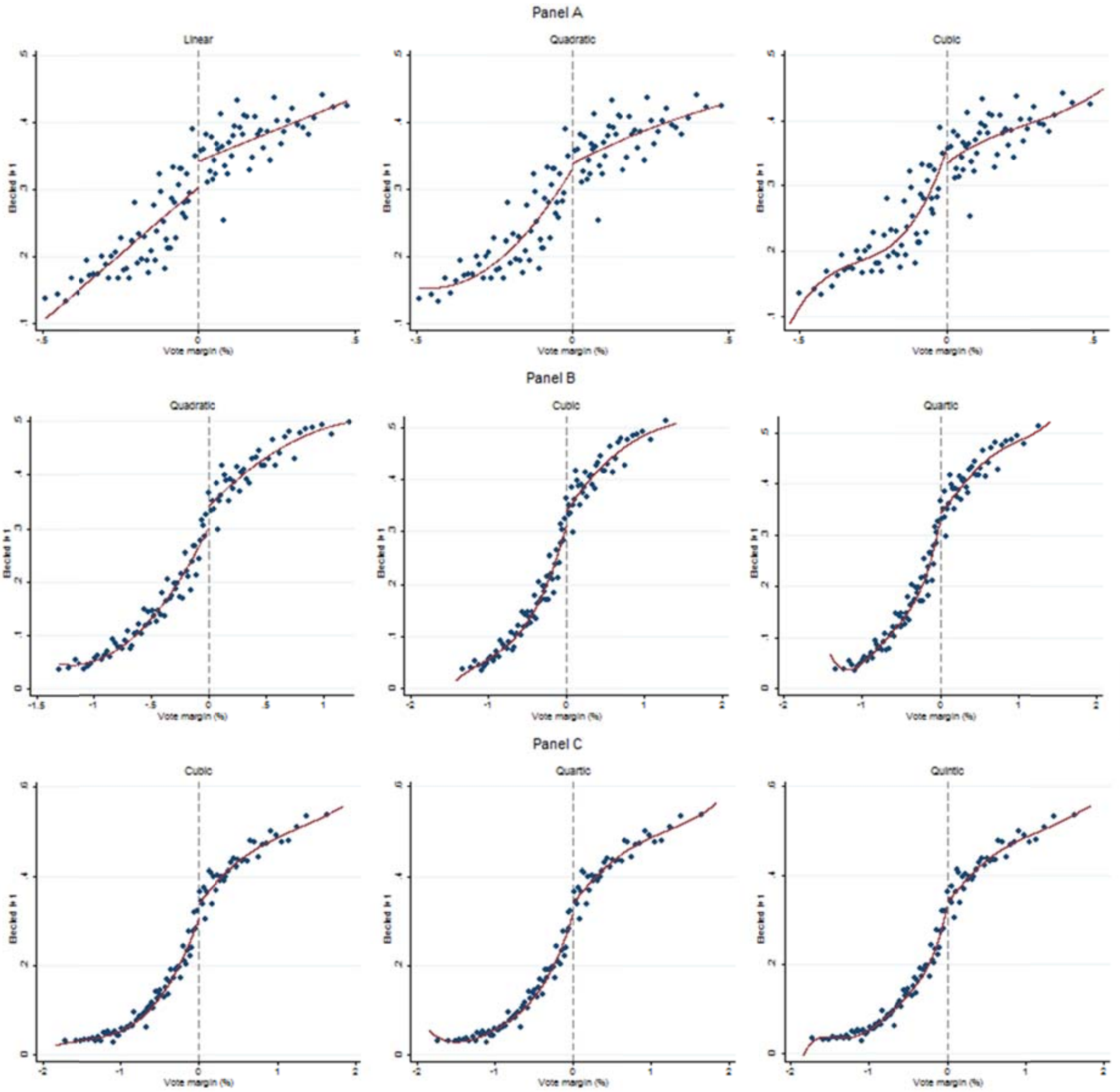


Figure 1. Curvature between the forcing variable and the outcome

Notes: Figure shows local polynomial fits with a triangular kernel within the optimal Imbens-Kalyanaraman (2012) bandwidth optimized for the linear specification in Panel A, quadratic specification in Panel B and cubic specification in Panel C. On left side, the graphs display the fits that are based on the same p (order of local polynomial specification) as the optimal bandwidths are calculated for. In the midmost graph, the fit uses a p+1 specification and on the right side, the graphs are based on a p+2 specification. Blue dots mark binned averages within 100 bins with equal number of observations.

Figure 2 displays RDD estimates for a large number of bandwidths using the three local polynomial regressions. The vertical bars indicate the location of the optimal bandwidth, which varies with the order of the polynomial. The figure provides us with two main findings. First, the bias relative to the experimental benchmark estimate (zero) seems to be almost monotonic in the size of the bandwidth. The approximation gets worse, as more and more data are included in the RDD sample. Second, when bandwidths narrower than the optimal ones are used, RDD reproduces the null result that the data from the random election outcomes suggest, irrespectively of which polynomial is used. How much narrower the bandwidth needs to be depends on the specification, but as a general rule, half the optimal is conservative enough to work well.

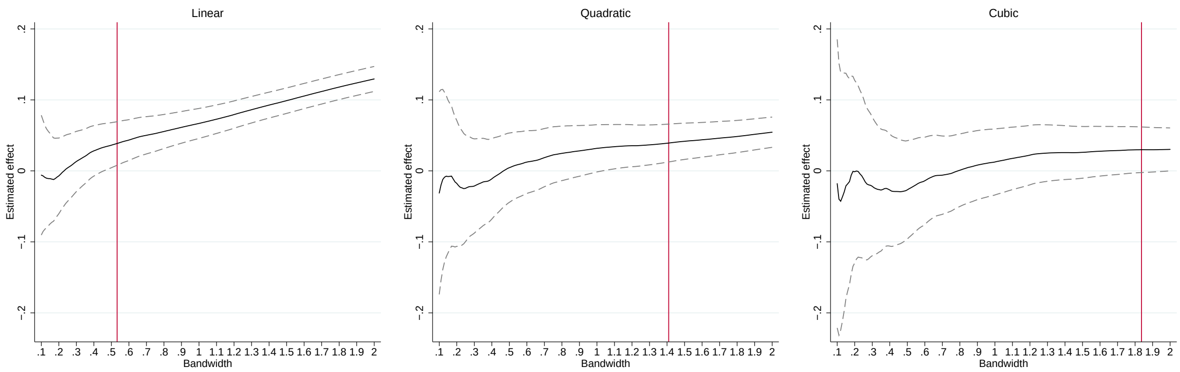


Figure 2. Conventional RDD estimates using various bandwidths.

Notes: Figure displays point estimates from local polynomial regressions with triangular kernel using various bandwidths. Dashed lines show 95 % confidence intervals computed using standard errors clustered at the municipality level. Red vertical line marks the optimal IK bandwidth.

Bias-corrected RDD estimations

To evaluate how the recently proposed bias-correction and robust inference method of CCT works (CCT-correction), we report in Table 4 a number of RDD estimates using the CCT-correction. In this method, a p^{th} order local polynomial is used to estimate the main effect whereas a $(p+1)^{th}$ order local polynomial is used to estimate the (potential) bias. The bandwidth that is used to estimate the bias is called a pilot bandwidth.

Table 4 consists of three panels. In Panel A, we use bandwidths optimized for the linear specification, but report the estimates from linear, quadratic and cubic local polynomial specifications. For this panel we choose the pilot bandwidth used to estimate the bias either by the data-driven method suggested by CCT (using the default option in the rdrobust Stata-package; see Calonico et al. 2014b) or by using the IK method. When the pilot bandwidth is chosen by the data-driven method of CCT, the main bandwidth is determined to be MSE optimal, based on the CCT bandwidth selection method. When the pilot bandwidth is chosen by the IK method, so is the main bandwidth. The results of this panel show that the CCT-correction is able to meet the replication standard, in the sense that when the CCT corrected estimates and standard errors are used, we do not, in general, reject the null hypothesis of no effect. The important exception to this result is the data-driven pilot bandwidth calculation suggested by CCT. It apparently leads to too wide pilot bandwidths. When the pilot and main bandwidths are chosen by the IK method, the CCT-correction meets the replication standard, irrespectively which local polynomial specification is used.

In Panel B, we again report the estimates from linear, quadratic and cubic local polynomial specifications but choose the bandwidths differently. We optimize the main bandwidths for the linear specification using the CCT and IK bandwidth selection methods. We then impose the pilot bandwidth to be the same as the main bandwidth. From the perspective of the point estimate, CCT-correction with the same main and pilot bandwidth amounts to using the conventional local polynomial approach, but with the twist that the main effect is estimated using a one order higher polynomial specification (p+1) than the specification for which the bandwidth is selected (p). It follows that the point estimate (but not the standard error) is the same in columns (4) and (5) of Table 3 as here in columns (7) and (8) of Panel B of Table 4. The results of this panel show that when implemented in this way, the CCT-correction is able to meet the replication standard.

In Panel C, we use the bandwidths optimized for the quadratic and cubic local specifications. They are chosen as in Panel A. We again find that the CCT-correction is able to meet the replication standard, provided that the pilot and main bandwidths are chosen by the IK method. The data-driven method suggested by CCT again seems to lead to a too wide pilot bandwidth.

Table 4. CCT bias-corrected local polynomial RDD estimates with robust inference.

Outcome: Elected next election

Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(3)	(4)	(5)	(6)
	Linear		Quadratic		Cubic	
Elected (bias correction and robust inference)	0.0295	0.0455**	0.0059	0.0210	-0.0232	-0.0043
	(0.0156)	(0.0120)	(0.0232)	(0.0182)	(0.0333)	(0.0259)
N	19407	26999	19407	26999	19407	26999
Bandwidth	0.53	0.74	0.53	0.74	0.53	0.74
Pilot bandwidth	1.14	3.03	1.14	3.03	1.14	3.03
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Main and pilot bandwidths optimized for local linear specification						
	(7)	(8)	(9)	(10)	(11)	(12)
	Linear		Quadratic		Cubic	
Elected (bias correction and robust inference)	0.0077	0.0217	-0.0220	-0.0042	-0.0332	-0.0214
	(0.0223)	(0.0181)	(0.0327)	(0.0258)	(0.0453)	(0.0352)
N	19407	26999	19407	26999	19407	26999
Bandwidth	0.53	0.74	0.53	0.74	0.53	0.74
Pilot bandwidth	0.53	0.74	0.53	0.74	0.53	0.74
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel C: Main and pilot bandwidths optimized for each specification separately						
	(13)	(14)	(15)	(16)	(17)	(18)
	Linear		Quadratic		Cubic	
Elected (bias correction and robust inference)	0.0295	0.0455**	0.0262	0.0518**	0.0250	0.0515**
	(0.0156)	(0.0120)	(0.0170)	(0.0110)	(0.0189)	(0.0107)
N	19407	26999	54465	78470	70577	112399
Bandwidth	0.53	0.74	1.41	2.09	1.84	3.98
Pilot bandwidth	1.14	3.03	1.49	5.38	1.92	7.90
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT

Notes: Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. We report bias-corrected estimates and robust standard errors computed using rdrobust command in Stata. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t*.

To explore how the bias corrected and robust estimates vary with different bandwidths and how the two bandwidth choices interact, we display in Figure 3 the bias-corrected RDD estimates and their robust 95% confidence intervals for a fixed pilot bandwidth, but for different main bandwiths. We use the IK method to determine the pilot bandwidth, because it seemed to work well. The figure shows that when fixing the pilot bandwidth to be IK optimal, the estimated effect is quite robust to the choice of the main bandwidth and most of the time not significantly different from zero. This confirms that the CCT-correction seems to work well, when the pilot bandwithd is chosen by the IK method.

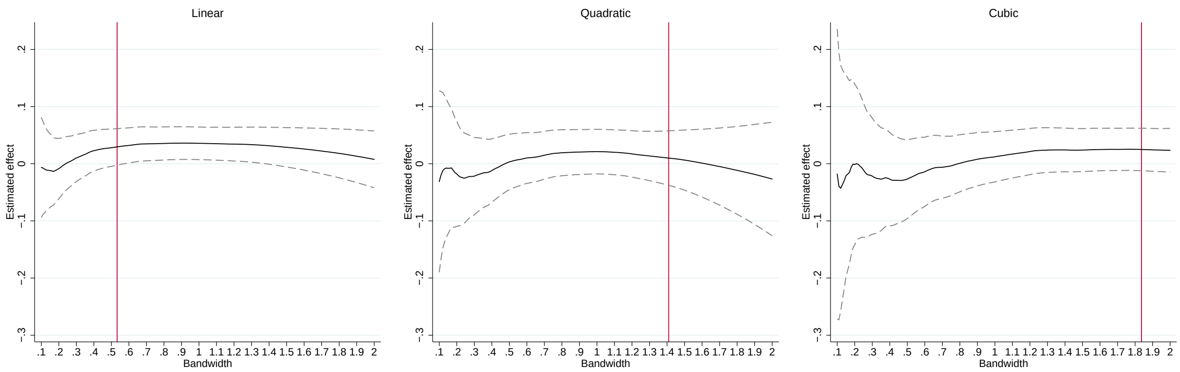


Figure 3. Bias-corrected RDD estimates, fixed pilot bandwidth.

Notes: Figure displays bias-corrected point estimates from local polynomial regressions with triangular kernel using various bandwidths. Dashed lines show 95 % confidence intervals computed using robust standard errors. Red vertical lines mark the optimal IK bandwidth. The pilot bandwidth for bias correction has been fixed to 1.14, 1.49 and 1.92 for linear, quadratic and cubic specifications, respectively. Estimations have been carried out using rdrobust command in Stata.

In Figure 4, we allow both bandwidths to vary and report the corresponding CCT-corrected estimates and their robust confidence intervals. The results resemble those we reported earlier (Figure 2) for the conventional RDD. The estimated effect mostly increases with the bandwiths, but now the replication standard is met across a wider range of bandwiths. The figure confirms that when the CCT-correction is used and the main bandwidth is chosen to be IK optimal or smaller, the null hypothesis of no effect is not rejected in any of the specifications.

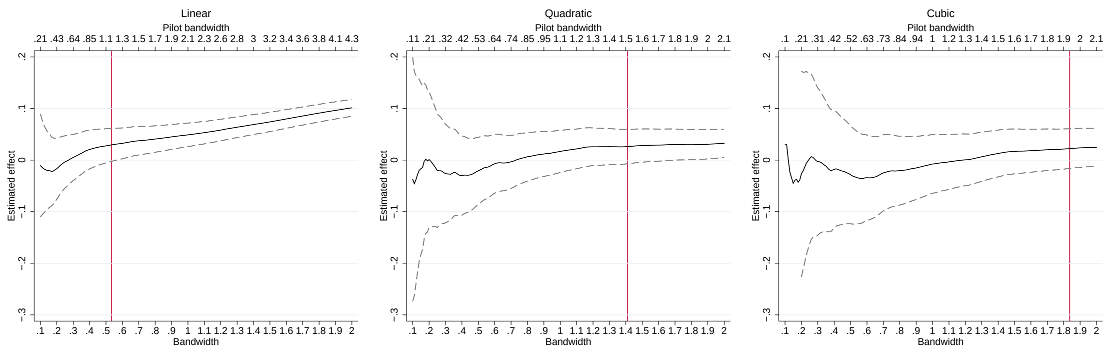


Figure 4. Bias-corrected RDD results, both bandwidths vary.

Notes: Figure displays bias-corrected point estimates from local polynomial regressions with triangular kernel using various bandwidths. Dashed lines show 95 % confidence intervals computed using robust standard errors. In the third graph, confidence intervals are omitted for bandwidths smaller than 0.2. Red vertical lines mark the optimal IK main and pilot bandwidths (both for estimation and bias correction). Estimations have been carried out using rdrobust command in Stata.

4 Discussion and robustness

4.1 RDD falsification and smoothness tests

The standard pattern of validity tests for the RDD includes the McCrary (2008) manipulation test, covariate balance tests, which are an indirect test of the smoothness assumption, and placebo tests, where the location of the cutoff is artificially redefined. We do not report the results of the validity tests in detail here. It suffices to note the following (see Appendix D for details).

First, there is no jump in the amount of observations at the cutoff of getting elected. Second, when testing for covariate balance, we allow for the possibility that the covariates have slopes (or even curvature) near the cutoff (e.g., Snyder et al. 2015 and Eggers et al. 2015) and estimate local linear specifications. We calculate the optimal bandwidths (and half the optimal ones) for different polynomials to address potential slope and curvature issues. We do this for each covariate separately. The covariate balance tests produce somewhat mixed evidence, but overall they suggest that RDD ought to work well in our application. This finding is somewhat in contrast with those of Caughey and Sekhon (2011), who mention the possibility that purposeful sorting by the candidates may invalidate the use of RDD also in the closest races. We find some evidence that there are fewer rejections of covariate balance when more flexible local polynomial specifications (or under-smoothing) are used.

Finally, the placebo cutoff tests provide signals that cast doubt on the appropriateness of standard local linear (and polynomial) RDD specifications with the MSE optimal bandwidths in our context. Moreover, the placebo tests do *not* suggest that under-smoothing procedures and use of higher degree local polynomials *without* adjusting the bandwidth accordingly would not work. This finding echoes the conclusion that when these bias-correction tools are used, RDD is able to reproduce the experimental estimate. In sum, this shows that the placebo cutoff tests can be useful in detecting too inflexible specifications.

4.2 When is RDD as good as randomly assigned?

One reason for the popularity of RDD is that close to the cutoff, variation in the treatment status may be "as good as randomized", provided that the forcing variable cannot be precisely manipulated (Lee 2008, p. 676). RDD is widely believed to meet the replication standard because of this particular feature. This feature may also be the reason why RDD has been used as a benchmark against which other non-experimental estimators have been compared (see, e.g., Lemieux and Milligan 2008).

This naturally leads to the question of whether we can identify a neighborhood around the cutoff where the randomization assumption is plausible. Knowing whether such a neighborhood can be found is useful, even though covariate balance in means is not a requirement for the validity of RDD. Our data are special, because we know that in a sample that includes the lotteries, the randomization assumption is satisfied if the neighborhood is degenerate at the cutoff.

Inspired by the approach proposed by Cattaneo et al. (2014), we explore the largest bandwidth in which the as-good-as-random assumption holds and then compare the sample means of the outcome variable across the cutoff. To determine the largest bandwidth in which the as-good-as-random assumption holds, we either look at the most important covariate or the minimum *p*-value among all the covariates. According to Eggers et al. (2015), incumbency status (elected at *t*-1) is a very good measure of candidate quality. If we use it, we find that bandwidths 0.04 or smaller are as-good-as-random at the 5% significance level (923 non-experimental observations). Based on the minimum *p*-value among all the covariates (but not correcting for multiple testing), it seems that bandwidths 0.02 or smaller would be as-good-as random at the 5% significance level (128 observations). These findings indicate that the approach proposed by Cattaneo et al. (2014) leads to rather conservative (small) samples in light of our other RDD findings. This is partly due to not correcting for multiple testing and partly due to the fact that in our election data, many covariates have rather steep slopes with respect to the forcing variable.[26]

The approach is, however, able to reproduce the experimental estimate. When we use these conservative bandwidths, there is no statistically significant difference in the means of getting elected at t+1 elections around the cutoff: The difference is 0.0101 (p-value 0.32) for the bandwidth of 0.04 and 0.0640 (p-value 0.75) for the bandwidth of 0.02.[27]

[26] One could consider generalizing the covariate-based bandwidth selection criterion by Cattaneo et al. (2014) to higher order local polynomial specifications (*p* = 0 in Cattaneo et al. 2014). This would involve conducting all the covariate balance tests with the same bandwidth and specification that is used to estimate the effect on the main outcome. Such covariate balance tests could be indicative of the issues (e.g. the curvature problem) in the particular specifications and bandwidth choices used to estimate the main outcome. However, a systematic analysis and formalization of this generalization is beyond the scope of this paper.

[27] Note that we do not have to resort to the randomization inference method proposed by Cattaneo et al. (2014a), because we have quite a lot of observations also within the two as-good-as-random bandwidths.

4.3 Robustness of RDD estimates

We have conducted a large number of tests to probe the robustness of our RDD findings. Taking each of them in turn (see Appendix F for details):

First, RDD is sometimes implemented using higher order *global* polynomials of the forcing variable. We have redone the RDD analysis using such parametric RDDs, using five different polynomials (1st-5th degree). These parametric RDD generates positive and statistically significant incumbency effects that are roughly similar in magnitude to those reported in Lee (2008). Consistent with what Gelman and Imbens (2014) argue, we find that this approach to implementing RDD provides misleading findings, as it does not allow us to recover the experimental estimate. The bias here is an order of magnitude larger than the one in the local polynomial specifications.

Second, we have considered the vote share in the subsequent elections as an alternative measure of incumbency advantage. As we reported earlier, the experimental estimate suggests no incumbency advantage when this alternative measure is used. In contrast, the RDD results suggest a positive effect when RDD is implemented in a standard fashion, using the local linear polynomial and various (MSE) optimal bandwidths.

Third, ties appear more often in elections in the smaller municipalities. As we reported earlier, the experimental estimate is quite precisely estimated and close to zero both in small and in large elections. However, our normalized forcing variable can get values really close to zero only when parties get a large amount of votes, which tends to happen in the elections in the larger municipalities. To check what this implies for our RDD findings, we have rerun parts of the RDD analysis separately for small and large municipalities. These estimations show that for both the larger and smaller municipalities, the bias increases with the bandwidth and decreases as the degree of local polynomial increases. It thus seems that the conclusions we draw are not driven by the size of the municipalities.

Fourth, another potential explanation for why the local linear RDD point estimates increase when the bandwidth gets wider is heterogeneity in the personal incumbency effect across municipalities (and party-lists). The use of wider bandwidths means that RDD identifies the effect for a different set of municipalities than what we have in the lottery sample. To rule out this explanation, we have repeated the RDD analysis using only those party-lists that were involved in the lotteries. In this case, increasing the bandwidth adds new candidates from the same lists, but does not add new lists or municipalities. Our main results remain unchanged.

Fifth, we have rerun the RDD estimations using alternative definitions for the forcing variable. The results show that our RDD findings are not driven by the choice of the forcing variable. For example, we get very similar results if the forcing variable is either the vote margin that is calculated in terms of the number of votes or vote shares. Moreover, we get similar results if the locations of the within-party thresholds are defined in an alternative way.

Sixth, we have studied whether there is heterogeneity in the effect between the parties. We found no evidence for substantial heterogeneity in the personal incumbency advantage between the parties.

Finally, we have already mentioned that the experimental estimate does not change if those who do not rerun are excluded from the lottery sample. We have replicated our baseline RDD analysis using the sample from which those who do not rerun are similarly excluded. Our results remain unchanged.

5 Conclusions

We have made use of elections data in which the electoral outcome was determined via a random seat assignment for a large number of candidates because of a tie in their vote count. These instances provide us with a randomized experiment against which we have benchmarked non-experimental RDD estimates of personal incumbency advantage. To our knowledge, the experiment is unique in the literature, because it takes place exactly at the cutoff. This means that both the experiment and RDD target the same treatment effect.

We find that there is no evidence of a personal incumbency advantage when data from the randomized elections is used. The point estimate of the incumbency advantage is close to zero and relatively precisely estimated. We have argued that this finding is neither surprising nor in conflict with the prior evidence, because we are looking at the effect of incumbency status on electoral success at a rather special context, in small local PR elections. It is possible that the randomized electoral victories that we study take place at a relatively unimportant margin, providing limited scope for the emergence and creation of personal incumbency advantage.

We also find that when RDD is applied in conventional fashion (i.e. using local linear regression with optimal bandwidths) to the same close elections, the estimates suggest a moderate and statistically significant personal incumbency effect. However, standard bias-correction tools, such as using narrower bandwidths than those suggested by the bandwidth algorithms, using higher degree local polynomials *without* adjusting the bandwidth accordingly or using the recent bias correction method of CCT with IK optimal bandwidths, bring the RDD results in line with those obtained using the randomized elections. The finding that higher order polynomials may lead to an improved estimate is in line with Card et al. (2014), but the MSE-based rule advocated by them seems not to work in our data.

Our results hence show that RDD can indeed meet the replication standard in the context of close elections. The lesson that the prior work on non-parametric methods has already stressed - and that our study clearly enforces - is that the bias induced by the local linear approximation can be large enough to severely distort the conclusions one draws. This possibility needs to be addressed carefully in applied RRD work.

References

Butler, Daniel M. 2009. "A Regression Discontinuity Design Analysis of the Incumbency Advantage and Tenure in the U.S. House." *Electoral Studies* 28(1): 123-28.

Black, Dan, Jose Galdo, and Jeffrey Smith. 2007. "Evaluating the Regression Discontinuity Design Using Experimental Data." Working paper.

Broockman, David E. 2009. "Do congressional candidates have reverse coattails? Evidence from a regression discontinuity design." *Political Analysis* 17(4): 418-34.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6): 2295–2326.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014b. "Robust Data-Driven Inference in the Regression Discontinuity Design." *Stata Journal* 14(4): 909-946.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2014. "Local Polynomial Order in Regressions Discontinuity Designs." Brandeis University Working Paper 81.

Cattaneo, Matias, Brigham R. Frandsen, and Rocio Titiunik. 2014. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." *Journal of Causal Inference* 3(1): 1-24.

Caughey, Devin, and Jasjeet S. Sekhon. 2011. "Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008." *Political Analysis* 19(4): 385–408.

Cook, Thomas D., and Vivian C. Wong. 2008. "Empirical tests of the validity of the regression discontinuity design." *Annals of Economics and Statistics* 91-92: 127-50.

Cook, Thomas D., William Shadish, and Vivian C. Wong. 2008. "Three conditions under which observational studies produce the same results as experiments." *Journal of Policy Analysis and Management* 27(4): 724–50.

Cox, Gary W., and Jonathan N. Katz. 1996. "Why did the incumbency advantage in U.S. House elections grow?" *American Journal of Political Science* 40(2): 478-97.

Dahlgaard, Jens Olav. 2013. "You just made it: Individual incumbency advantage under proportional representation." Working paper.

De Magalhaes, Leandro. 2014. "Incumbency Effects in a Comparative Perspective: Evidence from Brazilian Mayoral Elections." *Political Analysis* 23(1): 113-126.

Desposato, Scott W., and John R. Petrocik. 2003. "The Variable Incumbency Advantage: New Voters, Redistricting, and the Personal Vote." *American Journal of Political Science* 47(1): 18-32.

DiNardo, John, and David Lee. 2004. "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001." *The Quarterly Journal of Economics* 119(4): 1383-441.

Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder. 2015. "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(1): 259–274.

Erikson, Robert, and Rocio Titiunik. 2015. "Using regression discontinuity to uncover the personal incumbency advantage." *Quarterly Journal of Political Science*. Forthcoming.

Fan, Jianqing, and Gijbels, Irene. 1996. *Local Polynomial Modeling and its Applications*, New York, Chapham & Hall / CRC Press.

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

Fiva, Jon H., Olle Folke, and Rune J. Sørensen. 2013. "The Power of Parties." CESifo Working Paper Series 4119.

Folke, Olle, and James M. Snyder. 2012. "Gubernatorial midterm slumps." *American Journal of Political Science* 56(4): 931-48.

Folke, Olle. 2014. "Shades of Brown and Green: Party effects in Proportional Election Systems." *Journal of the European Economic Association* 12(5): 1361–1395.

Fowler, Anthony, and Andrew B. Hall. 2014. "Disentangling the Personal and Partisan Incumbency Advantages: Evidence from Close Elections and Term Limits." *Quarterly Journal of Political Science* 9(4): 501-531.

Fraker, Thomas, and Rebacca Maynard. 1987. "The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs." *Journal of Human Resources* 22(2): 194-227.

Freier, Ronny, and Christian Odendahl. 2012. "Do Parties Matter? Estimating the Effect of Political Power in Multi-Party Systems." DIW Discussion Papers 1205.

Gelman, Andrew, and Guido Imbens. 2014. "Why Higher Order Polynomials Should not be Used in Regression Discontinuity Designs." NBER Working Paper 20405.

Gelman, Andrew, and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34(4): 1142-64.

Golden, Miriam, and Lucio Picci. 2015. "Incumbency Effects under Proportional Representation: Leaders and Backbenchers in the Postwar Italian Chamber of Deputies." *Legislative Studies Quarterly*, forthcoming.

Gerber, Elisabeth R., and Daniel J. Hopkins. 2011. "When Mayors Matter: Estimating the Impact of Mayoral Partisanship on City Policy." *American Journal of Political Science* 55(2): 326-39.

Green, Donald P., Terence Y. Leong, Holger L. Kern, Alan S. Gerber, and Christopher W. Larimer. 2009. "Testing the Accuracy of Regression Discontinuity Analysis Using Experimental Benchmarks." *Political Analysis* 17(4): 400-17.

Grimmer, Justin, Eitan Hersh, Brian Feinstein, and Daniel Carpenter. 2012. "Are Close Elections Random?" Working paper.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1): 201-09.

Iltasanomat. 4/12/2011. "Tiesitkö tätä vaaleista? Näin käy, jos äänet menevät tasan." http://www.iltasanomat.fi/vaalit2011/art-1288381910312.html.

Imbens, Guido, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-35.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933-59.

Klašnja, Marko, and Rocio Titiunik. 2015. "The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability." Working paper.

Kotakorpi, Kaisa, Marko Terviö, and Panu Poutvaara. 2013. "Returns to Office in National and Local Politics." CESifo Working Paper Series 4542.

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review* 76(4): 604-20.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142(2): 675-97.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-355.

Lee, David S., Enrico Moretti, and Matthew J. Butler. 2004. "Do Voters Affect Or Elect Policies? Evidence from the U. S. House." *Quarterly Journal of Economics* 119(3): 807-59.

Lemieux, Thomas and Milligan, Kevin 2008. “Incentive effects of social assistance: A regression discontinuity approach.” *Journal of Econometrics* 142: 807–28.

Ludwig, Jens and Douglas L. Miller. 2007. “Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design.” *Quarterly Journal of Economics* 122(1): 159-208.

Lundqvist, Heléne. 2011. “Is It Worth It? On the Returns to Holding Political Office.” Working paper.

McCrary, Justin. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 142(2): 698-714.

Porter, Jack. 2003. “Estimation in the Regression Discontinuity Model.” Working paper.

Redmond, Paul and Regan, John. 2015. “Incumbency Advantage in a Proportional Electoral System: A Regression Discontinuity Analysis of Irish Elections”, *European Journal of Political Economy* 38: 244-256.

Shadish, William R., Rodolfo Galindo, Vivian C. Wong, Peter M. Steiner, and Thomas D. Cook. 2011. “A randomized experiment comparing random to cutoff-based assignment.” *Psychological Methods* 16(2): 179-91.

Snyder, James M., Olle Folke, and Shigeo Hirano. 2015. “Partisan Imbalance in Regression Discontinuity Studies Based on Electoral Thresholds.” *Political Science Research and Methods*. Forthcoming.

Snyder, Jason. 2005. “Detecting manipulation in US House elections.” Working paper.

Stone, Peter. 2011. *The Luck of Draw: The Role of Lotteries in Decision-Making*. Oxford University Press.

The Atlantic. 11/19/2012. "When a State Election Can Be Literally Determined by a Coin Toss." http://www.theatlantic.com/politics/archive/2012/11/when-a-state-election-can-be-literally-determined-by-a-coin-toss/265413/.

The Telegraph India. 2/7/2014. "In election tie, pick winner by lottery: SC." http://www.telegraphindia.com/1140207/jsp/nation/story_17909869.jsp#.VCUt8hCfjjV.

Thistlethwaite, Donald L., and Donald T. Campbell. 1960. "Regression-discontinuity analysis: An alternative to the ex post facto experiment." *Journal of Educational Psychology* 51(6): 309-17.

Time. 5/17/2013. "Coin Toss Determines Mayor in Philippine Town." http://newsfeed.time.com/2013/05/17/coin-toss-determines-mayor-in-philippine-town/.

Trounstine, Jessica. 2011. "Evidence of a Local Incumbency Advantage." *Legislative Studies Quarterly* 36(2): 255-80.

UPI. 7/14/2014. "New Mexico judicial election decided by coin toss." http://www.upi.com/Top_News/US/2014/07/14/New-Mexico-judicial-election-decided-by-coin-toss/6161405361068/.

Uppal, Yogesh. 2009. "The disadvantaged incumbents: estimating incumbency effects in Indian state legislatures." *Public Choice* 138(1): 9-27.

Uppal, Yogesh. 2010. "Estimating Incumbency Effects In U.S. State Legislatures: A Quasi-Experimental Study." *Economics and Politics* 22(2): 180-99.

Vogl, Tom S. 2014. “Race and the politics of close elections.” *Journal of Public Economics* 109(1) 101–13.

Wing, Coady, and Thomas D. Cook. 2013. “Strengthening the regression discontinuity design using additional design elements: A within-study comparison.” *Journal of Policy Analysis and Management* 32(4): 853–77.

Does Regression Discontinuity Design Work?

Evidence from Random Election Outcomes

ONLINE SUPPLEMENT

Ari Hyytinen, Jaakko Meriläinen, Tuukka Saarimaa, Otto Toivanen and Janne Tukiainen

18 September 2015

This is an online supplement to Hyytinen, Meriläinen, Saarimaa, Toivanen and Tukiainen (2015, HMSTT hereafter). We report here the additional empirical analyses to which HMSTT refers in the main text.

The supplement consists of Appendices A – F. Appendix A reports summary statistics for our data. In Appendix B, we describe a number of empirical results for the lottery sample. Appendix C characterizes graphically the forcing variable used in the regression discontinuity design (RDD). In Appendix D, we evaluate the validity of the RDD. Appendix E reports additional covariate balance tests for various RDD samples, determined by different bandwidth choices. Finally, a large battery of robustness checks is reported in Appendix F.

Appendix A: Supplementary information to HMSTT Section 2.2 (Data)

In this appendix, we report summary statistics for our data.

Table A1: This table reports descriptive statistics for the individual candidates. Overall it seems that the elected candidates have on average higher values in variables that can be seen as measuring candidate quality. Like we report in the main text, the table shows, for example, that the elected candidates have higher income, are more often university-educated and are less often unemployed. The difference is particularly striking when we look at the row which refers to the incumbency status: 58% of the elected candidates were incumbents, whereas only 6% of those who were not elected were incumbents.

Table A1. Descriptive statistics for individual candidates.

Individual characteristics

	All data (N = 198121)			Elected (N = 56734)			Not elected (N = 141384)		
Variable	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Elected next election (only re-runners)	82949	0.38	0.48	32070	0.79	0.41	50879	0.12	0.32
Elected next election (all candidates)	160730	0.19	0.40	46982	0.54	0.50	113748	0.05	0.22
Running next election	160730	0.52	0.50	46982	0.68	0.47	113748	0.45	0.50
Number of votes next election	82949	76	180	32070	131	268	50879	41	65
Vote share next election	82949	1.14	1.31	32070	2.05	1.54	50879	0.57	0.68
Vote share	198120	0.97	1.20	56734	2.22	1.50	141386	0.46	0.47
Number of votes	198120	61	149	56734	127	257	141386	34	45
Female	198121	0.39	0.49	56734	0.35	0.48	141387	0.40	0.49
Age	198120	46.75	12.64	56734	48.15	11.15	141386	46.18	13.15
Incumbent	198121	0.21	0.41	56734	0.58	0.49	141387	0.06	0.24
Municipal employee	160996	0.23	0.42	47060	0.27	0.44	113936	0.22	0.41
Wage income	98360	20372	26054	27398	24302	44222	70962	18855	13328
Capital income	98360	1857	22867	27398	3388	38750	70962	1266	11991
High professional	198025	0.19	0.40	56721	0.24	0.43	141304	0.18	0.38
Entrepreneur	198025	0.15	0.36	56721	0.23	0.42	141304	0.12	0.33
Student	198025	0.04	0.20	56721	0.02	0.13	141304	0.05	0.22
Unemployed	198025	0.07	0.25	56721	0.03	0.18	141304	0.08	0.27
University degree	159440	0.16	0.37	46711	0.20	0.40	112729	0.14	0.35
Coalition Party	198121	0.15	0.36	56734	0.15	0.35	141387	0.16	0.36
Social Democrats	198121	0.18	0.38	56734	0.18	0.38	141387	0.18	0.38
Center Party	198121	0.22	0.42	56734	0.30	0.46	141387	0.19	0.40
True Finns	198121	0.02	0.15	56734	0.01	0.12	141387	0.03	0.16
Green Party	198121	0.04	0.19	56734	0.02	0.15	141387	0.04	0.20
Socialist Party	198121	0.09	0.29	56734	0.07	0.26	141387	0.10	0.30
Swedish Party	198121	0.03	0.17	56734	0.04	0.20	141387	0.02	0.16
Christian Party	198121	0.04	0.18	56734	0.03	0.16	141387	0.04	0.19
Other parties	198121	0.23	0.42	56734	0.20	0.40	141387	0.24	0.43

Notes: Income data are not available for 2012 elections, and in 1996 elections they are available only for candidates who run also in 2000, 2004 and 2008 elections. Income is expressed in euros. Municipal employee status is not available for 2012 elections.

Table A2: This table reports descriptive statistics for the municipalities. The panel on the left shows, e.g., that there are three major parties in Finland. The three largest parties seat shares total to over 70%. There are two main reasons why there are differences in the variables related to elections between the elected candidates' municipalities (the panel in the middle) and the not-elected candidate's municipalities (the panel on the right). First, a larger share of all running candidates is elected in smaller municipalities. For example, the Center Party has a larger vote share in smaller municipalities. Second, there are more candidates in the larger municipalities. The table also shows that in a number of dimensions, like income, age and unemployment rate, there are no major differences in the municipal characteristics between elected and non-elected candidates.

Table A2. Descriptive statistics for municipalities.

Municipality characteristics									
	All data (N = 198121)			Elected (N = 56734)			Not elected (N = 141384)		
Variable	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Total number of votes	198120	19935	43682	56734	10607	26431	141386	23677	48421
Coalition Party seat share	198121	19.58	10.10	56734	17.61	10.52	141387	20.38	9.81
Social Democrats seat share	198121	21.88	10.21	56734	20.62	10.88	141387	22.38	9.88
Center Party seat share	198121	30.58	20.52	56734	35.20	21.14	141387	28.73	19.97
True Finns seat share	198121	3.77	5.87	56734	3.49	5.87	141387	3.88	5.86
Green Party seat share	198121	4.25	5.41	56734	2.89	4.30	141387	4.79	5.70
Socialist Party seat share	198121	8.57	7.37	56734	8.14	7.72	141387	8.74	7.22
Swedish Party seat share	198121	4.39	13.87	56734	5.19	16.80	141387	4.07	12.49
Christian Party seat share	198121	3.41	3.56	56734	3.24	3.79	141387	3.48	3.47
Other parties' seat share	198121	3.45	6.74	56734	3.50	7.56	141387	3.43	6.39
Voter turnout	196332	62.20	6.28	56174	63.40	6.28	140158	61.72	6.21
Population	197310	43407	95692	56581	22944	58177	140729	51634	106027
Share of 0-14-year-olds	196388	17.84	3.28	56331	17.96	3.47	140057	17.79	3.20
Share of 15-64-year-olds	196388	64.41	3.48	56331	63.49	3.27	140057	64.78	3.49
Share of over-65-year-olds	196388	17.75	4.82	56331	18.55	4.99	140057	17.43	4.72
Income per capita	196388	21204	5876	56331	20364	5634	140057	21543	5937
Unemployment	197310	13.50	5.71	56581	13.77	5.85	140729	13.39	5.65

Notes: Income per capita is expressed in euros.

Appendix B: Supplementary information to HMSTT Section 3.1 (Experimental estimates)

In this appendix, we report a number of empirical results obtained using the lottery sample (i.e. the sample which only includes the candidates that had a tie). These results bear on the robustness of the experimental estimate.

Table B1: This table shows additional balance checks for party affiliation and municipality characteristics in the lottery sample. These characteristics should be balanced by construction, as we construct the forcing variable within party lists. The table shows that the samples are, indeed, almost identical. The small and insignificant differences in means are likely due to the fact that in some lotteries there are more than two candidates.

Table B1. Additional balance checks.

Individual characteristics

Variable	Elected (N = 671) N	Mean	Std. Dev.	Not elected (N = 680) N	Mean	Std. Dev.	Difference
Coalition Party	671	0.20	0.40	680	0.20	0.40	0.00
Social Democrats	671	0.18	0.39	680	0.18	0.39	0.00
Center Party	671	0.42	0.49	680	0.42	0.49	0.00
True Finns	671	0.02	0.13	680	0.02	0.13	0.00
Green Party	671	0.01	0.11	680	0.01	0.11	0.00
Socialist Party	671	0.08	0.27	680	0.08	0.27	0.00
Swedish Party	671	0.03	0.18	680	0.04	0.19	-0.01
Christian Party	671	0.02	0.15	680	0.02	0.15	0.00
Other parties	671	0.03	0.18	680	0.03	0.18	0.00

Municipality characteristics

Variable	Elected (N = 671) N	Mean	Std. Dev.	Not elected (N = 680) N	Mean	Std. Dev.	Difference
Total number of votes	671	4467	12006	680	4395	11921	71
Coalition Party seat share	671	16.88	11.08	680	16.76	10.88	0.13
Social Democrats seat share	671	19.70	10.76	680	19.63	10.95	0.07
Center Party seat share	671	41.46	19.98	680	41.57	20.17	-0.11
True Finns seat share	671	1.92	4.79	680	1.89	4.59	0.02
Green Party seat share	671	1.72	3.29	680	1.73	3.31	-0.01
Socialist Party seat share	671	7.55	7.91	680	7.56	7.82	0.00
Swedish Party seat share	671	3.70	14.42	680	3.97	14.95	-0.27
Christian Party seat share	671	2.87	3.92	680	2.83	3.92	0.04
Other parties' seat share	671	3.76	8.59	680	3.63	8.48	0.13
Voter turnout	664	65.23	5.90	673	65.38	6.02	-0.15
Population	671	9316	25430	680	9145	25241	171
Share of 0-14-year-olds	667	18.31	3.31	676	18.42	3.33	-0.11
Share of 15-64-year-olds	667	62.97	2.87	676	62.89	2.90	0.07
Share of over-65-year-olds	667	18.72	4.69	676	18.69	4.68	0.03
Income per capita	667	18457	5372	676	18413	5372	44
Unemployment	671	14.85	6.75	680	14.80	6.69	0.05

Notes: Differences in means have been tested using t test adjusted for clustering at municipality level. Sample includes only candidates running in 1996-2008 elections. Income data are not available for 2012 elections, and in 1996 elections they are available only for candidates who run also in 2000, 2004 and 2008 elections. Income and income per capita are expressed in euros.

Table B2: This table reports experimental results for the alternative outcome of vote shares in the next elections. The regressions use the entire lottery sample. They provide no evidence of personal incumbency advantage. Out of interest, we have also checked that the effect is close to zero and not significant also if running in the next election or the absolute number of votes in the next election is used as the outcome variable.

Table B2. Experimental results the alternative outcome vote share.

Outcome: Vote share next election

	(9)	(10)	(11)	(12)
Elected	0.0116	0.0063	-0.0203	-0.0138
	(0.0579)	(0.0584)	(0.0670)	(0.0747)
N	1351	1351	1351	1351
R^2	0.00	0.06	0.37	0.52
Controls	No	Yes	Yes	Yes
Municipality fixed effects	No	No	Yes	No
Municipality-year fixed effects	No	No	No	Yes

Notes: Only actual lotteries are included in the regressions. Vote share is set to zero for those candidates that do not run in the next election. Set of controls includes age, gender, party affiliation, socio-economic status and incumbency status of a candidate, and total number of votes. Some specifications include also municipality or municipality-year fixed effects. All standard errors are clustered at the municipality level. Unit of observation is a candidate *i* at year *t*.

Table B3: In this table, we look at elections in small and large municipalities separately. We split the sample based on the median number of total votes in the municipality in the lottery sample. This median is 2422. The median is slightly higher (2662) in the entire sample. The regressions reported in the table below do not include any controls. They should therefore be compared to the result in column (1) in Table 2 in the main text of HMSTT. As can be seen from the table, we do not find evidence for an incumbency advantage in either sub-sample.

Table B3. Experimental results for small and large elections.

Outcome: Elected next election		
	(1)	(2)
Elected	0.0016	0.0057
	(0.0336)	(0.0363)
N	687	664
R^2	0.00	0.00
Sample	Small elections	Large elections

Notes: An election is considered small (large), if at most (more than) 2422 votes are cast. Only actual lotteries are included in the regressions. All standard errors are clustered at the municipality level. Unit of observation is a candidate *i* at year *t*.

Table B4: We have reproduced the experimental estimate using a sample from which those who do not rerun are excluded. We report these results for our main outcome and the alternative outcome (the vote share). As above, these results provide no evidence of a personal incumbency advantage.

Table B4. Experimental estimates for rerunners.

Outcome: Elected next election				
	(1)	(2)	(3)	(4)
Elected	-0.0026	-0.0022	0.0251	0.0345
	(0.0350)	(0.0360)	(0.0503)	(0.0638)
N	820	820	820	820
R^2	0.00	0.04	0.41	0.64
Outcome: Vote share next election				
	(5)	(6)	(7)	(8)
Elected	-0.0118	-0.0092	0.0510	0.0209
	(0.0684)	(0.0680)	(0.0823)	(0.1045)
N	820	820	820	820
R^2	0.00	0.17	0.67	0.80
Controls	No	Yes	Yes	Yes
Municipality fixed effects	No	No	Yes	No
Municipality-year fixed effects	No	No	No	Yes
Controls	No	Yes	Yes	Yes
Municipality fixed effects	No	No	Yes	No
Municipality-year fixed effects	No	No	No	Yes

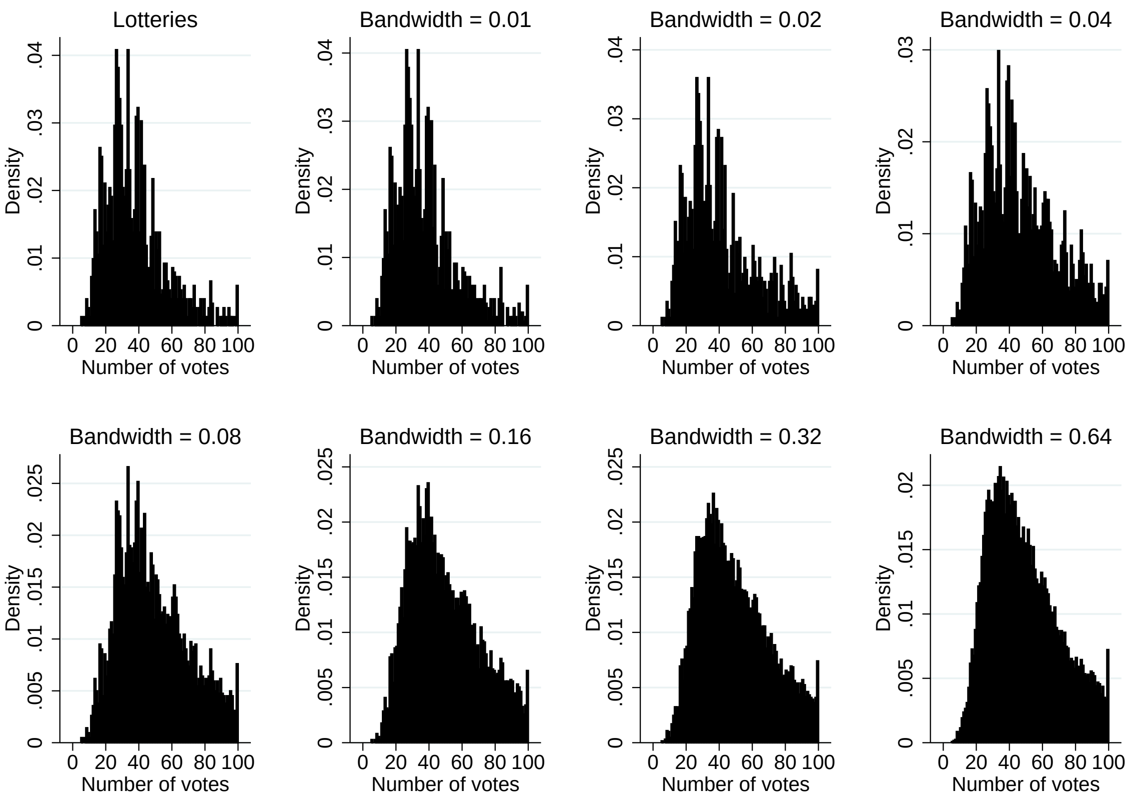
Notes : Only actual lotteries and rerunning candidates are included in the regressions. Set of controls includes age, gender, party affiliation, socio-economic status and incumbency status of a candidate, and total number of votes. Some specifications include also municipality or municipality-year fixed effects. Unit of observation is a candidate *i* at year *t* .

Appendix C: Supplementary information to HMSTT Section 3.2 (Non-experimental estimates)

This appendix provides additional figures to characterize our forcing variable, v_{it}. We call our forcing variable "Vote margin (%)" in some of the graphs below, where the margin refers to the distance to the cutoff. The forcing variable is reported in percentage points. For example, a value 0.5 refers to 5 votes out of 1000.

Figure C1: In this figure, we graph the distribution of the number of votes within different bandwidths in the forcing variables. The figures show how many votes the candidates involved in close elections receive. The distribution gets a large amount of mass around 30–50 votes.

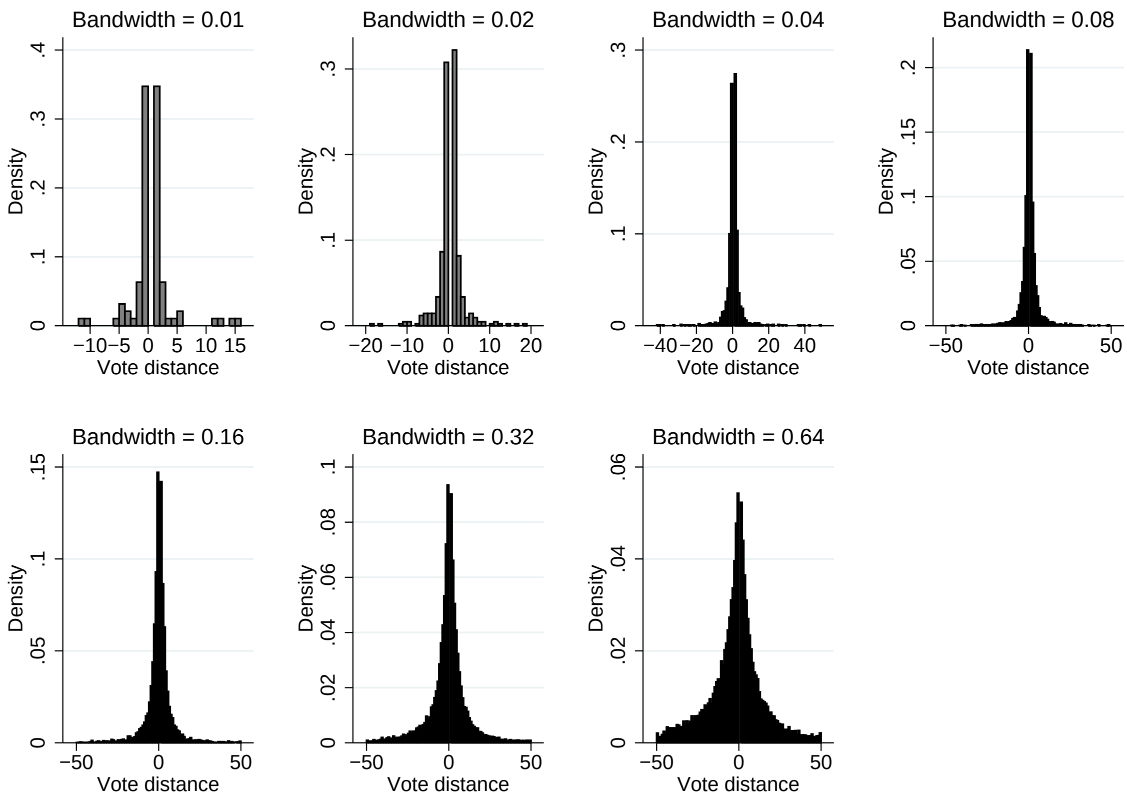
Figure C1. The distribution of the number of votes for different bandwidths.



Notes: Figure shows the distribution of number of votes within one bandwidth on both sides of the cutoff for different bandwidths. Bin size is 1 vote. x-axis is restricted to 100 votes.

Figure C2: This figure displays the relationship between the forcing variable and the distance to cutoff (vote distance), as measured by the absolute number of votes. The density graphs show that, as expected, the candidates are further away from the cutoff in terms of absolute number of votes as the bandwidth becomes wider. For all reported bandwidths, the most common distance is only one or two votes.

Figure C2. Distribution of the distance to cutoff in absolute votes for different bandwidths of the forcing variable.



Notes: Figure shows the distribution of distance measured in votes for different bandwidths. Bin size is 1 vote. x-axis is restricted to 50 votes. Lotteries are excluded.

Figure C3: This figure maps the relationship between the forcing variable (vote margin, x-axis) and the distance to cutoff measured in the absolute number of votes (y-axis). It shows that, overall, the two are positively correlated within the reported bandwidth. There are fairly many observations also on or close by the horizontal line. This means that, within the reported bandwidth, for each value of the forcing variable there are many observations that are only one or two votes from the cutoff. This echoes what Figure C2 showed.

Figure C3. Relationship between the forcing variable and the distance to cutoff measured in absolute votes.

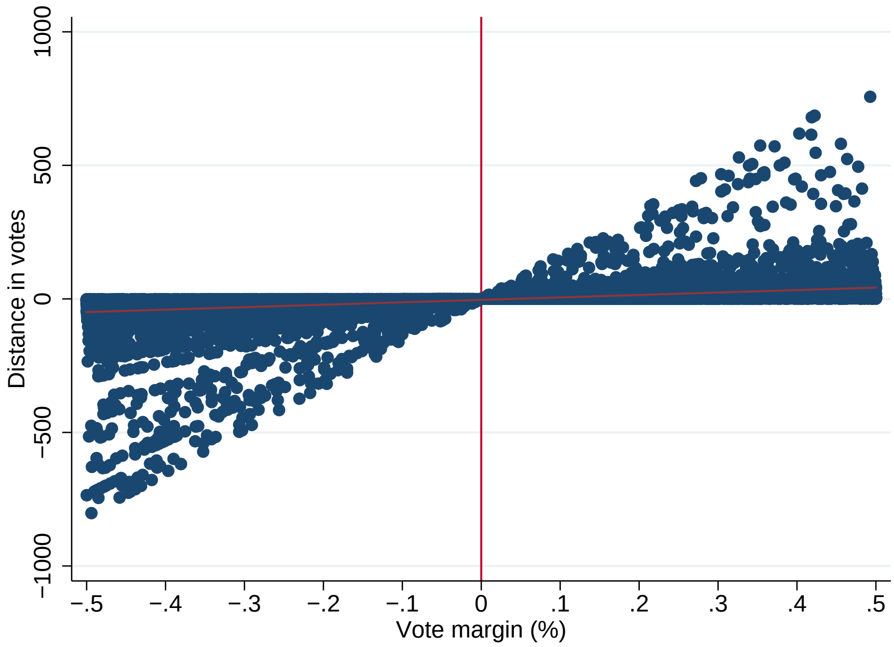
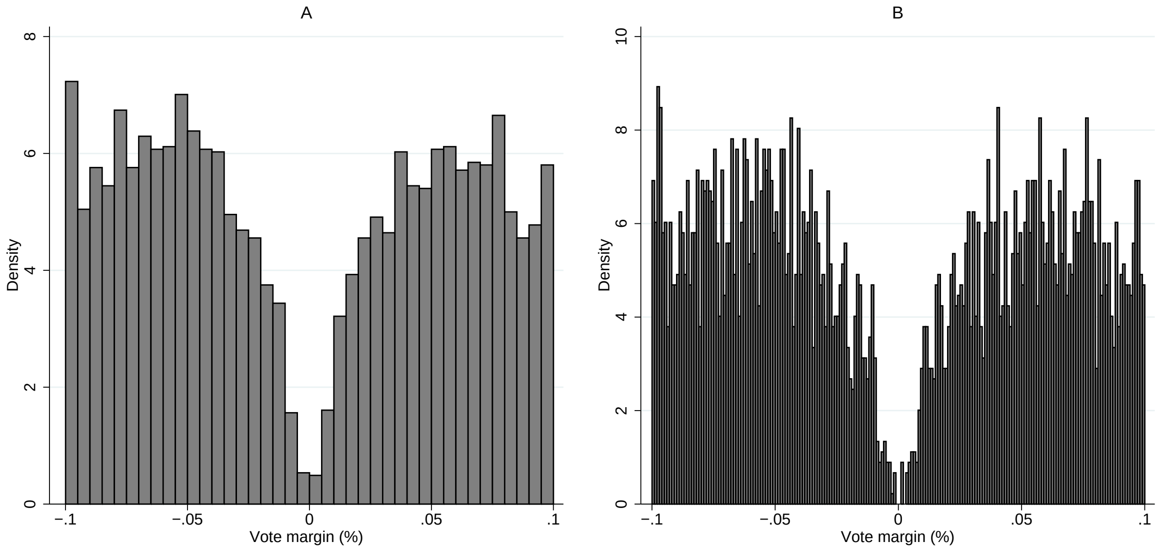


Figure C4: These histograms show the distribution of the forcing variable within two very small bandwidths nearby the RDD cutoff. The histograms suggest that the forcing variable can be treated as continuous for the purposes of RDD. The dip in the density for forcing variable values between -0.01 and 0.01 is related to the fact that the forcing variable can obtain such small values only when the party lists are large. For example, a value of 0.01 refers to one vote out of ten thousand. Lists that get more than ten thousand votes exist only in the larger municipalities.

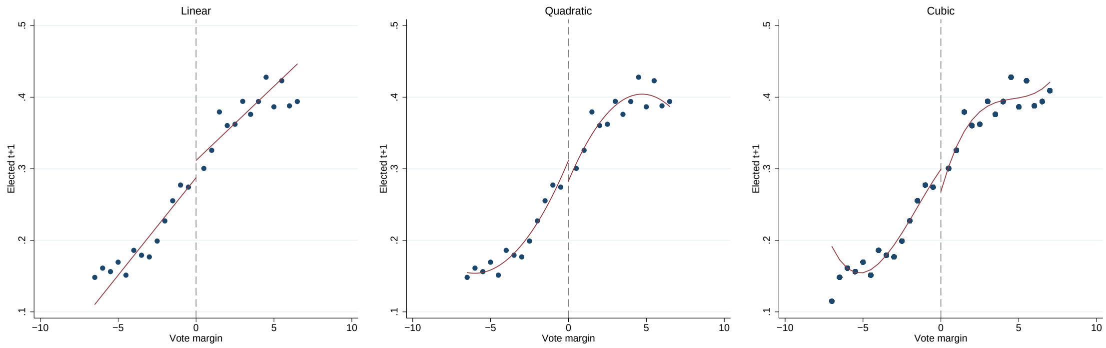
Figure C4. Histogram of the forcing variable close to the cutoff.



Notes: Figure A shows histogram of the forcing variable with bins of 0.005, and figure B uses bins of 0.001. Values of the forcing variable are limited between -0.1 and 0.1. Lotteries have been excluded.

Figure C5: These figures display RDD fit and a scatter of plot of observation bins around the cutoff when the forcing variable is defined as the (non-normalized) number of votes. The main purpose of these figures is to show that curvature issues in the relationship between the forcing variable and outcome are not unique to the way define the forcing variable. This indeed appears not to be the case: As the figures show, there is a clear jump at the cutoff in the figure on the left and evidence of curvature in the middle and on the right.

Figure C5. Curvature between the non-scaled forcing variable (number of votes) and the outcome.



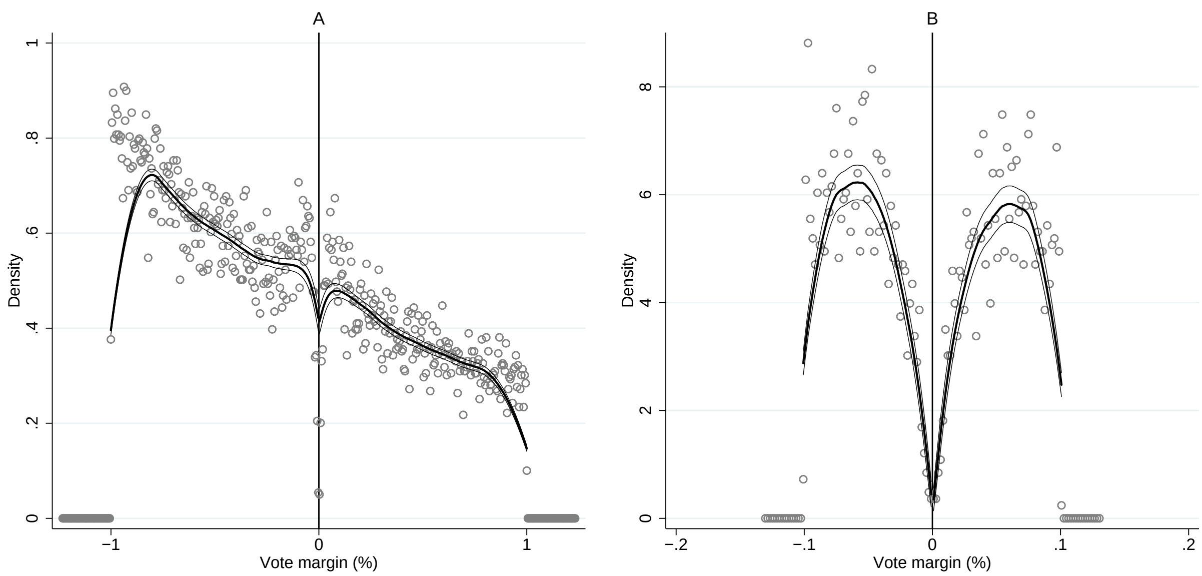
Notes: Figure shows local polynomial fits with triangular kernel within the optimal Imbens-Kalyanaraman (2012) bandwidth optimized for the linear specification. The fit is based on the local linear (Panel A), quadratic (Panel B) and cubic (Panel C) specifications. Blue dots marks bins within each value of the discrete forcing variable measured as the distance from the cutoff in the absolute number of votes (bins of 0.5).

Appendix D: Supplementary information to HMSTT Section 4.1 (RDD falsification and smoothness tests)

In this appendix, we report validity tests to for the regression discontinuity design. The standard pattern of validity tests for the RDD includes i) the McCrary (2008) manipulation test, ii) covariate balance tests, and iii) placebo tests where the location of the cutoff is artificially redefined.

Figure D1: This figure reports the McCrary (2008) tests. The test asks whether there is a jump in the amount of observations at the cutoff of getting elected. Such jump would indicate that some candidates have been able to manipulate into getting the treatment. There is no jump. The estimated difference in height is -0.0140 (standard error 0.0474) in graph A (the values of the forcing variable restricted between -1 and 1), and -0.5701 (standard error 0.6616) in graph B (the values of the forcing variable restricted between -0.1 and 0.1). This is not surprising, since there cannot be a jump in the amount of candidates elected: The number of council seats available is fixed. If one candidate is able to manipulate into getting elected, another candidate will not be elected.

Figure D1. McCrary density test.



Notes: Graph A shows the McCrary (2008) density test with the forcing variable within -1 and 1. Graph B shows the density test with forcing variable within -0.1 and 0.1.

Table D1: The main identification assumption in RDD is that covariates develop smoothly over the cutoff. The recent literature (e.g. Snyder et al. 2015 and Eggers et al. 2015) argues that especially in close election applications, balance tests based on the comparisons of means across the cutoff are likely to (wrongly) signal imbalance, because the covariates vary strongly with the forcing variable near the cutoff. One should, therefore, control for this co-variation (“slopes”) when implementing the balance tests. Panel A of Table D1 uses therefore the optimal bandwidth for the local linear specification computed for each covariate separately. When testing for covariate smoothness, bandwidth needs to be optimized for each covariate separately, because they are each unique in their relation to the forcing variable. We report in Panel B of Table D1 also the results that use half the optimal bandwidth. We do so to check how under-smoothing influences the covariance balance tests and to make sure that curvature issues (similar to those we report for our main outcome) do not lead to wrong conclusions about the covariate balance. If some of the covariates have a lot of curvature nearby the cutoff, one might wrongly infer that there is imbalance unless under-smoothing is used.

As can be seen from Panel A and B, there are some significant estimates. We cannot rule out that the few imbalances are due to multiple testing, because Panel A and B are not completely in line with each other in this regard. It is also possible that the estimated jumps are due to substantial curvature in the relationship between the given covariate and the forcing variable near the cutoff. This seems to be at least partly the case, since many of the jumps are no longer statistically significant when more flexible specifications (smaller bandwidths for a given local polynomial or higher order polynomials for a given bandwidth) are used. This means that there are fewer rejections of covariate balance when more flexible local polynomial specifications (or under-smoothing) are used.

We conclude that, taken together, the covariate balance tests provide somewhat mixed evidence. Overall, they do not cast clear doubt on the validity of RDD.

Table D1. Covariate smoothness test.

Panel A: Optimal bandwidths for covariates

	(1)	(2)	(3)		(4)	(5)	(6)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Vote share	-0.0053	-0.0094	-0.0039	Number of votes	3.0601**	2.7446	-1.5009
	(0.0088)	(0.0109)	(0.0136)		(1.1197)	(1.4207)	(3.0460)
N	27125	27125	27125	N	37155	37155	37155
Bandwidth	0.74	0.74	0.74	Bandwidth	0.99	0.99	0.99
	(7)	(8)	(9)		(10)	(11)	(12)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Female	-0.0368**	-0.0182	-0.0069	Age	0.3489	-0.2478	-0.2245
	(0.0083)	(0.0104)	(0.0130)		(0.2157)	(0.3227)	(0.4691)
N	94186	94186	94186	N	71197	71197	71197
Bandwidth	2.74	2.74	2.74	Bandwidth	1.85	1.85	1.85
	(13)	(14)	(15)		(16)	(17)	(18)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Incumbent	0.0192	0.0219	0.0230	Municipal employee	0.0002	0.0006	0.0054
	(0.0121)	(0.0194)	(0.0265)		(0.0060)	(0.0079)	(0.0101)
N	27450	27450	27450	N	107962	107962	107962
Bandwidth	0.75	0.75	0.75	Bandwidth	3.60	3.60	3.60
	(19)	(20)	(21)		(22)	(23)	(24)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Wage income	-310	1358*	1678**	Capital income	425	180	-561
	(415)	(552)	(639)		(228)	(301)	(415)
N	44971	44971	44971	N	64677	64677	64677
Bandwidth	1.96	1.96	1.96	Bandwidth	3.50	3.50	3.50
	(25)	(26)	(27)		(28)	(29)	(30)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
High professional	-0.0269**	-0.0055	0.0097	Entrepreneur	0.0156*	0.0011	-0.0106
	(0.0077)	(0.0088)	(0.0131)		(0.0076)	(0.0112)	(0.0141)
N	93022	93022	93022	N	60120	60120	60120
Bandwidth	2.69	2.69	2.69	Bandwidth	1.55	1.55	1.55
	(31)	(32)	(33)		(34)	(35)	(36)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Student	-0.0047	-0.0046	-0.0073	Unemployed	0.0038	0.0057	0.0071
	(0.0030)	(0.0039)	(0.0051)		(0.0032)	(0.0046)	(0.0060)
N	77124	77124	77124	N	78964	78964	78964
Bandwidth	2.04	2.04	2.04	Bandwidth	2.10	2.10	2.10
	(37)	(38)	(39)				
	Linear	Quadratic	Cubic				
University	-0.0168	0.0089	0.0287*				
	(0.0097)	(0.0096)	(0.0137)				
N	71648	71648	71648				
Bandwidth	2.45	2.45	2.45				

Panel B: 0.5 * Optimal bandwidths for covariates

	(79)	(80)	(81)		(82)	(83)	(84)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Vote share	0.0000	0.0033	-0.0027	Number of votes	0.6235	-0.2545	2.5474
	(0.0091)	(0.0134)	(0.0195)		(1.4369)	(1.4309)	(5.7319)
N	13555	13555	13555	N	18077	18077	18077
Bandwidth	0.19	0.19	0.19	Bandwidth	0.25	0.25	0.25
	(85)	(86)	(87)		(88)	(89)	(90)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Female	-0.0195*	-0.0093	-0.0137	Age	-0.0958	-0.0945	-0.2721
	(0.0097)	(0.0143)	(0.0194)		(0.3160)	(0.4752)	(0.6505)
N	52884	52884	52884	N	34347	34347	34347
Bandwidth	0.69	0.69	0.69	Bandwidth	0.46	0.46	0.46
	(91)	(92)	(93)		(94)	(95)	(96)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Incumbent	0.0219	0.0188	0.0106	Municipal employee	0.0024	0.0074	0.0014
	(0.0181)	(0.0296)	(0.0451)		(0.0073)	(0.0106)	(0.0144)
N	13686	13686	13686	N	69386	69386	69386
Bandwidth	0.19	0.19	0.19	Bandwidth	0.90	0.90	0.90
	(97)	(98)	(99)		(100)	(101)	(102)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Wage income	979*	1692**	1312	Capital income	1	-439	-126
	(441)	(637)	(830)		(281)	(417)	(597)
N	21842	21842	21842	N	40606	40606	40606
Bandwidth	0.49	0.49	0.49	Bandwidth	0.88	0.88	0.88
	(103)	(104)	(105)		(106)	(107)	(108)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
High professional	-0.0076	0.0066	0.0104	Entrepreneur	0.0012	-0.0112	-0.0183
	(0.0085)	(0.0133)	(0.0161)		(0.0106)	(0.0146)	(0.0199)
N	51758	51758	51758	N	28379	28379	28379
Bandwidth	0.67	0.67	0.67	Bandwidth	0.39	0.39	0.39
	(109)	(110)	(111)		(112)	(113)	(114)
	Linear	Quadratic	Cubic		Linear	Quadratic	Cubic
Student	-0.0046	-0.0088	-0.0159*	Unemployed	0.0059	0.0077	0.0112
	(0.0037)	(0.0054)	(0.0070)		(0.0044)	(0.0064)	(0.0083)
N	38230	38230	38230	N	39557	39557	39557
Bandwidth	0.51	0.51	0.51	Bandwidth	0.53	0.53	0.53
	(115)	(116)	(117)				
	Linear	Quadratic	Cubic				
University	0.0072	0.0287*	0.0244				
	(0.0089)	(0.0135)	(0.0181)				
N	38404	38404	38404				
Bandwidth	0.61	0.61	0.61				

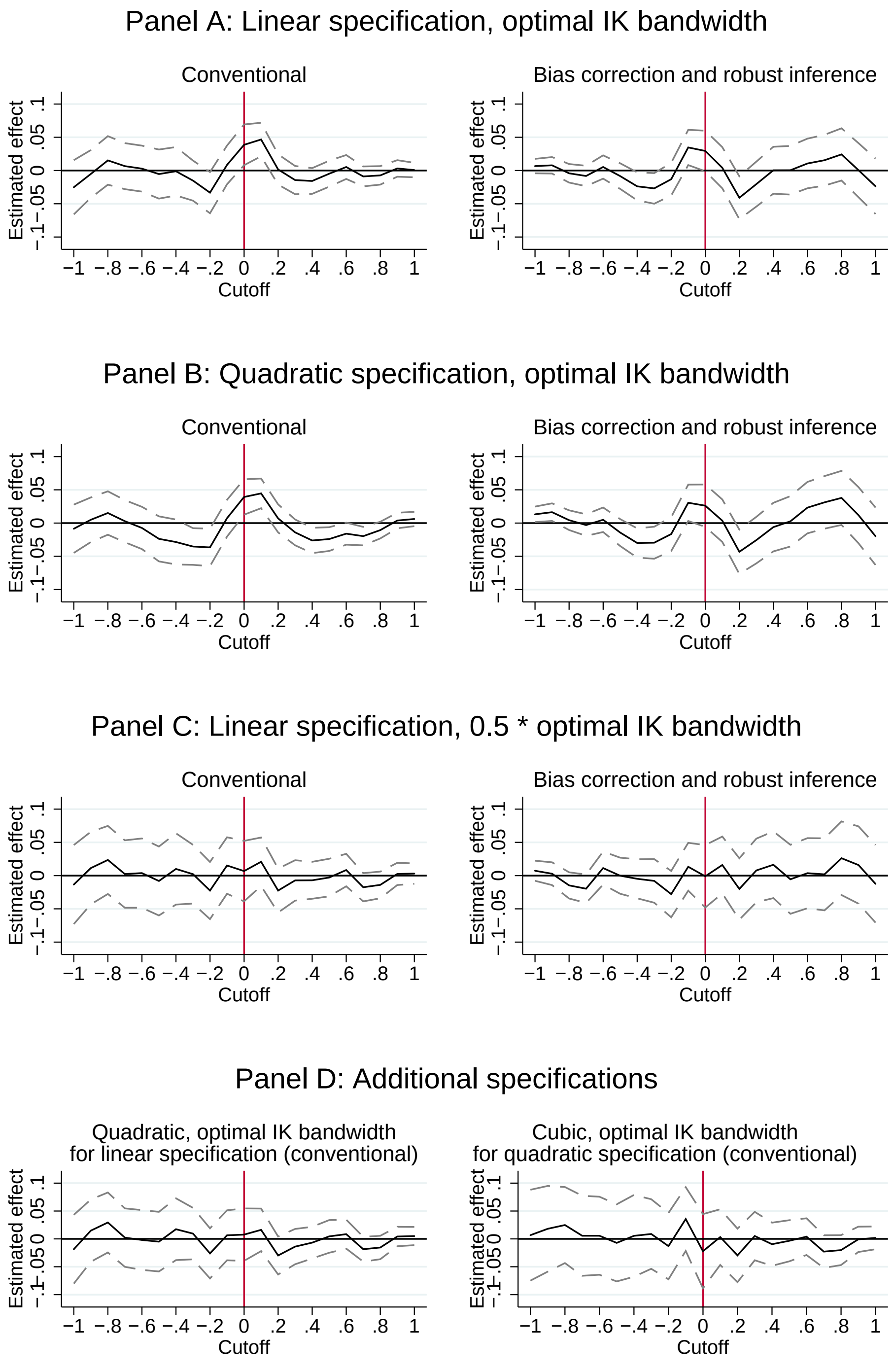
Notes: Panel A shows estimated discontinuities in covariates using local polynomial regressions within the optimal IK bandwidth. Panel B uses bandwidth half of the optimal. Bandwidth has been chosen for local linear specification. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t*.

Figure D2: Figure D2 reports a series of placebo tests where the location of the cutoff is artificially redefined. If there are jumps in locations other than the true cutoff, it would suggest that strong nonlinearities or discontinuities in the relationship between the forcing variable and the outcome may be driving the RDD result (instead of a causal effect at the cutoff). Typically, these tests are used in applications where there is a documented effect at the cutoff (that is statistically different from zero) and the researcher wants to show that this statistically significant jump is unique (or, at least, that only 5% of the placebo cutoffs show jumps that are significant at the 5% level).

In Panel A and B, we display the placebo RDD estimates that are based on the conventional local linear and quadratic specification, using the corresponding IK optimal bandwidths. As we report in the main text, the RDD estimates produced by these specifications indicate that there would be a positive jump at the true cutoff. This is in contrast to what our experimental estimate suggests. As the placebo estimates on the left of these panels show, there also are statistically significant jumps at some of the placebo cutoffs located close by the true cutoff. Some of these jumps are even larger than the one found at the true cutoff. These placebo tests are thus indicative of these RDD specifications not working properly. The placebo graphs on the right have been produced using the same specifications as on the left, but with the CCT-correction. They, too, are indicative of these specifications not working as expected.

In Panels C and D, we explore whether those RDD specifications that in our context seem to work are problematic in the light of the placebo tests. Panel C reports the results for half the optimal (IK) bandwidths: On the left, we use the conventional local linear specification for this under-smoothing approach. The corresponding estimates based on the CCT-correction are displayed on the right. In Panel D we explore whether a polynomial of order p+1 is flexible enough for the bandwidth that has been optimized for a polynomial of order p. The panel reports these results for the quadratic and cubic local polynomials. As the two panels show, there are no jumps at any of the placebo cutoffs, implying that these specifications work appropriately. In sum, the placebo tests reported in Panel C and D do suggest that the under-smoothing procedure or the use of higher degree local polynomials without adjusting the bandwidth accordingly work. These findings thus suggest that the placebo cutoff tests seem to be of use in detecting too inflexible specifications.

Figure D2. RDD estimates at the artificial cutoffs.



Notes: The figure shows the RDD point estimates and the 95% confidence intervals from specifications using local polynomial regression with a triangular kernel. All the left hand graphs and also the right hand graph in Panel D use conventional approach with optimal IK bandwidths and confidence intervals constructed using standard errors clustered by municipality. All the right hand graphs in Panels A-C use IK bandwidth and bias-correction and robust inference by Calonico et al. (2014a). We report the results at various artificial (placebo) cutoffs where the location of the artificial cutoff relative to the true cutoff is reported in the x-axis. In Panel A, bandwidth is optimized for the linear specification, In Panel B, bandwidth is half the one in Panel A and in Panel C, bandwidth is optimized for the quadratic specification. In Panel D, bandwidth is optimized for p-order polynomial specification whereas the fit is based on p+1 order. Optimal bandwidth is based on the specification and sample at the real cutoff. Vertical red line marks the real cutoff.

Appendix E: Supplementary information to HMSTT Section 4.2 (When is RDD as good as randomly assigned?)

This appendix reports means tests of covariate balance within small bandwidths near the cutoff. These tests do not control for the slopes (or curvature) of the forcing variable nearby the cutoff. They are not tests of whether the covariates develop smoothly over the cutoff, but rather tests for whether the treatment is as good as randomly assigned.

Table E1: This table looks at the covariate balance of candidate characteristics. It reports the means of the candidate characteristics for small bandwidths on both sides of the cutoff as well as a *t*-test for the difference of the means. As we report in the main text, when incumbency status (elected at *t*-1) is used, we find that bandwidths 0.04 or smaller are as-good-as-random at the 5% significance level (923 observations). Based on the minimum *p*-value among all the covariates (but not correcting for multiple testing), it seems that bandwidths 0.02 or smaller would be as-good-as random at the 5% significance level (128 observations). These numbers are obtained by starting from the zero bandwidth and widening the bandwidth until the first statistically significant coefficient is found. This is a conservative approach in the sense that if we started from wider bandwidths and decreased their length until no significant differences are found, we would get somewhat larger bandwidth estimates. For example, based on Table E2, a bandwidth of 0.05 would be as-good-as-random (but 0.10 or larger would not).

Table E2: This table reproduces the analysis of Table E1 for municipality-level covariates. As the table shows, they are balanced, as they should be by construction.

Table E1. Covariate balance within small bandwidths (candidate characteristics).

Bandwidth = 0.01

Variable	Elected (N = 37) N	Mean	Std. dev.	Not elected (N = 38) N	Mean	Std. dev.	Difference
Vote share	37	0.49	0.20	38	0.46	0.18	0.02
Number of votes	37	291	230	38	282	208	9
Female	37	0.49	0.51	38	0.47	0.51	0.01
Age	37	47.62	9.91	38	49.74	11.86	-2.12
Incumbent	37	0.41	0.50	38	0.55	0.50	-0.15
Municipal employee	37	0.43	0.50	38	0.37	0.49	0.06
Wage income	22	32287	14415	19	30766	18607	1521
Capital income	22	1769	7168	19	1039	2937	729
High professional	37	0.49	0.51	38	0.34	0.48	0.14
Entrepreneur	37	0.08	0.28	38	0.05	0.23	0.03
Student	37	0.00	0.00	38	0.03	0.16	-0.03
Unemployed	37	0.03	0.16	38	0.00	0.00	0.03
University degree	33	0.30	0.47	32	0.22	0.42	0.08

Bandwidth = 0.05

Variable	Elected (N = 729) N	Mean	Std. dev.	Not elected (N = 761) N	Mean	Std. dev.	Difference
Vote share	729	0.90	0.44	761	0.85	0.43	0.05
Number of votes	729	102	111	761	102	111	0
Female	729	0.41	0.49	761	0.37	0.48	0.04
Age	729	47.04	11.79	761	47.01	12.43	0.03
Incumbent	729	0.39	0.49	761	0.34	0.48	0.05
Municipal employee	729	0.30	0.46	761	0.26	0.44	0.03
Wage income	424	23779	14561	419	23257	14778	522
Capital income	424	2349	19179	419	1560	5466	790
High professional	729	0.26	0.44	761	0.27	0.44	-0.01
Entrepreneur	729	0.17	0.37	761	0.18	0.38	-0.01
Student	729	0.02	0.15	761	0.03	0.17	-0.01
Unemployed	729	0.05	0.21	761	0.03	0.17	0.02
University degree	597	0.19	0.39	641	0.17	0.38	0.02

Bandwidth = 0.10

Variable	Elected (N = 1778) N	Mean	Std. dev.	Not elected (N = 1906) N	Mean	Std. dev.	Difference
Vote share	1778	1.02	0.53	1906	0.95	0.50	0.07
Number of votes	1778	90	104	1906	87	101	3
Female	1778	0.38	0.49	1906	0.39	0.49	0.00
Age	1778	46.45	11.79	1906	46.47	11.94	-0.03
Incumbent	1778	0.36	0.48	1906	0.31	0.46	0.05**
Municipal employee	1778	0.27	0.45	1906	0.27	0.44	0.01
Wage income	1015	23629	15358	1082	22551	14898	1079
Capital income	1015	2152	13581	1082	2092	12732	60
High professional	1778	0.25	0.43	1906	0.25	0.43	0.00
Entrepreneur	1778	0.19	0.39	1906	0.19	0.39	0.00
Student	1778	0.02	0.14	1906	0.03	0.18	-0.01**
Unemployed	1778	0.04	0.20	1906	0.03	0.18	0.01
University degree	1466	0.19	0.39	1576	0.17	0.38	0.02

Bandwidth = 0.20

Variable	Elected (N = 3508) N	Mean	Std. dev.	Not elected (N = 3891) N	Mean	Std. dev.	Difference
Vote share	3508	1.11	0.58	3891	0.98	0.54	0.12
Number of votes	3508	85	101	3891	81	96	5
Female	3508	0.36	0.48	3891	0.40	0.49	-0.03**
Age	3508	46.51	11.59	3891	46.17	12.02	0.34
Incumbent	3508	0.36	0.48	3891	0.29	0.45	0.07**
Municipal employee	3508	0.26	0.44	3891	0.27	0.44	0.00
Wage income	2053	23214	15043	2284	22068	14576	1146
Capital income	2053	2032	10591	2284	2139	10670	-107
High professional	3508	0.24	0.43	3889	0.24	0.43	0.00
Entrepreneur	3508	0.21	0.40	3889	0.19	0.39	0.02
Student	3508	0.02	0.15	3889	0.03	0.18	-0.01*
Unemployed	3508	0.04	0.20	3889	0.04	0.19	0.00
University degree	2858	0.18	0.39	3201	0.17	0.38	0.01

Bandwidth = 0.30

Variable	Elected (N = 5172) N	Mean	Std. dev.	Not elected (N = 5879) N	Mean	Std. dev.	Difference
Vote share	5172	1.16	0.62	5879	0.98	0.56	0.18*
Number of votes	5172	86	109	5879	77	91	9
Female	5172	0.36	0.48	5879	0.39	0.49	-0.03**
Age	5172	46.55	11.66	5879	46.28	12.05	0.27
Incumbent	5172	0.38	0.49	5879	0.28	0.45	0.10*
Municipal employee	5172	0.26	0.44	5879	0.26	0.44	0.00
Wage income	3026	23005	15161	3453	22287	14672	718
Capital income	3026	2222	10719	3453	2090	10650	132
High professional	5172	0.23	0.42	5877	0.24	0.43	-0.01
Entrepreneur	5172	0.21	0.41	5877	0.19	0.39	0.03
Student	5172	0.02	0.15	5877	0.03	0.17	-0.01
Unemployed	5172	0.04	0.20	5877	0.04	0.20	0.00
University degree	4232	0.18	0.38	4832	0.17	0.38	0.01

Bandwidth = 0.40

Variable	Elected (N = 6628) N	Mean	Std. dev.	Not elected (N = 7949) N	Mean	Std. dev.	Difference
Vote share	6628	1.19	0.63	7949	0.95	0.56	0.24**
Number of votes	6628	87	114	7949	76	89	11
Female	6628	0.36	0.48	7949	0.39	0.49	-0.03**
Age	6628	46.74	11.62	7949	46.19	12.18	0.55
Incumbent	6628	0.39	0.49	7949	0.26	0.44	0.13**
Municipal employee	6628	0.26	0.44	7949	0.25	0.43	0.01
Wage income	3933	23038	16227	4711	22171	14583	867
Capital income	3933	2283	10198	4711	2051	10129	232
High professional	6628	0.23	0.42	7947	0.24	0.42	-0.01
Entrepreneur	6628	0.21	0.41	7947	0.18	0.38	0.03*
Student	6628	0.02	0.15	7947	0.03	0.17	-0.01*
Unemployed	6628	0.04	0.20	7947	0.04	0.20	0.00
University degree	5402	0.18	0.38	6534	0.17	0.38	0.00

Bandwidth = 0.55

Variable	Elected (N = 8710) N	Mean	Std. dev.	Not elected (N = 11348) N	Mean	Std. dev.	Difference
Vote share	8710	1.23	0.65	11348	0.92	0.56	0.32**
Number of votes	8710	87	113	11348	74	85	13
Female	8710	0.37	0.48	11348	0.40	0.49	-0.03**
Age	8710	46.77	11.56	11348	46.06	12.22	0.71*
Incumbent	8710	0.40	0.49	11348	0.24	0.43	0.16**
Municipal employee	8710	0.26	0.44	11348	0.25	0.43	0.01
Wage income	5132	22752	15775	6786	22342	14972	410
Capital income	5132	2444	15064	6786	2112	13538	332
High professional	8710	0.23	0.42	11345	0.24	0.42	-0.01
Entrepreneur	8710	0.22	0.41	11345	0.17	0.38	0.05**
Student	8710	0.02	0.15	11345	0.03	0.18	-0.01**
Unemployed	8710	0.04	0.20	11345	0.04	0.21	0.00
University degree	7091	0.17	0.38	9323	0.18	0.38	0.00

Bandwidth = 0.70

Variable	Elected (N = 10632) N	Mean	Std. dev.	Not elected (N = 14888) N	Mean	Std. dev.	Difference
Vote share	10632	1.27	0.66	14888	0.87	0.56	0.40**
Number of votes	10632	87	112	14888	72	80	15
Female	10632	0.36	0.48	14888	0.40	0.49	-0.04**
Age	10632	46.84	11.50	14888	45.89	12.27	0.95**
Incumbent	10632	0.42	0.49	14888	0.21	0.41	0.21**
Municipal employee	10632	0.26	0.44	14888	0.25	0.43	0.01
Wage income	6267	22702	15692	8955	22320	15013	383
Capital income	6267	2469	14437	8955	1986	12443	483*
High professional	10631	0.22	0.42	14884	0.24	0.43	-0.01
Entrepreneur	10631	0.22	0.42	14884	0.17	0.38	0.06**
Student	10631	0.02	0.15	14884	0.03	0.18	-0.01**
Unemployed	10631	0.04	0.19	14884	0.04	0.21	-0.01
University degree	8634	0.17	0.37	12252	0.18	0.38	-0.01

Bandwidth = 1.00

Variable	Elected (N = 14138) N	Mean	Std. dev.	Not elected (N = 23320) N	Mean	Std. dev.	Difference
Vote share	14138	1.34	0.69	23320	0.76	0.56	0.58**
Number of votes	14138	88	114	23320	66	71	22*
Female	14138	0.36	0.48	23320	0.41	0.49	-0.05**
Age	14138	47.01	11.40	23320	45.56	12.29	1.44**
Incumbent	14138	0.44	0.50	23320	0.17	0.37	0.27**
Municipal employee	14138	0.26	0.44	23320	0.25	0.43	0.01
Wage income	8322	22738	15865	14023	22077	14841	661
Capital income	8322	2738	17936	14023	1769	11484	969**
High professional	14136	0.22	0.42	23314	0.24	0.42	-0.01
Entrepreneur	14136	0.23	0.42	23314	0.16	0.37	0.07**
Student	14136	0.02	0.15	23314	0.04	0.19	-0.01**
Unemployed	14136	0.04	0.19	23314	0.05	0.21	-0.01**
University degree	11490	0.17	0.38	19252	0.18	0.38	-0.01

Notes: * and ** denote 5 % and 1 % statistical significance of difference in means respectively. The significance of differences is tested using t test adjusted for clustering by municipality. 0.55 bandwidth equals roughly the optimal bandwidth chosen using Imbens and Kalyanaraman's (2008) algorithm. Sample includes only candidates running in 1996-2008 elections. Lotteries have been excluded. In 1996 elections income data are available only for candidates who run also in 2000, 2004 and 2008 elections. Income is expressed in euros.

Table E1 (continued). Covariate balance within small bandwidths (candidate characteristics).

Bandwidth = 0.01

	Elected (N = 37)			Not elected (N = 38)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	37	0.32	0.47	38	0.34	0.48	-0.02
Social Democrats	37	0.32	0.47	38	0.32	0.47	0.01
Center Party	37	0.11	0.31	38	0.11	0.31	0.00
True Finns	37	0.00	0.00	38	0.00	0.00	0.00
Green Party	37	0.11	0.31	38	0.11	0.31	0.00
Socialist Party	37	0.08	0.28	38	0.08	0.27	0.00
Swedish Party	37	0.05	0.23	38	0.05	0.23	0.00
Christian Party	37	0.00	0.00	38	0.00	0.00	0.00
Other parties	37	0.00	0.00	38	0.00	0.00	0.00

Bandwidth = 0.05

	Elected (N = 729)			Not elected (N = 761)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	729	0.18	0.38	761	0.18	0.38	0.00
Social Democrats	729	0.27	0.44	761	0.28	0.45	-0.01
Center Party	729	0.40	0.49	761	0.40	0.49	0.01
True Finns	729	0.00	0.04	761	0.00	0.04	0.00
Green Party	729	0.02	0.14	761	0.02	0.14	0.00
Socialist Party	729	0.04	0.20	761	0.04	0.19	0.00
Swedish Party	729	0.07	0.26	761	0.07	0.25	0.00
Christian Party	729	0.00	0.05	761	0.00	0.06	0.00
Other parties	729	0.02	0.12	761	0.01	0.12	0.00

Bandwidth = 0.10

	Elected (N = 1778)			Not elected (N = 1906)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	1778	0.19	0.39	1906	0.18	0.39	0.00
Social Democrats	1778	0.24	0.43	1906	0.25	0.43	0.00
Center Party	1778	0.41	0.49	1906	0.41	0.49	0.00
True Finns	1778	0.00	0.06	1906	0.00	0.06	0.00
Green Party	1778	0.02	0.15	1906	0.02	0.15	0.00
Socialist Party	1778	0.06	0.23	1906	0.05	0.22	0.01
Swedish Party	1778	0.06	0.23	1906	0.05	0.23	0.00
Christian Party	1778	0.00	0.06	1906	0.00	0.07	0.00
Other parties	1778	0.02	0.14	1906	0.02	0.14	0.00

Bandwidth = 0.20

	Elected (N = 3508)			Not elected (N = 3891)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	3508	0.18	0.39	3891	0.19	0.39	-0.01
Social Democrats	3508	0.24	0.43	3891	0.25	0.43	-0.01
Center Party	3508	0.41	0.49	3891	0.39	0.49	0.01
True Finns	3508	0.01	0.07	3891	0.01	0.07	0.00
Green Party	3508	0.02	0.15	3891	0.02	0.15	0.00
Socialist Party	3508	0.06	0.23	3891	0.05	0.23	0.00
Swedish Party	3508	0.06	0.24	3891	0.06	0.24	0.00
Christian Party	3508	0.01	0.08	3891	0.01	0.08	0.00
Other parties	3508	0.02	0.14	3891	0.02	0.14	0.00

Bandwidth = 0.30

	Elected (N = 5172)			Not elected (N = 5879)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	5172	0.19	0.39	5879	0.19	0.40	-0.01
Social Democrats	5172	0.24	0.42	5879	0.25	0.43	-0.01
Center Party	5172	0.40	0.49	5879	0.38	0.49	0.02
True Finns	5172	0.01	0.08	5879	0.01	0.08	0.00
Green Party	5172	0.02	0.15	5879	0.02	0.15	0.00
Socialist Party	5172	0.06	0.23	5879	0.06	0.23	0.00
Swedish Party	5172	0.06	0.24	5879	0.06	0.23	0.00
Christian Party	5172	0.01	0.09	5879	0.01	0.10	0.00
Other parties	5172	0.02	0.14	5879	0.02	0.15	0.00

Bandwidth = 0.40

	Elected (N = 6628)			Not elected (N = 7949)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	6628	0.18	0.39	7949	0.19	0.39	-0.01
Social Democrats	6628	0.23	0.42	7949	0.26	0.44	-0.02
Center Party	6628	0.41	0.49	7949	0.38	0.48	0.03
True Finns	6628	0.01	0.09	7949	0.01	0.08	0.00
Green Party	6628	0.02	0.14	7949	0.03	0.16	0.00
Socialist Party	6628	0.06	0.24	7949	0.06	0.23	0.00
Swedish Party	6628	0.06	0.23	7949	0.05	0.23	0.00
Christian Party	6628	0.01	0.10	7949	0.01	0.10	0.00
Other parties	6628	0.02	0.15	7949	0.02	0.15	0.00

Bandwidth = 0.55

	Elected (N = 8710)			Not elected (N = 11348)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	8710	0.18	0.39	11348	0.20	0.40	-0.01
Social Democrats	8710	0.23	0.42	11348	0.26	0.44	-0.03
Center Party	8710	0.40	0.49	11348	0.36	0.48	0.04
True Finns	8710	0.01	0.09	11348	0.01	0.09	0.00
Green Party	8710	0.02	0.14	11348	0.03	0.16	-0.01
Socialist Party	8710	0.06	0.24	11348	0.06	0.23	0.00
Swedish Party	8710	0.06	0.24	11348	0.05	0.22	0.01
Christian Party	8710	0.01	0.10	11348	0.01	0.10	0.00
Other parties	8710	0.02	0.15	11348	0.02	0.15	0.00

Bandwidth = 0.70

	Elected (N = 10632)			Not elected (N = 14888)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	10632	0.18	0.38	14888	0.20	0.40	-0.02
Social Democrats	10632	0.23	0.42	14888	0.27	0.44	0.04
Center Party	10632	0.41	0.49	14888	0.35	0.48	0.06
True Finns	10632	0.01	0.09	14888	0.01	0.08	0.00
Green Party	10632	0.02	0.14	14888	0.03	0.17	-0.01
Socialist Party	10632	0.06	0.24	14888	0.06	0.23	0.01
Swedish Party	10632	0.06	0.24	14888	0.05	0.22	0.01
Christian Party	10632	0.01	0.10	14888	0.01	0.10	0.00
Other parties	10632	0.02	0.15	14888	0.02	0.15	0.00

Bandwidth = 1.00

	Elected (N = 14138)			Not elected (N = 23320)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Coalition Party	14138	0.17	0.38	23320	0.21	0.41	-0.04
Social Democrats	14138	0.22	0.42	23320	0.29	0.45	-0.06*
Center Party	14138	0.42	0.49	23320	0.33	0.47	0.09*
True Finns	14138	0.01	0.09	23320	0.01	0.09	0.00
Green Party	14138	0.02	0.14	23320	0.03	0.18	-0.01
Socialist Party	14138	0.06	0.24	23320	0.06	0.23	0.01
Swedish Party	14138	0.06	0.24	23320	0.05	0.22	0.01
Christian Party	14138	0.01	0.11	23320	0.01	0.09	0.00
Other parties	14138	0.02	0.15	23320	0.02	0.14	0.00

Notes: * and ** denote 5 % and 1 % statistical significance of difference in means respectively. The significance of differences is tested using t test adjusted for clustering by municipality. 0.55 bandwidth equals roughly the optimal bandwidth chosen using Imbens and Kalyanaraman's (2008) algorithm. Sample includes only candidates running in 1996-2008 elections. Lotteries have been excluded. In 1996 elections income data are available only for candidates who run also in 2000, 2004 and 2008 elections. Income is expressed in euros.

Table E2. Covariate balance within small bandwidths (municipality characteristics).

Bandwidth = 0.01							
	Elected (N = 37)			Not elected (N = 38)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	37	85715	91807	38	81938	84156	3777
Coalition Party seat share	37	25.45	8.06	38	25.88	7.72	-0.42
Social Democrats seat share	37	24.72	7.91	38	24.62	7.11	0.11
Center Party seat share	37	13.84	17.27	38	12.52	15.01	1.32
True Finns seat share	37	1.68	3.14	38	1.46	2.71	0.21
Green Party seat share	37	10.72	7.85	38	10.48	7.44	0.24
Socialist Party seat share	37	9.68	6.17	38	10.19	6.11	-0.50
Swedish Party seat share	37	6.99	14.64	38	7.41	14.76	-0.41
Christian Party seat share	37	3.47	1.83	38	3.56	1.73	-0.08
Other parties' seat share	37	3.46	4.57	38	3.93	4.64	-0.47
Voter turnout	36	57.33	5.43	37	57.17	5.10	0.16
Population	37	190585	204484	38	182109	190372	8477
Share of 0-14-year-olds	36	16.85	2.72	37	16.73	2.39	0.12
Share of 15-64-year-olds	36	68.61	2.55	37	68.75	2.27	-0.14
Share of over-65-year-olds	36	14.54	3.01	37	14.52	2.68	0.02
Income per capita	36	23360	6682	37	23337	6347	23
Unemployment	37	13.63	7.06	38	13.34	6.58	0.29

Bandwidth = 0.05							
	Elected (N = 729)			Not elected (N = 761)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	729	20538	38928	761	21816	40254	-1278
Coalition Party seat share	729	19.88	10.11	761	20.28	10.02	-0.40
Social Democrats seat share	729	23.57	11.09	761	23.80	10.72	-0.23
Center Party seat share	729	27.84	21.17	761	27.21	20.81	0.63
True Finns seat share	729	1.67	3.13	761	1.70	3.41	-0.03
Green Party seat share	729	4.73	5.32	761	4.90	5.37	-0.16
Socialist Party seat share	729	9.42	7.76	761	9.25	7.46	0.17
Swedish Party seat share	729	5.58	17.11	761	5.61	16.78	-0.04
Christian Party seat share	729	3.72	3.47	761	3.92	3.73	-0.20
Other parties' seat share	729	3.34	6.16	761	3.22	5.97	0.11
Voter turnout	711	60.49	6.20	746	60.27	6.40	0.22
Population	720	44984	87091	750	48014	89200	-3030
Share of 0-14-year-olds	711	18.25	3.38	741	18.13	3.22	0.12
Share of 15-64-year-olds	711	65.48	3.10	741	65.64	3.09	-0.16
Share of over-65-year-olds	711	16.27	4.10	741	16.23	4.12	0.04
Income per capita	711	20848	5875	741	21072	5823	-224
Unemployment	720	14.05	6.10	750	13.79	5.96	0.27

Bandwidth = 0.10							
	Elected (N = 1778)			Not elected (N = 1906)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	1778	17428	36257	1906	18328	37762	-901
Coalition Party seat share	1778	19.47	10.24	1906	19.68	10.22	-0.20
Social Democrats seat share	1778	22.61	10.89	1906	22.80	10.65	-0.19
Center Party seat share	1778	30.78	21.57	1906	30.38	21.31	0.40
True Finns seat share	1778	1.78	3.63	1906	1.86	3.81	-0.08
Green Party seat share	1778	4.14	5.10	1906	4.20	5.20	-0.07
Socialist Party seat share	1778	9.03	7.82	1906	8.84	7.66	0.19
Swedish Party seat share	1778	4.99	15.85	1906	5.03	15.87	-0.04
Christian Party seat share	1778	3.54	3.60	1906	3.70	3.74	-0.16
Other parties' seat share	1778	3.52	6.39	1906	3.43	6.23	0.09
Voter turnout	1747	61.36	6.41	1872	61.38	6.47	-0.02
Population	1765	37731	79644	1885	39837	82836	-2105
Share of 0-14-year-olds	1750	18.34	3.31	1871	18.25	3.25	0.09
Share of 15-64-year-olds	1750	64.99	3.20	1871	65.05	3.22	-0.06
Share of over-65-year-olds	1750	16.67	4.36	1871	16.70	4.38	-0.03
Income per capita	1750	20478	5760	1871	20594	5769	-116
Unemployment	1765	13.91	6.00	1885	13.75	5.94	0.16

Bandwidth = 0.20							
	Elected (N = 3508)			Not elected (N = 3891)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	3508	15371	33726	3891	17537	38594	-2166
Coalition Party seat share	3508	19.01	10.43	3891	19.25	10.42	-0.24
Social Democrats seat share	3508	22.31	10.97	3891	22.64	10.87	-0.33
Center Party seat share	3508	31.80	21.41	3891	30.83	21.29	0.97
True Finns seat share	3508	1.65	3.48	3891	1.73	3.72	-0.08
Green Party seat share	3508	3.79	4.95	3891	3.99	5.13	-0.20
Socialist Party seat share	3508	8.82	7.84	3891	8.88	7.77	-0.06
Swedish Party seat share	3508	5.31	16.44	3891	5.38	16.51	-0.08
Christian Party seat share	3508	3.49	3.78	3891	3.52	3.72	-0.03
Other parties' seat share	3508	3.56	6.61	3891	3.62	6.65	-0.06
Voter turnout	3455	61.93	6.48	3833	61.76	6.47	0.18
Population	3490	33380	74698	3865	38006	84731	-4626
Share of 0-14-year-olds	3462	18.38	3.25	3836	18.28	3.23	0.10
Share of 15-64-year-olds	3462	64.73	3.20	3836	64.89	3.24	-0.16
Share of over-65-year-olds	3462	16.89	4.42	3836	16.84	4.41	0.06
Income per capita	3462	20213	5801	3836	20328	5637	-115
Unemployment	3490	13.98	6.06	3865	13.93	6.03	0.05

Bandwidth = 0.30							
	Elected (N = 5172)			Not elected (N = 5879)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	5172	15217	34682	5879	17398	38557	-2180
Coalition Party seat share	5172	18.89	10.39	5879	19.24	10.40	-0.35
Social Democrats seat share	5172	22.15	10.93	5879	22.59	10.89	-0.44
Center Party seat share	5172	32.38	21.35	5879	31.12	21.28	1.26
True Finns seat share	5172	1.66	3.54	5879	1.72	3.70	-0.06
Green Party seat share	5172	3.69	4.93	5879	3.93	5.09	-0.24
Socialist Party seat share	5172	8.69	7.80	5879	8.79	7.75	-0.11
Swedish Party seat share	5172	5.36	16.74	5879	5.32	16.39	0.04
Christian Party seat share	5172	3.46	3.72	5879	3.49	3.65	-0.03
Other parties' seat share	5172	3.53	6.57	5879	3.64	6.63	-0.11
Voter turnout	5100	62.12	6.53	5799	61.87	6.53	0.25
Population	5145	33041	76681	5842	37734	84822	-4693
Share of 0-14-year-olds	5103	18.34	3.25	5798	18.26	3.22	0.08
Share of 15-64-year-olds	5103	64.64	3.23	5798	64.83	3.28	-0.19
Share of over-65-year-olds	5103	17.02	4.48	5798	16.90	4.45	0.11
Income per capita	5103	20078	5719	5798	20259	5655	-181
Unemployment	5145	13.99	6.05	5842	13.97	6.01	0.02

Bandwidth = 0.40							
	Elected (N = 6628)			Not elected (N = 7949)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	6628	15007	35009	7949	18142	40267	-3135
Coalition Party seat share	6628	18.62	10.40	7949	19.30	10.37	-0.68
Social Democrats seat share	6628	21.89	10.94	7949	22.74	10.91	-0.84
Center Party seat share	6628	33.09	21.44	7949	31.06	21.28	2.03
True Finns seat share	6628	1.68	3.60	7949	1.70	3.68	-0.02
Green Party seat share	6628	3.60	4.92	7949	4.01	5.19	-0.41
Socialist Party seat share	6628	8.68	7.81	7949	8.80	7.69	-0.12
Swedish Party seat share	6628	5.28	16.68	7949	5.10	15.90	0.18
Christian Party seat share	6628	3.43	3.71	7949	3.51	3.64	-0.07
Other parties' seat share	6628	3.53	6.67	7949	3.64	6.69	-0.11
Voter turnout	6536	62.26	6.54	7836	61.82	6.51	0.44
Population	6597	32631	77601	7899	39354	88720	-6722
Share of 0-14-year-olds	6543	18.38	3.27	7838	18.25	3.22	0.12
Share of 15-64-year-olds	6543	64.58	3.24	7838	64.87	3.28	-0.30
Share of over-65-year-olds	6543	17.05	4.47	7838	16.88	4.41	0.17
Income per capita	6543	19989	5725	7838	20310	5702	-320
Unemployment	6597	14.05	6.12	7899	13.96	6.00	0.08

Bandwidth = 0.55							
	Elected (N = 8710)			Not elected (N = 11348)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	8710	14284	33359	11348	19657	43620	-5373
Coalition Party seat share	8710	18.43	10.47	11348	19.54	10.39	-1.11
Social Democrats seat share	8710	21.76	11.00	11348	22.90	10.90	-1.13
Center Party seat share	8710	33.42	21.52	11348	30.68	21.35	2.74
True Finns seat share	8710	1.68	3.70	11348	1.70	3.67	-0.02
Green Party seat share	8710	3.51	4.82	11348	4.19	5.39	-0.68
Socialist Party seat share	8710	8.67	7.79	11348	8.82	7.60	-0.14
Swedish Party seat share	8710	5.45	17.10	11348	4.95	15.36	0.49
Christian Party seat share	8710	3.37	3.71	11348	3.48	3.62	-0.11
Other parties' seat share	8710	3.53	6.68	11348	3.60	6.62	-0.07
Voter turnout	8590	62.39	6.57	11183	61.76	6.51	0.63
Population	8672	30986	73934	11284	42703	95855	-11717
Share of 0-14-year-olds	8606	18.37	3.30	11191	18.20	3.22	0.17
Share of 15-64-year-olds	8606	64.48	3.25	11191	64.94	3.34	-0.46
Share of over-65-year-olds	8606	17.14	4.52	11191	16.86	4.45	0.28
Income per capita	8606	19941	5669	11191	20339	5731	-398
Unemployment	8672	14.03	6.09	11284	13.96	6.02	0.06

Bandwidth = 0.70							
	Elected (N = 10632)			Not elected (N = 14888)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	10632	13755	32233	14888	22049	48155	-8295
Coalition Party seat share	10632	18.25	10.51	14888	19.75	10.40	-1.50
Social Democrats seat share	10632	21.61	11.04	14888	22.98	10.83	-1.37
Center Party seat share	10632	33.86	21.66	14888	30.06	21.42	3.81
True Finns seat share	10632	1.70	3.76	14888	1.70	3.63	0.00
Green Party seat share	10632	3.42	4.75	14888	4.45	5.61	-1.03
Socialist Party seat share	10632	8.64	7.83	14888	8.89	7.55	-0.24
Swedish Party seat share	10632	5.43	17.13	14888	4.91	15.05	0.52
Christian Party seat share	10632	3.36	3.69	14888	3.48	3.59	-0.11
Other parties' seat share	10632	3.52	6.67	14888	3.62	6.64	-0.10
Voter turnout	10477	62.49	6.57	14669	61.59	6.49	0.90
Population	10588	29837	71416	14805	47990	105723	-18153
Share of 0-14-year-olds	10498	18.38	3.29	14680	18.18	3.22	0.20
Share of 15-64-year-olds	10498	64.42	3.23	14680	65.09	3.37	-0.67
Share of over-65-year-olds	10498	17.20	4.51	14680	16.73	4.42	0.47
Income per capita	10498	19848	5652	14680	20432	5798	-585
Unemployment	10588	14.07	6.09	14805	13.94	6.01	0.13

Bandwidth = 1.00							
	Elected (N = 14138)			Not elected (N = 23320)			
Variable	N	Mean	Std. dev.	N	Mean	Std. dev.	Difference
Total number of votes	14138	13094	30932	23320	26432	53737	-13338*
Coalition Party seat share	14138	18.16	10.48	23320	20.46	10.25	-2.31
Social Democrats seat share	14138	21.46	11.11	23320	23.39	10.63	-1.93
Center Party seat share	14138	34.37	21.75	23320	28.52	21.24	5.85
True Finns seat share	14138	1.72	3.80	23320	1.69	3.55	0.03
Green Party seat share	14138	3.28	4.66	23320	4.99	5.97	-1.72*
Socialist Party seat share	14138	8.57	7.80	23320	8.94	7.47	-0.38
Swedish Party seat share	14138	5.42	17.23	23320	4.79	14.45	0.63
Christian Party seat share	14138	3.36	3.73	23320	3.47	3.41	-0.11
Other parties' seat share	14138	3.48	6.84	23320	3.60	6.56	-0.12
Voter turnout	13937	62.66	6.59	22978	61.12	6.41	1.54
Population	14075	28404	68622	23190	57594	118321	-29190*
Share of 0-14-year-olds	13966	18.39	3.29	23001	18.08	3.18	0.31
Share of 15-64-year-olds	13966	64.32	3.22	23001	65.40	3.43	-1.08*
Share of over-65-year-olds	13966	17.29	4.51	23001	16.51	4.36	0.77
Income per capita	13966	19758	5631	23001	20716	5904	-958
Unemployment	14075	14.08	6.09	23190	13.82	5.98	0.26

Notes: * and ** denote 5 % and 1 % statistical significance of difference in means respectively. The significance of differences is tested using t test adjusted for clustering by municipality. 0.55 bandwidth equals roughly the optimal bandwidth chosen using Imbens and Kalyanaraman's (2008) algorithm. Sample includes only data for 1996-2008 elections. Income per capita is expressed in euros.

Appendix F: Supplementary information to HMSTT Section 4.3 (Robustness tests)

This appendix discusses the robustness tests (#1–#6) that we have conducted.

Robustness test #1: Global polynomial RDD

Table F1: In this table we report results for a parametric RDD specification using higher order *global* polynomials (1^{st}-5^{th} degree) of the forcing variable on both sides of the cutoff. As the table shows, the treatment effect estimates tend to get smaller when the degree of the polynomial increases, but even for the 5^{th} degree polynomial, they are positive, large and highly significant. The bias using global polynomials seems to an order of magnitude larger than the one obtained using local polynomials. This approach generates incumbency effects that are roughly similar in magnitude to those reported in Lee (2008). It should be noted, however, that his estimates refer to an amalgam of party and personal incumbency effects and apply to a very different institutional context.

Table F1. Parametric RDD with 1^{st} – 5^{th} order polynomials.

Outcome: Elected next election

	(1)	(2)	(3)	(4)	(5)
Elected	0.4320**	0.3862**	0.3418**	0.2959**	0.2552**
	(0.0053)	(0.0062)	(0.0069)	(0.0077)	(0.0085)
N	154545	154545	154545	154545	154545
R^2	0.33	0.33	0.33	0.34	0.34
Order of control polynomial	1st	2nd	3rd	4th	5th

Notes: Each specification uses the whole range of data. All standard errors are clustered at the municipality level.* and ** denote 5% and 1% statistical significance levels respectively. Unit of observation is a candidate *i* at year *t*.

Robustness test #2: Alternative measure of incumbency advantage

Table F2: In this table, we look at the effect of being elected in election at time *t* on the vote share in the election at time *t*+1. As we reported earlier (Table B2 in Appendix B), the effect is not statistically different from zero in the lottery sample when this variable is used as an alternative outcome. As the table below shows, the conventional RDD using optimal bandwidths and local linear specification produces a positive and significant effect. The more flexible specifications reproduce the experimental estimate: The estimates suggest that the under-smoothing procedure and the use of higher degree local polynomials without adjusting the bandwidth accordingly work. It is, however, important to point out that some of the estimates in Panel B are negative and quite large in the absolute value.

Table F2. RDD results, incumbency advantage in vote share in the next election.

Outcome: Vote share next election

Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(4)	(5)	(7)	(8)
	Linear		Quadratic		Cubic	
Elected	0.0491**	0.0361	0.0063	-0.0011	-0.0190	-0.0338
	(0.0189)	(0.0207)	(0.0269)	(0.0306)	(0.0364)	(0.0396)
N	36834	28925	36834	28925	36834	28925
Bandwidth	0.99	0.79	0.99	0.79	0.99	0.79
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Bandwidth optimized for local linear specification * 0.5						
	(10)	(11)	(13)	(14)	(16)	(17)
	Linear		Quadratic		Cubic	
Elected	0.0160	0.0074	-0.0261	-0.0523	-0.0857	-0.1003
	(0.0257)	(0.0283)	(0.0377)	(0.0427)	(0.0517)	(0.0575)
N	17930	14348	17930	14348	17930	14348
Bandwidth	0.49	0.39	0.49	0.39	0.49	0.39
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT

Notes : Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t* .

Robustness test #3: Small vs. large municipalities

Tables F3 and F4: These tables reports RDD results separately for small (Table F3) and large (Table F4) municipalities and thus for small and large elections. We use the median number of votes in the municipality in the lottery sample as the point of division (i.e., 2422 votes). As is noted in the main text of HMSTT (and in Appendix B), ties usually appear in elections held in slightly smaller municipalities (those with a small number of voters). This means that our experimental estimate may mostly apply to such elections. As we reported earlier, the experimental estimate is very close to zero both in small and in large elections. However, our forcing variable, v_{it}, can get values really close to zero only when parties get a large amount of votes. This tends to happen in larger elections. The RDD estimates, which use the narrowest bandwidths, may thus mostly apply to them. To check whether the discrepancy between the experimental and the RDD estimates is driven by the size of the municipalities, Tables F3 and F4 reports parts of our RDD analysis separately for small and large municipalities. The results show that our conclusions are not driven by the size of the elections.

Table F3. RDD results for small municipalities.

Outcome: Elected next election

Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(3)	(4)	(5)	(6)
	Linear		Quadratic		Cubic	
Elected	0.1123**	0.0356	0.0340*	0.0133	0.0112	0.0015
	(0.0114)	(0.0187)	(0.0166)	(0.0293)	(0.0224)	(0.0394)
N	23967	10611	23967	10611	23967	10611
Bandwidth	4.01	1.41	4.01	1.41	4.01	1.41
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Bandwidth optimized for local linear specification * 0.5						
	(7)	(8)	(9)	(10)	(11)	(12)
	Linear		Quadratic		Cubic	
Elected	0.0511**	0.0183	0.0170	0.0071	0.0102	0.0390
	(0.0156)	(0.0272)	(0.0239)	(0.0433)	(0.0328)	(0.0707)
N	14563	5598	14563	5598	14563	5598
Bandwidth	2.00	0.71	2.00	0.71	2.00	0.71
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT

Notes: Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t*. Sample includes only small elections in which at most 2422 votes were given.

Table F4. RDD results for large municipalities.

Outcome: Elected next election

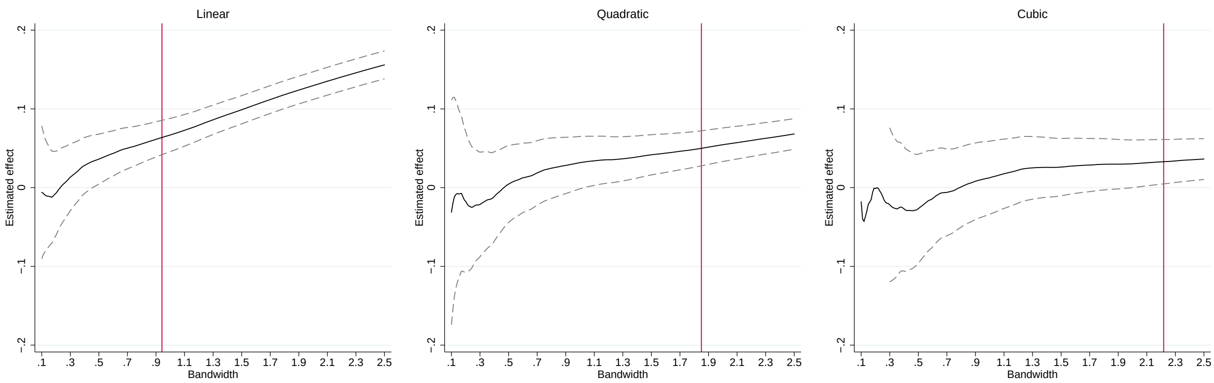
Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(3)	(4)	(5)	(6)
	Linear		Quadratic		Cubic	
Elected	0.0506**	0.0637**	0.0099	0.0235	-0.0259	-0.0071
	(0.0162)	(0.0139)	(0.0244)	(0.0220)	(0.0326)	(0.0284)
N	17665	22917	17665	22917	17665	22917
Bandwidth	0.62	1.11	0.62	1.11	0.62	1.11
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Bandwidth optimized for local linear specification * 0.5						
	(7)	(8)	(9)	(10)	(11)	(12)
	Linear		Quadratic		Cubic	
Elected	0.0102	0.0277	-0.0346	-0.0255	-0.0313	-0.0391
	(0.0233)	(0.0203)	(0.0349)	(0.0311)	(0.0499)	(0.0419)
N	8945	11344	8945	11344	8945	11344
Bandwidth	0.31	0.55	0.31	0.55	0.31	0.55
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT

Notes : Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t* . Sample includes only large elections in which more than 2422 voters voted.

Robustness test #4: Heterogeneity in the personal incumbency effect

Figure F1: This figure shows RDD point estimates and their 95 % confidence intervals for a wide range of bandwidths, obtained using only those party-lists that were involved in the lotteries. When these party-lists are used, increasing the bandwidths adds new candidates from the same lists, but does not add new lists or municipalities to the sample. The reason for reporting these results is that, besides the bias caused by the potentially incorrect linear approximation, the point estimates may increase due to heterogeneity in the personal incumbency effect across municipalities (and thus party-lists). The use of wider bandwidths means that our baseline RDD identifies the effect for a different set of municipalities than what we have in the experimental sample. The findings reported in Figure F1 do not support the explanation of heterogeneous treatment effects, as the patterns that we find here are similar to those reported in the main text of HMSTT (Figure 2).

Figure F1. RDD estimates using only party lists with lotteries.

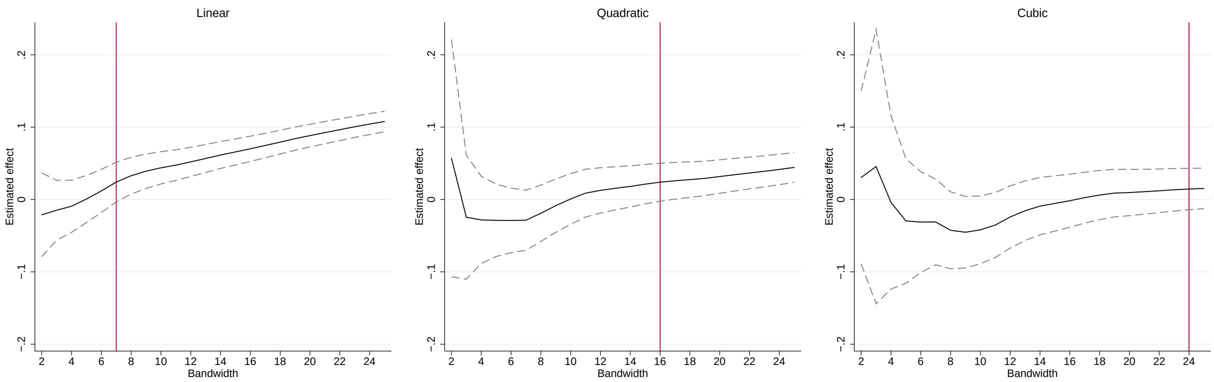


Notes: The graph displays the point estimates of incumbency advantage for various bandwidths. Dashed lines mark the 95 % confidence intervals. In the third graph, confidence intervals have been omitted for bandwidths smaller than 0.3. Red vertical line marks the optimal bandwidth chosen using IK method. The sample includes only candidates from party lists that have lotteries.

Robustness test #5: Alternative definitions for the forcing variable.

Figure F2: This figure reports RDD results when a non-scaled version of our forcing variable is used. The forcing variable is defined as in the main text of HMSTT, but is not scaled with the total number of votes the party got. We display the RDD estimates for linear, quadratic and cubic local polynomial specifications. As the figure shows, the results that we obtain using this alternative forcing variable echo our baseline RDD results. The local linear polynomial produces biased results, but the higher order polynomials and bandwidths smaller than optimal work better.

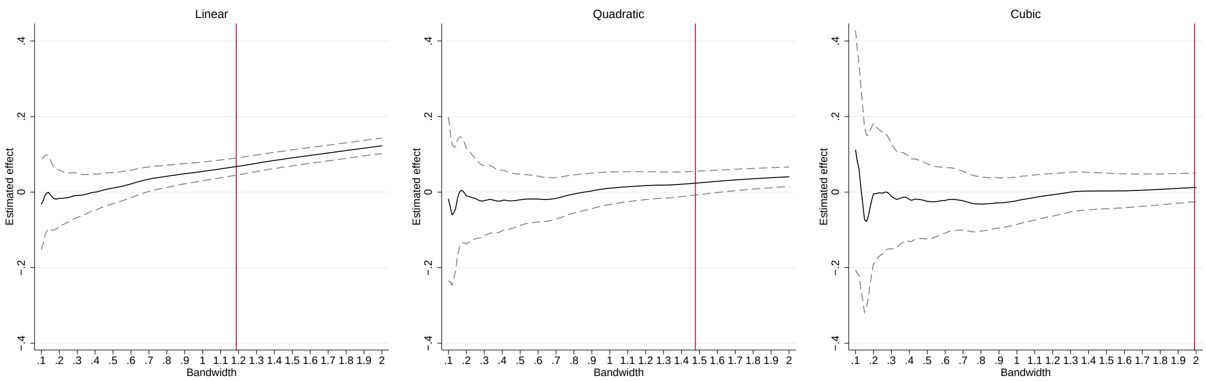
Figure F2. RDD estimates using absolute vote margin, measured in number of votes, as the forcing variable.



Notes: The graph displays the point estimates of incumbency advantage for various bandwidths. The bars below and above the point estimates show the 95 % confidence intervals. Red vertical line marks the optimal bandwidth chosen using IK method. The forcing variable is as in the main text but not scaled with the total number of votes the party got.

Figure F3: This figure reports RDD results when another alternative version of our forcing variable is used. For this figure we define the cutoff as the number of votes of the first non-elected (last elected) candidate of the ordered party list for the elected (non-elected) candidates. The forcing variable is then the distance from this cutoff multiplied by 100 and divided by the number of party's votes. As the figure shows, the results echo our baseline RDD results.

Figure F3. RDD estimates using the distance to the first non-elected (or last elected) candidate as the forcing variable.



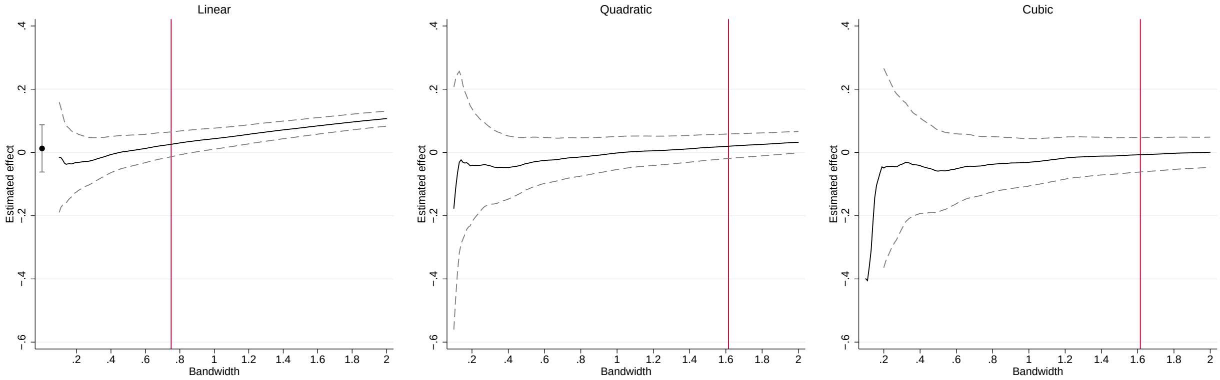
Notes: The graph displays the point estimates of incumbency advantage for various bandwidths. The bars below and above the point estimates show the 95 % confidence intervals. Red vertical line marks the optimal bandwidth chosen using IK method. The forcing variable is defined as follows. For elected (non-elected) candidates, the cutoff is the number of votes of the first non-elected (last elected) candidate of the ordered party list. The forcing variable is then the distance from this cutoff multiplied by 100 and divided by the number of party's votes.

Robustness test #5: Heterogeneity in the effect between parties.

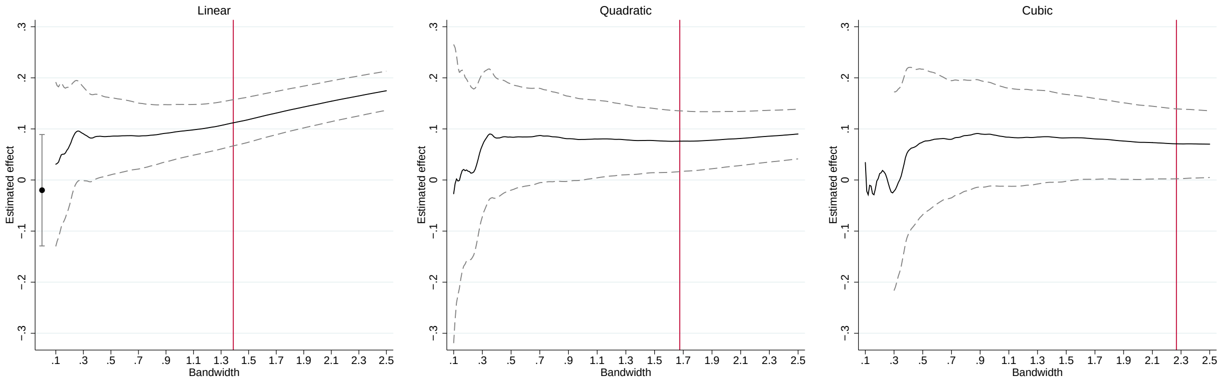
Figure F4: This figure reports graphically the RDD results separately for each of the three large parties (Panel A: Center Party, Panel B: National Coalition Party and Panel C: Social Democratic Party). The graphs allow us to study whether there is heterogeneity in the effect between the parties. Our motivation to look at such heterogeneity is that it could be an alternative explanation for the disparity between the experimental estimate and non-experimental RDD estimates. Suppose, for example, that there is no incumbency advantage within party A but a positive advantage within party B. Then if party A is more often involved in lotteries and if for some reason party B is overrepresented in the RDD samples (that are based on larger bandwidths), we might observe that the experimental estimate is zero and that RDD estimates produce a positive effect, especially when larger bandwidths are used. Figure F4 allows us to rule out such explanations. It seems that there is no substantial heterogeneity in the within party personal incumbency advantage between parties.

Figure F4. RDD estimates for different parties.

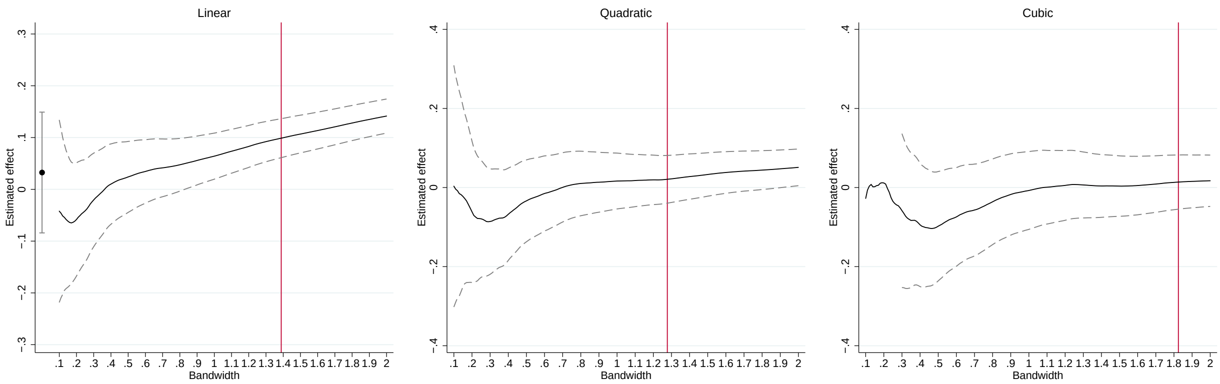
Panel A: Center Party.



Panel B: National Coalition Party.



Panel C: Social Democratic Party.



Notes: The graph displays the point estimates of incumbency advantage for various bandwidths. Dashed lines show the 95 % confidence intervals. In the third graph, confidence intervals have been omitted for bandwidths smaller than 0.2 in Panel A and smaller than 0.3 in Panels B and C. Red vertical line marks the optimal bandwidth chosen using IK method. The figure for linear specification also displays the estimate from the lottery sample and its 95 % confidence interval.

Robustness test #6: Excluding from the sample those who do not rerun

Table F5: These tables reports RDD results for a sample from which those who do not rerun are excluded. In Panel A, elected next election is the outcome variable, whereas in Panel B we use the vote share in the next election again as the alternative outcome. As we reported earlier (in Appendix B), the experimental estimates suggest no effect on these outcome variables when the sample from which those who do not rerun are excluded. Our motivation to report these results is that the previous literature is mixed on how those who do not rerun should be treated: For instance, Uppal (2010) report the results for a sample that includes all candidates and for a sample that only includes those who rerun, whereas de Magalhaes (2014) argues in favor of including all the candidates.

The results of the two panels show that our conclusions are unaffected by the sample restriction: We again find that the standard implementation (local linear with IK optimal bandwidth) of RDD generates a positive and significant effect. We also find that undershooting appears to work (with one exception) and that the use of higher degree local polynomials without adjusting the bandwidth accordingly reproduces the experimental estimate in the sense that we do no reject the null hypothesis of no effect. These insignificant findings are largely, but not in each case, due to greater standard errors, as the estimated effects do not systematically become closer to zero as the more flexible approaches are used.

Table F5. RDD estimates for using rerunners only.

Panel A: Elected next election.

Outcome: Elected next election

Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(3)	(4)	(5)	(6)
	Linear		Quadratic		Cubic	
Elected	0.0671**	0.0751**	0.0505	0.0531	0.0369	0.0428
	(0.0212)	(0.0188)	(0.0310)	(0.0279)	(0.0429)	(0.0369)
N	12058	15079	12058	15079	12058	15079
Bandwidth	0.54	0.69	0.54	0.69	0.54	0.69
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Bandwidth optimized for local linear specification * 0.5						
	(7)	(8)	(9)	(10)	(11)	(12)
	Linear		Quadratic		Cubic	
Elected	0.0484	0.0574*	0.0342	0.0351	0.0563	0.0339
	(0.0299)	(0.0264)	(0.0457)	(0.0403)	(0.0682)	(0.0564)
N	6209	7745	6209	7745	6209	7745
Bandwidth	0.27	0.34	0.27	0.34	0.27	0.34
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT

Notes : Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t* . Sample includes only rerunning candidates.

Panel B: Outcome: Vote share next election.

Outcome: Vote share next election

Panel A: Bandwidth optimized for local linear specification						
	(1)	(2)	(4)	(5)	(7)	(8)
	Linear		Quadratic		Cubic	
Elected	0.0490*	0.0494*	0.0471	0.0470	0.0488	0.0518
	(0.0239)	(0.0244)	(0.0328)	(0.0337)	(0.0437)	(0.0453)
N	16668	15697	16668	15697	16668	15697
Bandwidth	0.76	0.72	0.76	0.72	0.76	0.72
Bandwidth selection method	IK	CCT	IK	CCT	IK	CCT
Panel B: Bandwidth optimized for local linear specification * 0.5						
	(10)	(11)	(13)	(14)	(16)	(17)
	Linear		Quadratic		Cubic	
Elected	0.0578	0.0601	0.0367	0.0264	-0.0278	-0.0278
	(0.0309)	(0.0318)	(0.0459)	(0.0471)	(0.0598)	(0.0614)
N	16668	15697	16668	15697	16668	15697
Bandwidth	0.38	0.36	0.38	0.36	0.38	0.36
Bandwidth selection method	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT	0.5 * IK	0.5 * CCT

Notes: Table shows estimated incumbency advantage using local polynomial regressions within various bandwidths. The standard errors are clustered at municipality level. * and ** denote 5% and 1% statistical significance levels, respectively. Unit of observation is a candidate *i* at year *t*. Sample includes only rerunning candidates.

References

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014a. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6): 2295–2326.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014b. "Robust Data-Driven Inference in the Regression Discontinuity Design." *Stata Journal* 14(4): 909-946. (CITE?)

De Magalhaes, Leandro. 2014. "Incumbency Effects in a Comparative Perspective: Evidence from Brazilian Mayoral Elections." *Political Analysis* 23(1): 113-126.

Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder. 2015. "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(1): 259–274.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142(2): 675-97.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2): 698-714.

Snyder, James M., Olle Folke, and Shigeo Hirano. 2015. "Partisan Imbalance in Regression Discontinuity Studies Based on Electoral Thresholds." *Political Science Research and Methods*. Forthcoming.

Uppal, Yogesh. 2010. "Estimating Incumbency Effects In U.S. State Legislatures: A Quasi-Experimental Study." *Economics and Politics* 22(2): 180-99.