Leandro De Magalhaes, Dominik Hangartner, Salomo Hirvonen, Jaakko Meriläinen, Nelson A. Ruiz, and Janne Tukiainen

How Much Should We Trust Regression Discontinuity Design Estimates? Evidence from Experimental Benchmarks of the Incumbency Advantage

Aboa Centre for Economics

Discussion paper No. 135
Turku 2020

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.

Leandro De Magalhaes, Dominik Hangartner, Salomo Hirvonen, Jaakko Meriläinen, Nelson A. Ruiz, and Janne Tukiainen

How Much Should We Trust Regression Discontinuity Design Estimates? Evidence from Experimental Benchmarks of the Incumbency Advantage

ABSTRACT

Regression discontinuity designs (RDD) are widely used in the social sciences to estimate causal effects from observational data. Scholars can choose from a range of methods that implement different RDD estimators, but there is a paucity of research on the performance of these different estimators in recovering experimental benchmarks. Leveraging exact ties in local elections in Colombia and Finland, which are resolved by random coin toss, we find that RDD estimation using bias-correction and robust inference (CCT) performs better in replicating experimental estimates of the individual incumbency advantage than local linear regression with conventional inference (LLR). We assess the generalizability of our results by estimating incumbency effects across different subsamples, and in other countries. We find that CCT consistently comes closer to the experimental benchmark, produces smaller estimates than LLR, and that incumbency effects are highly heterogeneous, both in magnitude and sign, across countries with similar open-list PR systems.

Contact information

Leandro De Magalhaes, University of Bristol, Department of Economics, 12 Priory Road, 0C1, Bristol BS8 1TU, UK.
Leandro.DeMagalhaes@bristol.ac.uk.

Dominik Hangartner, Department of Government, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK, and Center for Comparative and International Studies, ETH Zurich, 8092 Zurich, Switzerland.
D.Hangartner@lse.ac.uk.

Salomo Hirvonen, University of Bristol, Department of Economics, 12 Priory Road, 0C1, Bristol BS8 1TU, UK.
salomo.hirvonen@bristol.ac.uk.

Jaakko Meriläinen, ITAM, Department of Economics and Center for Economic Research, Av. Camino Santa Teresa 930, Col. Héroes de Padierna, Del. Magdalena Contreras, 10700 México, D.F., Mexico.
jaakko.merilainen@itam.mx.

Nelson A. Ruiz, Department of Politics and International Relations, University of Oxford, Manor Road Building Manor Rd, Oxford OX1 3UQ, UK. nelson.ruiz@politics.ox.ac.uk.

Janne Tukiainen (Corresponding author), University of Turku, Department of Economics, Rehtorinpellonkatu 3, FI-20014 University of Turku, Finland, and VATT Institute for Economic Research, Arkadiankatu 7, FI-00101, Helsinki, Finland. janne.tukiainen@utu.fi.

Regression discontinuity designs (RDD) are widely used in political science and neighbouring disciplines to learn about a broad range of policy effects, running the gamut from forced labor (Dell 2010) to citizenship (Hainmueller et al. 2017) to trickle-up political socialization (Dahlgaard 2018). The popularity of RDD is not surprising given that it is often heralded as one of the few observational study designs that is successful in approximating experimental benchmarks (Green et al. 2009). Over the last few years, a range of RDD estimators and implementations have been proposed (see, e.g., Cattaneo et al. 2020, Calonico et al. 2014), and researchers have become increasingly aware that not only the validity of the research design is important, but also the details of the RDD implementation matter (Gelman and Imbens 2019, Hyytinen et al. 2018). Yet, there has been little effort in examining which RDD estimator and implementation is best able to replicate experimental estimates. This question is of fundamental importance for the trust we can put in RDD estimates, and particularly relevant in the electoral context, where numerous RDD applications have leveraged close elections to estimate the effects of holding office on various outcomes including incumbency advantage (for a recent review, see De la Cuesta and Imai 2016).

Over the last few years, researchers have broadly converged to using local polynomial specifications to estimate treatment effects from RD designs (Gelman and Imbens 2019). This implementation involves several steps: choosing a bandwidth, a weighting scheme for observations closer and farther away from the threshold, the order of the polynomial for the locally weighted least squares regression, and a method for statistical inference (see, e.g., Hahn et al. 2001, Cattaneo et al. 2020). To date, the dominant implementation selects the bandwidth by minimizing the mean-squared-error (MSE), uses a triangular kernel to weight the sample, a (linear) polynomial of order one, and conducts inference using OLS approximations (Cattaneo et al. 2020). We thus refer to this implementation as the conventional local linear approach (LLR).

However, Calonico et al. (2014) show that, while optimal for point estimation, LLR leads to biased confidence intervals due to the approximation error of the local polynomial estimator. One suggestion for how to correct for the bias is to use smaller than optimal bandwidths combined with OLS approximation (Cattaneo et al. 2020). However, such ad hoc undersmoothing leads to a

loss of statistical power (Calonico et al. 2014, 2020). As an alternative approach, Calonico et al. (2014) propose bias-corrected and robust confidence intervals that provide valid inferences without undersmoothing. Their estimator first estimates the bias using higher order polynomials and then subtracts the estimated bias from the local linear RD point estimate. Robust inference is achieved by incorporating the contribution of the bias-correction step to the variability of the bias-corrected point estimator. We refer to this implementation as the CCT approach. Calonico et al. (2014) and Calonico et al. (2020) provide both theoretical results and Monte Carlo simulations that suggest that CCT has smaller coverage errors than LLR. With this paper, we complement these results by evaluating the performance of LLR and CCT implementations against experimental benchmarks.

Building on and extending the approach of Hyytinen et al. (2018), our validation analysis leverages electoral ties that are decided by a lottery to estimate the effect of being the incumbent (versus being the runner-up) on getting elected in the next election in Colombia and in Finland. The lottery-based estimates of the individual incumbency advantage are very close to zero, precisely estimated, and not statistically significant in both countries. Since random assignment between incumbents and runner-ups in tied elections occurs right at the within-party cutoff for winning office, RDD should, if properly implemented, yield an estimate that matches the experimental benchmark.

We find that CCT works better than the more conventional LLR in replicating the experimental benchmark: For both countries, CCT produces small and insignificant estimates of the incumbency advantage, whereas LLR produces larger and statistically significant estimates. The results are qualitatively similar when we expand the analysis across several subsamples. We also extend the analysis to less frequently used global polynomial estimation approaches and find that CCT robustly outperforms them in recovering the experimental benchmarks.

To assess the generalizabilty of this finding, we also estimate the individual incumbency advantage in Brazil and Denmark—neighboring countries of Colombia and Finland, respectively, with similar open-list PR electoral systems. The pattern we observe when comparing CCT and LLR estimates is similar across all countries. This suggests that the upward bias in the LLR

4

estimates might be a more general issue for estimates of the incumbency advantage. The differences between LLR and CCT, and LLR and the lottery estimates, respectively, are sometimes statistically significant, but their magnitudes are fairly small. This suggests that the details of the RDD implementation matter, but correcting for the bias does not fundamentally change our extant understanding of the incumbency advantage in open-list PR systems.

Building on the insights from our validation analysis, we substantively compare the individual incumbency (dis-)advantage across the four countries. We document considerable variation in incumbency effects, both with respect to magnitude and sign. The incumbency effect is negative in Brazil, close to zero in Colombia and Finland, and positive in Denmark. The finding that the incumbency (dis-)advantage varies substantially across countries with similar open-list PR systems implies that differences in the electoral systems are certainly not the sole explanation for the observed variation in incumbency advantage. This is in contrast with previous work, which has argued that the election system is a key moderator of the incumbency advantage (Redmond and Regan 2015; Ariga 2015; Dettman et al. 2017). We discuss the substantive implications of this finding in the conclusion.

Institutions and Data

Our data cover local government (municipal council) elections in Brazil (2000–2008), Colombia (2003–2015), Denmark (2005–2014), and Finland (1996–2012).[1] Brazil and Finland feature pure open-list electoral systems. In Colombia, parties can choose between open or closed lists (Hangartner et al. 2019), while in Denmark, parties can choose between open and semi-open lists. In both Colombia and Denmark, the vast majority of parties choose open lists, and our sample focuses on those cases. For the RDD analysis, we focus on party lists that nominate at least two candidates, and elect at least one and fewer than all listed candidates. This results in at least one winner and one non-elected runner-up for each local party list. The resulting data consist of 586,706 candidate-election year observations for Brazil, 147,558 for Colombia, 12,633 for Denmark, and 154,543 for Finland. The data from Colombia and Finland include a number of tied

elections that are resolved via lottery.[2] Ties happen fairly frequently: we observe 463 such candidate-election year observations for Colombia and 1,351 for Finland. These lottery samples are sufficiently large to provide us with a reliable experimental benchmark for the RDD estimates.

Estimates of the Individual Incumbency Advantage

We leverage electoral ties in Finland and Colombia, which are resolved by random coin toss, to construct an experimental benchmark of the personal incumbency advantage. Because the candidates involved in tied elections have exactly the same number of votes, the average treatment effect estimated from the lottery sample is a local estimate at the cutoff point that determines whether or not a candidate gets elected. This implies that in addition to focusing on the same institutional context and population, the lottery and RDD also target the same estimand. These lotteries are therefore an ideal benchmark to evaluate the performance of the RDD estimator.

To construct the benchmark estimates, we focus on the sample of tied candidates and regress the indicator variable for getting elected in the next election $(t+1)$, the outcome, on a binary indicator for getting elected in the current election (t), the treatment, using OLS. We do not condition on running in election $t+1$, because this decision might well be endogenous to getting elected in t. The results are shown in the first row of Table 1. The incumbency estimates for both Colombia and Finland are close to zero in magnitude, -0.030 and 0.004, and not statistically significant at conventional levels ($p = 0.371$ and $p = 0.860$, respectively). Thus, there is no evidence that being the winner in election t (versus being the runner-up) increases the probability of getting elected in the next election $t+1$. Moreover, with 95% confidence intervals of $[-0.097, 0.037]$ for Colombia and $[-0.044, 0.053]$ for Finland, we can rule out all but relatively small incumbency effects.

We then estimate the same quantities using RDD. We construct the running variable from the winning margin for candidates within the same party list. For elected candidates, this equals their within-party vote share minus the within-party vote share of the first non-elected candidate. For the non-elected, this equals their within-party vote share minus the within-party vote share of the last elected candidate. This allows us to compare candidates who barely won a seat to those who run on

6

the same list but barely failed with respect to their propensity of getting elected in the next election. The RDD results are reported in Table 1.[3] The LLR approach yields estimates that significantly diverge from the experimental benchmark for both Colombia and Finland.

**Table 1.** Effect of incumbency on winning next election.

	Colombia	Finland	Brazil	Denmark
	(1)	(2)	(3)	(4)
Lottery	-0.030	0.004		
	[-0.097,0.037]	[-0.044,0.053]		
N	463	1351		
LLR	0.044**	0.037*	-0.048**	0.239**
	[0.022,0.066]	[0.006,0.068]	[-0.068,-0.028]	[0.195,0.283]
CCT	0.030	-0.014	-0.066**	0.152**
	[-0.004,0.063]	[-0.065,0.038]	[-0.098,-0.034]	[0.092,0.212]
N	26210	21875	41349	6134
Bandwidth	2.46	0.73	1.08	2.98

Notes: The dependent variable is equal to one if a candidate wins the next election, and zero otherwise. The 95% confidence intervals are based on standard errors that are clustered by municipality and reported in brackets. The RDD estimations uses a triangular kernel and CER-optimal bandwidths. * and ** denote statistical significance at 5% and 1%, respectively.

The estimates reported in columns (1) and (2) suggest an incumbency effect of 0.044 ($p < 0.01$) in Colombia and 0.037 ($p < 0.01$) in Finland. The difference between the lottery and LLR estimates is statistically significant for Colombia ($p = 0.037$) but not for Finland ($p = 0.270$). In contrast, CCT estimates are closer to and not significantly different from the lottery estimates ($p = 0.113$ for Colombia and $p = 0.617$ for Finland), and themselves not significantly different from zero at the 5% level ($p = 0.080$ for Colombia and $p = 0.603$ for Finland).[4] For Colombia and Finland, LLR overestimates the incumbency advantage, and CCT brings the point estimates closer to zero. For Denmark, the CCT estimates are also smaller than the LLR estimates, while for Brazil, the CCT estimates are slightly larger in absolute terms (Columns (4) and (3) in Table 1). Online Appendix Table 4 shows that CCT also robustly outperforms global polynomial estimation approaches in recovering the experimental benchmarks.

In order to assess whether this pattern documented for the main effect of the incumbency advantage holds more broadly, we split the Colombian and Finnish samples by the median values of the covariates in our dataset: the absolute number of votes that define the cutoff value; the number of incumbent candidates; and the number of total candidates on a given party list.

Figure 1. Subsample estimates of the incumbency effect on winning the next election.
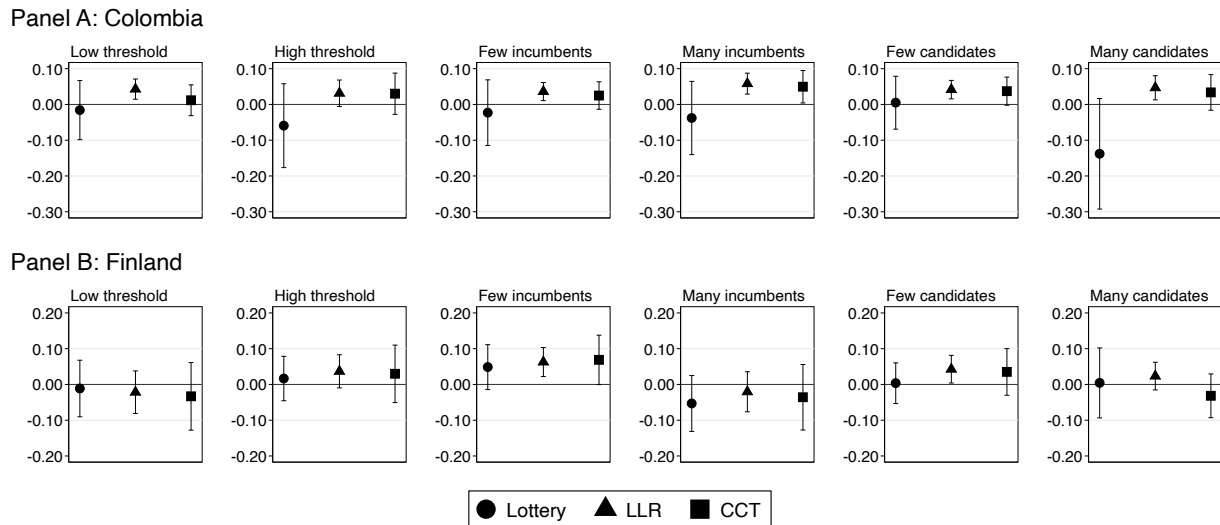
Notes: The dependent variable is equal to 1 if a candidate wins the next election, and 0 otherwise. The 95% confidence intervals are based on standard errors clustered at the municipality level. We use a triangular kernel and CER-optimal bandwidths in the RDD estimations. Samples are split at median values of covariates.

Figure 1 compares the incumbency effects using lotteries, LLR and CCT for the twelve subsamples. While the lottery estimates show that there are no statistically significant effects in any of the subsamples, LLR finds a substantively small but at the 0.05 level statistically significant incumbency advantage in seven out of twelve subsamples. The CCT estimates are again more in line with the lottery estimates: only one out of twelve estimates is statistically significant at the 0.05 level, and in nine out of twelve comparisons are the CCT point estimates closer to the lottery than the LLR estimates.[5] The Appendix reports further analyses. Figure A3 shows that the CCT estimates are more stable than LLR estimates to varying the bandwidth. This suggests that the choice between CCT and LLR may be more consequential than the choice of the bandwidth selection algorithm with CCT. Figure A4 shows placebo effects estimated from placebo cutoffs for

CCT and LLR. In both countries, LLR estimates a higher number of statistically significant placebo effects than CCT. While 23% of 60 placebo estimates are significant at the 0.05 level with LLR, this fraction decreases to 7% for the CCT. Note, however, that the higher number of significant placebo effects with LLR is not a bad thing: After all, this implies that the significance of the incumbency effect at the actual cutoff might be similarly inflated.[6]

Comparing Incumbency Effects across Countries

Next, we compare the personal incumbency (dis-)advantage across the four countries. Table 1 Columns (1)–(4) compare winners and runners-up in terms in election t regarding their propensity of getting elected in $t + 1$. Note that this estimate of the unconditional incumbency advantage has several limitations. First, this estimate does not directly correspond to the theoretical concept of the (personal) incumbency advantage since the latter is interested in a comparison of the incumbent and challenger and therefore conditions on both candidates running in elections t and $t + 1$. Second, the unconditional estimate assumes the same outcome value (i.e. zero) for candidates who re-ran and lost as for those who did not re-run at all. While this is less of an issue when comparing RDD estimates within the same country, it does complicate comparisons across countries with different re-running rates. In the Online Appendix, we further discuss this issue and two solutions using bounds (following Anagol and Fujiwara 2016 and De Magalhães and Hirvonen 2019). Despite their differences, both analyses relying on bounds confirm the main findings reported in Table 1: We find a small and statically not significant incumbency advantage in Colombia, no advantage in Finland, a small but statistically significant disadvantage in Brazil, and a large and statistically significant advantage in Denmark.

Concluding Remarks

Leveraging tied elections resolved by random coin toss in Colombia and Finland as benchmarks, we find that CCT performs better than LLR or global polynomial approaches in recovering the individual incumbency advantage. Extending this analysis to neighboring countries with similar

open-list PR systems, we find further evidence that LLR estimates would lead to false conclusions about the electoral benefits of incumbency.

These findings have implications for methodological research on RDD. At the very least, this study offers two benchmarks where CCT outperforms other RDD implementations. More broadly, these findings echo other studies that show how RDD estimates, even from larger samples, respond sensitively to the details of the implementation (Gelman and Imbens 2019; Hyytinen et al. 2018). Clearly, further research is needed to explore whether the bias documented here for estimates of the incumbency advantage is also present in other RDD contexts.

Our findings also have implications for our understanding of the role that electoral systems play in explaining differences in the incumbency advantage across countries. First, our findings that CCT-based estimates of the incumbency effect are fairly small and that previous studies, which mostly relied on LLR, have likely reported somewhat inflated point and biased uncertainty estimates, suggest that the personal incumbency effect is probably smaller than what has been documented to date. Second, a sizeable body of work has suggested that plurality and majoritarian systems, where typically fewer candidates are running in elections, generate larger incumbency advantages compared to PR systems with multi-seat constituencies and longer party lists (Redmond and Regan 2015; Ariga 2015; Ariga et al. 2016; Dettman et al. 2017). Our estimates show that even among countries with similar PR systems, there is substantial heterogeneity in terms of both magnitude and sign of the incumbency advantage. This implies that variation in electoral systems is certainly not the sole, and possibly not even the main, explanation for the observed cross-country differences in incumbency advantage. Future research should aim at providing a comprehensive set of reliable incumbency effect estimates from a broad range of democratic countries to investigate how contextual factors and electoral systems shape the incumbency advantage.

Notes

[1] The data for Brazil, Colombia and Finland are obtained from the respective electoral authorities. The election results for Denmark come from Dahlgaard (2016).

[2]Ties also exist in Brazil, but they are not resolved via lottery. Instead, the oldest tied candidate gets elected. Danish elections also use lotteries to resolve ties but they do not occur frequently enough for statistical analysis.

[3]The LLR and CCT estimates use the same coverage error probability (CER)-optimal bandwidth. Bandwidth selection and standard errors are adjusted for clustering at the municipal level. Following Calonico et al. (2018), the main and bias bandwidths are set at the CER-optimal main bandwidth level for CCT. Importantly, our findings are the same when using MSE-optimal bandwidths. In addition, the results are also robust to different choices of kernels (see Figure A2).

[4]We find that the details of the CCT implementation also matter: Both robust inference and bias-correction are needed as bias-correction alone is not sufficient to recover the lottery estimates. We use a first order local polynomial to estimate the main effect and second order local polynomial to estimate the bias.

[5]We provide further subsample analyses in Online Appendix Figures A6-A8, which compare LLR and CCT estimates across the four countries. The pattern documented here holds more broadly.

[6]Tables A1, A2 and A3, and Figure A5 report the results from covariate smoothness and density tests. We find no evidence of sorting at the threshold.

References

Anagol, Santosh and Thomas Fujiwara (2016). The runner-up effect. *Journal of Political Economy 124*(4), 927–991.

Ariga, Kenichi (2015). Incumbency Disadvantage under Electoral Rules with Intraparty Competition: Evidence from Japan. *Journal of Politics 77*(3), 874–887.

Ariga, Kenichi, Yusaku Horiuchi, Roland Mansilla, and Michio Umeda (2016). No sorting, no advantage: Regression discontinuity estimates of incumbency advantage in Japan. *Electoral Studies 43*, 21–31.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2018). On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statictical Association 113*(522), 767–779.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal 23*(2), 192–210.

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

Cattaneo, Matias, Nicolas Idrobo, and Rocio Titiunik (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge University Press.

Dahlgaard, Jens Olav (2016). You just made it: Individual incumbency advantage under proportional representation. *Electoral Studies 44*, 319–328.

Dahlgaard, Jens Olav (2018). Trickle-Up Political Socialization: The Causal Effect on Turnout of Parenting a Newly Enfranchised Voter. *American Political Science Review 112*(3), 698–705.

De la Cuesta, Brandon and Kosuke Imai (2016). Misunderstandings about the RD design in the study of close elections. *Annual Review of Political Science 19*, 375–396.

De Magalhães, Leandro and Salomo Hirvonen (2019). The Incumbent-Challenger Advantage and the Winner-Runner-up Advantage. Discussion Paper 19/710, Department of Economics, University of Bristol. Available at `https://ideas.repec.org/p/bri/uobdis/19-710.html`.

Dell, Melissa (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*(6), 1863–1903.

Dettman, Sebastian, Thomas B. Pepinsky, and Jan H. Pierskalla (2017). Incumbency advantage and candidate characteristics in open-list proportional representation systems: Evidence from Indonesia. *Electoral Studies 48*, 111–120.

Gelman, Andrew and Guido Imbens (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

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), 400–417.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Hainmueller, Jens, Dominik Hangartner, and Giuseppe Pietrantuono (2017). Catalyst or Crown: Does Naturalization Promote the Long-Term Social Integration of Immigrants? *American Political Science Review 111*, 256–276.

Hangartner, Dominik, Nelson Ruiz, and Janne Tukiainen (2019). Open or Closed? How List Type Affects Electoral Performance, Candidate Selection, and Campaign Effort. VATT Working Paper 120–2019. Available at `http://dx.doi.org/10.2139/ssrn.3418767`.

Hyytinen, Ari, Jaakko Meriläinen, Tuukka Saarimaa, Otto Toivanen, and Janne Tukiainen (2018). When does regression discontinuity design work? Evidence from random election outcomes. *Quantitative Economics 9*(2), 1019–1051.

Redmond, Paul and John Regan (2015). Incumbency advantage in a proportional electoral system: A regression discontinuity analysis of Irish elections. *European Journal of Political Economy 38*, 244–256.

How Much Should We Trust Regression Discontinuity

Design Estimates? Evidence from Experimental

Benchmarks of the Incumbency Advantage

Online Appendix

July 22, 2020

This appendix contains supplementary information and results to the paper "How Much Should We Trust RDD Estimates of the Individual Incumbency Advantage?".

Robustness and Validity Analysis

Figure A1 shows the standard RDD plots. We fit both linear and quadratic polynomial specifications on both sides of the cutoff within the CER-optimal bandwidths. The specification with the linear polynomial corresponds to the LLR approach in the main paper, and the quadratic polynomials corresponds to the CCT implementation approach. The graphs visualize what we already saw in the main results. Using the local linear specification gives a lower estimate of the effect of getting elected on the re-election probability in Brazil than the local quadratic specification. Similarly, using the local quadratic specification decreases the discontinuity gap in all the other countries in our data compared to the local linear specification. It is well known that the local linear specification may be subject to bias if there is curvature in the underlying data close to the cutoff.

This notion of bias is also reflected in Figures A2 and A3 in which we examine the robustness of our findings to alternative bandwidths. Figure A2 reports RD estimates for Colombia and Finland from specifications that use either MSE- or CER-optimal bandwidths. This choice does not matter for the magnitude of the regression coefficients as much as the choice between LLR and CCT. For the CCT estimations, we fix the main and bias bandwidths to be the same as proposed by Calonico et al. (2018). The figure also demonstrates that the choice of kernel appears to be less important for the estimation results. In Figure A3 we show RD estimates for each of the countries using the LLR (left-hand side graphs) and CCT approach (right-hand side graphs). In all the cases, the estimate monotonically increases with the bandwidth (aside the noisy very small bandwidths) and CCT approach is more robust.

We then estimate our RDD model using placebo cutoffs. Here we move the cutoff artificially to the left or right of the real cutoff and repeat the estimation. The first purpose of these tests is to provide an indirect test of key assumption of the continuity (of the conditional expectation

of potential outcomes at the cutoff). While this smoothness assumption is not directly testable at the cutoff, because the treatment status changes there, it should also hold in other locations of the forcing variable. The second purpose of the placebo cutoff tests is to analyze whether the implementation is appropriate. As the issues of misspecifying the relationship between the forcing variable and the outcome are most likely not specific to the cutoff, the placebo cutoff test are informative of implementing (specification and inference) the RDD in a wrong way. These results are reported in Figure A4 where the left-hand side graphs show RDD estimation results using the LLR, varying the location of the cutoff, and the right-hand side graphs show results using the CCT. The figures plotting the results using the CCT approach rarely show any statistically significant effects where they should not appear, indicating that the CCT specification is more appropriate for all the countries (and that the design is valid). In contrast, the LLR approach often fails the placebo cutoff analysis indicating the specification is wrong.

To assess further the validity of our findings, we conduct the following two tests. First, Tables A1 and A2 show the effect of getting elected on the lagged dependent variable using the lotteries in Colombia and Finland, and an RDD approach. We see no effect of getting elected on incumbency, as should be the case if no electoral manipulation based on candidate type is present in the close elections. Second, we verify that the density of the running variable evolves smoothly at the cutoff. This is, indeed, the case. Figure A5 reports the conventional McCrary (2008) test graphically, and Table A3 shows results from the density test proposed by Cattaneo et al. (2018).

Parametric RDD Estimates

Table A4 illustrates the problems with fitting global polynomials to the data. We use 1st-5th degree polynomials on both sides of the cutoff and show that such specifications over-estimate the magnitude of the incumbency advantage compared to what we find using the conventional and CCT approaches. The bias in these global estimates is considerably larger than the bias in LLR.

Heterogeneous Effects

We have also tried to shed some more light on the discrepancy between the LLR and CCT approaches by exploring effect heterogeneity. First, we look at the RDD estimation results splitting the sample in four based on quartiles (based on the full sample) of the minimum within-party vote share among the elected (Figure A6). Second, Figures A7 and A8 show the RDD estimates by the number of candidates and number of incumbent candidates, respectively. While the figures reveal some underlying heterogeneities in the point estimates, the differences between the conventional and CCT RD estimates are systematic. Thus, they do not give us any hints about the origins of the differences between the two estimation approaches.

Comparing Incumbency Effects across Countries

Finally, we note that in the main paper we compare winners versus runners-up on how they fare in a subsequent election. This estimate has been called the individual and unconditional incumbency advantage. It closely related to, but still differs from the usual theoretical definitions of the incumbency advantage.[1] The latter focuses on the comparison between incumbent and challenger and relate rather to effects that are conditional on re-running, such as visibility, ability to campaign and attract campaign funding, scare-off effects, voter information, selection based on higher ability etc. (King 1991; Cox and Katz 1996; Ashworth and Bueno de Mesquita 2008). Therefore, RDD as implemented in Table 1 in the main paper does not identify directly either the personal nor the partisan incumbency advantage. However, unlike in a typical majoritarian two-party system RDD (Fowler and Hall 2014; Erikson and Titiunik 2015), we can rule out partisan incumbency advantage by the within-party design.

We first follow De Magalhães and Hirvonen (2019) to estimate bounds for the individual incumbency advantage using the CCT approach with CER-optimal bandwidths. The lower bound assumes all the decisions not to re-run are strategic, that is, candidates choose not re-run because

---

[1] De Magalhães and Hirvonen (2019) discuss this issue in detail.

they would not win. This is the object in Table 1. The upper bound assumes that: i) non-rerunning in the winner's (treatment) group is due to random attrition; ii) the same share of the population suffered from random attrition in the runner-up (control) group, which should be the case if the control and treatment groups are balanced; and iii) the remaining non-rerunners in the runner-up (control) group must have done so for strategic reasons. In other words, the upper bound is the RDD estimate of the the object estimated in Table 1 scaled by the re-running rate of the winners. We assess the robustness of the bounds to alternative specifications in Figure A9 where we re-estimate the bounds using alternative bandwidths. The graphs follow the same pattern as Figure A3. The CCT approach appears to be more robust than the conventional local linear approach.

There are also alternative ways of conducting the bounding exercise. We use the bounds proposed in Anagol and Fujiwara (2016) in Figure A11. Again, we can think of there being four types of candidates: "always takers", who re-run whether they are incumbents or runners-up; "never takers", who do not re-run under any circumstances; "compliers", who re-run if they become incumbents but do not re-run if they are runners-up; and "defiers", who re-run if they are runners-up but do not re-run if they are incumbents. To generate the bounds we must assume there are no "defiers". This is a natural assumption in the Anagol and Fujiwara (2016) application, but not in our case, because some candidates may want to be in office only one term, but not many. The upper bound of the personal incumbency advantage assumes that, among "compliers", all runners-up who chose not to re-run would have lost their election. In other words, every decision to retire by "compliers" was strategic as they were sure to lose. Thus, upper bound is estimated by dividing the estimates in Table 1 by the estimated rerunning rates of incumbents at the cutoff. To compute the lower bound, we assume that the runners-up in the "compliers" group would have done as well as the incumbents who chose to re-run as a best possible scenario for the "compliers". Thus the lower bound is equal to the following term subtracted from the upper bound: estimated effect of incumbency on re-running multiplied by the estimate of the probability of victory for the incumbents conditional on re-running.
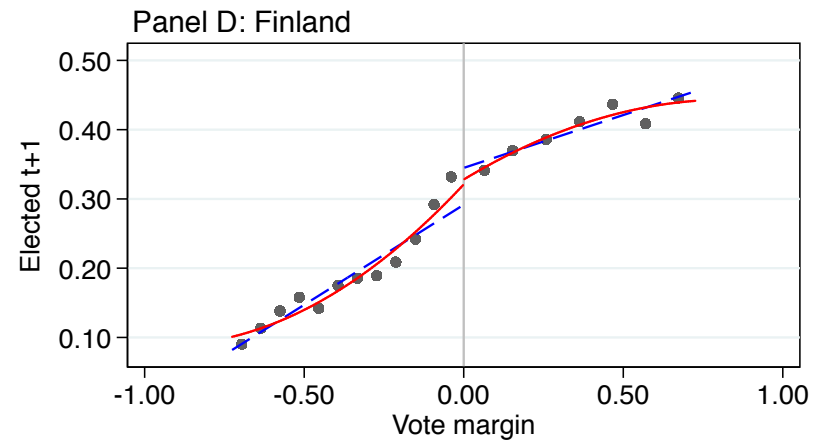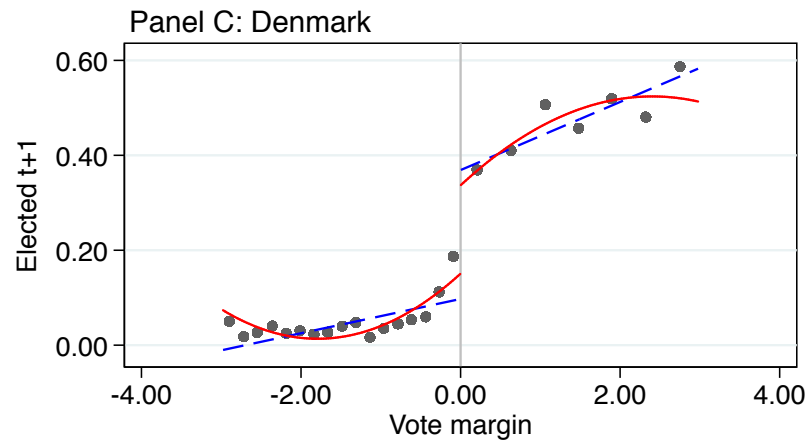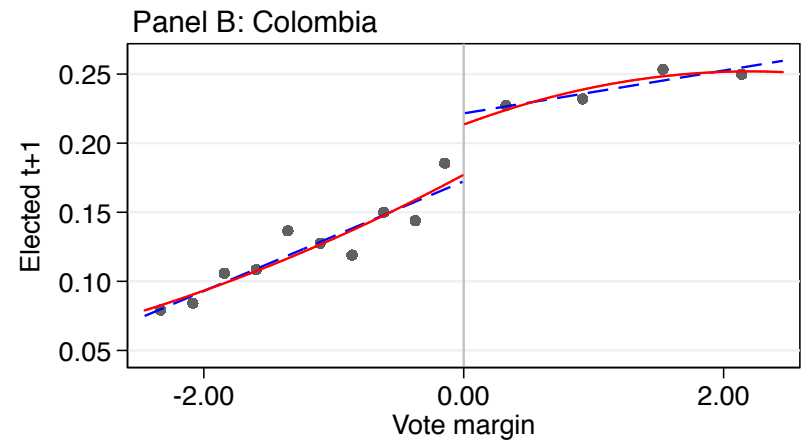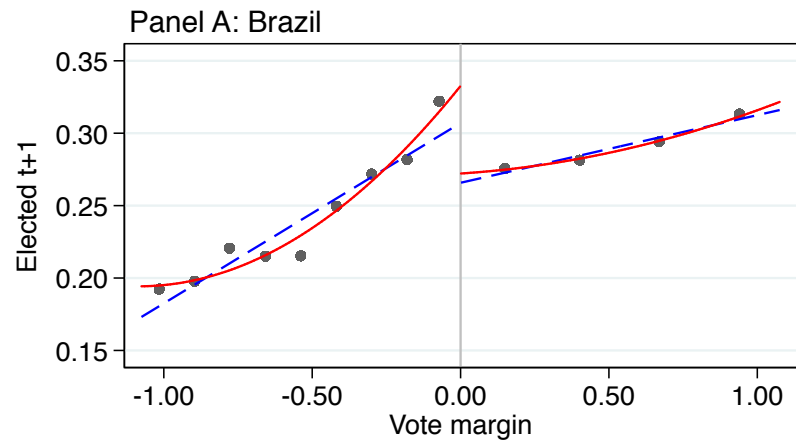
Figure A1. RDD graphs.

Notes: Dashed lines show local linear fits within CER-optimal bandwidths, and solid lines show a quadratic fit. Binned averages have been chosen using IMSE-optimal evenly-spaced method of Calonico et al. (2015).
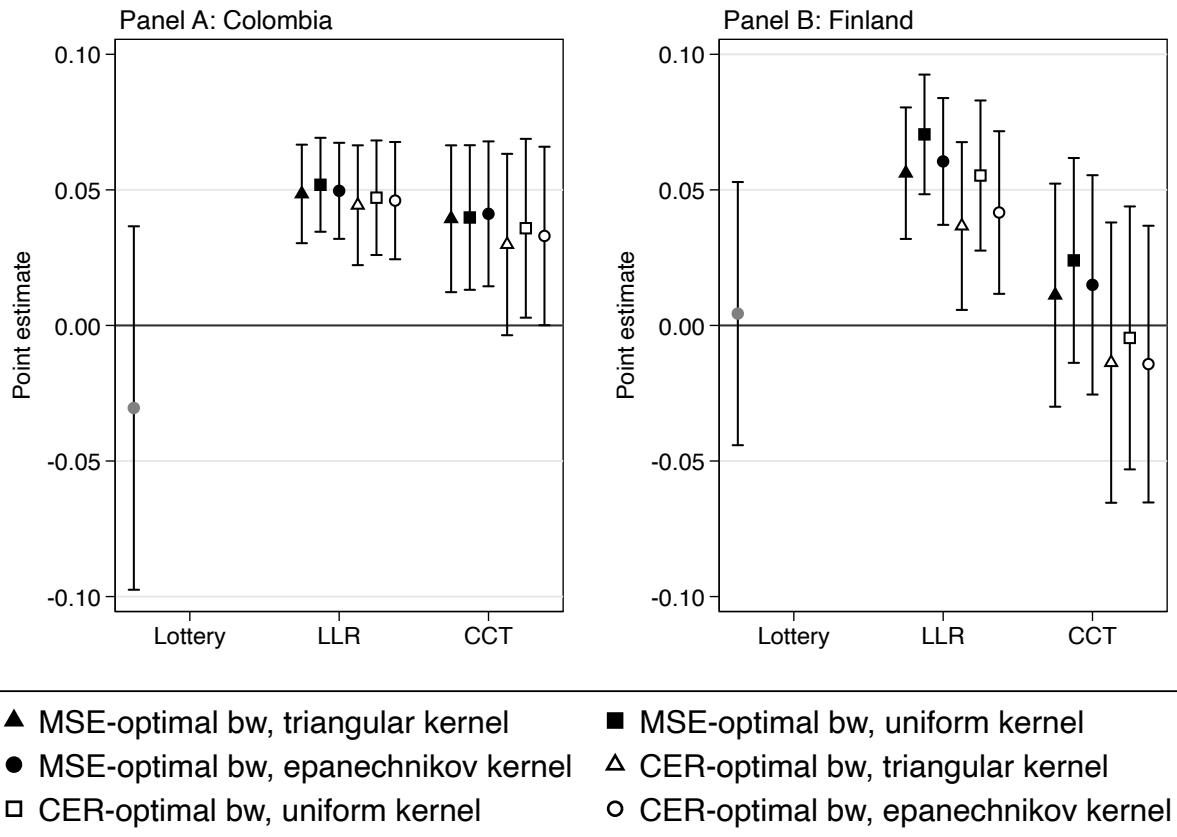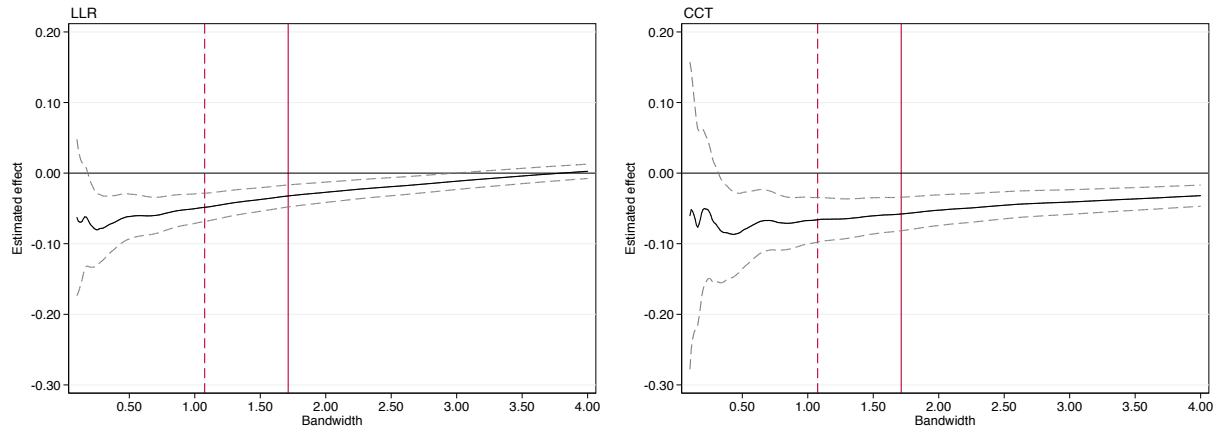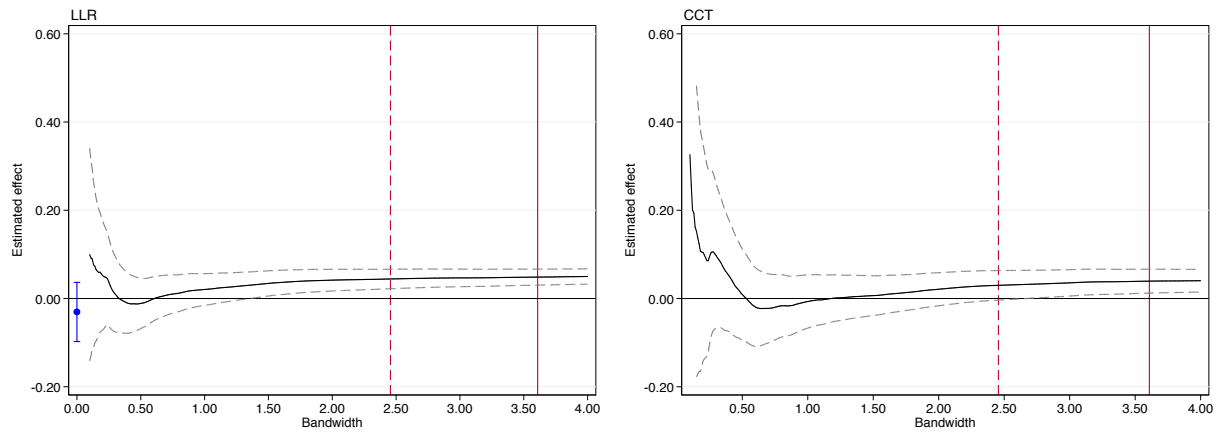
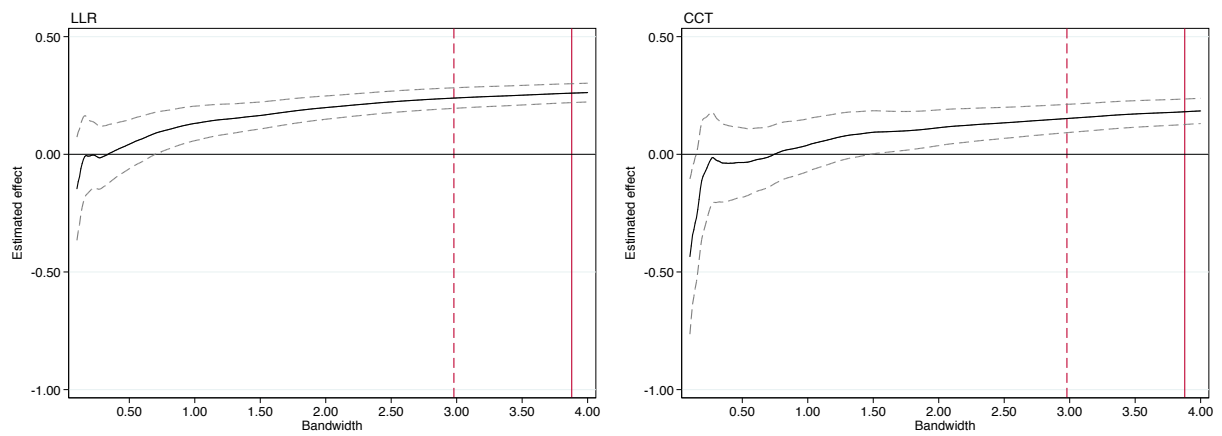**Figure A2.** RDD estimates using different optimal bandwidths and kernels.

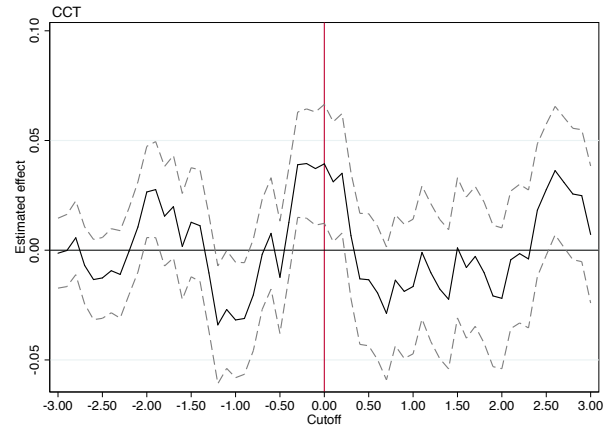**Figure A3.** RDD estimates using varying bandwidths.

Panel D: Finland
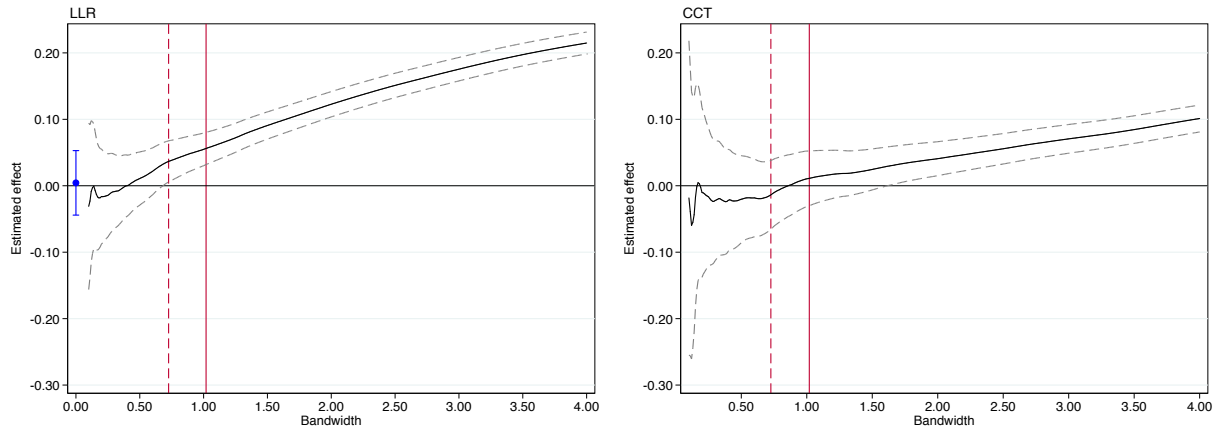
Figure A3 (continued). RDD estimates using varying bandwidths.

Notes: The dependent variable is a dummy for getting elected in $t+1$. Figures show RD estimates using varying bandwidths, and corresponding 95% confidence intervals constructed using standard errors clustered at the municipality level. The left-hand side figures use the conventional approach, whereas the right-hand side figures use the CCT approach. Dashed vertical lines mark the CER-optimal bandwidths, and solid lines the MSE-optimal bandwidths. We use a triangular kernel. The figures for Colombia and Finland also report the lottery estimates and their 95% confidence intervals.

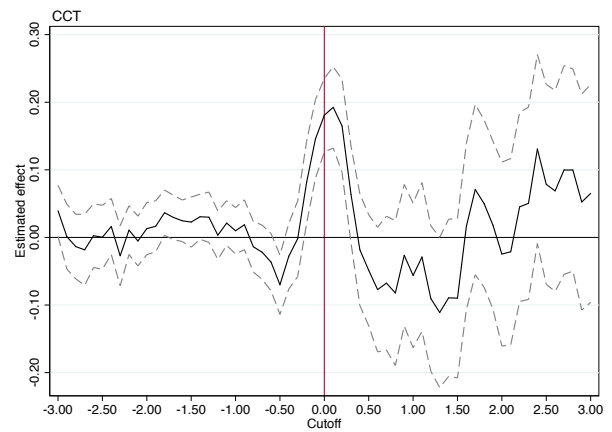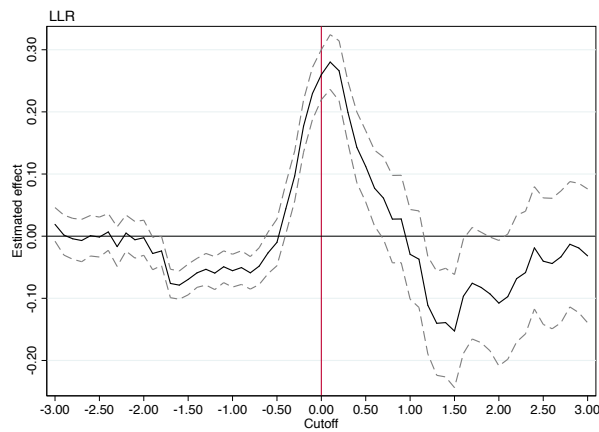Panel A: Brazil
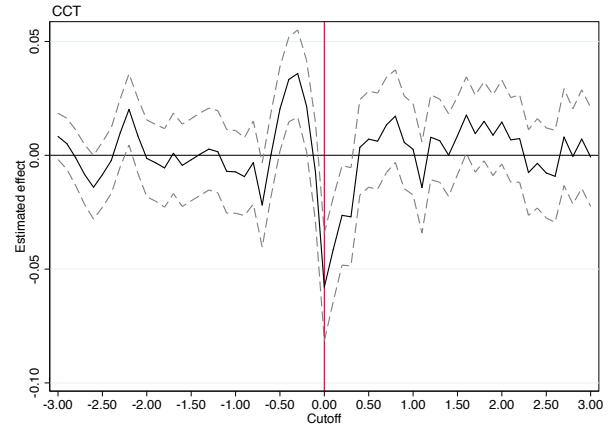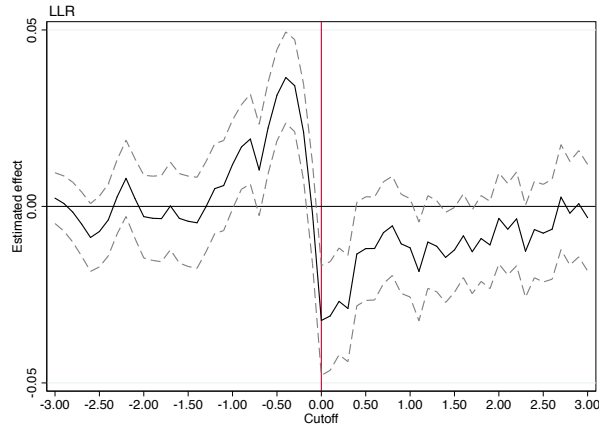
Panel B: Colombia
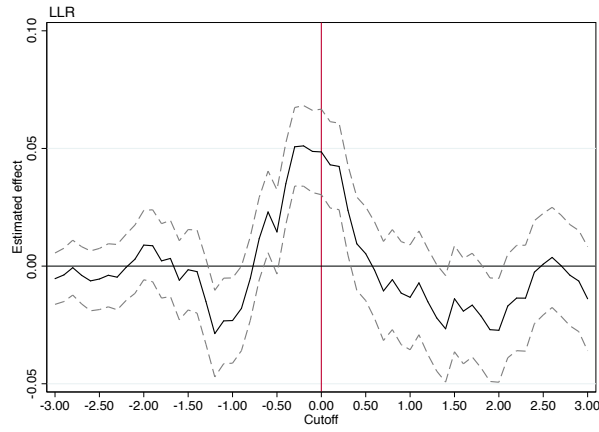
Panel C: Denmark

Figure A4. RDD estimates using artificial cutoffs.
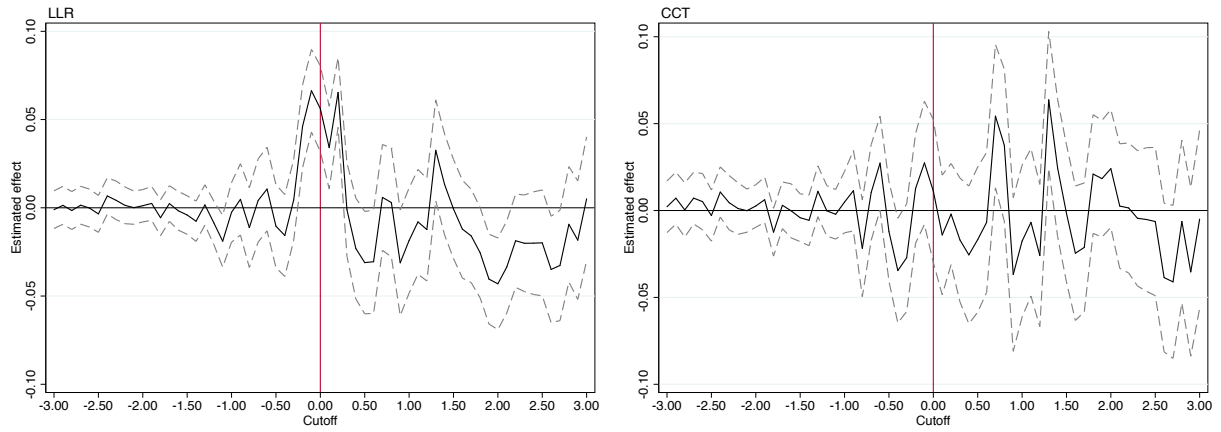
Panel D: Finland

Figure A4 (continued). RDD estimates using artificial cutoffs.

Notes: The dependent variable is a dummy for getting elected in $t + 1$. Figures show RD estimates using artificial cutoffs, and corresponding 95% confidence intervals constructed using standard errors clustered at the municipality level. The left-hand side figures use the conventional approach, whereas the right-hand side figures use the CCT approach. We use optimal bandwidths estimated at the true cutoff, and a triangular kernel.

Table A1. Effect on the lagged dependent variable, lotteries.

	Colombia	Finland
	(1)	(2)
Treatment	0.059	-0.024
	[-0.020,0.139]	[-0.074,0.026]
N	365	1351

Notes: The dependent variable is a dummy for being an incumbent. 95% confidence intervals based on standard errors clustered at the municipality level are reported in brackets. We use a triangular kernel. * and ** denote statistical significance at 5% and 1% levels, respectively.

Table A2. Effect on the lagged dependent variable, RDD.

	Brazil	Colombia	Denmark	Finland
	(1)	(2)	(3)	(4)
LLR	-0.020	0.004	0.054**	0.011
	[-0.045,0.006]	[-0.020,0.028]	[0.030,0.078]	[-0.023,0.045]
CCT	-0.031	-0.025	0.012	0.011
	[-0.070,0.009]	[-0.060,0.011]	[-0.022,0.047]	[-0.045,0.067]
N	28768	23375	7113	20036
Bandwidth	1.68	3.13	4.83	0.67

Notes: The dependent variable is a dummy for being an incumbent. Table shows RDD estimates using CER-optimal bandwidths. 95% confidence intervals based on standard errors clustered at the municipality level are reported in brackets. We use a triangular kernel. * and ** denote statistical significance at 5% and 1% levels, respectively.
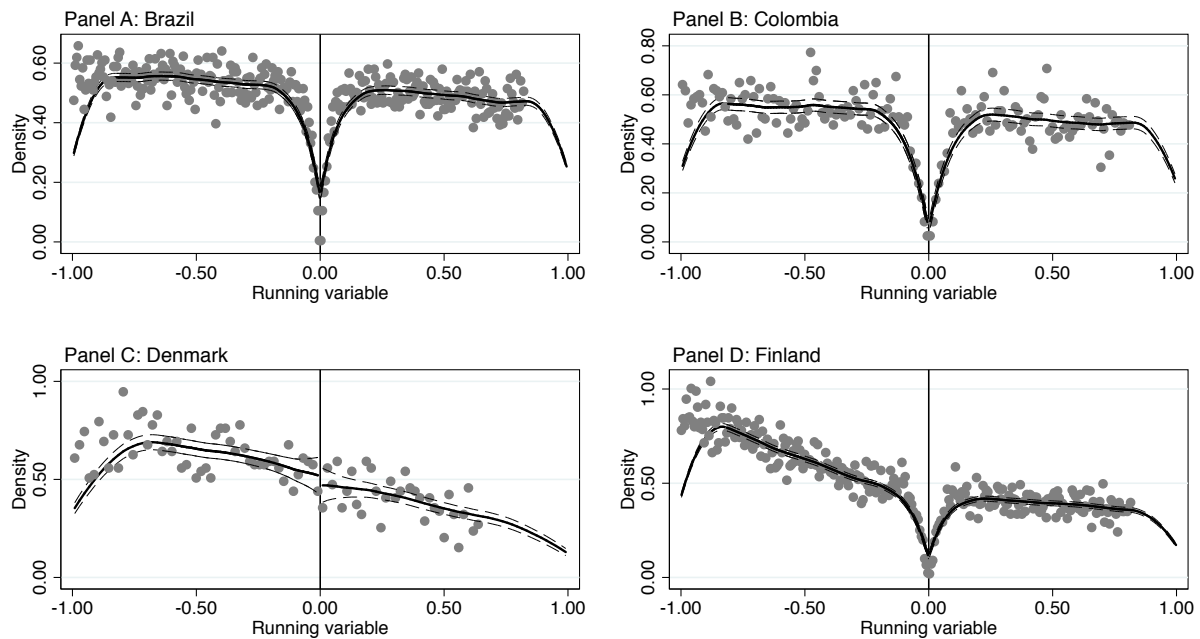
Figure A5. McCrary density test.

Notes: The figure illustrates McCrary (2008) density test graphically. Dashed lines are the 95% confidence intervals constructed using bootstrapped standard errors. We restrict the running variable between -1 and 1 and omit electoral ties at the cutoff.

Table A3. Density test.

Country	Conventional		Robust	
	T	*p*	*T*	*p*
Brazil	0.81	0.42	0.25	0.81
Colombia	-0.19	0.85	0.02	0.98
Denmark	-0.95	0.34	-0.31	0.75
Finland	3.21	0.00	0.22	0.83

Notes: Table shows the density test statistics and respective *p*-values from Cattaneo et al. (2018) test. We restrict the running variable between -1 and 1 and omit electoral ties at the cutoff.
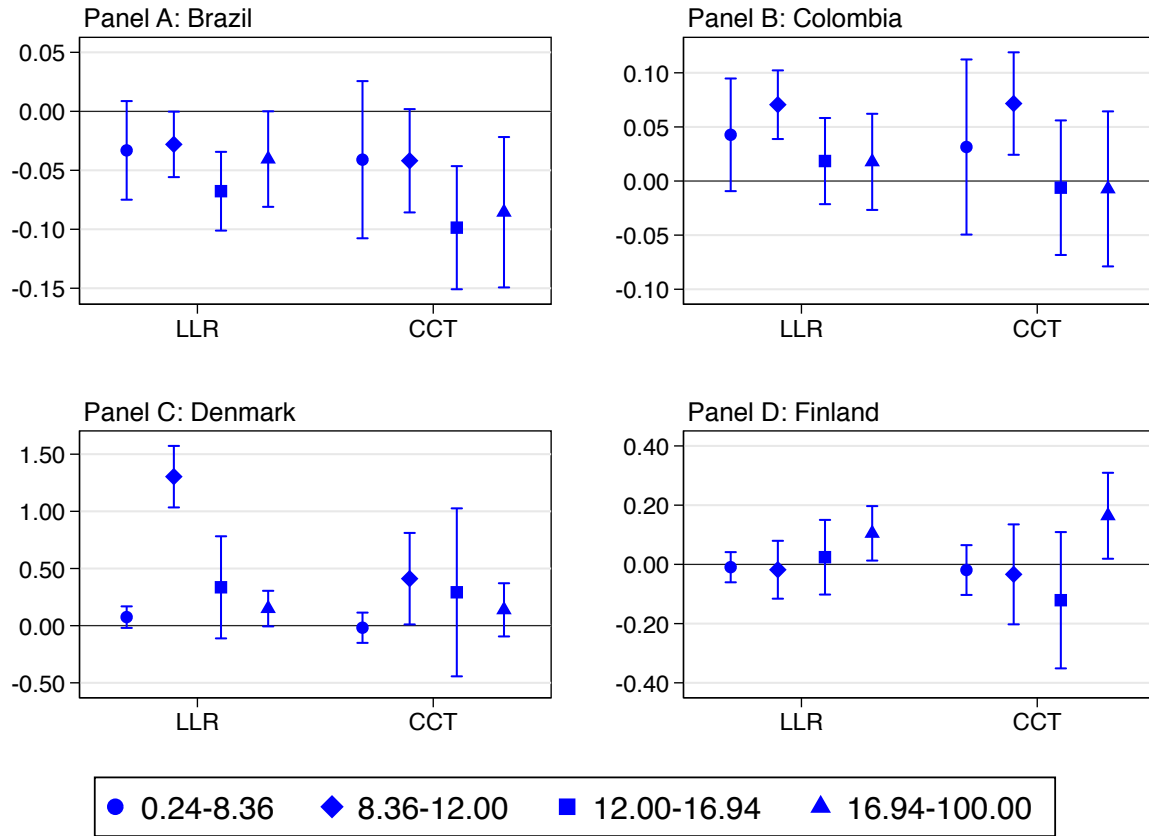
Figure A6. Heterogeneous effects by the cutoff value.

Notes: The dependent variable is a dummy for getting elected in $t + 1$. Figure shows conventional and robust RDD estimates using CER-optimal bandwidths, and their respective 95% confidence intervals. The confidence intervals are based on standard errors clustered at the municipality level. We use a triangular kernel. The sample is split in quartiles of the minimum within-party vote share among the elected.
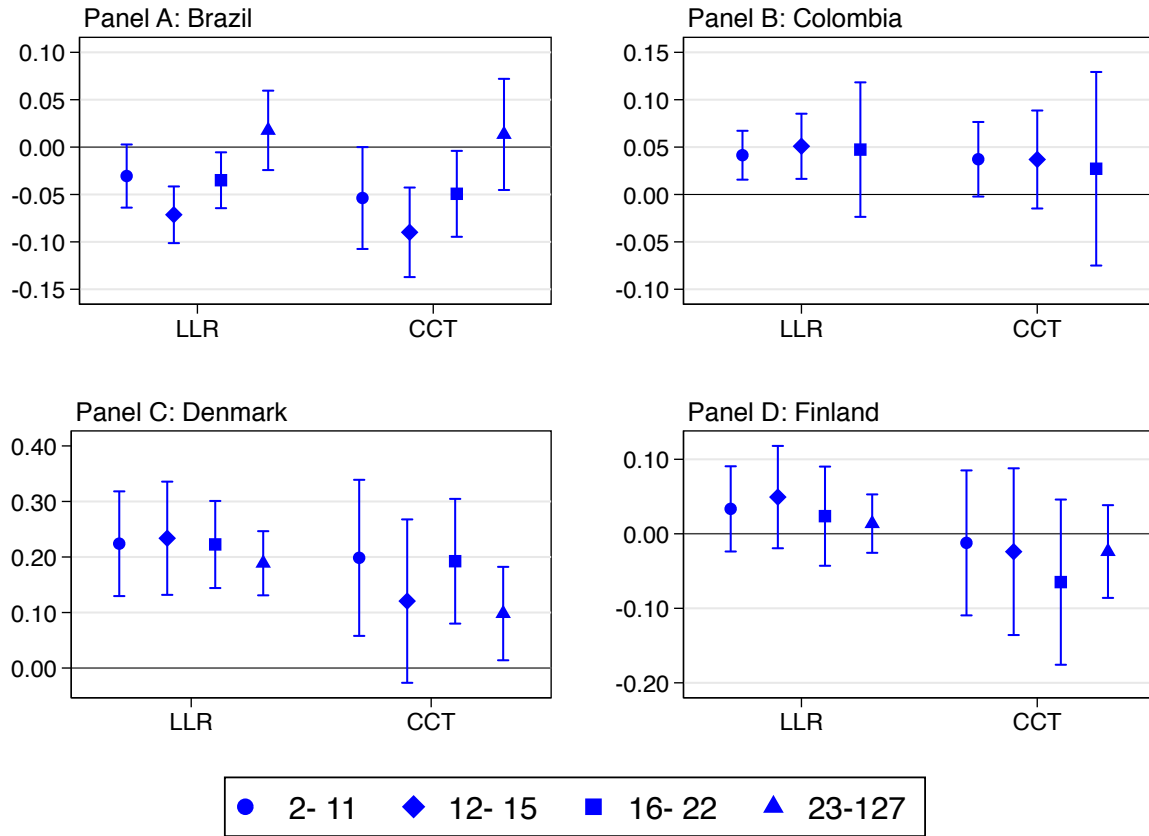
Figure A7. Heterogeneous effects by the number of candidates.

Notes: The dependent variable is a dummy for getting elected in $t+1$. Figure shows conventional and robust RDD estimates using CER-optimal bandwidths, and their respective 95% confidence intervals. The confidence intervals are based on standard errors clustered at the municipality level. We use a triangular kernel. The sample is split in quartiles of the number of candidates (at the party level). For Colombia, the third group pools together parties with 16-45 candidates.
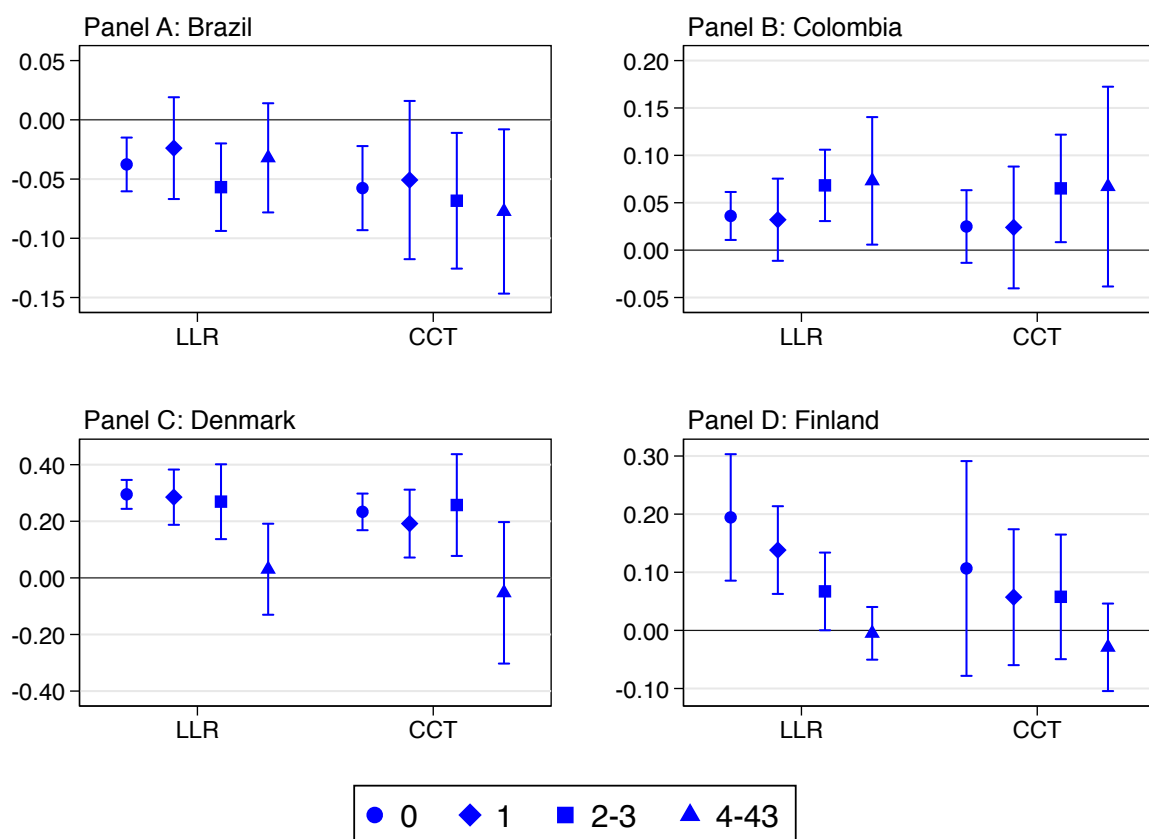
Figure A8. Heterogeneous effects by the number of incumbent candidates.

Notes: The dependent variable is a dummy for getting elected in $t+1$. Figure shows conventional and robust RDD estimates using CER-optimal bandwidths, and their respective 95% confidence intervals. The confidence intervals are based on standard errors clustered at the municipality level. We use a triangular kernel.

Table A4. Parametric RDD estimates.

	Panel A: Global linear control function			
	Brazil	Colombia	Denmark	Finland
	(1)	(2)	(3)	(4)
Elected	0.247**	0.201**	0.415**	0.452**
	[0.241,0.252]	[0.193,0.210]	[0.389,0.440]	[0.442,0.462]
N	585706	147558	12633	154543
	Panel B: Global quadratic control function			
	Brazil	Colombia	Denmark	Finland
	(5)	(6)	(7)	(8)
Elected	0.180**	0.153**	0.391**	0.413**
	[0.172,0.187]	[0.143,0.162]	[0.363,0.419]	[0.402,0.425]
N	585706	147558	12633	154543
	Panel C: Global cubic control function			
	Brazil	Colombia	Denmark	Finland
	(9)	(10)	(11)	(12)
Elected	0.124**	0.115**	0.381**	0.375**
	[0.116,0.132]	[0.103,0.126]	[0.351,0.411]	[0.363,0.388]
N	585706	147558	12633	154543
	Panel D: Global quartic control function			
	Brazil	Colombia	Denmark	Finland
	(13)	(14)	(15)	(16)
Elected	0.078**	0.085**	0.361**	0.342**
	[0.069,0.086]	[0.073,0.098]	[0.330,0.392]	[0.328,0.356]
N	585706	147558	12633	154543
	Panel E: Global quintic control function			
	Brazil	Colombia	Denmark	Finland
	(17)	(18)	(19)	(20)
Elected	0.044**	0.065**	0.339**	0.308**
	[0.035,0.053]	[0.051,0.078]	[0.306,0.371]	[0.293,0.324]
N	585706	147558	12633	154543

Notes: The dependent variable is a dummy for getting elected in $t +$ 1. 95% confidence intervals based on standard errors clustered at the municipality level are reported in brackets. * and ** denote statistical significance at 5% and 1% levels, respectively.

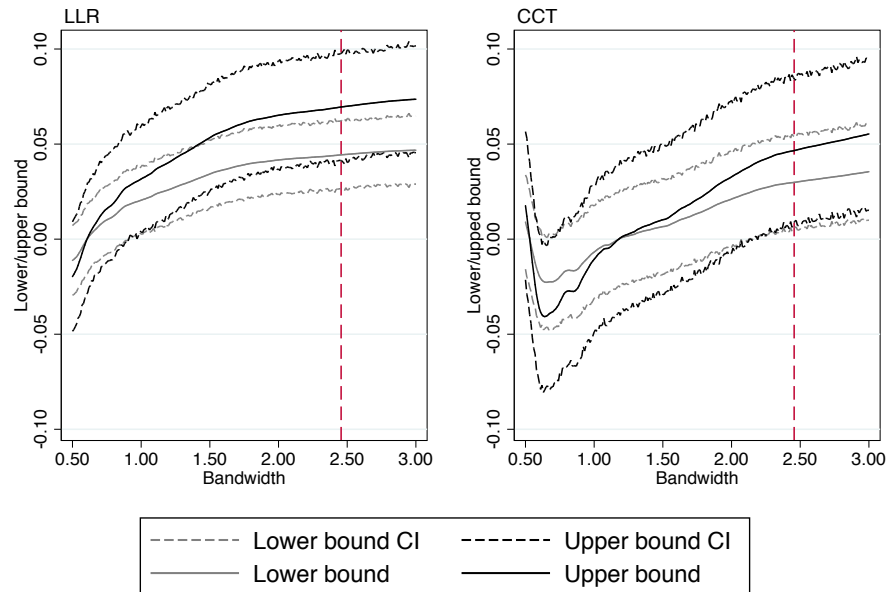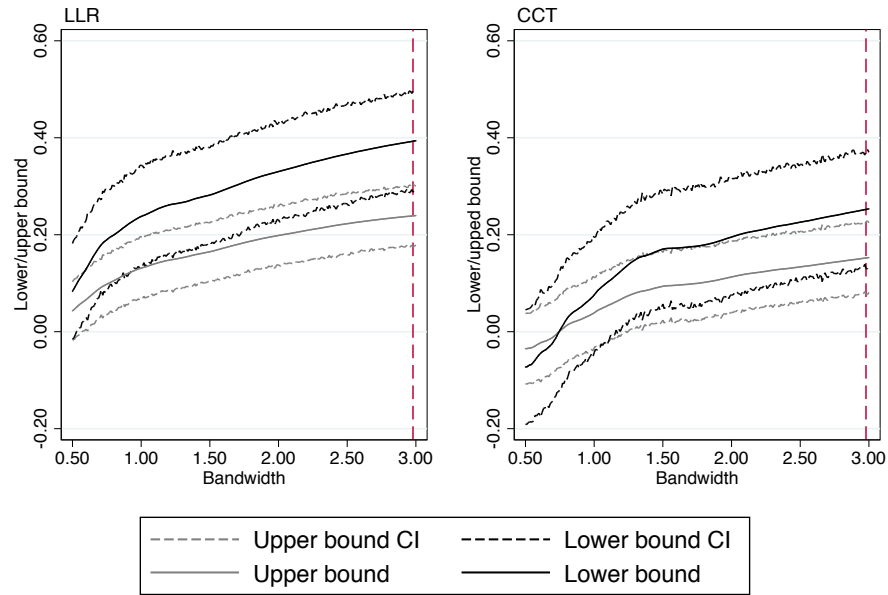Panel A: Brazil
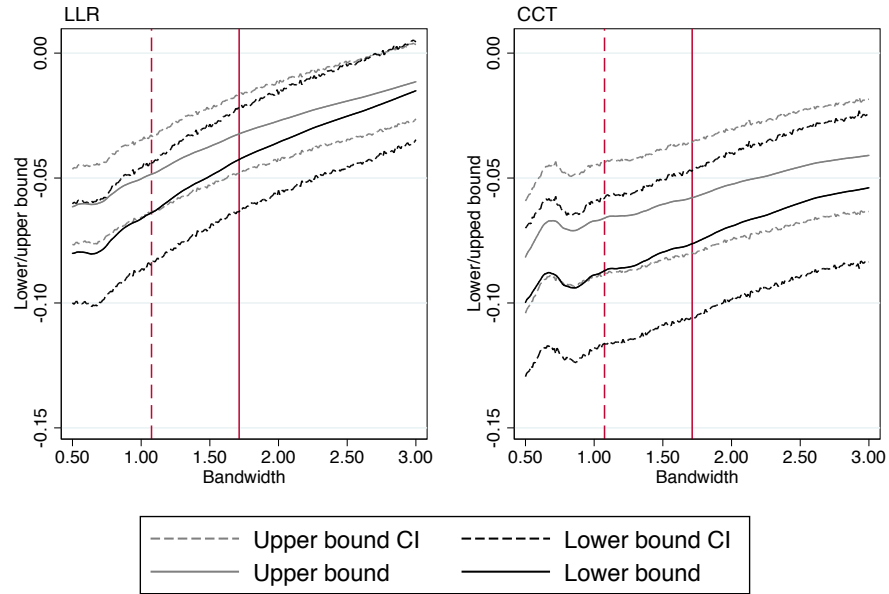
Panel B: Colombia

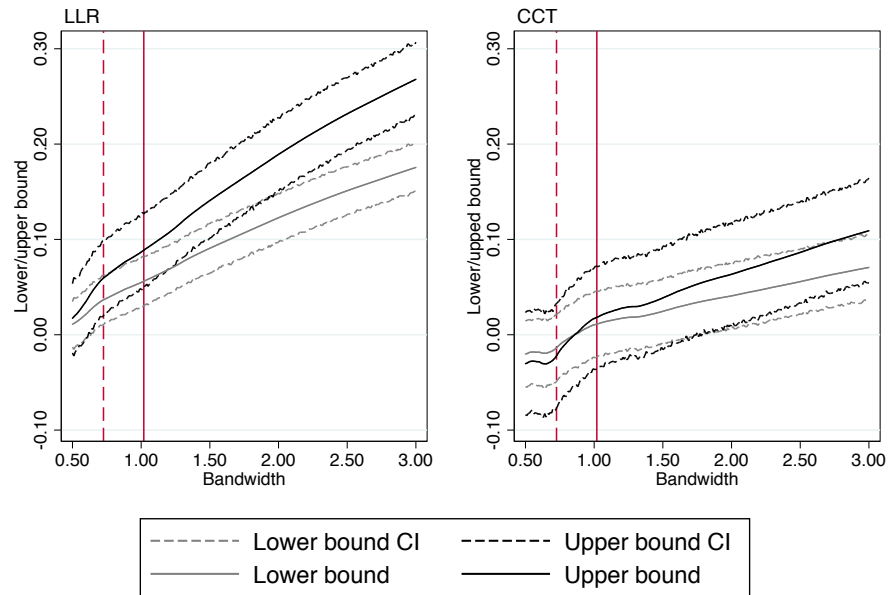**Figure A9.** Bounds estimated using alternative bandwidths.

Figure A9 (continued). Bounds estimated using alternative bandwidths.

Notes: Figures show De Magalhaes and Hirvonen (2019) bounds for the incumbency advantage and the corresponding 95% confidence intervals constructed using bootstrapped standard errors. The left-hand side figures use the conventional approach, whereas the right-hand side figures use the CCT approach. All regressions use a triangular kernel.
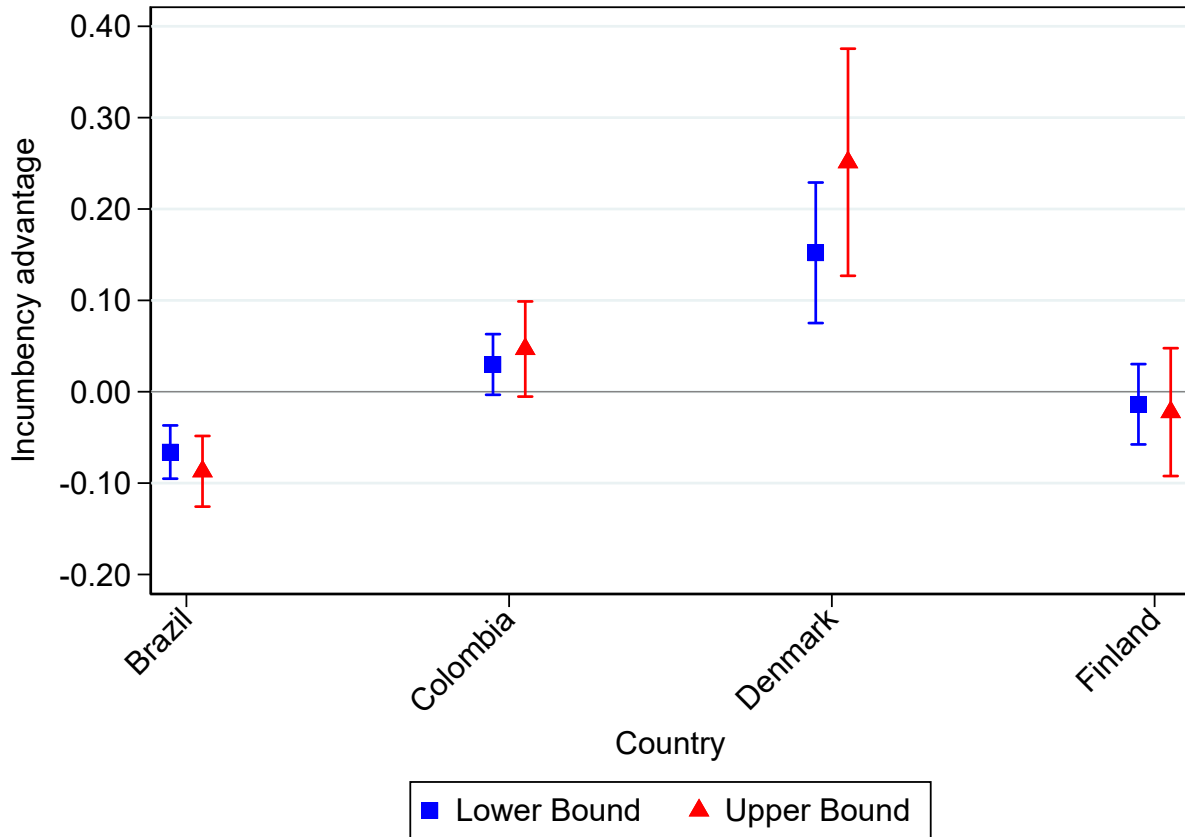
Figure A10. De Magalhães and Hirvonen (2019) bounds for the personal incumbency advantage (CCT approach).

Notes: The figure shows the upper and lower bound estimates of the individual incumbency advantage and their bootstrapped 95% confidence intervals. The bounds are estimated following De Magalhães and Hirvonen (2019). The upper bound is equal to the estimated individual unconditional incumbency advantage divided by the estimated probability of re-running for incumbents. The lower bound is the estimated individual and unconditional incumbency advantage. To compute the bounds, we use the CCT approach and CER-optimal bandwidths.
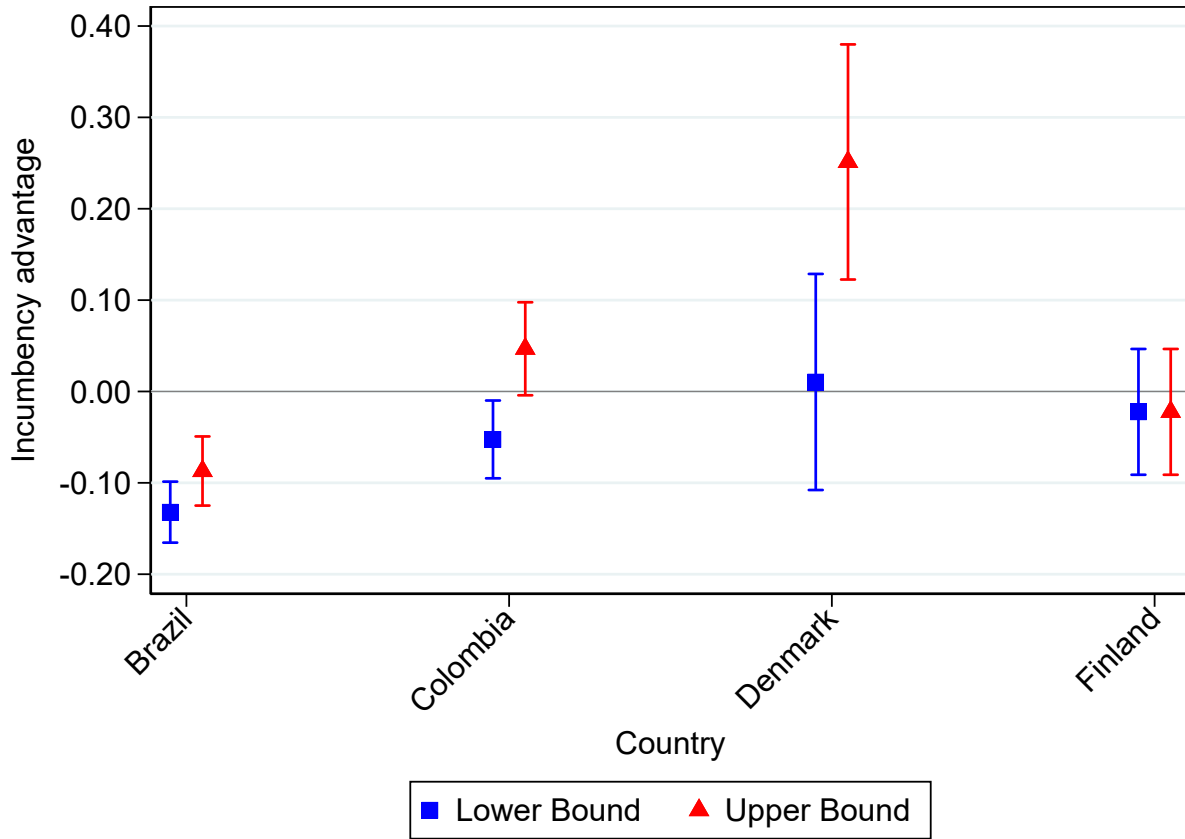
Figure A11. Anagol and Fujiwara (2016) bounds for the personal incumbency advantage (CCT approach).

Notes: Figure shows the upper and lower bound estimates of the personal incumbency advantage and their bootstrapped 95% confidence intervals. The bounds are estimated following Anagol and Fujiwara (2016) The upper bound is equal to the estimated unconditional incumbency advantage divided by the estimated probability of re-running for incumbents. The lower bound is equal to the upper bound subtracted by the multiplication of the estimated effect of incumbency on re-running and the estimated probability of incumbents getting elected conditional on re-running. To compute the bounds, we use the CCT approach and CER-optimal bandwidths.

References

Anagol, Santosh and Thomas Fujiwara (2016). The runner-up effect. *Journal of Political Economy 124*(4), 927–991.

Ashworth, Scott and Ethan Bueno de Mesquita (2008). Electoral selection, strategic challenger entry, and the incumbency advantage. *Journal of Politics 70*(4), 1006–1025.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2018). On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference. *Journal of the American Statictical Association 113*(522), 767–779.

Calonico, Sebastian, Matias D. Cattaneo, and Roío Titiunik (2015). Optimal Data-Driven Regression Discontinuity Plots. *Journal of the American Statistical Association 110*(512), 1753–1769.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma (2018). Manipulation testing based on density discontinuity. *Stata Journal 18*(1), 234–261.

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), 478–497.

De Magalhães, Leandro and Salomo Hirvonen (2019). The Incumbent-Challenger Advantage and the Winner-Runner-up Advantage. Discussion Paper 19/710, Department of Economics, University of Bristol. Available at `https://ideas.repec.org/p/bri/uobdis/19-710.html`.

Erikson, Robert and Rocio Titiunik (2015). Using Regression Discontinuity to Uncover the Personal Incumbency Advantage. *Quarterly Journal of Political Science 10*(1), 101–119.

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*, 501–531.

King, Gary (1991). Constituency service and incumbency advantage. *British Journal of Political Science 21*(1), 119–128.

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

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi