Do Open Lists Increase Turnout? Probably Not: Evidence from the Re-Analysis of a Study on Spanish Local Elections

Leonardo Carella[a]

[a]*Nuffield College, University of Oxford, New Rd, Oxford, OX1 1NF, United Kingdom*

Abstract

This research note replicates and critiques the analysis of a notable paper (Carlos Sanz, "The effect of electoral systems on voter turnout: Evidence from a natural experiment", *PSRM*, 2017), presenting evidence against its main finding: that open-list ballot structures increase voter turnout. The study employs a regression-discontinuity (RD) design on data from Spanish municipal elections (1979–2011), where municipalities with fewer than 250 residents employ open lists whereas larger towns employ closed lists. Through a series of statistical tests and the inspection of alternative data sources, I show that the threshold's effect on turnout recovered in the original paper results from (1) *non-random missing data*, due to inconsistencies in how non-valid votes were recorded above and below the threshold, and (2) *compound treatment bias*, due to changes in list-length requirements around the threshold. Further analysis suggests that, rather than improving participation, the more complex open-list ballot used in Spain actually hinders voters' ability to express their preferences, by making it easier to cast mistakenly null votes. To support this interpretation, I present additional analysis of the causal effects of ballot complexity on null voting from Polish local elections. I conclude discussing practical and substantive implications for conducting population-based RD studies and evaluating electoral system effects.

Keywords: Electoral systems, turnout, ballot structure, open lists, regression discontinuity

Email address: leonardo.carella@nuffield.ox.ac.uk (Leonardo Carella)

1. Introduction

The relationship between electoral systems and turnout is a well-trodden line of research in political science. Dimensions of electoral institutions that affect incentives for party entry and voter mobilization — such as the electoral formula and the district magnitude — have been shown to lead to participation gains (Blais and Aarts, 2006; Grofman and Selb, 2011; Eggers, 2015; Cox et al., 2016; Figueroa, 2024). However, there is less consensus (Söderlund, 2018) on the turnout effects of *ballot structure*, the set of rules that govern "the physical operation that the voter has to do in order to exercise suffrage: crossing the name of a candidate, choosing a list, ordering several names, and so on" (Bosch and Orriols, 2014, p. 493). Political scientists traditionally describe variation along this dimension in terms of the party- or candidate-centered nature of vote choice. 'Closed' or 'non-preferential' ballot structures constrain the voter to party choice, whereas 'open' or 'preferential' systems allow for the expression of categorical or ordinal preferences for *candidates* within or across parties (Blais, 1988; Farrell and McAllister, 2006).

In an influential paper, Sanz (2017) argues that open-list (OL) systems, which allow voters to choose candidates within a party-list, are conducive to higher turnout than closed-list (CL) systems, where voters can only select a party. The paper — which to my knowledge is the only causally identified study on the effects of ballot structure on turnout — support this claim with a regression-discontinuity (RD) analysis, leveraging a discontinuity in Spanish local election rules. Drawing on data from 9 election-years (1979–2011), the author concludes that OL, used in municipalities with fewer than 250 residents, is associated with a 1–2 percentage point increase in turnout relative to CL, used in municipalities with over 250 residents. The hypothesized mechanism is that personalized competition makes it more likely for popular candidates from small parties to get elected under OL, leading to increased competition relative to CL. This research note contends that the effect of OL on turnout recovered by the author is due to two sources of bias: (1) *non-random missing data*, which affects observations in the 1983 election-year, and (2) *compound treatment bias* (Eggers et al., 2018), which affects all analysis but particularly 1979 municipal elections.

In section 2, I replicate the analysis subsetting the data by election-year, which suggests that the effect recovered by Sanz (2017) is largely driven by observations in the first two election-years in the sample: 1979 and 1983. In the data for 1983, virtually no blanks and nulls were recorded *for municipalities above the threshold*. As shown in section 3, these 'missing' votes were genuinely cast, but not recorded in official records: once we account for data missingness, the estimated effect of OL is non-significant. In section 4, I examine the 1979 election data: here, the threshold *does* affect both party entry and turnout, consistently with the paper's theoretical expectations, but this effect is likely unrelated to ballot structure. In fact, the same threshold also coincides with a large increase in the number of candidates legally required to field a list, making it harder for parties to *find enough candidates* to compete, especially in towns of few hundred residents and in the first local elections after the return of democracy in Spain. I support this conclusion by showing similar effects at a 'placebo' threshold (5,000 residents), where list-length requirements change but ballot structure does not.

In section 5, I reconsider the question of ballot structure effects on electoral participation from a different perspective: the degree of cognitive effort required to cast a valid vote under different electoral rules (ballot complexity). I estimate a significant effect of OL on the share of null votes, and argue that this effect is attributable directly to ballot structure: it is easier for voters to make mistakes in the OL system, where they have to express 1–4 candidate preferences, than in the CL system, where they simply have to enclose a party ballot in an envelope. I support this interpretation by showing that the effect of OL's relative complexity is stronger for municipality with lower levels of education and by presenting original analysis from Polish municipal elections (2010–2024), which also present a population-based discontinuity in local electoral rules. In the conclusion (section 6), I discuss methodological and substantive implications of the analysis presented for conducting population-based RD designs and for the study of electoral system effects.

2. Replication and Visual Inspection of the Data

Municipality size	100-250 residents	251-1000 residents
Electoral Formula	**Multi-member majoritarian**: the most voted candidates from *any* party are elected.	**PR (D'Hondt)** with 5% threshold: seats are attributed proportionally *across party lists*.
Ballot Structure	**Open List**, with limited vote and *panachage*: each voter can choose up to four candidates, from all parties.	**Closed List**: voters choose one party-list; candidates are elected according to their list position, which voters cannot change.
Council Size	5 members	7 members
List length	**Maximum** of 5 candidates	**Minimum** of 7 candidates

Table 1: Spanish municipal election rules.

Table 1 summarizes the differences in electoral systems employed in Spanish local elections for municipalities with fewer and more than 250 residents, the population threshold that allows for causal identification of the effects of variation in electoral rules. The main results in Sanz (2017, p. 693, table 3) are derived from two-way fixed-effect parametric RD models[1] with the functional form:

$$\text{Turnout}_{m,t} = \alpha_m + \gamma_t + \beta_1 D_{m,t} + \beta_2(\log(\text{Population}_{m,t}) - \log(250)) +$$
$$\beta_3 \left[D_{m,t} \times \left(\log(\text{Population}_{m,t}) - \log(250)\right)\right] + \epsilon_{m,t}$$

[1]This design choice make the paper closer to a difference-in-difference design than a RD design: "the regressions do not compare municipalities just above and just below the threshold but municipalities that switch from one system to another with those that remain in the same system" (Sanz, 2017, p. 695).

Where m indexes the municipality, t indexes the election-year, D is a 'discontinuity' binary variable taking the value of 1 if the municipality uses open-list (i.e. has fewer than 250 residents) and 0 otherwise. α_m represents municipality fixed effects and γ_t election-year fixed-effects. The sample is restricted with the Imbens-Kalyanaraman bandwidth selection procedure (Imbens and Kalyanaraman, 2012), and fractions thereof. The results of the main analysis were successfully replicated and are shown in table 2 below. The estimates for D correspond to β_1 in the model's equation, i.e. the estimated causal effect of crossing the 250-resident threshold, *going from closed-list PR to open-list majoritarian*.

	Dependent Variable: Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
D (Under 250)	1.031**	1.237**	1.156**	1.547***	1.844***	1.308*
	(0.381)	(0.395)	(0.407)	(0.431)	(0.491)	(0.641)
Num.Obs.	15954	13556	11226	8798	6404	3976
Bandwidth	1×IK	0.85×IK	0.7×IK	0.55×IK	0.4×IK	0.25×IK

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: Replication of Sanz (2017), p. 693, table 3.

The estimates of table 2, however, represent weighted averages of estimated treatment effects in each fixed-effect group. Two-way fixed-effect estimates are consistent and unbiased only under the assumption of homogeneous treatment effects across groups (Gibbons et al., 2019). To investigate potential heterogeneities across elections-years, I fit RD models subsetting the data for each election-year (table 3).

	Dependent Variable: Turnout								
Year	1979	1983	1987	1991	1995	1999	2003	2007	2011
D (Under 250)	2.619**	5.640***	−0.812	−1.939*	1.124+	−0.868	−0.821	−0.181	−1.188*
	(0.972)	(1.107)	(1.223)	(0.902)	(0.668)	(0.708)	(0.627)	(0.586)	(0.589)
Num. Obs.	3373	2681	1817	2763	3400	2982	3050	3445	3079
Bandwidth	1×IK	1×IK	1×IK	1×IK	1×IK	1×IK	1×IK	1×IK	1×IK

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: RD estimates of 250-resident threshold effect on turnout, by election-year.

There is a strong positive effect of the threshold on turnout in 1979 (+2.6 percentage points) and, especially, in 1983 (+5.6 points), while in following years the effect is negative in six out of seven cases. To get a sense of what might be driving this pattern, in figures 1 and 2, I plot binned averages of the dependent variable across ranges of values of the running variable (logged population). I do so across election-year subsamples, and for three dependent variables: turnout (voters for 100 residents), share of nulls (nulls for 100 voters) and share of blanks (blanks for 100 voters). Three striking patterns emerge from 'eyeballing' the data:

1. There are virtually no blanks or nulls recorded in the data for municipalities employing CL in 1983, unlike all other elections. This suggests that the effect of the discontinuity on turnout may be affected by non-random data missingness, as shown in section 3.
2. The data for 1979 does present a marked discontinuity in turnout around the 250-resident threshold, which cannot be readily explained by inconsistencies in non-valid ballot data (although blanks and nulls for this year are almost invariably grouped into a single category). However, there is also a discontinuity in the effect of turnout at 5,000 residents (dotted red line), which closes in the following elections, *at the same time as* the discontinuity associated with the 250-resident threshold also disappears. As the same closed-list ballot structure is used either side of the 5,000-resident threshold, this suggests that rule changes associated with population cutoffs *other than ballot structure* may be driving turnout differences around both thresholds in early election-years. As discussed in section 4, changes in list-length requirements are the most likely suspect.
3. There is a small but noticeable decrease in the share of null votes as municipalities' population size crosses the 250-resident threshold in elections after 1987, when null ballots are recorded consistently either side of the threshold and separately from blanks. This suggests that the threshold may have implications for the probability of casting null votes, rather than casting a vote at all. As discussed in section 5, there is reasonable evidence that this is actually an effect of open/closed variation in ballot structure.

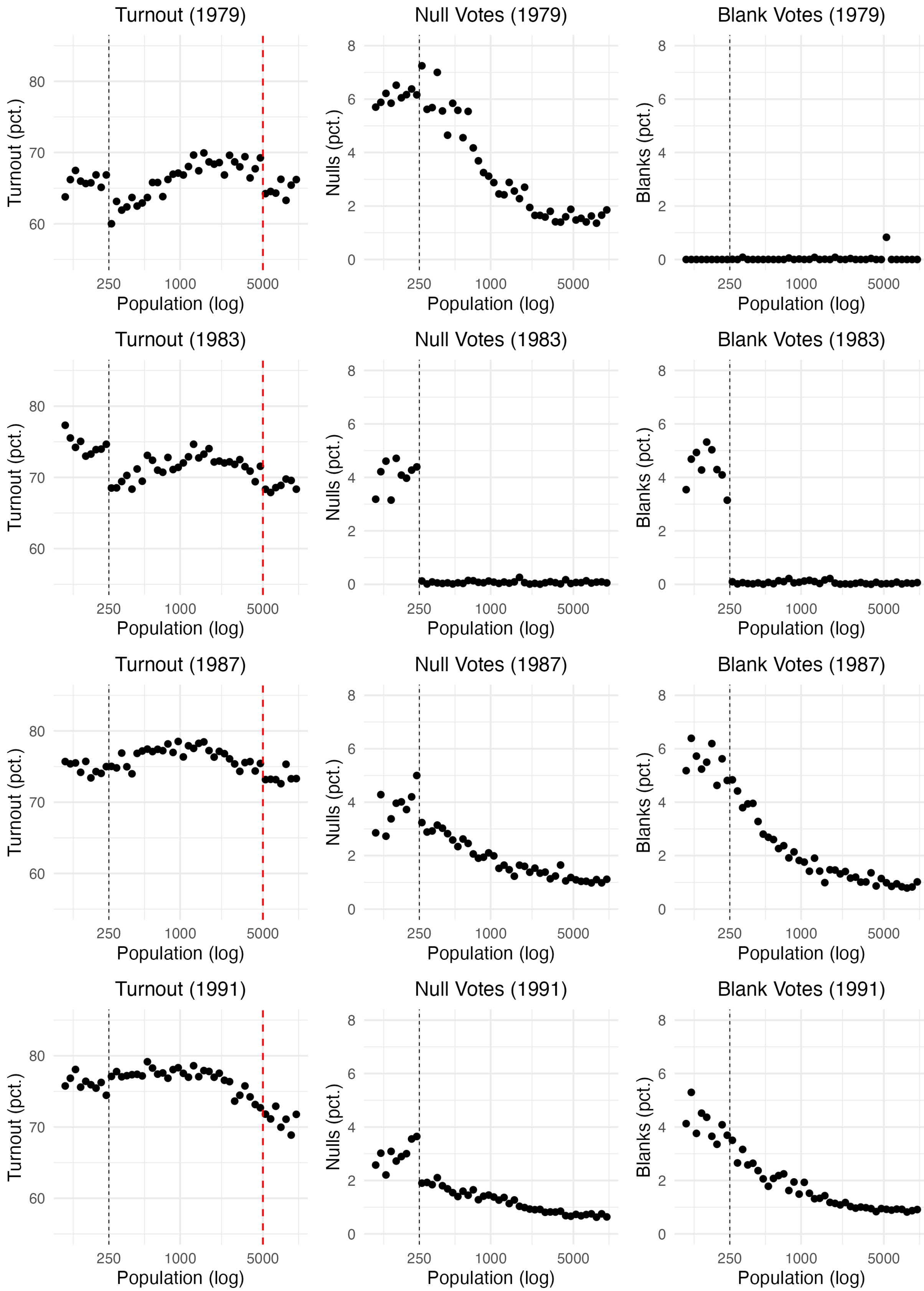

Figure 1: RD plots (1979–1991).

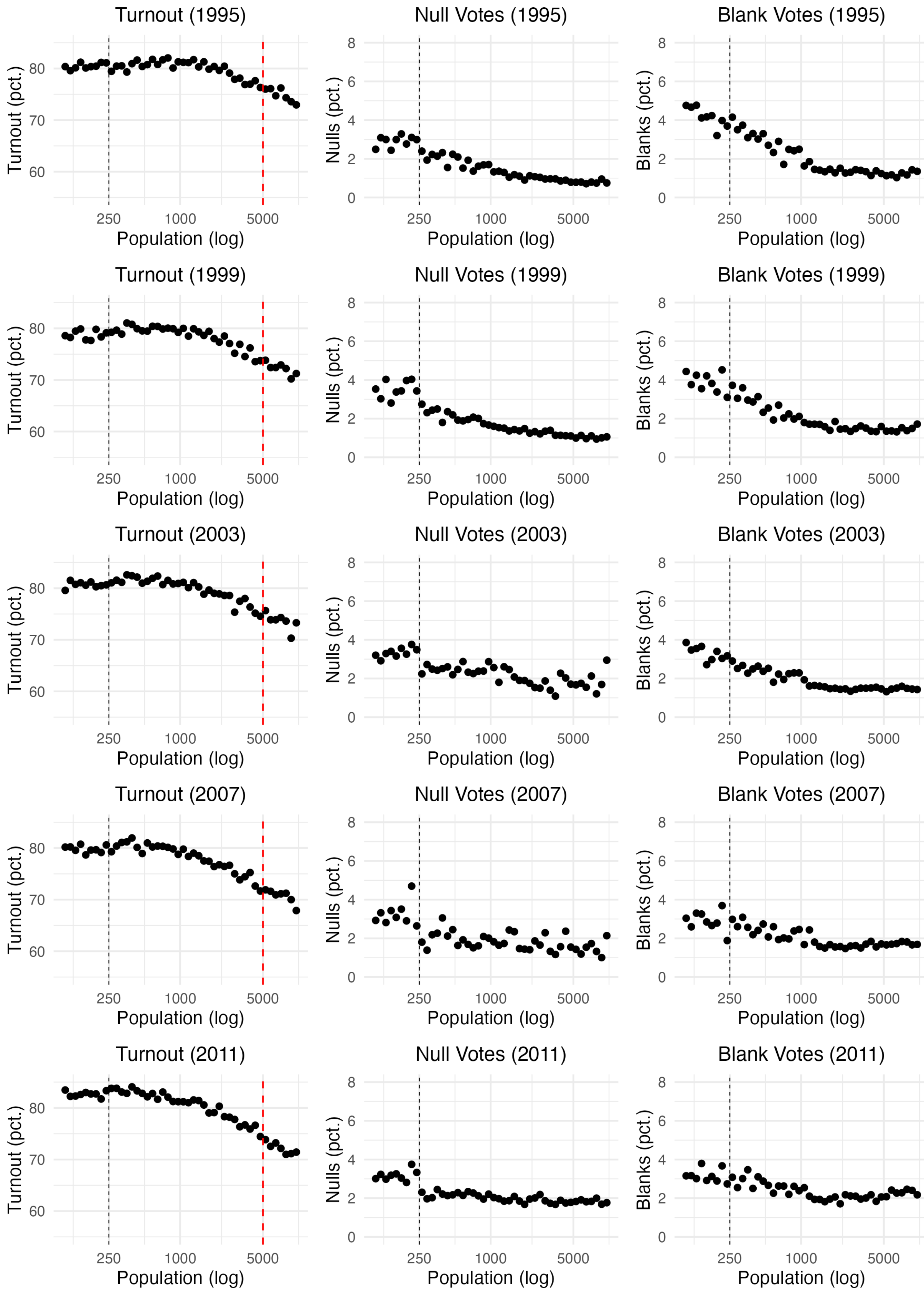

Figure 2: RD plots (1995–2011).

3. Evidence of Non-Random Missing Data (1983)

It is clear that, in the 1983 elections, null and blanks votes were excluded from electoral records for almost all municipalities with more than 250 residents. Prior to a 1985 reorganization of the electoral code,[2] publication of official results of local elections was not centralized, but delegated to local election boards (*Juntas Electorales de Zona*). The 1983 data used by Sanz (2017) are those published by the Spanish Interior Ministry, which in turn collated data from these bodies: this process is likely to be responsible for the data loss.[3] To corroborate this suspicion, I consulted publications on Spanish regional politics which drew on data directly from the provincial bodies, rather than via the intermediary of the Ministry's records.

Data Source		Interior Ministry			Election Boards			
Municipality	Pop.	Valid	Blanks	Nulls	Valid	Blanks	Nulls	Same Total?
Population 251–400								
Ábalos	303	210	0	0	210	2	7	No
Berceo	294	204	0	0	204	4	2	No
Bergasa	261	181	0	0	181	0	2	No
Camprovín	306	192	0	0	192	5	2	No
Cárdenas	319	146	0	0	146	22	16	No
Cordovín	315	151	0	0	151	25	14	No
Population 100–250								
Alesón	171	131	0	0	131	0	0	Yes
Baños de Rioja	171	109	0	10	109	0	10	Yes
Bobadilla	222	93	2	39	93	2	39	Yes
Briñas	225	159	1	4	159	1	4	Yes
Canales de la Sierra	102	68	0	0	52	15	1	Yes
Canillas de Rio Tuerto	102	44	6	6	44	6	6	Yes

Table 4: 1983 local election returns (La Rioja).

Table 4 compares the Interior Ministry data with those collected by Fernández Ferrero (1995, pp. 375–469) from the election boards of Calahorra, Logroño and Haro in La Rioja

[2]Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General, also known as LOREG

[3]The overall percentage of blanks and nulls recorded in 1983 is a highly implausible 0.18%, against e.g. 2.3% for 1987 and 4.2% for 2011.

region (I report here results for the first municipalities within 150 residents of the threshold in alphabetical order, but these issues affect the entire sample). As suspected, null and blank votes were cast in municipalities with more than 250 residents, but they were not recorded in the Ministry's dataset. In smaller municipalities, the total number of ballot cast from the Ministry's data is the same as the alternative data sources, and thus likely accurate: null and blanks are correctly recorded, though sometimes they are counted as valid votes (e.g. 'Canales de la Sierra' in table 4). A similar problem emerges from a comparison with election returns from the province of Albacete, compiled by Collado and de Dios (1984) (table 5). In the appendix, I present further evidence of missing data in the 1983 election-year, leveraging turnout data in concurrent regional elections.

Data Source		Interior Ministry			Election Boards			
Municipality	Population	Valid	Blanks	Nulls	Valid	Blanks	Nulls	Same Total?
Population 251–400								
Balsa de Ves	374	244	0	0	244	2	0	No
Masegoso	344	224	0	0	224	0	1	No
Villatoya	252	120	0	0	120	3	1	No
Villavaliente	330	250	0	0	250	1	3	No
Population 100–250								
Golosalvo	139	105	1	1	105	1	1	Yes
Montalvos	206	160	0	1	156	0	1	Yes

Table 5: 1983 Local election returns (Albacete).

Accounting for the problem of missing data in 1983 is sufficient to overturn the main substantive inference from the estimation strategy adopted by Sanz (2017). Crossing the 250-resident threshold is not associated with an increase in valid-vote turnout (votes for candidates as a percentage of eligible residents) in 1983 or across the sample (models 1–2 in table 6). Moreover, when 1983 elections are removed from the sample, the effect of the threshold on turnout is non-significant (model 3).

	Dependent Variable:		
	Valid-Vote Turnout		Turnout
	(1)	(2)	(3)
D (Under 250)	0.301	−0.115	0.545
	(1.191)	(0.465)	(0.382)
Num.Obs.	2722	14 146	14 546
Bandwidth	1×IK	1×IK	1×IK
Election-Year Sample	1983	All	All but 1983
TWFE	No	Yes	Yes
	Note: *p<0.1; **p<0.05; ***p<0.01		

Table 6: RD estimates of 250-resident threshold effects on turnout/valid-vote turnout.

4. Evidence of Compound Treatment Bias (1979)

Missing data cannot explain the positive effect of the 250-resident threshold on turnout recovered in the subsample of elections that took place in 1979. In this section, I argue that this effect is likely genuine, but *it is not attributable to ballot structure*: rather, it is the consequence of a different aspect of electoral rules that varies at the threshold ('compound treatment'). In fact, when a municipality's population crosses 250 residents, not only does the size of the council change — from 5 to 7 — but electoral regulations[4] stipulate that, in municipalities with over 250 residents

> "the lists that run in each district must include, as a *minimum*, as many candidate names as the number of councillors to elect" [my emphasis and translation]

while in municipalities with fewer than 250 residents

> "each party, coalition, federation or local group might present a list with a *maximum* of five names." [my emphasis and translation]

[4]Ley 39/1978, de 17 de julio, de elecciones locales, Article 11 and Disposiciones Transitorias, article 8; these regulations are kept in subsequent legislation, including the 1985 LOREG.

In practice, this means that, in a municipality with 251 residents, for a minimally competitive election (two lists of seven candidates), there must be at least 14 people willing to run for the council, or 5.6% of the population. In a municipality with 249 residents, there can be a minimally competitive election if only *six* people wish to run for five council seats — for instance, in two lists of three — amounting to 2.4% of the population. In sum, as hypothesized by Sanz (2017), the effect on turnout recovered in 1979 is likely driven by reduced list entry in municipalities above 250 residents, but it is not *closed* lists that constrains entry, it is the fact that these lists must have — by law — many more candidates. According to Spanish Interior Ministry data, the average number of candidates in municipalities just below the threshold (200–250 residents) in 1979 was 7.74: about half of what would be required to have a two-party election if they had fallen just above the threshold. In sum, it may not be the fact that below the threshold lists are *open* that reduces competition (and thus turnout), but the fact that they can be *incomplete*.

	Dependent Variable:			
	No. Lists	Turnout	No. Lists	Turnout
D (Under 250), from ≥ 7 to ≤ 5 cands.	0.209***	2.619**		
	(0.059)	(0.972)		
D (Under 5,000), from ≥ 21 to ≥ 17 cands.			0.173*	3.673**
			(0.076)	(1.179)
Num.Obs.	2246	3373	3186	2074
Bandwidth	1×IK			
				Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: RD estimates of 250- and 5,000-resident threshold effects on turnout and number of lists (1979).

As a test of the plausibility of this alternative explanation, I consider a different threshold where list-length requirements vary, *but ballot structure does not*, as suggested by Eggers et al. (2018, p. 214). At 5,000 residents, the size of Spanish councils increases from 17 to 21: this means that — assuming a single-district election — the number of candidates required to hold an election between n lists increases by a fraction of the population equal to $4n/5000$.

This is much less demanding than the increase associated with the 250-resident discontinuity, as complete lists are required either side of the 5,000-resident threshold. However — as shown in table 7 — the 5,000-resident discontinuity is also associated with a significant increase in number of lists *and* in turnout in 1979.[5]

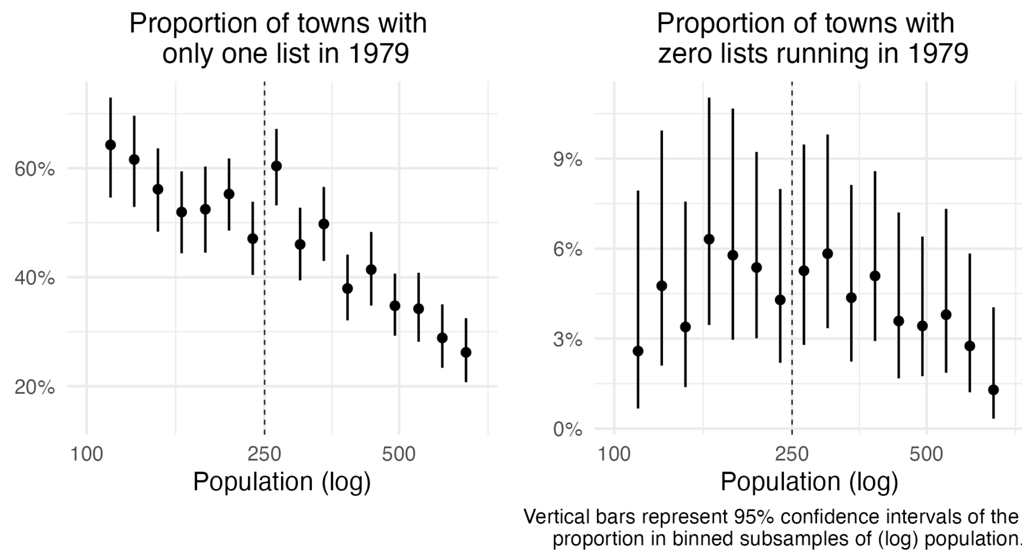

Figure 3: Binned proportion of single- or no-list municipal elections (1979).

The problem of "not finding enough people to run a list" may seem trivial, but it should be understood in the context of very small townships holding elections in the immediate aftermath of the return of democracy in Spain. When the first municipal elections were held, parties had only been legalized three years earlier and their organization was limited: the Christian-Democratic UCD and the Social-Democratic PSOE, which won 60% of the almost 70,000 councilors elected that year, had memberships of 101,000 and 50,000 respectively (Linz and Montero, 1999, p. 36). In 1979, over 27% of *all* local elections — and over 50% of local elections in municipalities with population between 225 and 275 residents — were effectively uncompetitive, with only one list running. In almost 5% of these, not even *one*

[5]I refrain from using causal language, as other aspects of municipal legislation change at 5,000 residents (e.g. municipal services provision). Moreover, a larger district magnitude can be related to turnout through causal pathways other than list-length requirements (e.g. it may increase proportionality in seat allocation).

list was presented, requiring a repeat election (figure 3).[6]

4.1. Why no Effect in Later Elections?

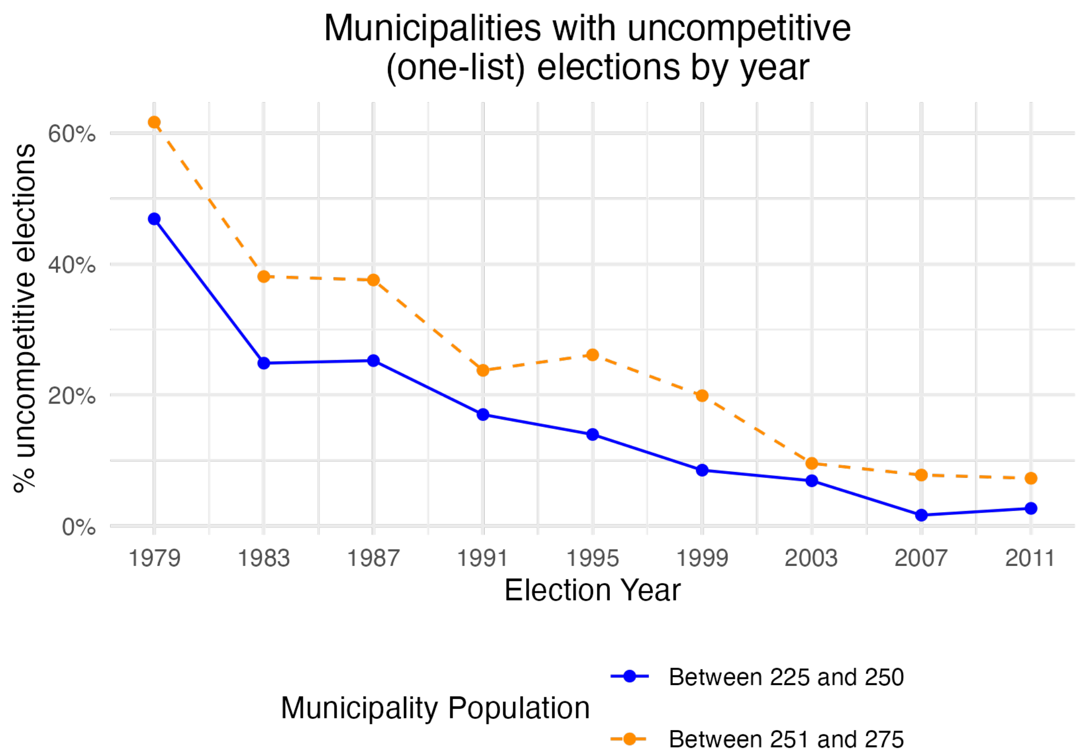

Figure 4: Proportion of uncompetitive (single-list) elections (1979–2011).

It is worth asking why the effect of the 250-resident threshold identified in 1979 (which, I argued, is likely due to list-length requirements rather than ballot structure) 'disappears' in following election-years. The development of party organizations at local level probably played a role in making elections minimally competitive even in small locales. As shown in figure 4, the share of single-list elections in municipalities with population within 25 residents from the threshold drops by an order of magnitude between 1979 and 2011. But perhaps more importantly, after 1983 other elections were held concurrently with municipal elections

[6]Information on municipalities that did not hold an election in 1979 were drawn from the official act 'Corrección de errores del Real Decreto 814/1979, de 20 de abril, por el que se convocan Elecciones Municipales parciales (BOE núm. 100, de 26 de abril de 1979).'

in most of Spain: regional elections (in 14 out of 17 regions)[7] and EU elections (everywhere, in 1987 and 1999). Conversely, in 1979 concurrent elections only took place in two regions.

As electoral rules and party supply in regional and EU elections are uniform across municipalities regardless of size, it is possible that competition in these concurrent election is driving voters' turnout decisions, so that turnout coattails (Magar, 2012) 'wash out' any effect of local election rules. To illustrate heterogeneities due to concurrent elections, I re-estimate the effect of the 250-resident discontinuity on municipal election turnout, interacting the discontinuity with a vector that takes the value of 1 if the election took place concurrently with other contests. Effectively, what yields variation on the moderator is that three regions have non-coincident local and regional election cycles, and that in 1987 and 1999 EU elections coincided with local elections across the entire country. The model thus is:

$$\begin{aligned}
&\text{Turnout}_{m,t} = \alpha_m + \gamma_t + \beta_1 D_{m,t} + \beta_2 \text{Concurrent Election}_{m,t} + \\
&\beta_3 (D_{m,t} \times \text{Concurrent Election}_{m,t}) + \beta_4 (\log(\text{Population}_{m,t}) - \log(250)) + \\
&\beta_5 \left[D_{m,t} \times \left(\log(\text{Population}_{m,t}) - \log(250)\right)\right] + \epsilon_{m,t}
\end{aligned}$$

I report in table 8 the estimate for β_1 (the effect of the discontinuity in municipal elections *without* concurrent contests) and β_3 (the difference in effect between municipalities *with* concurrent contests and those without). I show results for samples in the Imbens-Kalyanaraman optimal bandwidth, 1.5 times the bandwidth, and 75% of the bandwidth. For comparison, I repeat the procedure with the 5,000-resident discontinuity.

[7] 'Regional elections' here refers to *elecciones autonomicas* in 13 *comunidades autonomas* and to *elecciones a las Juntas Generales* in the three provinces of the Basque Country, which are held concurrently with municipal elections.

	Dependent Variable: Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
D (Under 250)	1.846+	1.394	2.687*			
	(1.040)	(0.851)	(1.164)			
D × Concurrent	−2.032*	−1.880*	−2.950**			
	(1.008)	(0.803)	(1.142)			
D (Under 5,000)				0.516	0.787*	0.445
				(0.335)	(0.315)	(0.349)
D × Concurrent				−1.087***	−1.259***	−0.997**
				(0.279)	(0.244)	(0.305)
Num.Obs.	11 496	16 826	8635	10 983	16 649	8121
Bandwidth	1×IK	1.5×IK	0.75×IK	1×IK	1.5×IK	0.75×IK

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: Effect of the threshold on turnout (1987–2011), conditional on concurrent elections

The results show that in municipal elections held after 1979, we still observe an effect of the 250-resident threshold on turnout, *but only for elections did not take place concurrently with other contests*. (The size of this effect is smaller than in 1979 and highly sensitive to bandwidth specification.) Conversely, for municipalities that did *not* hold concurrent election — and after 1979 this is the case for the vast majority of observations — the effect of the threshold is negative and close to zero. The fact that we also observe heterogeneous treatment effect for the 5,000-resident threshold is consistent with the interpretation that list-length requirements are driving the effect of the threshold (conditional on concurrent elections), rather than OL.

Summing up the findings of the analysis so far, the positive effect of OL on turnout found in the original paper seems to be the results of three different components:

1. The positive effect of the discontinuity among the subsample of elections held in 1983, which is due to measurement error in the data.
2. The positive effect of the discontinuity among the subsample of elections that were *not* held concurrently with other ballots (i.e. most of the 1979 elections, and a small minority of the rest), which is genuine, but compatible with an alternative explanation

that it is due to list-length requirements rather than OL.

3. The null effect of the population threshold among the subsample elections held concurrently with other ballots (i.e. the majority of elections after 1979).

Evidence for a positive effect of OL on turnout is, therefore, inconclusive at best.

5. Ballot Structure, Complexity and Null Votes

However, there is an effect of the discontinuity that can be attributed to open/closed variation in ballot structure with reasonable confidence: *the open-list system increases the occurrence of null voting*, because it is a relatively more complex ballot than closed lists. In fact, he Spanish CL system places a very low cognitive burden on the voter: she simply has to enclose the ballot paper of her party of choice in an envelope, without making any markings. Conversely, in OL elections, all candidate lists are printed on a single ballot paper, and the voter has to tick one to four boxes next to candidates' names: the ballot is null if there are more than four ticks or additional markings (e.g. a cross on the party's symbol). It stands to reason that the relatively higher complexity involved in voting with an open-list ballot increases rates of null voting (Cox and Le Foulon, 2024).

	Dependent Variable: Percentage Null Votes					
	(1)	(2)	(3)	(4)	(5)	(6)
D (Under 250)	1.250***	1.258***	1.308***	1.252***	1.263***	1.036***
	(0.148)	(0.155)	(0.164)	(0.175)	(0.189)	(0.253)
Num.Obs.	18 199	15 567	12 982	10 308	7554	4659
Bandwidth	1×IK	0.85×IK	0.7×IK	0.55×IK	0.4×IK	0.25×IK

Note: *p<0.1; **p<0.05; ***p<0.01

Table 9: RD estimates of the 250-resident threshold effect on the percentage of null votes (1987–2011).

As shown in table 9, the open-list system used below the 250-resident threshold is associated with a 1–1.3 percentage point increase in null votes relative to the one used above the threshold, as already identified in the RD plots. These estimates are from models using the same two-way fixed effects set-up as the main 'turnout' models, but considering only the election-years where null votes are recorded consistently and separately from blanks (1987–2011).[8] Of course, compound treatment remains an obstacle to causal inference for *all* 'effects' of the discontinuity, not just for turnout. However, two additional pieces of evidence suggest that the threshold's effect on null vote share may be attributed to the relatively higher complexity involved in filling out correctly the OL ballot. First, I consider a testable implication of the argument: if the effect on nulls is due to ballot complexity, it should be stronger in locales with lower levels of education. Second, I probe the argument's external validity, showing that population-based discontinuities in Polish municipal electoral systems associated with increases in ballot complexity also have effects on null shares consistent with the theory.

5.1. Heterogeneous Treatment Effects by Education

I collected data from the 1991 and 2001 Spanish census on municipalities' *share of residents without educational qualifications* (the numerator is the sum of illiterates and literates who did not complete primary school; in 1991, the denominator is population aged 10 and over; in 2001, population aged 16 and over). I assigned the values from the 1991 census to municipal elections held before 1995, and the values from the 2001 census to those held after 1999. Models 1 and 2 in table 10 present the estimated effect of the discontinuity on null votes for 'high education' and 'low education' subsamples, defined — respectively — as the bottom and top terciles of the 'share of residents without educational qualifications' variable (computed for municipalities comprised between 100 and 1000 residents). In model 4, I interact the discontinuity with the share of residents without qualifications.[9] The results lend some

[8] As shown in the appendix, I find no effect of the threshold on blank votes, which suggests that variation in null votes across electoral systems is unlikely to be driven by changes in incentives for 'protest' voting.

[9] Because the education variable is different for the two census years and I do not have the data to observe changes in the education profile of the municipalities across elections, I cannot use municipality fixed-effects, which would be collinear with the moderator. Thus, I employ region × year fixed effects, to account for the

credence to the argument that the complexity of OL ballots increases null vote share: the effect of OL is consistently higher in the low-education subsample than in the high-education subsample, and the interaction factor is positive, suggesting that the effect of the discontinuity is more marked in places where more voters may be prone to make mistakes with a complex ballot (though it reaches 95% confidence level only in the 1987–1995 subsample). Pooled sample estimates, which return substantively similar results, are reported in the appendix.

	DV: Percentage of Null Votes (1987–1995)			
	(1) High Educ.	(2) Low Educ.	(3) Full Sample	(4) Full Sample
D (Under 250)	0.995***	1.771***	1.444***	1.087***
	(0.249)	(0.364)	(0.211)	(0.242)
Share No Qualifications			0.406	−0.225
			(0.268)	(0.320)
D × Share No Qualifications				1.246**
				(0.442)
Num.Obs.	3300	1944	5499	5499
	DV: Percentage of Null Votes (1999–2011)			
	(1) High Educ.	(2) Low Educ.	(3) Full Sample	(4) Full Sample
D (Under 250)	1.166***	1.547***	1.453***	1.360***
	(0.244)	(0.206)	(0.153)	(0.175)
Share No Qualifications			−0.222	−0.498+
			(0.203)	(0.260)
D × Share No Qualifications				0.510
				(0.375)
Num.Obs.	4111	3938	9266	9266
Bandwidth	1×IK			
Fixed Effects	Region × Year			

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: Effect of the threshold on null votes, conditional on municipality education.

fact that the occurrence and nature of concurrent elections may affect turnout levels and correlate with the education characteristics of municipalities. This has also the advantage of avoiding ‘forbidden comparisons’ between earlier- and later-treated observations, which produce bias in presence of heterogeneous treatment effect (which, as shown, is pronounced across years and regions). I employ heteroskedasticity-robust standard errors throughout.

5.2. External Validity

	Polish Municipal Elections (1998–2011)	
Population	*Below 20,000*	*Above 20,000*
Electoral Formula	**Multi-member majoritarian**: the most voted candidate(s) from *any* party are elected.	**PR (D'Hondt)** with 5% threshold: seats are attributed proportionally *across party lists*.
Ballot Structure	**Open List**, with multiple votes: voters can choose as many candidates as there are seats, across party-lists.	**Open List** with single vote: voters must choose exactly one candidate, out of multi-member party-lists.
District Magnitude	1–5	5–8
Council Size	15 members	21 members
Ballot Complexity	Open-List PR is **much more complex** than Open-List Majoritarian: below the threshold (OL Majoritarian), multiple votes are allowed; above the threshold (OL PR), they result in a null vote.	
	Polish Municipal Elections (2011–)	
Population	*Below 20,000*	*Above 20,000*
Electoral Formula	**Plurality**: the most voted candidate is elected.	**PR (D'Hondt)** with 5% threshold: seats are attributed proportionally *across party lists*.
Ballot Structure	**Single-member election**: parties present one candidate each, and voters must choose one.	**Open List** with single vote: voters must choose exactly one candidate, out of multi-member party-lists.
District Magnitude	1	5–10
Council Size	15 members	21 members
Ballot Complexity	Open-List PR is **somewhat more complex** than Single-member Plurality: a single vote is allowed in both systems; but below the threshold (Single-member Plurality) parties only present one candidate, above the threshold (Open-List PR), each list has multiple candidates.	

Table 11: Description of electoral systems for Polish municipal councils.

The electoral system of Polish municipal councils (*rady gmin*) varies according to a population-based discontinuity at 20,000 residents. Above the threshold, councilors are elected via districted open-list PR with a single candidate vote; below the threshold, the system was open-list majoritarian in multi-member district with multiple votes between 1998 and 2010 and single-member district plurality (i.e. 'first-past-the-post') after a 2011 reform.[10] As summarized in table 11, the two electoral systems vary across a number of dimensions, including ballot complexity. Specifically, prior to 2011 the OL-PR system is much more complex than the OL majoritarian system: the former requires voters to choose exactly *one* candidate, while the latter allows them to choose as many as there are councilors to be elected in the district. After 2011, both systems restrict choice to one candidate, but the OL ballot is still somewhat more conducive to voter error: in these, parties run multi-candidate lists, while each party only fields one candidate in the plurality elections.

I was able to collect the percentage of null votes cast in three Polish elections: 2010 (held under OL majoritarian/OL-PR rules), 2018 and 2024 (held under single-member plurality/OL-PR rules). Figure 5 shows the estimated effect of crossing the 20,000 resident threshold, going from the simpler ballot structure (multi-member majoritarian with multiple votes, or single-member plurality) to the more complex ballot structure (OL-PR). Regression tables and details on model specification are reported in the appendix. In short, a similar pattern to what observed in Spain emerges: (1) the more complex open-list ballot structure is associated with higher rates of null votes; (2) the effect is higher in 2010, when the system below the threshold is more 'forgiving' (voters may make multiple markings) and lower thereafter; (3) the effect is higher when computed in subsamples of municipalities with lower levels of education.

[10] Electoral Code of 5 January 2011. Note that, prior to 2011, the majoritarian system could be used in districts of magnitude one, so that effectively the 1998–2010 system comprised a mixture of single- and multi-member majoritarian elections.

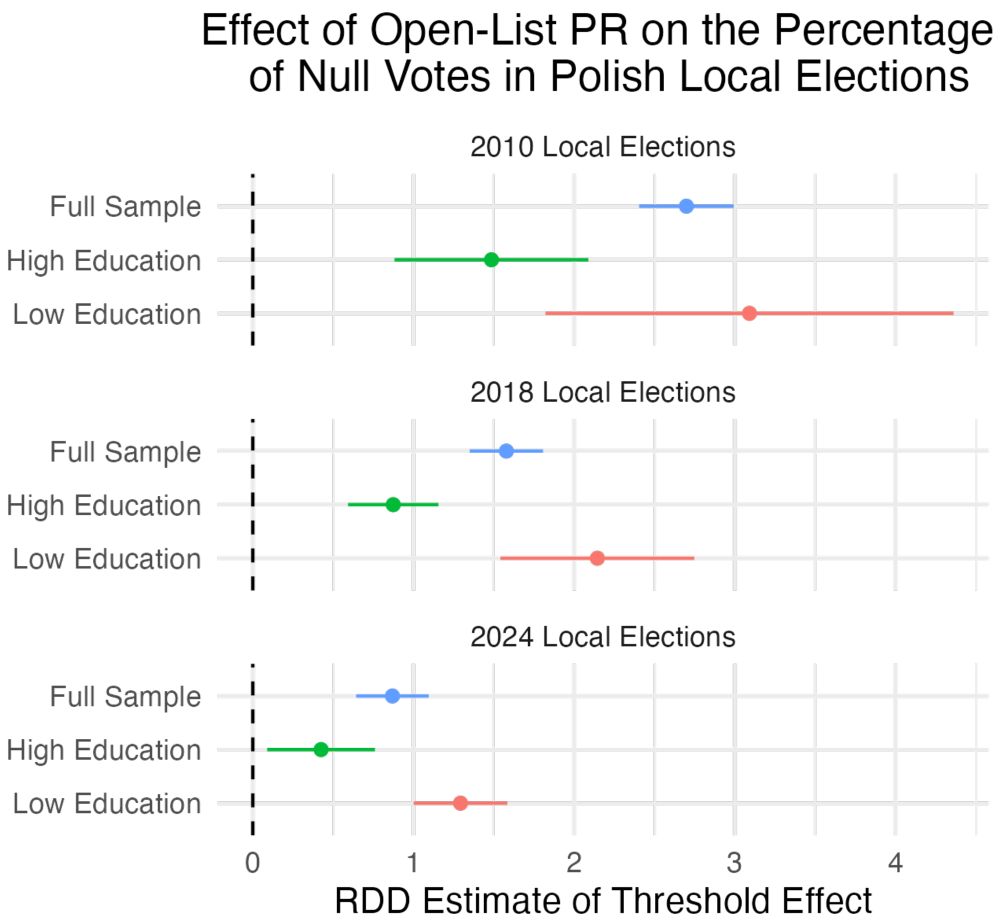

Figure 5: RD estimates of 20,000-resident threshold on turnout (Poland).

6. Conclusion

Do open lists increase turnout? This re-analysis of a key case of sub-national variation in electoral rules suggests that the answer is 'probably not, but they may responsible for increases in the share of people who mistakenly spoil their ballot'. While, in the case study under consideration, compound treatment issues remain an obstacle to causal identification (even of a null result), the analysis presented makes contributions of practical and substantive relevance for the study of electoral system effects. First, it illustrates how variation in electoral rules may correlate with measurement error: as administrative records usually publish sub-national election returns held under different systems in separate datasets, inconsistencies in whether and how votes are recorded in these sources may bias pooled estimates. In so doing, this contribution also alerts researchers of a missing data problem in a widely used source on Spanish local elections (Solé-Ollé, 2006; Costas-Pérez et al., 2012; Vallbé and Ferran, 2017).

Second, this research note illustrates some strategies (subset analysis, 'placebo' treatment models, cross-case comparison) to address suspected cases of compound treatment (Eggers et al., 2018) in RD designs. Overall, this replication exercise suggests that, when it comes to highly multi-dimensional institutions like electoral systems, population-based RD estimates should be interpreted with caution and may have limited external validity. In particular, electoral system effects on outcomes that are interesting substantively and normatively (turnout) may be driven by aspects of local election rules that are unlikely to matter beyond the specific case setting (list-length requirements). Conversely, variation on a dimension that is interesting in a comparative perspective (open-list *vs* closed-list ballot structure) may matter primarily for ballot complexity and null votes, which — while relevant to the fairness of electoral institutions — are rather less 'exciting' for researchers than outcomes like party competition and turnout.

Third, the analysis suggests that concurrent elections may pose a problem to the estimation of electoral system effects in local elections. Concurrent elections are common in settings that have become popular with researchers due to population-based discontinuities in municipal election rules (Barone and De Blasio, 2013; Eggers, 2015; Kantorowicz, 2017): Italy (provincial, regional and occasionally EU elections), France (cantonal elections, prior to 2010) and Poland (county councils and regional elections). To advance our understanding of electoral system effects on turnout, future research may wish to engage more explicitly with the issue of turnout coattails (Bracco and Revelli, 2018) in these contexts.

Finally, by providing evidence for a normatively undesirable effect of OL (more null votes) and casting doubt on their relevance for a normatively desirable one (higher turnout), the analysis has implications for institutional design and electoral system choice. Contrary to an earlier consensus among political scientists in favor of preferential, candidate-centered systems (Bowler et al., 2005), these findings chime with a more recent wave of scholarship that highlights how the presumed benefits of open ballot structures may be overstated (Bosch and Orriols, 2014; Bowler et al., 2018) and their drawbacks underappreciated (Gonzalez-Eiras and Sanz, 2021; Bieber and Wingerter, 2022; Ellenbroek, 2024).

Data Availability Statement

Data and replication files will be made available on a public repository upon acceptance.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Barone, G., De Blasio, G., 2013. Electoral rules and voter turnout. International Review of Law and Economics 36, 25–35.

Bieber, I., Wingerter, L., 2022. Is it all a question of the electoral system? The effects of electoral system types on the representation of women in German municipal councils. German politics 31, 532–557.

Blais, A., 1988. The classification of electoral systems. European Journal of Political Research 16, 99–110.

Blais, A., Aarts, K., 2006. Electoral systems and turnout. Acta politica 41, 180–196.

Bosch, A., Orriols, L., 2014. Ballot structure and satisfaction with democracy. Journal of Elections, Public Opinion & Parties 24, 493–511.

Bowler, S., Farrell, D.M., Pettitt, R.T., 2005. Expert opinion on electoral systems: So which electoral system is “best”? Journal of Elections, Public Opinion & Parties 15, 3–19.

Bowler, S., McElroy, G., Müller, S., 2018. Voter preferences and party loyalty under cumulative voting: Political behaviour after electoral reform in Bremen and Hamburg. Electoral Studies 51, 93–102.

Bracco, E., Revelli, F., 2018. Concurrent elections and political accountability: Evidence from Italian local elections. Journal of Economic Behavior & Organization 148, 135–149.

Collado, I., de Dios, J., 1984. Las elecciones de la transición en Castilla-La Mancha, Vol 1, Tomo 1: Albacete. Confederación Española de Centros de Estudios Locales, Albacete.

Costas-Pérez, E., Solé-Ollé, A., Sorribas-Navarro, P., 2012. Corruption scandals, voter information, and accountability. European journal of political economy 28, 469–484.

Cox, G.W., Fiva, J.H., Smith, D.M., 2016. The contraction effect: How proportional representation affects mobilization and turnout. The Journal of Politics 78, 1249–1263.

Cox, L., Le Foulon, C., 2024. More options, but less willing to cast a valid vote: Evidence from electoral reform in chile. Comparative Political Studies , 00104140241237480.

Eggers, A.C., 2015. Proportionality and turnout: Evidence from French municipalities. Comparative Political Studies 48, 135–167.

Eggers, A.C., Freier, R., Grembi, V., Nannicini, T., 2018. Regression discontinuity designs based on population thresholds: Pitfalls and solutions. American Journal of Political Science 62, 210–229.

Ellenbroek, V., 2024. The effect of more choice on voter turnout causal evidence from Germany. German Politics , 1–27.

Farrell, D.M., McAllister, I., 2006. Voter satisfaction and electoral systems: Does preferential voting in candidate-centred systems make a difference? European Journal of Political Research 45, 723–749.

Fernández Ferrero, M.A., 1995. Los procesos electorales en La Rioja: 1977-1996. Ph.D. thesis. Universidad de La Rioja.

Figueroa, V., 2024. Does proportionality increase turnout? A study of adaptation to oscillating electoral systems. Comparative Political Studies , 00104140241290209.

Gibbons, C.E., Suárez Serrato, J.C., Urbancic, M.B., 2019. Broken or fixed effects? Journal of Econometric Methods 8, 20170002.

Gonzalez-Eiras, M., Sanz, C., 2021. Women's representation in politics: The effect of electoral systems. Journal of Public Economics 198, 104399.

Grofman, B., Selb, P., 2011. Turnout and the (effective) number of parties at the national and district levels: A puzzle-solving approach. Party Politics 17, 93–117.

Imbens, G., Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. The Review of economic studies 79, 933–959.

Kantorowicz, J., 2017. Electoral systems and fiscal policy outcomes: Evidence from Poland. European Journal of Political Economy 47, 36–60.

Linz, J.J., Montero, J.R., 1999. The party systems of Spain: old cleavages and new challenges. volume 138. Instituto Juan March de estudios e investigaciones Madrid.

Magar, E., 2012. Gubernatorial coattails in Mexican congressional elections. The Journal of Politics 74, 383–399.

Sanz, C., 2017. The effect of electoral systems on voter turnout: Evidence from a natural experiment. Political Science Research and Methods 5, 689–710.

Söderlund, P., 2018. Candidate-centred electoral systems and voter turnout, in: Electoral Rules and Electoral Behaviour. Routledge, pp. 14–31.

Solé-Ollé, A., 2006. The effects of party competition on budget outcomes: Empirical evidence from local governments in Spain. Public Choice 126, 145–176.

Vallbé, J.J., Ferran, J.M., 2017. The road not taken. Effects of residential mobility on local electoral turnout. Political Geography 60, 86–99.

Appendix

Appendix A. Further Evidence of Missing Data in 1983

As an additional check to validate the hypothesis that the 1983 election data underestimate turnout above the threshold, I compare turnout in the 1983 municipal elections with turnout in the concurrent regional elections. I was able to gather municipality-level data on regional election turnout for three regions (*comunidades autónomas*): Valencia, Asturias and Castilla y León.[11] The effect of the discontinuity on *turnout discrepancy* (turnout in local elections *minus* turnout in the regional election) is around 5.9 points. (table A.12). It is, of course, possible that voters decided to take part in one election but not the other, and that they were put off by restricted competition in CL elections. But it is highly improbable that this estimate is almost exactly identical to (1) the its effect on *turnout* in municipal elections (+5.6) (2) its effect on *non-valid votes* (+5.8), i.e. blanks and nulls as a percentage of eligible voters. Moreover, if the same analysis is conducted on turnout discrepancy between local and regional elections in the following (1987) election year, the threshold has no discernible effect (figure A.6). Taken together with the evidence from alternative primary sources presented in the paper, missing blanks and nulls in the data for municipal elections in towns with over 250 residents is by far the simplest explanation.

[11] These are available at the website `https://www.datoselecciones.com/`.

	Dependent Variable:		
	Turnout Discrepancy	Turnout (Local)	Non-Valid (Local)
D (Under 250)	5.877***	5.609***	5.783***
	(0.604)	(1.448)	(0.569)
Num.Obs.	1471	1091	1196
Bandwidth	1×IK	1×IK	1×IK
Sample	Valencia, Asturias and Castilla y León, 1983		
			Note: *p<0.1; **p<0.05; ***p<0.01

Table A.12: RD estimates of 250-resident threshold effects on 1983 turnout discrepancy between regional and local elections, 1983 local election turnout, and non-valid vote percentage in the 1983 local elections.

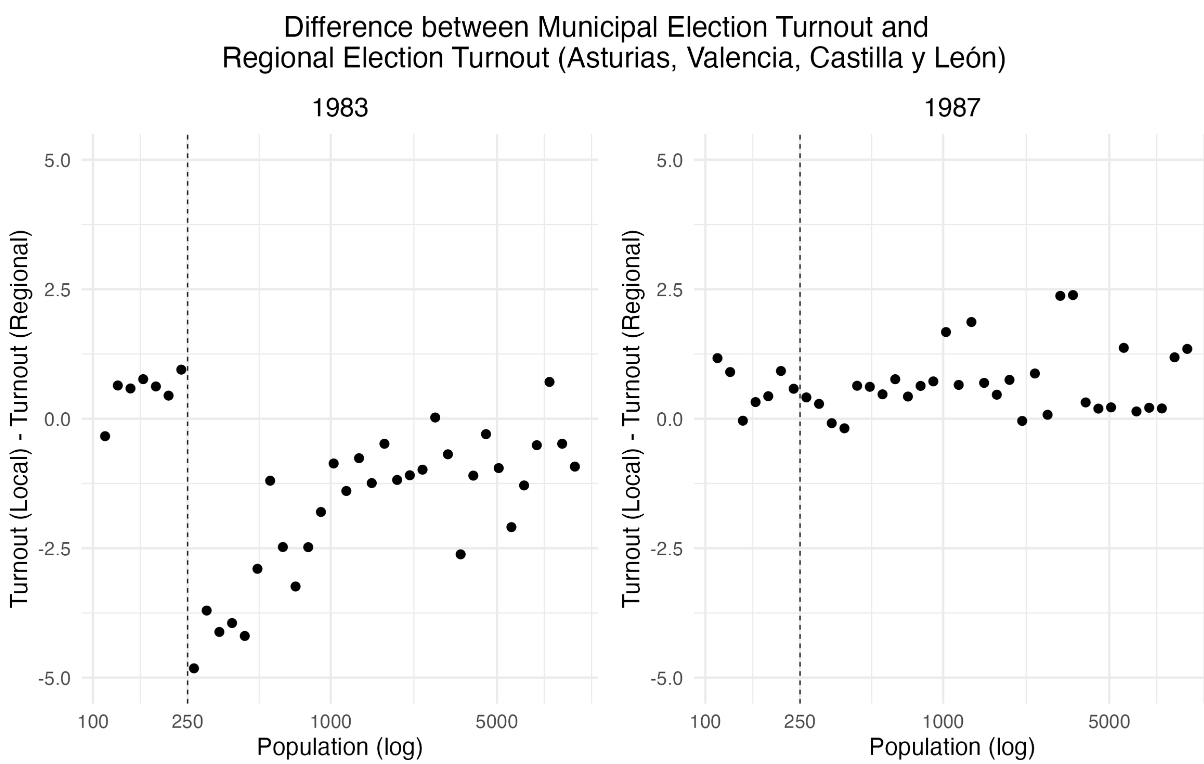

Figure A.6: Discrepancy in turnout between local and regional elections (1983 and 1987).

Appendix B. No Effect of the Threshold on Blank Votes

The table below reproduces the same model specifications used to estimate threshold effects on turnout (table 2) and null votes (table 9), but using the percentage of blank votes as dependent variable. I find no significant effects.

	Dependent Variable: Percentage Blank Votes					
	(1)	(2)	(3)	(4)	(5)	(6)
D (Under 250)	−0.111	−0.163	−0.139	−0.173	−0.068	−0.668+
	(0.177)	(0.186)	(0.199)	(0.207)	(0.226)	(0.289)
Num.Obs.	20 044	17 232	14 397	11 385	8319	5167
Bandwidth	1×IK	0.85×IK	0.7×IK	0.55×IK	0.4×IK	0.25×IK

Note: *p<0.1; **p<0.05; ***p<0.01

Table B.13: RD estimates of the 250-resident threshold effect on the percentage of blank votes (1987–2011).

Appendix C. Heterogeneous Effects by Education, Pooled Sample Analysis

Table C.14 shows heterogeneities by educational profile of the municipality in the effect of the 250-resident discontinuity on the percentage of null votes, computed across *all* elections between 1987 and 2011. The 'high education' subset is defined as the bottom tercile of the variable 'share no qualifications' (SNQ), computed for observations with population between 100 and 1000 residents; the 'low education' subset is defined as the upper tercile. In columns 2–5, SNQ is from the 1991 census; in columns 6–9, it is from the 2001 census.

	Dependent Variable: Pct. Null Votes							
	High Edu	Low Edu	Full Sample		High Edu	Low Edu	Full Sample	
D (under 250)	1.360***	1.452***	1.458***	1.283***	1.465***	1.897***	1.463***	1.345***
	(0.209)	(0.193)	(0.137)	(0.172)	(0.236)	(0.224)	(0.137)	(0.154)
SNQ			0.105	−0.203			−0.153	−0.493+
			(0.167)	(0.205)			(0.193)	(0.250)
D × SNQ				0.596*				0.631+
				(0.299)				(0.345)
Num.Obs.	6611	4639	12 321	12 321	4527	4055	12 329	12 329
SNQ		From 1991 Census				From 2001 Census		

Note: *p<0.1; **p<0.05; ***p<0.01

Table C.14: RD estimates of 250-resident threshold effect on percentage of null votes: heterogeneous effects by educational profile of the municipality (1987–2011), with the moderator/subgroup variable 'Share No Qualification' measured either with 1991 or 2001 census data.

Appendix D. Polish Data Analysis

This section describes the data and analysis used to produce figure 5 in section 5.2.

Appendix D.1. Data

I obtained data for turnout, null votes and legal population in the 2010, 2014 and 2018 local government elections from the Polish National Election Office. I merged municipality-level data with data on educational qualifications from the 2002 and 2021 census data from the Polish statistical office[12], using the former for the 2010 election and the latter for the 2018 and 2021 elections (I was not able to find comparable data at municipality level in the intervening 2011 census). The main education variable used to subset the main corresponds to the share of residents aged over 13 without qualifications beyond primary school leaving certificate. Descriptive statistics are reported in table D.15.

Table D.15: Descriptive Statistics

Population	*min*	*max*	*mean*	*median*	*N (%) <20,000*	*N (%) ≥ 20,000*
2010	1305	704277	14607	7508	2152 (87%)	326 (13%)
2018	1343	702867	14271	7478	2145 (87%)	331 (13%)
2024	1306	706244	14423	7359	2143 (86%)	336 (14%)
Percentage of Nulls	*min*	*max*	*mean*	*median*	*mean <20,000*	*mean ≥ 20,000*
2010	0	7.41	1.2	0.91	0.92	2.98
2018	0	3.86	0.73	0.58	0.57	1.79
2024	0	3.59	0.68	0.48	0.6	1.16
Share Primary or Less	*min*	*max*	*mean*	*median*	*mean <20,000*	*mean ≥ 20,000*
2010	0.17	0.68	0.43	0.44	0.45	0.31
2018	0.8	0.36	0.2	0.19	0.2	0.13
2024	0.8	0.36	0.2	0.19	0.2	0.13

I defined the 'low' and 'high' education subgroups as municipalities with a share of residents without post-primary education above the top tercile and below the bottom tercile of the variable, *computed for municipalities in the proximity of the threshold* (15,000-25,000 residents). This is because the educational profile of municipalities is strongly correlated with

[12]Electoral return data are publicly available at `https://danewyborcze.kbw.gov.pl/indexc4fa.html?title=Strona_g%C5%82%C3%B3wna`. Municipality-level census data are publicly available at `https://bdl.stat.gov.pl/bdl/metadane/grupy/31?back=True`.

their size, so that computing the cutoffs on the whole population would return an insufficient number of observations either side of the discontinuity in the subgroup analysis.

Appendix D.2. Model and Results

The model specification is a standard separate-slope regression discontinuity:

$$\begin{aligned}&\text{Percent Nulls}_i = \alpha_p + \beta_1 D_i + \beta_2(\log(\text{Population}_i) - \log(20,000)) + \\&\beta_3\left[D_i \times (\log(\text{Population}_i) - \log(20,000))\right] + \epsilon_i\end{aligned}$$

Where α_p represents county (*powiat*) fixed effects: these are 368 territorial units between the municipality (*gmina*) and the region (*województwo*), and correspond to the lowest subdivision of the Polish territory that elects representative bodies in local government elections, bar the municipality. D_i is a vector taking the value of 1 if the municipality has a population above the 20,000 resident threshold. For consistency with the main analysis, I reduce the sample with the same Imbens-Kalyanaraman bandwidth selection procedure; I cluster standard errors at the *powiat* level. Estimates are reported in table D.16. Figure D.7 shows binned averages of the dependent variable in standard RD plots, as well as sensitivity plots for the full-sample models, where the effect of the threshold is re-estimated letting the bandwidth vary from a 0.25 to 1.25 times the IK bandwidth.

Table D.16:

2010 Local Council Elections	*Dependent Variable*: Pct. Null Votes		
	Full Sample	Low Educat.	High Educat.
D (Population > 20,000)	2.698***	3.091***	1.485***
	(0.150)	(0.639)	(0.303)
$\log(\text{Population}) - \log(20,000)$	0.063	−0.450	−0.074
	(0.075)	(1.522)	(1.049)
$D \times (\log(\text{Population}) - \log(20,000))$	−1.173***	2.336	3.636+
	(0.218)	(3.586)	(2.103)
Num.Obs.	1160	109	103
2018 Local Council Elections			
	Full Sample	Low Educat.	High Educat.
D (Population > 20,000)	1.578***	2.144***	0.874***
	(0.116)	(0.305)	(0.142)
$\log(\text{Population}) - \log(20,000)$	0.071	−0.014	0.241
	(0.083)	(0.227)	(0.202)
$D \times (\log(\text{Population}) - \log(20,000))$	−0.724**	−0.173	0.021
	(0.256)	(1.073)	(0.448)
Num.Obs.	751	247	201
2024 Local Council Elections			
	Full Sample	Low Educat.	High Educat.
D (Population > 20,000)	0.869***	1.293***	0.425*
	(0.115)	(0.148)	(0.169)
$\log(\text{Population}) - \log(20,000)$	0.116	0.144*	0.356+
	(0.121)	(0.062)	(0.190)
$D \times (\log(\text{Population}) - \log(20,000))$	−0.629*	−0.979*	−0.337
	(0.291)	(0.484)	(0.338)
Num.Obs.	579	577	194
Powiat fixed effects + clustered S.E.	✓	✓	✓
Bandwidth		1 × IK	

Figure D.7: Binned Average and Bandwidth sensitivity plot (Polish local election data).