(online) **Appendix** (ver. 1.02). Supporting information *for* "Magnitude Matters: Voter Turnout Under Different Electoral Systems"

Michał Pierzgalski*[1], Maciej A. Górecki[2], and Paweł Stępień[1]

[1]University of Lodz, Lodz, POLAND
[2]University of Warsaw, Warsaw, POLAND

June 6, 2025

*Corresponding author: pierzgal@gmail.com

Contents

Estimation and Inference **3**
Main RDD Results for the Polish Municipal Elections (1998-2010 Pooled) . . 8
RDD Results for the 1998 Polish Municipal Elections 10
Additional Models 11
"Donut Hole" RDD Approach 16
RDD Plots 18

Assessing RDD Validity: Falsification Tests **19**
Placebo Test: The 2014 Local Elections with Temporarily Suspended Population Threshold 19
Testing Balance in Covariates and Placebo Outcomes 21
Absence of Manipulation in the Forcing Variable 23
Supplementary Manipulation Testing Results 27
Power Analysis for Main Models 31

Data and Variables **35**

Replication Materials **35**

Estimation and Inference

Municipality Type	1998 Assembly Elections	2002–2010	2014	2018
		Concurrent Assembly/Executive Elections		
County-status cities (GM)	OLPR	OLPR	OLPR	OLPR
Medium-sized (GP)	OLPR	OLPR	SMDP	OLPR
Small (GW)	SMDP, MMDP	SMDP, MMDP	SMDP	SMDP

Table 1: Voting systems in Polish municipal elections (1998–2018)

Notes: GM: county-status cities; GP: medium-sized municipalities ($> 20,000$ inhabitants); GW: small municipalities (at most 20,000 inhabitants). Electoral thresholds: 5% required in 2018 and 2006 (all communes); 2002–2014 (county-status communes only); no threshold in 1998. OLPR: Open-list proportional representation; SMDP: Single-member district plurality; MMDP: Multi-member district plurality.
Source: Sokół (2015: 165)

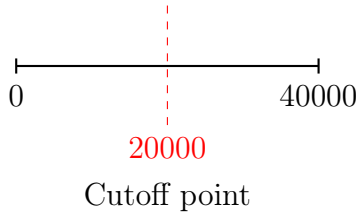


Figure 1: The running variable interval (population of municipalities) with the cutoff point at 20,000.

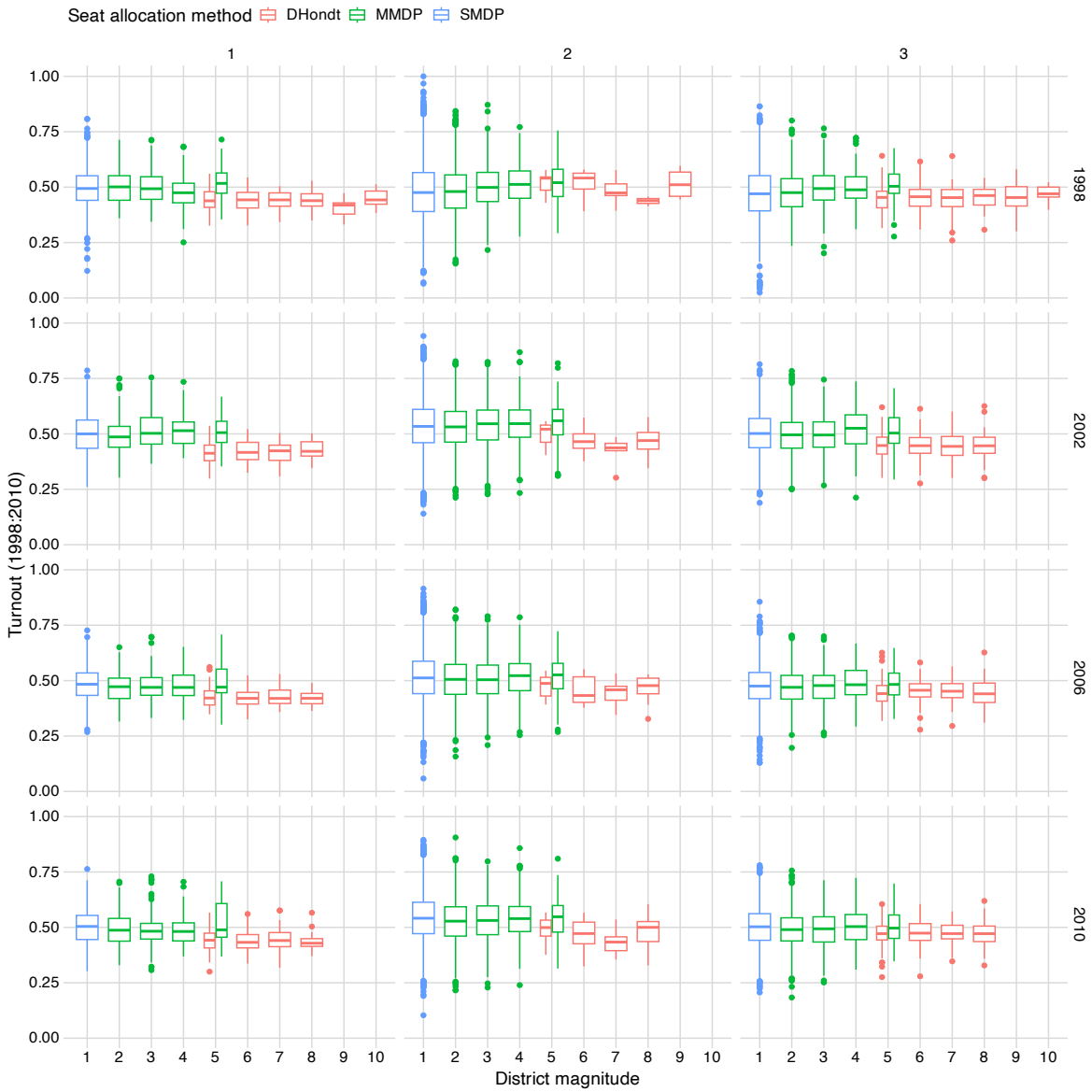


Figure 2: The 1998-2010 municipal elections: descriptive analysis using box plots for municipalities with populations ranging from 0 to 40000.

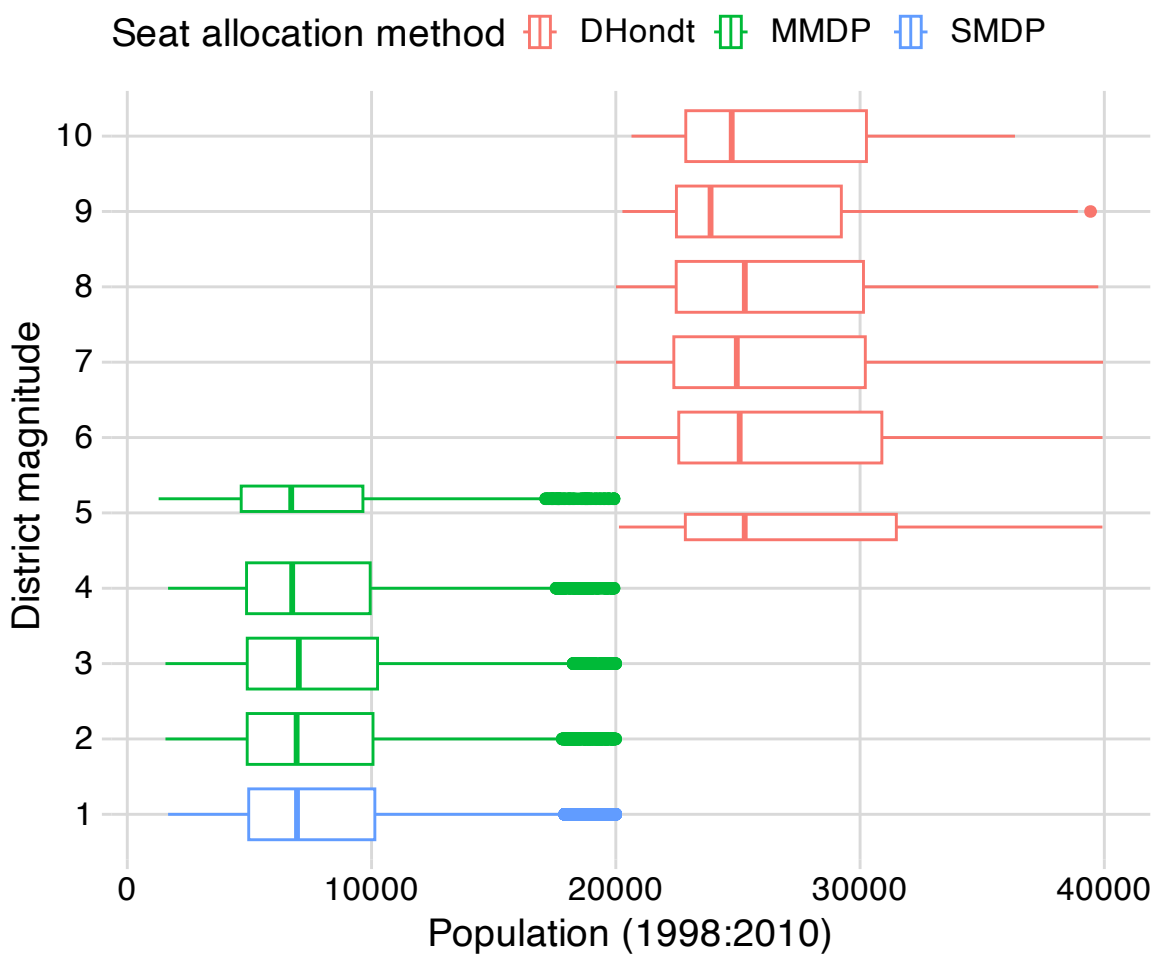


Figure 3: The 1998-2010 municipal elections (municipalities with populations ranging from 0 to 40000): district magnitude *versus* municipal population

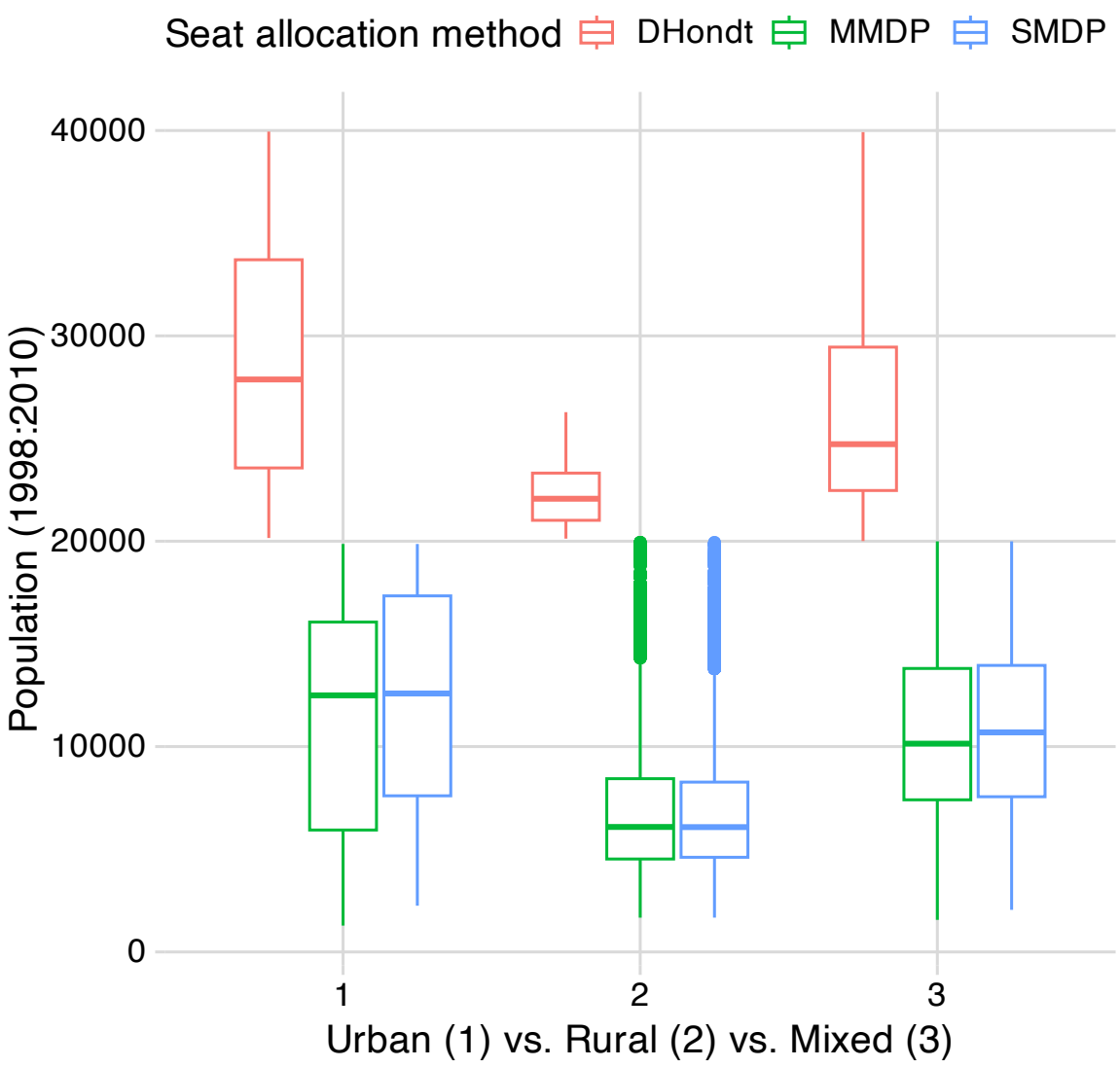


Figure 4: The 1998-2010 municipal elections (municipalities with populations ranging from 0 to 40000): municipality type *versus* municipal population

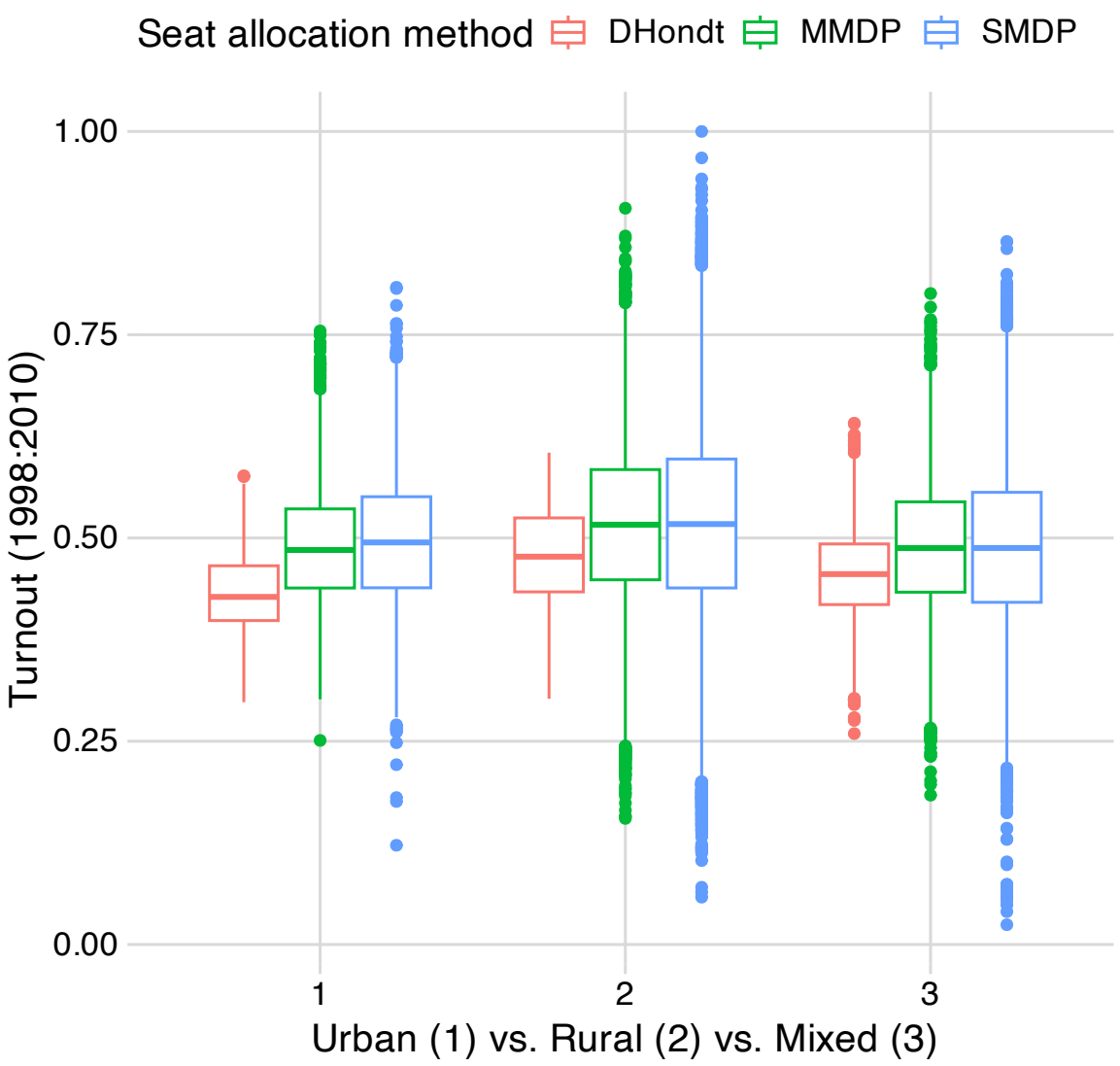


Figure 5: The 1998-2010 municipal elections (municipalities with populations ranging from 0 to 40000): municipality type *versus* turnout

Main RDD Results for the Polish Municipal Elections (1998-2010 Pooled)

Table 2: Sharp RD Estimates from the Main District-Level Analysis, 1998-2010 Pooled

Model	Type	Estimate (95% CI)	p-value	N Total	N Effective	Bandwidth
				(Left, Right)	(Left, Right)	(Left, Right)
M1. SMDP-PR [clustervar]	Bias-Corrected	0.031 (0.022, 0.040)	0.000	(53540, 2786)	(1405, 761)	(1924.31, 2862.55)
M1. SMDP-PR [clustervar]	Robust	0.031 (0.021, 0.042)	0.000	(53540, 2786)	(1405, 761)	(1924.31, 2862.55)
M2. SMDP-PR [CER, clustervar]	Bias-Corrected	0.033 (0.023, 0.042)	0.000	(53540, 2786)	(1249, 698)	(1759.42, 2617.25)
M2. SMDP-PR [CER, clustervar]	Robust	0.033 (0.022, 0.043)	0.000	(53540, 2786)	(1249, 698)	(1759.42, 2617.25)
M3. MMDP-PR [clustervar]	Bias-Corrected	0.002 (-0.006, 0.011)	0.606	(28707, 2786)	(406, 743)	(1434.75, 2835.93)
M3. MMDP-PR [clustervar]	Robust	0.002 (-0.007, 0.012)	0.638	(28707, 2786)	(406, 743)	(1434.75, 2835.93)
M4. MMDP-PR [CER, clustervar]	Bias-Corrected	0.002 (-0.007, 0.010)	0.682	(28707, 2786)	(384, 691)	(1311.80, 2592.91)
M4. MMDP-PR [CER, clustervar]	Robust	0.002 (-0.008, 0.011)	0.714	(28707, 2786)	(384, 691)	(1311.80, 2592.91)
M5. SMDP-PR	Bias-Corrected	0.017 (0.008, 0.025)	0.000	(53540, 2786)	(4090, 940)	(4600.00, 3372.48)
M5. SMDP-PR	Robust	0.017 (0.007, 0.026)	0.001	(53540, 2786)	(4090, 940)	(4600.00, 3372.48)
M6. SMDP-PR [CER]	Bias-Corrected	0.019 (0.008, 0.031)	0.001	(53540, 2786)	(2037, 480)	(2662.09, 1951.70)
M6. SMDP-PR [CER]	Robust	0.019 (0.008, 0.031)	0.001	(53540, 2786)	(2037, 480)	(2662.09, 1951.70)
M7. MMDP-PR	Bias-Corrected	-0.000 (-0.009, 0.008)	0.970	(28707, 2786)	(1339, 940)	(3526.32, 3383.42)
M7. MMDP-PR	Robust	-0.000 (-0.010, 0.010)	0.974	(28707, 2786)	(1339, 940)	(3526.32, 3383.42)
M8. MMDP-PR [CER]	Bias-Corrected	0.002 (-0.009, 0.013)	0.740	(28707, 2786)	(648, 501)	(2100.93, 2015.79)
M8. MMDP-PR [CER]	Robust	0.002 (-0.010, 0.014)	0.752	(28707, 2786)	(648, 501)	(2100.93, 2015.79)
M9. MMDP-PR [DM = 5, clustervar]	Bias-Corrected	-0.023 (-0.039, -0.008)	0.003	(2074, 1084)	(97, 493)	(3336.67, 4752.29)
M10. MMDP-PR [DM = 5, clustervar]	Robust	-0.023 (-0.040, -0.007)	0.005	(2074, 1084)	(97, 493)	(3336.67, 4752.29)
M11. MMDP-PR [DM = 5, CER, clustervar]	Bias-Corrected	-0.022 (-0.037, -0.006)	0.006	(2074, 1084)	(86, 460)	(3050.74, 4345.06)
M12. MMDP-PR [DM = 5, CER, clustervar]	Robust	-0.022 (-0.038, -0.006)	0.008	(2074, 1084)	(86, 460)	(3050.74, 4345.06)
M13. MMDP-PR [DM = 5]	Bias-Corrected	-0.006 (-0.022, 0.010)	0.485	(2074, 1084)	(54, 572)	(2180.64, 5772.14)
M14. MMDP-PR [DM = 5]	Robust	-0.006 (-0.024, 0.012)	0.536	(2074, 1084)	(54, 572)	(2180.64, 5772.14)
M15. MMDP-PR [DM = 5, CER]	Bias-Corrected	-0.001 (-0.019, 0.017)	0.939	(2074, 1084)	(40, 384)	(1457.52, 3858.04)
M16. MMDP-PR [DM = 5, CER]	Robust	-0.001 (-0.020, 0.018)	0.942	(2074, 1084)	(40, 384)	(1457.52, 3858.04)

Notes: *N* represents the number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with the triangular kernel. clustervar: municipality type (urban, rural, mixed) is used for cluster-robust variance estimation. Election year is included as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

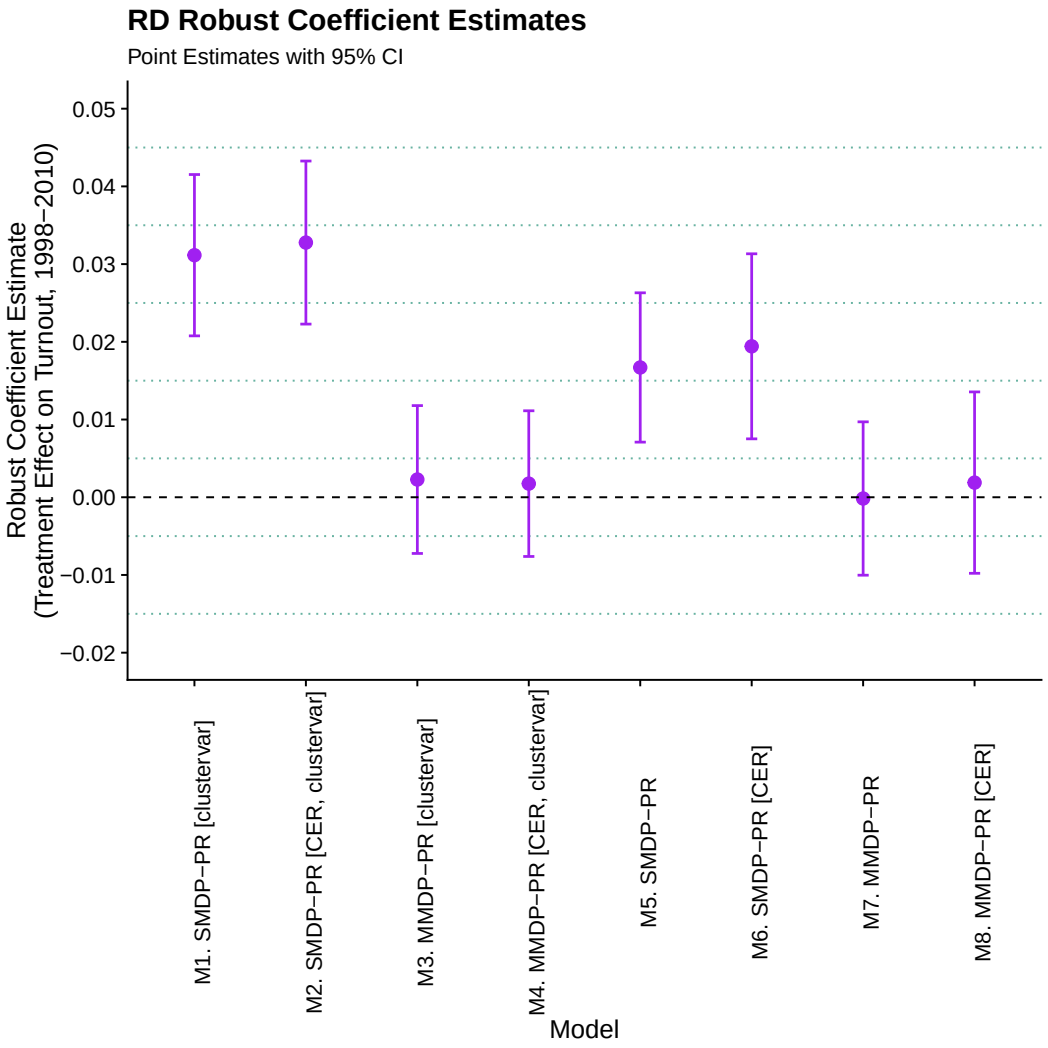


Figure 6: 1998-2010 elect. (controlling for election years)

Figure 7: The plots summarize the main RDD estimations for district-level turnout (In models 1-4, we applied cluster-robust variance estimation; clustering by municipality type: urban, rural, mixed).

RDD Results for the 1998 Polish Municipal Elections

Table 3: Sharp RD estimates using local polynomial regression (1998)

Model	Estimates		p-value	Sample Size
	Estimate	95% CI		N eff. (Left, Right)
M1. SMDP-PR	0.040	(0.020, 0.059)	0.000	(862, 269)
M2. SMDP-PR [clustervar]	0.037	(0.032, 0.042)	0.000	(176, 74)
M3. SMDP-PR [CER]	0.038	(0.015, 0.061)	0.001	(515, 150)
M4. SMDP-PR [CER, clustervar]	0.024	(0.018, 0.030)	0.000	(167, 70)
M5. MMDP-PR	-0.003	(-0.020, 0.013)	0.692	(573, 269)
M6. MMDP-PR [clustervar]	-0.013	(-0.016, -0.010)	0.000	(117, 94)
M7. MMDP-PR [CER]	0.000	(-0.018, 0.019)	0.974	(282, 157)
M8. MMDP-PR [CER, clustervar]	-0.018	(-0.020, -0.015)	0.000	(70, 82)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with the triangular kernel. clustervar: municipality type (urban, rural, mixed) is used for cluster-robust variance estimation. Election year is included as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

Additional Models

Our supplementary analyses strengthen the main findings across different periods, municipality categories, and levels of data aggregation. Figure 8 presents comprehensive visualizations of our findings, with panel (a) showing district-level results and panel (b) displaying municipality-level patterns.

In further analyses, we focus only on urban and mixed urban-rural municipalities, excluding rural ones. This choice is based on several factors. Most rural municipalities have populations below 20,000. In 1998, for instance, the distribution of constituencies in rural municipalities was highly skewed: a mere 33 exceeded the threshold, while a staggering 17,007 fell below it. By excluding them, we improve the comparability of our sample and address the tendency for higher voter turnout in small rural municipalities. This approach also helps reduce the impact of rural-specific factors that might affect our results, allowing for a more accurate estimation of how electoral rules affect turnout.

First, our focused analysis of the 1998 elections—the only ones without concurrent executive contests—shows stronger effects than the pooled analysis. At the district level (Table 4 and Figure 8(a) - left panel), PR outperforms SMDP with a 4.0 percentage point increase in turnout (95% CI: [0.020, 0.059], $p < 0.001$). This effect strengthens to 4.7 percentage points when excluding rural municipalities. The MMDP-PR comparison remains insignificant across all specifications, which is consistent with our main findings. When using CER-optimal bandwidth, these patterns persist, with effects ranging from 3.8 to 5.7 percentage points for SMDP-PR comparisons.

For concurrent elections (2002-2010), the district-level analysis (Table 5 and Figure 8(a) - right panel) shows a weaker but still significant PR advantage over SMDP when excluding rural municipalities (2.6 percentage points, $p < 0.001$), increasing to 4.0 percentage points under CER-optimal bandwidth. The MMDP-PR comparison remains insignificant across all specifications, reinforcing the importance of district magnitude in shaping turnout[1].

Our municipality-level analysis, comparing PR with plurality systems jointly (Tables 6–8 and Figure 8(b)), shows more modest effects. In the pooled analysis (1998-2010), PR shows a positive but statistically insignificant effect on turnout (1.7 percentage

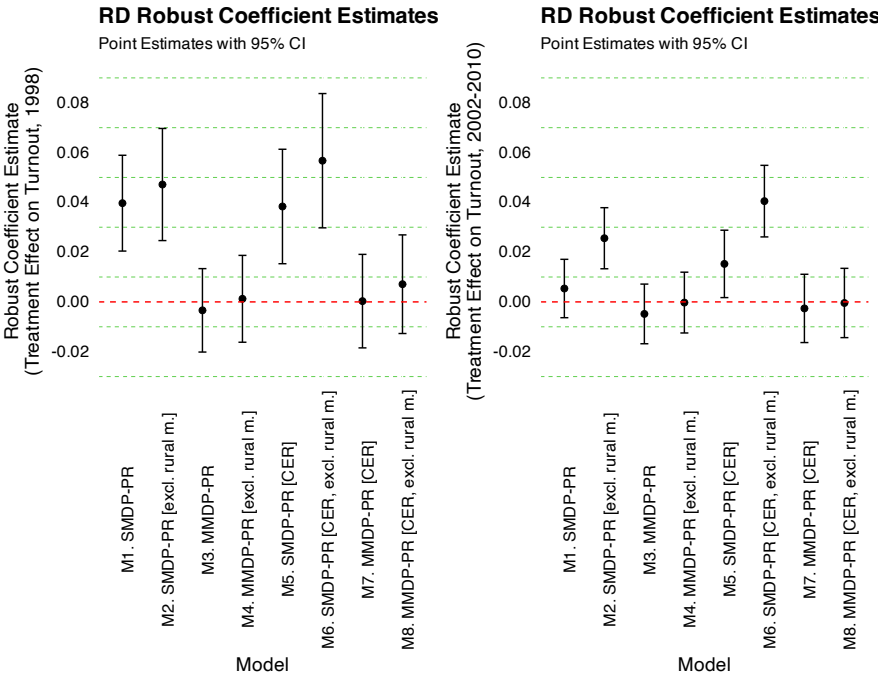
[1]Besides, we found no substantial evidence of increasing OLPR impact when comparing the 1998 non-concurrent elections to the 2002-2010 concurrent elections. This divergence likely stems from Poland's institutional context, where plurality rules were already established through the 1990 and 1994 elections, potentially dampening any temporal learning effects. Analysis of concurrent elections (2002, 2006, and 2010) at the district level showed that the increase in voter turnout under open-list proportional representation, compared to single-member district plurality, was less pronounced than in the 1998 RDD analysis. Consistent with our main findings, comparisons between multi-member district plurality and PR systems (for the 2002-2010 period) show no significant differences across all specifications, reinforcing the importance of district magnitude in shaping turnout.

points, 95% CI: [-0.007, 0.041], p = 0.164). This pattern persists under CER-optimal bandwidth (1.8 percentage points, p = 0.156). The 1998 non-concurrent election analysis (Figure 8(b) - left panel) yields marginally significant positive effects under certain specifications, notably a 3.8 percentage point increase with CER-optimal bandwidth (p = 0.017). For the 2002-2010 period (Figure 8(b) - right panel), effects remain positive but insignificant, ranging from 1.1 to 2.1 percentage points across specifications.

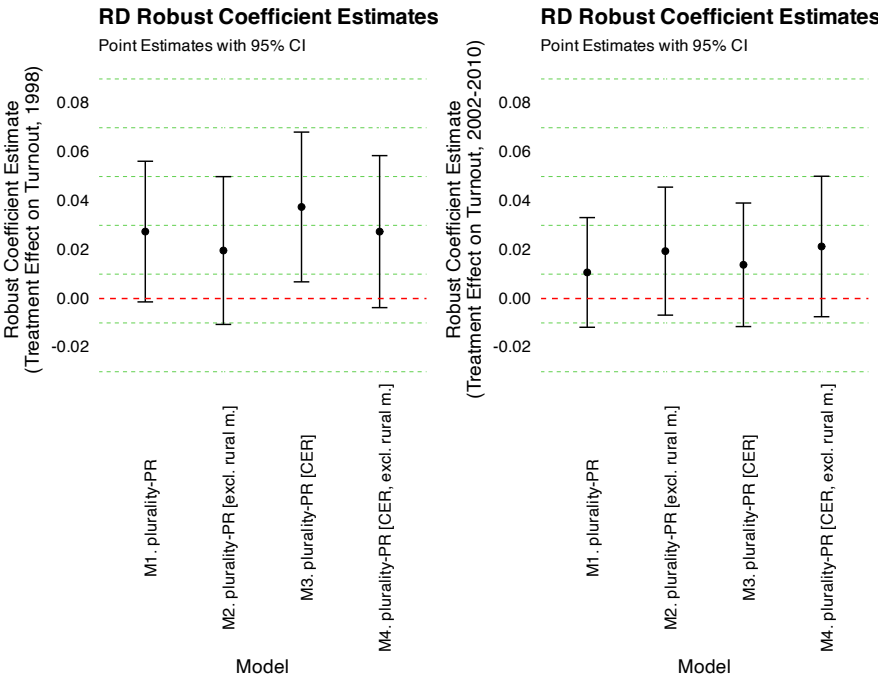
These supplementary analyses reinforce four key aspects of our findings:

1. Electoral system effects are more pronounced in non-concurrent elections (1998)
2. The PR advantage over SMDP is stronger in non-rural areas
3. The MMDP-PR comparison consistently shows no significant differences
4. Municipality-level effects are generally weaker than district-level effects, suggesting that mixing SMDP and MMDP within municipalities may dilute the PR effect

It is important to note certain limitations in these analyses. The 1998 election analysis, particularly at the municipality level, may be underpowered due to smaller sample sizes, warranting cautious interpretation of effect magnitudes. This is especially evident in the reduced effective sample sizes for CER-optimal bandwidth specifications (see Tables 4 and 7). Additionally, while excluding rural municipalities helps isolate effects in more comparable contexts, it further reduces sample sizes. However, the consistency of findings across multiple specifications and time periods, despite these limitations, strengthens our confidence in the overall patterns observed.



(a) District-level Turnout: 1998 vs. 2002-2010; Rural Units Excluded vs. Included



(b) Municipality-level turnout: 1998 vs. 2002-2010; Rural Units Excluded vs. Included

Figure 8: Separate examination of voting turnout patterns in the 1998 election and the 2002-2010 electoral cycles.

Table 4: Sharp RD Estimates Using Local Polynomial Regression (1998): Rural Units Excluded vs. Included

Model	Estimates		p-value	Sample Size
	Estimate	95% CI		N eff. (Left, Right)
M1. SMDP-PR	0.040	(0.020, 0.059)	0.000	(862, 269)
M2. SMDP-PR [excl. rural m.]	0.047	(0.025, 0.070)	0.000	(499, 300)
M3. MMDP-PR	-0.003	(-0.020, 0.013)	0.692	(573, 269)
M4. MMDP-PR [excl. rural m.]	0.001	(-0.016, 0.019)	0.889	(357, 313)
M5. SMDP-PR [CER]	0.038	(0.015, 0.061)	0.001	(515, 150)
M6. SMDP-PR [CER, excl. rural m.]	0.057	(0.030, 0.084)	0.000	(273, 186)
M7. MMDP-PR [CER]	0.000	(-0.018, 0.019)	0.974	(282, 157)
M8. MMDP-PR [CER, excl. rural m.]	0.007	(-0.013, 0.027)	0.482	(171, 201)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ RDD with a triangular kernel. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

Table 5: Sharp RD Estimates Using Local Polynomial Regression (2002, 2006, 2010): Rural Units Excluded vs. Included

Model	Estimate (95% CI)	p-value	Sample Size
			N eff. (Left, Right)
M1. SMDP-PR	0.005 (-0.006, 0.017)	0.369	(1746, 912)
M2. SMDP-PR [excl. rural m.]	0.026 (0.013, 0.038)	0.000	(982, 971)
M3. MMDP-PR	-0.005 (-0.017, 0.007)	0.429	(727, 920)
M4. MMDP-PR [excl. rural m.]	-0.000 (-0.012, 0.012)	0.965	(590, 965)
M5. SMDP-PR [CER]	0.015 (0.002, 0.029)	0.027	(874, 504)
M6. SMDP-PR [CER, excl. rural m.]	0.040 (0.026, 0.055)	0.000	(526, 649)
M7. MMDP-PR [CER]	-0.003 (-0.016, 0.011)	0.708	(388, 530)
M8. MMDP-PR [CER, excl. rural m.]	-0.000 (-0.014, 0.013)	0.951	(342, 653)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with the triangular kernel, using election year as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

Table 6: Sharp RD estimates using local polynomial regression (1998, 2002, 2006, 2010)

Model	Estimates		p-value	Sample Size
	Estimate	95% CI		N eff. (Left, Right)
M1. plurality-PR (1998:2010)	0.017	(-0.007, 0.041)	0.164	(513, 292)
M2. plurality-PR (1998:2010, CER)	0.018	(-0.007, 0.042)	0.156	(452, 264)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with triangular kernel and cluster-robust variance estimation (clustering by municipality type: urban, rural, mixed), using election year as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

Table 7: Sharp RD estimates using local polynomial regression (1998)

Model	Estimate (95% CI)	p-value	Sample Size
			N eff. (Left, Right)
M1. plurality-PR (98)	0.027 (-0.001, 0.056)	0.062	(113, 92)
M2. plurality-PR (98) [excl. rural m.]	0.020 (-0.011, 0.050)	0.203	(111, 93)
M3. plurality-PR (98) [CER]	0.038 (0.007, 0.068)	0.017	(69, 66)
M4. plurality-PR (98) [CER, excl. rural m.]	0.027 (-0.004, 0.059)	0.084	(74, 74)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ RDD with a triangular kernel. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

Table 8: Sharp RD estimates using local polynomial regression; municipal-level (2002, 2006, 2010)

Model	Estimate (95% CI)	p-value	Sample Size
			N eff. (Left, Right)
M1. plurality-PR (2002-2010)	0.011 (-0.012, 0.033)	0.351	(357, 303)
M2. plurality-PR (2002-2010) [excl. rural m.]	0.019 (-0.007, 0.046)	0.147	(263, 300)
M3. plurality-PR (2002-2010) [CER]	0.014 (-0.011, 0.039)	0.284	(190, 195)
M4. plurality-PR (2002-2010) [CER, excl. rural m.]	0.021 (-0.007, 0.050)	0.146	(161, 225)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with the triangular kernel, using election year as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

"Donut Hole" RDD Approach

We implement the "donut hole" RDD approach, which shows that our main estimates are robust to dropping the observations near the population threshold where strategic sorting is likely to occur (Barreca et al. 2011, Cattaneo et al. 2019: 104-106, Sanz 2020) – see Table 9 for the summary of the results with varying "donut sizes". Specifically, we repeat the estimation and inference procedures after removing observations within different ranges around the threshold (1%, 2%, 3%, 4%, 5%, and 6% of observations around the cutoff). In implementing the "donut hole" approach, we note that varying the excluded range can affect the bandwidth selected by data-driven methods such as MSE-optimal or CER-optimal bandwidth selectors. To ensure consistent comparisons and isolate the effect of excluding observations near the threshold, we apply the MSE- and CER-optimal bandwidths obtained from the full sample across all "donut" specifications. This approach prevents conflating the effects of changing "donut" sizes with changes in the estimation window that would occur if bandwidth were re-optimized for each specification. The overall stability of our results across these specifications suggests our findings are not driven by potential manipulation around the threshold.

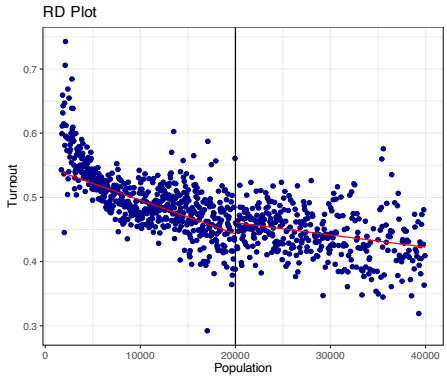
Using both MSE- and CER-optimal bandwidths from the full sample, we found that the SMDP-PR effect remained positive and statistically significant across most specifications, ranging from 0.031 (95% CI: 0.021, 0.042) with no exclusion to 0.104 (95% CI: 0.074, 0.133) with a 3% exclusion zone. The effect only diminished when excluding 5% or more of observations around the threshold. In contrast, the MMDP-PR comparison consistently showed no significant effects across most specifications until the 5-6% exclusion range. This pattern of results suggests our findings are not driven by potential manipulation around the threshold and are robust to reasonable adjustments in the estimation window.

Table 9: Sharp RD "Donut Hole" Robust Estimates from District-Level Analysis, 1998-2010 Pooled

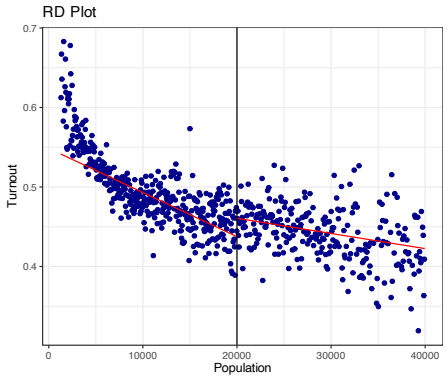
Model	Estimate (95% CI)	p-value	Sample Size
			N eff. (Left, Right)
SMDP-PR [D0%]	0.031 (0.021, 0.042)	0.000	(1405, 761)
SMDP-PR [CER, D0%]	0.033 (0.022, 0.043)	0.000	(1249, 698)
MMDP-PR [D0%]	0.002 (-0.007, 0.012)	0.638	(406, 743)
MMDP-PR [CER, D0%]	0.002 (-0.008, 0.011)	0.714	(384, 691)
SMDP-PR [D1%]	0.049 (0.037, 0.061)	0.000	(1310, 724)
SMDP-PR [CER, D1%]	0.031 (0.012, 0.050)	0.001	(1236, 692)
MMDP-PR [D1%]	-0.012 (-0.033, 0.009)	0.266	(380, 737)
MMDP-PR [CER, D1%]	-0.019 (-0.041, 0.003)	0.092	(358, 685)
SMDP-PR [D2%]	0.096 (0.075, 0.118)	0.000	(1310, 724)
SMDP-PR [CER, D2%]	0.103 (0.079, 0.126)	0.000	(1154, 661)
MMDP-PR [D2%]	0.018 (-0.006, 0.042)	0.132	(339, 706)
MMDP-PR [CER, D2%]	0.006 (-0.020, 0.032)	0.660	(317, 654)
SMDP-PR [D3%]	0.104 (0.074, 0.133)	0.000	(1280, 690)
SMDP-PR [CER, D3%]	0.113 (0.079, 0.146)	0.000	(1124, 627)
MMDP-PR [D3%]	0.019 (-0.015, 0.052)	0.272	(332, 672)
MMDP-PR [CER, D3%]	0.003 (-0.034, 0.041)	0.857	(310, 620)
SMDP-PR [D4%]	0.046 (0.018, 0.074)	0.001	(1195, 655)
SMDP-PR [CER, D4%]	0.050 (0.015, 0.085)	0.005	(1039, 592)
MMDP-PR [D4%]	0.017 (-0.009, 0.043)	0.195	(302, 637)
MMDP-PR [CER, D4%]	-0.013 (-0.046, 0.021)	0.450	(280, 585)
SMDP-PR [D5%]	0.002 (-0.020, 0.025)	0.832	(1158, 623)
SMDP-PR [CER, D5%]	-0.005 (-0.038, 0.028)	0.759	(1002, 560)
MMDP-PR [D5%]	0.242 (0.202, 0.282)	0.000	(263, 605)
MMDP-PR [CER, D5%]	0.258 (0.225, 0.292)	0.000	(241, 553)
SMDP-PR [D6%]	-0.044 (-0.078, -0.009)	0.014	(1131, 595)
SMDP-PR [CER, D6%]	-0.066 (-0.105, -0.028)	0.001	(975, 532)
MMDP-PR [D6%]	0.092 (0.025, 0.160)	0.007	(240, 577)
MMDP-PR [CER, D6%]	0.071 (0.004, 0.138)	0.039	(218, 525)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All "donut hole" (1%–6%) estimations employ covariate-adjusted RDD with triangular kernel and cluster-robust variance estimation (clustering by municipality type: urban, rural, mixed), using election year as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

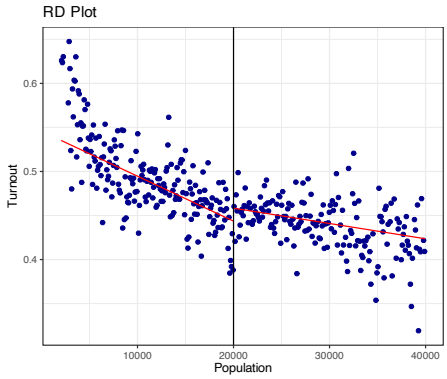
RDD Plots



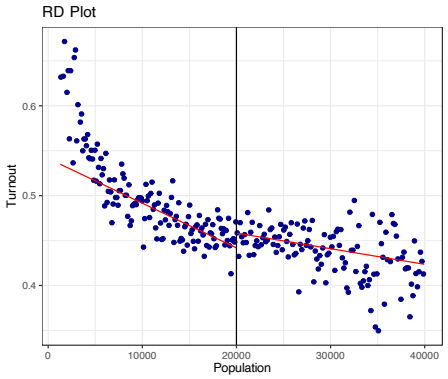
(a) 1998-2010 elections (district-level), SMDP vs. OLPR



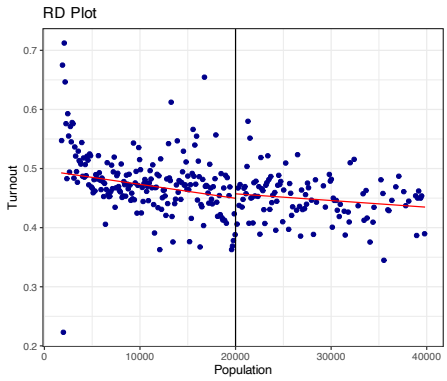
(b) 1998-2010 elections (district-level), MMDP vs. OLPR



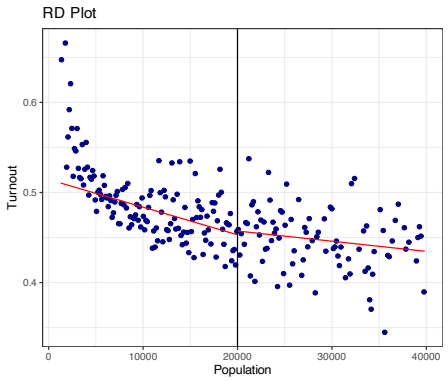
(c) 1998-2010 elections (district-level, excluding rural units), SMDP vs. OLPR



(d) 1998-2010 elections (district-level, excluding rural units), MMDP vs OLPR



(e) 1998 elections (district-level), SMDP vs. OLPR



(f) 1998 elections (district-level), MMDP vs OLPR

Figure 9: RDD plots (the procedure to select the number of bins: mimicking variance evenly-spaced method using spacings estimators; the order of the global-polynomial used to approximate the population conditional mean functions for control and treated units: $p = 1$)

Assessing RDD Validity: Falsification Tests

Regression discontinuity analysis is typically accompanied by various validation checks designed to assess the plausibility of the identification assumptions. While these assumptions are fundamentally untestable, several empirical tests can help establish whether they are likely to be met (for a detailed discussion, see Cattaneo et al. 2019: 848-849). These tests include:

- *Artificial cutoff values.* We execute a placebo test using a fake cut-off in the 2014 local elections with a temporarily suspended 20,000 population threshold;

- *Testing balance in covariates and placebo outcomes.* The treatment should not affect pre-intervention covariates or placebo outcomes. As Lee and Lemieux (2010) emphasized, the covariates must be determined prior to the realization of the current period's running and treatment variables. If empirical evidence indicates that important covariates or placebo outcomes differ systematically at the cutoff, the RDD assumption of continuity becomes implausible. Specifically, as Sanz (2017: 702) notes: "municipalities just above and just below the threshold should not differ, conditional on the fixed effects, in other variables that may themselves have an effect on voter turnout".

- *Manipulation* or *sorting* tests based on the density discontinuity of a running variable. If units cannot manipulate the value of the score that they receive, we should expect as many observations just above the cutoff as just below it;

- *Alternative bandwidth and window choices.* In our study, we use both MSE-optimal and CER-optimal data-driven bandwidth selection methods. Data-driven methods are strongly recommended in modern RDD analysis because they provide an objective, automated approach to bandwidth selection, reducing researcher discretion and potential bias. Specifically, using *rdrobust* R package, we employ two different MSE-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator; this approach allows us to account for potential asymmetry in the data distribution around the cutoff. The MSE-optimal selector minimizes the mean squared error of the estimator, while the CER-optimal selector aims to minimize coverage errors in confidence intervals.

Placebo Test: The 2014 Local Elections with Temporarily Suspended Population Threshold

To further validate our research design, we conducted a placebo test using data from the 2014 local elections in Poland. This test is particularly valuable because, in 2014,

the population threshold that typically determines the electoral system was temporarily suspended. As a result, all municipalities in the relevant population range used single-member district plurality constituencies, regardless of their population size. In a well-designed regression discontinuity study, we would expect no significant difference in outcomes at a threshold where no actual change in treatment occurs. This is exactly what we observed in our analysis of the 2014 elections. Our models show that the difference in voter turnout between municipalities just below and just above the 20,000 population cutoff is neither substantial nor statistically significant (see Table 10). We present evidence of this lack of discontinuity in Figure 10. The 2014 placebo test provides valuable evidence supporting our overall research methodology. It demonstrates that our analytical approach does not artificially create discontinuities where none should exist, thereby increasing confidence in the results we observed for years when the population threshold did determine the electoral system.

Table 10: Sharp RD estimates using local polynomial regression - placebo cutoff 2014

Model	Estimate (95% CI)	p-value	N eff. (Left, Right)
M1. SMDP-SMDP (2014)	0.004 (-0.023, 0.030)	0.787	(119, 52)
M2. SMDP-SMDP (2014) [CER]	0.006 (-0.023, 0.035)	0.666	(102, 50)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ RDD with a triangular kernel and cluster-robust variance estimation (clustering by municipality type: urban, rural, mixed). Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations.

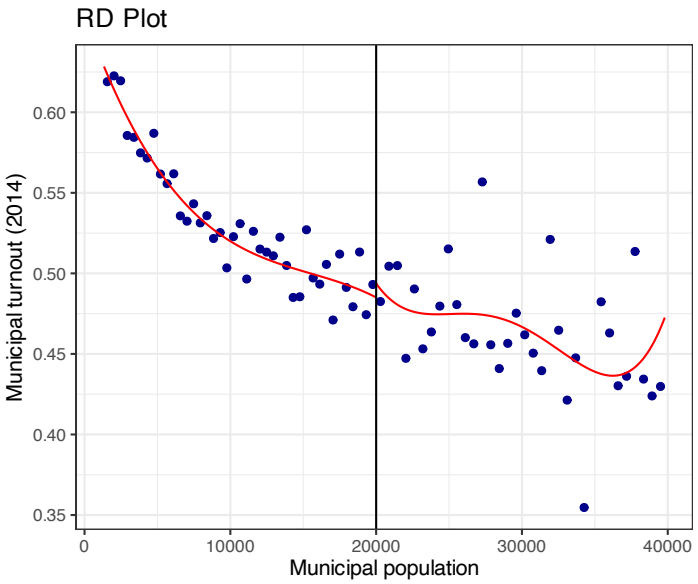


Figure 10: RDD plot for 2014 elections (no change of electoral rules at 20k cut-off)

Testing Balance in Covariates and Placebo Outcomes

In a well-designed regression discontinuity study, neither predetermined covariates nor placebo outcomes—variables known to be unaffected by the treatment—should show discontinuities at the threshold. As Cattaneo et al. (2022: 843) argue, balance analysis tests "the null hypothesis that the mean of pre-determined covariates is the same in the treatment and control groups." Rejection of this hypothesis would suggest a lack of comparability between groups, potentially undermining the validity of the results.

Our analysis examines potential discontinuities at the 20,000-inhabitant threshold across multiple dimensions (Table 11). We focus on two categories of variables: electoral outcomes from general elections (national parliament and presidential) and demographic characteristics, both measured at the municipal level.

The results demonstrate a strong balance across the threshold. For electoral outcomes, we find no significant differences in either presidential election turnout (estimate: 0.013, 95% CI: [-0.050, 0.077], p = 0.681) or national parliament (Sejm) election turnout (estimate: -0.004, 95% CI: [-0.067, 0.059], p = 0.894). Vote share analysis reveals similar patterns of balance, with no significant discontinuities in either winning party share (estimate: 0.010, 95% CI: [-0.019, 0.038], p = 0.503) or presidential race winner share (estimate: -0.018, 95% CI: [-0.050, 0.014], p = 0.273).

Further supporting these findings, the proportion of invalid ballots shows no significant differences in either legislative (estimate: 0.002, 95% CI: [-0.006, 0.011], p = 0.590) or presidential elections (estimate: 0.000, 95% CI: [-0.001, 0.001], p = 0.778). These results collectively suggest that the treatment does not correlate with broader

patterns of political participation or preferences.

Demographic variables similarly support the comparability of municipalities around the threshold. We find no significant discontinuity in either the proportion of the female population (estimate: -0.001, 95% CI: [-0.008, 0.006], p = 0.764) or municipal population density (estimate: 0.138, 95% CI: [-4.930, 5.206], p = 0.958).

These results collectively strengthen the validity of our RDD in two ways. First, they demonstrate that municipalities just above and below the threshold are comparable in terms of broader political behaviour patterns. Second, they show balance in fundamental demographic characteristics. This comprehensive balance supports our attribution of local election turnout differences to the change in the electoral system rather than to pre-existing differences between municipalities.

Table 11: Sharp RD Estimates Using Local Polynomial Regression: Analysis of Pre-Intervention Covariates and Placebo Outcomes

Model	Estimate (95% CI)	p-value	Sample Size
			N eff. (Left, Right)
Presidential Turnout	0.013 (-0.050, 0.077)	0.681	(1003, 441)
Presidential Turnout [CER]	0.013 (-0.051, 0.077)	0.687	(846, 416)
Sejm Turnout	-0.004 (-0.067, 0.059)	0.894	(1225, 284)
Sejm Turnout [CER]	-0.004 (-0.066, 0.058)	0.904	(1028, 262)
Winning party share	0.010 (-0.019, 0.038)	0.503	(219, 511)
Winning party share [CER]	0.011 (-0.017, 0.038)	0.446	(200, 493)
Winner share	-0.018 (-0.050, 0.014)	0.273	(325, 363)
Winner share [CER]	-0.019 (-0.051, 0.014)	0.260	(285, 340)
Invalid ballots (sejm)	0.002 (-0.006, 0.011)	0.590	(755, 390)
Invalid ballots (sejm) [CER]	0.003 (-0.007, 0.012)	0.583	(679, 360)
Invalid ballots (presid)	0.000 (-0.001, 0.001)	0.778	(245, 410)
Invalid ballots (presid) [CER]	0.000 (-0.001, 0.001)	0.768	(222, 394)
Female population	-0.001 (-0.008, 0.006)	0.764	(349, 430)
Female population [CER]	-0.001 (-0.008, 0.006)	0.752	(307, 415)
Density	0.138 (-4.930, 5.206)	0.958	(439, 353)
Density [CER]	0.198 (-4.986, 5.381)	0.940	(385, 326)

Notes: N represents the effective number of observations (Left, Right) around the cutoff. All estimations employ covariate-adjusted RDD with triangular kernel and cluster-robust variance estimation (clustering by municipality type: urban, rural, mixed), using a year as a discrete covariate. Bandwidth selection: "msetwo" indicates two different MSE-optimal bandwidth selectors (below and above the cutoff); "certwo" indicates two different CER-optimal bandwidth selectors. VCE method: heteroskedasticity-robust nearest neighbour variance estimator. Models are adjusted for "masspoints" to ensure preliminary bandwidths contain a minimal number of unique observations. Municipal density is calculated as population divided by geodetic area. All variables are aggregated at the municipality level. The data spans elections from 1997 to 2011, with demographic indicators measured in election years (autumn) and electoral data corresponding to presidential or parliamentary general election years.

Absence of Manipulation in the Forcing Variable

A crucial assumption in regression discontinuity designs (RDD) is the absence of precise manipulation of the running variable. In our context, this means ensuring that municipalities cannot precisely control their population size to influence their electoral system assignment. To validate this assumption, we employ density discontinuity testing, a method widely used to detect potential manipulation or sorting around the cutoff where treatment assignment changes (Cattaneo, Jansson, and Ma 2020, 2022).

Following McCrary (2008), density testing is based on the principle that if units could precisely manipulate their position relative to the cutoff, we would expect to observe a discontinuity in the density of the running variable at the threshold. Such a discontinuity would suggest that units are sorting themselves on different sides of the cutoff, potentially undermining the as-good-as-random assignment assumption crucial for causal inference in RDD.

We implement a formal density test based on local polynomials, following the approach proposed by Cattaneo et al. (2022). This test improves upon earlier methods by providing robust and bias-corrected inference and optimal bandwidth selection. Figure 11 visually represents the results of this test, illustrating the density of municipalities around the 20,000 population threshold.

Our analysis reveals no detectable manipulation at or near the 20,000 population cutoff, suggesting that municipalities are not systematically altering their population figures to influence their electoral system assignment. This finding strengthens our research design, implying that the assignment of electoral systems (SMDP, MMDP, or PR) based on the population threshold can indeed be treated as locally randomized.

Tables 12–15 present two sets of manipulation tests—one for the full sample and one excluding rural units. The density tests (Tables 12–14) show no significant manipulation at conventional levels. However, the binomial tests (Tables 13–15) indicate some imbalance in observations around the cutoff for larger window sizes, with p-values becoming significant (< 0.05) starting at a window of approximately 1,188 inhabitants.

To address these findings and ensure robust results, we implement several safeguards:

- We employ both MSE-optimal and CER-optimal bandwidth selectors to balance efficiency and robustness;
- We cluster standard errors at the municipality-type level to account for potential within-group correlation.

In conclusion, while our manipulation testing reveals some imbalance in observation counts at larger bandwidths, we find no substantial evidence of sorting or manipulation at or near the 20,000 population cutoff. This finding, combined with our methodological

safeguards, enhances the internal validity of our causal estimates of electoral system effects on voter turnout in Polish municipalities.

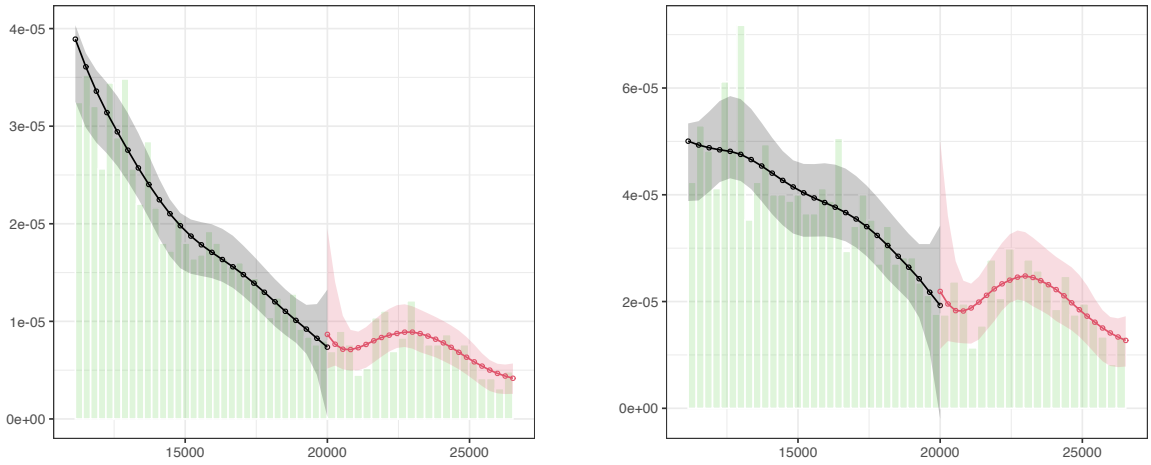


Figure 11: Manipulation Testing (a.k.a McCrary density test) for the population of municipalities (1998-2010); (a – left panel) all studied municipalities; (b – right panel) excluding rural units

Table 12: Manipulation Testing Using Local Polynomial Density Estimation

	Left of c	Right of c
Number of obs	8537	761
Eff. Number of obs	304	158
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	2954.315	2175.256
Method	T	$P \geq \|T\|$
Conventional	0.6085	0.5428
Robust	1.6668	0.0956

Table 13: P-values of Binomial Tests ($H_0 : p = 0.5$)

Window Length	<c	≥c	$P \geq \|T\|$
305.000 + 305.000	22	20	0.8776
599.368 + 512.806	43	39	0.7407
893.737 + 720.612	66	51	0.1953
1188.105 + 928.419	97	68	0.0290
1482.473 + 1136.225	127	79	0.0010
1776.842 + 1344.031	161	86	0.0000
2071.210 + 1551.837	191	96	0.0000
2365.578 + 1759.643	222	118	0.0000
2659.946 + 1967.450	264	138	0.0000
2954.315 + 2175.256	304	158	0.0000

Table 14: Manipulation Testing Using Local Polynomial Density Estimation (excluding rural units)

	Left of c	Right of c
Number of obs	2262	711
Eff. Number of obs	323	277
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	3597.43	4218.077
Method	T	$P > \|T\|$
Conventional	-0.9817	0.3262
Robust	-0.0403	0.9679

Table 15: P-values of Binomial Tests ($H_0 : p = 0.5$) (excluding rural units)

Window Length	<c	≥c	$P > \|T\|$
362.000 + 362.000	22	20	0.8776
721.492 + 724.000	45	42	0.8304
1080.984 + 1086.000	67	63	0.7926
1440.477 + 1448.000	101	74	0.0491
1799.969 + 1810.000	134	102	0.0434
2159.461 + 2172.000	165	128	0.0353
2518.953 + 2534.000	202	152	0.0091
2878.445 + 2896.000	242	179	0.0025
3237.938 + 3258.000	277	212	0.0038
3597.430 + 3620.000	323	240	0.0005

Supplementary Manipulation Testing Results

The supplementary manipulation testing analysis (Tables 16–19) yields the following findings:

- For the 1998 elections, we fail to reject the null hypothesis of continuity in the density at the cutoff, suggesting no evidence of manipulation around the threshold during this initial period.
- For the pooled 2002-2010 elections, we again find no significant discontinuity in the density at the cutoff, indicating that the continuity of the density functions for control and treatment units persists in subsequent elections.

Table 16: 1998 Elections: Manipulation Testing Results

	Left of c	Right of c
Number of obs	2116	188
Eff. Number of obs	92	73
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	3450.632	3729.236

Method	T	$P > \|T\|$
Conventional	−0.0172	0.9863
Robust	0.4907	0.6236

Window Length	$< c$	$\geq c$	$P > \|T\|$
1069.000 + 1069.000	20	21	1.0000
1333.626 + 1364.582	29	24	0.5831
1598.252 + 1660.164	35	28	0.4500
1862.877 + 1955.745	41	35	0.5666
2127.503 + 2251.327	47	44	0.8341
2392.129 + 2546.909	54	48	0.6208
2656.755 + 2842.491	65	55	0.4114
2921.381 + 3138.073	74	65	0.4976
3186.007 + 3433.655	82	69	0.3288
3450.632 + 3729.236	92	73	0.1609

Table 17: 1998 Elections (excluding rural municipalities): Manipulation Testing Results

	Left of c	Right of c
Number of obs	543	180
Eff. Number of obs	125	90
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	5221.953	5656.356

Method	T	$P > \|T\|$
Conventional	0.0323	0.9743
Robust	−0.2254	0.8217

Window Length	$< c$	$\geq c$	$P > \|T\|$
1413.000 + 1413.000	25	20	0.5515
1836.217 + 1884.484	36	30	0.5386
2259.434 + 2355.968	43	37	0.5764
2682.651 + 2827.452	55	47	0.4884
3105.868 + 3298.936	68	58	0.4228
3529.085 + 3770.420	82	68	0.2885
3952.302 + 4241.904	94	74	0.1425
4375.519 + 4713.388	109	81	0.0499
4798.736 + 5184.872	116	86	0.0410
5221.953 + 5656.356	125	90	0.0202

Table 18: 2002-2010 Elections: Manipulation Testing Results

	Left of c	Right of c
Number of obs	6421	573
Eff. Number of obs	250	118
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	3111.415	2294.308

Method	T	$P > \|T\|$
Conventional	0.133	0.8942
Robust	0.9865	0.3239

Window Length	$< c$	$\geq c$	$P > \|T\|$
446.000 + 446.000	26	20	0.4614
742.157 + 651.368	41	36	0.6488
1038.314 + 856.735	63	44	0.0814
1334.472 + 1062.103	85	56	0.0181
1630.629 + 1267.470	108	59	0.0002
1926.786 + 1472.838	136	66	0.0000
2222.943 + 1678.206	157	81	0.0000
2519.100 + 1883.573	184	92	0.0000
2815.257 + 2088.941	214	108	0.0000
3111.415 + 2294.308	250	118	0.0000

Table 19: 2002-2010 Elections (excluding rural municipalities): Manipulation Testing Results

	Left of c	Right of c
Number of obs	1719	531
Eff. Number of obs	247	220
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	3730.951	4496.672

Method	T	$P > \|T\|$
Conventional	−0.9698	0.3321
Robust	−0.2271	0.8203

Window Length	$< c$	$\geq c$	$P > \|T\|$
503.000 + 503.000	23	20	0.7608
861.661 + 946.741	39	42	0.8243
1220.322 + 1390.483	63	52	0.3511
1578.984 + 1834.224	84	74	0.4741
1937.645 + 2277.965	112	95	0.2661
2296.306 + 2721.707	133	123	0.5739
2654.967 + 3165.448	162	152	0.6116
3013.628 + 3609.189	191	177	0.4980
3372.289 + 4052.931	214	192	0.2973
3730.951 + 4496.672	247	220	0.2289

Power Analysis for Main Models

Table 20: Power Analysis for 1998-2010 District-Level Analysis (SMDP vs OLPR)

Parameter	Value
Number of obs	56326
BW type	msetwo
Kernel type	Triangular
VCE method	NN
Derivative	0
HA: τ	0.041
Cutoff c	20000

	Left of c	Right of c
Number of obs	53540	2786
Eff. number of obs	4234	923
BW loc. poly.	4732.429	3352.061
Order loc. poly.	1	1
Sampling BW	4732.429	3352.061
New sample	4234	923

Power against:	$H_0 : \tau = 0$	$0.2\tau = 0.008$	$0.5\tau = 0.02$	$0.8\tau = 0.032$	$\tau = 0.041$
Robust bias-corrected	0.05	0.391	0.988	1	1

Table 21: Power Analysis for 1998-2010 District-Level Analysis (MMDP vs OLPR)

Parameter	Value
Number of obs	31493
BW type	msetwo
Kernel type	Triangular
VCE method	NN
Derivative	0
HA: τ	0.031
Cutoff c	20000

	Left of c	Right of c
Number of obs	28707	2786
Eff. number of obs	1301	932
BW loc. poly.	3464.642	3365.976
Order loc. poly.	1	1
Sampling BW	3464.642	3365.976
New sample	1301	932

Power against:	$H_0 : \tau = 0$	$0.2\tau = 0.006$	$0.5\tau = 0.015$	$0.8\tau = 0.025$	$\tau = 0.031$
Robust bias-corrected	0.05	0.231	0.862	0.998	1

Table 22: Power Analysis for the 1998 District-Level Analysis (SMDP vs OLPR)

Parameter	Value
Number of obs	16007
BW type	msetwo
Kernel type	Triangular
VCE method	NN
Derivative	0
HA: τ	0.044
Cutoff c	20000

	Left of c	Right of c
Number of obs	15212	795
Eff. number of obs	862	269
BW loc. poly.	3023.549	3142.073
Order loc. poly.	1	1
Sampling BW	3023.549	3142.073
New sample	862	269

Power against:	$H_0 : \tau = 0$	$0.2\tau = 0.009$	$0.5\tau = 0.022$	$0.8\tau = 0.035$	$\tau = 0.044$
Robust bias-corrected	0.05	0.146	0.61	0.948	0.994

Table 23: Power Analysis for the 1998 District-Level Analysis (MMDP vs OLPR)

Parameter	Value
Number of obs	9266
BW type	msetwo
Kernel type	Triangular
VCE method	NN
Derivative	0
HA: τ	0.033
Cutoff c	20000

	Left of c	Right of c
Number of obs	8471	795
Eff. number of obs	573	269
BW loc. poly.	4310.06	3184.568
Order loc. poly.	1	1
Sampling BW	4310.06	3184.568
New sample	573	269

Power against:	$H_0 : \tau = 0$	$0.2\tau = 0.007$	$0.5\tau = 0.016$	$0.8\tau = 0.026$	$\tau = 0.033$
Robust bias-corrected	0.05	0.119	0.477	0.861	0.967

Data and Variables

Our analyses are based on the official data retrieved from the National Electoral Commission (https://danewyborcze.kbw.gov.pl/indexc4fa.html?title=Strona_g%C5%82%C3%B3wna). Regarding the socioeconomic control variables, the data come from the Local Data Bank (https://bdl.stat.gov.pl/bdl/start) - Poland's largest database of socioeconomic data.

Replication Materials

Replication data and R programming language code are available for review, and they can be retrieved from Harvard Dataverse: https://dataverse.harvard.edu/privateurl.xhtml?token=74272b82-390e-4de8-b548-2c2deb17a08f [Temporary URL]