Online Appendices

Appendix A: Model

There is a continuum of individuals who have to decide on two dimensions of policy. The first one is a primary policy dimension, which I will call ideological stance, or, simply, ideology, i. Ideology can be left wing, denoted by $\underline{i}$, or right wing, $\overline{i}$, so $i \in \{\underline{i}, \overline{i}\}$. We can think of ideology as a salient policy issue, such as the progressiveness of policy.[37] The second issue is targeted or special-interest spending, denoted by s. Spending also can take two values, $s \in \{\underline{s}, \overline{s}\}$, where $\underline{s}$ denotes low (maybe zero) spending and $\overline{s}$ denotes high spending.

Individuals' preferences differ over the two issues. With respect to ideology, individuals are divided into leftists and rightists, $v \in \{l, r\}$. A fraction γ_l of individuals are leftists and prefer $\underline{i}$. The rest, $\gamma_r = 1 - \gamma_l$, are rightists and prefer $\overline{i}$. Let $i^*(v)$ denote the optimal ideological policy from the perspective of an individual of type v (hence, $i^*(l) = \underline{i}$ and $i^*(r) = \overline{i}$). Without loss of generality, I assume that leftists are the majority, thus $\gamma_l > .5$. With respect to special-interest spending, individuals are divided (independently of their preferences in regard to ideology) into citizens and special interests, $w \in \{c, x\}$. A fraction γ_c of individuals are citizens, c, and prefer low spending, $\underline{s}$. The remaining individuals, $\gamma_x = 1 - \gamma_c$, are special interests and prefer high spending, $\overline{s}$. The special interests are those who benefit from special-interest spending, so they do not necessarily constitute an economic elite—they could be, for example, public employees who are favored by the government. Special interests are a minority, thus $\gamma_x < \gamma_r$. Let $s^*(w)$ denote the optimal ideological policy from the perspective of an individual of type w (hence, $s^*(c) = \underline{s}$ and $s^*(x) = \overline{s}$).

[37] For example, municipalities can use tax deductions to benefit poor individuals. Even if it may seem surprising that a main "ideological" issue determines the vote at the local level, the correlation between votes to the main right (left)-wing party in local and national elections is .63 (.57), suggesting that the determinants of voting are closely related at the two levels.

Individuals' utilities are given by:

$$\begin{aligned} u_{l,c}(i,s) &= b_{lc}\mathbb{1}[i=\underline{i}]+\theta_{lc}\mathbb{1}[s=\underline{s}], \\ u_{r,c}(i,s) &= b_{rc}\mathbb{1}[i=\bar{i}]+\theta_{rc}\mathbb{1}[s=\underline{s}], \\ u_{l,x}(i,s) &= b_{lx}\mathbb{1}[i=\underline{i}]+\theta_{lx}\mathbb{1}[s=\bar{s}], \\ u_{r,x}(i,s) &= b_{rx}\mathbb{1}[i=\bar{i}]+\theta_{rx}\mathbb{1}[s=\bar{s}], \end{aligned} \tag{3}$$

where $u_{v,w}$ denotes the utility of type (v,w). Citizens are more concerned about ideology, so that $b_{lc} > \theta_{lc}$ and $b_{rc} > \theta_{rc}$. The special interests, in contrast, care more about special-interest spending, as they benefit directly from it: $b_{lx} < \theta_{lx}$ and $b_{rx} < \theta_{rx}$.

Policy under Representative Democracy

Under representative democracy, policy is delegated to an elected representative, who is an individual and will always implement his or her preferred policy $i^*(v), s^*(w)$ —there is no possibility of commitment. Candidates in the election are put forward by two political parties P, denoted A and B, $P \in \{A, B\}$. Each party is comprised of member individuals bound together by their views on ideology. All members in Party A are leftists, and all members in B are rightists. Within each party, however, there can be any combination of citizens and special interests, so that, even if in the entire population citizens are a majority, that may or may not be the case within political parties. Each party selects a candidate that a majority of its members prefer. Because every individual within a party shares the same preferences toward ideology, the preferences of the majority in the spending dimension will determine which candidate the party proposes. Let s^*_P denote the preferences of the majority of Party P on spending. Parties are not restricted to proposing a member of their own party; thus, in principle, they could propose somebody with an opposing ideology, but that will not happen in equilibrium.

To introduce uncertainty into the election, Besley and Coate (2008) assume that there are some noise voters, a fraction of whom will vote for A's candidate according to the realization of some random variable. To keep the notation simple, I assume instead that the probability that Party A's candidate wins the election is given by the share of individuals who prefer Party A's candidate over Party B's candidate.[38]

[38]The insights and conclusions are the same with both approaches.

Naturally, an individual prefers Party A's candidate if the policy that Party A's candidate will implement gives him or her more utility than will the policy that Party B's candidate will implement, according to (3). More formally, an individual of type $\{v, w\}$ faced with candidates (v_A, w_A) and (v_B, w_B) will favor Party A's candidate if $u_{v,w}(i^*(v_A), s^*(w_A)) > u_{v,w}(i^*(v_B), s^*(w_B))$, and will favor Party B's candidate if $u_{v,w}(i^*(v_A), s^*(w_A)) < u_{v,w}(i^*(v_B), s^*(w_B))$. If both candidates give him or her the same utility, then he or she will favor each candidate with probability 1/2.

Party members know the election probabilities associated with different candidate pairs and take them into account when selecting candidates. An equilibrium is a pair of candidates, one for each party, such that each party's majority members do not want to deviate from their choice, given the other party's choice. More formally, a pair of candidates (v_A, w_A) and (v_B, w_B) is an equilibrium if type (l, s_A^*) individuals prefer a type (v_A, w_A) candidate to any other type of candidate, given that Party B is running a type (v_B, w_B) candidate and, conversely, type (r, s_B^*) individuals prefer a type (v_B, w_B) candidate to any other type of candidate, given that Party A is running a type (v_A, w_A) candidate. Any equilibrium results in a probability distribution over outcomes. The policy outcome will be that associated with Party P's candidate with a probability equal to the chance that Party P's candidate wins.

Equilibrium under Representative Democracy

Case 1: For each $\mathbf{P}, \mathbf{s}_\mathbf{P}^* = \bar{\mathbf{s}}$.

In this case, special interests are a majority in both political parties. This represents a situation in which political parties are captured by special interests. In this case, an equilibrium exists in which both parties run with a special interest of their preferred ideological position, and that equilibrium is unique. This is summarized in Proposition 1.

Proposition 1. *If, for each $P, s_P^* = \bar{s}$, then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is the unique equilibrium.*

Proof. I need to show that if, for each P, $s_P^* = \bar{s}$, then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is the unique equilibrium.

I first show that $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is an equilibrium.

In the proposed equilibrium, Party A wins the election with probability γ_l, and its median voter, who is a leftist special interest, gets utility $b_{lx} + \theta_{lx}$ if that happens.

If Party B wins, the median voter in Party A gets utility θ_{lx}. Therefore, the expected utility of the median voter in Party A, which I will denote by $u(A)$, is $u^*(A) = \gamma_l(b_{lx} + \theta_{lx}) + \gamma_r\theta_{lx} = \theta_{lx} + \gamma_l b_{lx}$ in the proposed equilibrium.

Party A has three possible deviations: (a) If it deviates to proposing ($v_A = l, w_A = c$), Party A reduces both the probability of winning (to $\gamma_c\gamma_l$) and the utility in case of a win (to b_{lx}), so $u(A) = \gamma_c\gamma_l b_{lx} + (1 - \gamma_c\gamma_l)\theta_{lx} < u^*(A)$. (b) If it deviates to ($v_A = r, w_A = c$), Party A increases the probability of winning from γ_l to γ_c but at the cost of getting zero utility if that happens (it is proposing its least preferred policy): $u(A) = \gamma_x\theta_{lx} < u^*(A)$. (c) If it deviates to ($v_A = r, w_A = x$), then both parties propose the same policy and win with 1/2 probability, and $u(A) = \theta_{lx} < u^*(A)$.

Now consider Party B. In the proposed equilibrium, its median voter, who is a rightist special interest, has expected utility $u^*(B) = \gamma_r(b_{rx} + \theta_{rx}) + \gamma_l\theta_{rx} = \theta_{rx} + \gamma_r b_{rx}$. Consider the three possible deviations: (a) If it deviates to proposing ($v_B = r, w_B = c$), $u(B) = \gamma_c\gamma_r b_{rx} + (1 - \gamma_c\gamma_r)\theta_{rx} < u^*(B)$. (b) If it deviates to ($v_B = l, w_B = c$), $u(B) = \gamma_x\theta_{rx} < u^*(B)$. (c) If it deviates to ($v_B = l, w_B = x$), $u(B) = \theta_{rx} < u^*(B)$. This completes the proof that ($v_A = l, w_A = x$) and ($v_B = r, w_B = x$) is an equilibrium.

I now show that the equilibrium ($v_A = l, w_A = x$) and ($v_B = r, w_B = x$) is unique. To do so, I will show that ($v_A = l, w_A = c$) and ($v_B = r, w_B = c$) is not an equilibrium (it is trivial to show that other possible proposals cannot be an equilibrium). Under ($v_A = l, w_A = c$) and ($v_B = r, w_B = c$), $u^*(A) = \gamma_l b_{lx}$. If A deviates to ($v_A = l, w_A = x$), $u(A) = (\gamma_c\gamma_l + \gamma_x)(b_{lx} + \theta_{lx}) > u^*(A)$. Thus, A prefers to deviate, and ($v_A = l, w_A = c$) and ($v_B = r, w_B = c$) cannot be an equilibrium.

□

The intuition is as follows. Party A will not want to switch to a leftist citizen, as it would lower the probability of winning (from γ_l to $\gamma_l\gamma_c$, as it loses all of the votes of the special interests), and it would lower the utility in case of a win for the median individual in the party, which is a special interest. The same happens if Party A switches to a rightist special interest—in this case, the probability of winning goes down to 1/2, as both parties propose the exact same types of candidates. Finally, a "radical" switch to a rightist citizen increases the probability of winning (to γ_l) but at the cost of sacrificing the preferred policy in both dimensions, so the utility in the case of winning is zero. The reasoning for Party B is analogous. Finally, note that both parties running with a citizen of their preferred ideological type cannot be an equilibrium, as both parties would want to deviate to a special interest of their preferred ideological type.

Case 2: For each $\mathbf{P}, \mathbf{s}^*_{\mathbf{P}} = \underline{\mathbf{s}}$.

In this case, citizens are a majority in political parties, as they are in the entire population. In this case as well, an equilibrium exists in which both parties run with a special-interest candidate of their preferred ideological position, under certain conditions that are indicated in the following proposition.

Proposition 2. *(i) If, for each P, $s^*_P = \underline{s}$, and, for each v,*

$$\frac{b_{vc}}{\theta_{vc}} > \frac{\gamma_c}{\gamma_x},$$

then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is an equilibrium.

(ii) If, additionally,

$$\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_x + \gamma_c\gamma_l}{\gamma_x + \gamma_c\gamma_l - \gamma_l}$$

or

$$\frac{b_{rc}}{\theta_{rc}} > \frac{\gamma_x + \gamma_c\gamma_r}{\gamma_x + \gamma_c\gamma_r - \gamma_r},$$

then the equilibrium is unique.

Proof. (i) I need to show that, if for each P, $s^*_P = \underline{s}$, and for each v,

$$\frac{b_{vc}}{\theta_{vc}} > \frac{\gamma_c}{\gamma_x},$$

then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is an equilibrium.

Under the proposed equilibrium, the median voter in Party A, who is a leftist citizen, has an expected utility of $u^*(A) = \gamma_l b_{lc}$. Consider his or her three possible deviations: (a) If it deviates to $(v_A = l, w_A = c)$, Party A increases the utility in case of a win (to $b_{lx} + \gamma_{lc}$) but at the cost of reducing the probability of winning (to $\gamma_c\gamma_l$). Because $\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_c}{\gamma_x}$, by assumption, $u(A) = \gamma_c\gamma_l(b_{lc} + \gamma_{lc}) < u^*(A)$. (b) If it deviates to $(v_A = r, w_A = c)$, $u(A) = \gamma_c\theta_{lc}$. Because $\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_c}{\gamma_x}$, by assumption, $u(A) < u^*(A)$. (c) If it deviates to $(v_A = r, w_A = x)$, $u(A) = 0 < u^*(A)$. For Party B, the argument is symmetric.

(ii) I need to show that if, additionally,

$$\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_x + \gamma_c\gamma_l}{\gamma_x + \gamma_c\gamma_l - \gamma_l}$$

or

$$\frac{b_{rc}}{\theta_{rc}} > \frac{\gamma_x + \gamma_c\gamma_r}{\gamma_x + \gamma_c\gamma_r - \gamma_r},$$

then the equilibrium is unique.

I will first show that $(v_A = l, w_A = c)$ and $(v_B = r, w_B = c)$ is not an equilibrium. Under $(v_A = l, w_A = c)$ and $(v_B = r, w_B = c)$, $u^*(A) = \gamma_l b_{lc} + \theta_{lc}$. Consider the best deviation possible deviation for A, which is $(v_A = l, w_A = x)$ (it is trivial to show that the two other possible deviations are worse). Then, $u(A) = (\gamma_c\gamma_l + \gamma_x)b_{lc} + (1 - \gamma_c\gamma_l - \gamma_x)\theta_{lc}$. If

$$\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_x + \gamma_c\gamma_l}{\gamma_x + \gamma_c\gamma_l - \gamma_l},$$

then $u(A) < u^*(A)$, so Party A will prefer to deviate. For Party B, the best possible deviation is $(v_B = r, w_B = x)$. By symmetry, if

$$\frac{b_{rc}}{\theta_{rc}} > \frac{\gamma_x + \gamma_c\gamma_r}{\gamma_x + \gamma_c\gamma_r - \gamma_r},$$

party B will prefer to deviate. Because it is sufficient that one of the two parties deviates, if

$$\frac{b_{lc}}{\theta_{lc}} > \frac{\gamma_x + \gamma_c\gamma_l}{\gamma_x + \gamma_c\gamma_l - \gamma_l}$$

or

$$\frac{b_{rc}}{\theta_{rc}} > \frac{\gamma_x + \gamma_c\gamma_r}{\gamma_x + \gamma_c\gamma_r - \gamma_r},$$

then $(v_A = l, w_A = c)$ and $(v_B = r, w_B = c)$ is not an equilibrium. □

Assumption (i) rules out that parties want to deviate by running with a citizen instead of a special interest. By proposing a citizen, the median voter in the parties sacrifices some probability of winning (for Party A, it goes down to $\gamma_c\gamma_l$ from γ_l) but obtains some additional utility if the party wins the election, $(\gamma_c l)$. The assumption guarantees that the first effect dominates the second by requiring that ideology, relative to spending, is sufficiently important to citizens, given the distribution of types in the population (or, stated differently, that there are enough special interests, given the preferences of citizens). Assumption (ii) rules out that both parties' choosing a citizen is an equilibrium.

Case 3: $\mathbf{s}^*_{\mathbf{P}} = \underline{\mathbf{s}}$ for some P, $\mathbf{s}^*_{\mathbf{P'}} = \overline{\mathbf{s}}$ for $\mathbf{P} \neq \mathbf{P'}$.

This case is a combination of the other two: Citizens are the majority in one party but not in the other.

Proposition 3. *If $s^*_l = \underline{s}$, $s^*_r = \overline{s}$, and $b_{lc}/\theta_{lc} > \gamma_c/\gamma_x$, then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is the unique equilibrium. If $s^*_l = \overline{s}$, $s^*_r = \underline{s}$, and $b_{rc}/\theta_{rc} > \gamma_c/\gamma_x$, then $(v_A = l, w_A = x)$ and $(v_B = r, w_B = x)$ is the unique equilibrium.*

Proof. The proof of Proposition 3 follows directly from the proof of Proposition 1 and the proof of Proposition 2. □

Following the logic of the first two cases, now the requirement that special interests are sufficiently large is required only for the party in which citizens are a majority.

Policy and Equilibrium under Direct Democracy

Under direct democracy, ideology and spending are voted on separately. The equilibrium outcomes are therefore $i = \underline{i}$ and $s = \underline{s}$. Intuitively, as issues are unbundled, the position preferred by the median voter (a leftist in ideology and a citizen in spending) prevails in both dimensions.

In sum, the model predicts that direct democracy will reduce spending if special interests have captured political parties, or if they are sufficiently large. In addition, the model predicts that spending in representative democracy should not depend on which party is in office: Both parties, A and B, converge to the overspending position. Finally, it is important to note that the model does not yield a clear-cut prediction in terms of welfare. Even if direct democracy leads to policy more in line with the median voter's preferences, a switch from representative to direct democracy creates winners and losers. Citizens, who are the majority, benefit from direct democracy as a result of reduced special-interest spending, while special interests lose. But special interests may feel more intensely about the issue. Hence, a measure of welfare—for example, utilitarian—could rise or fall.

Appendix B: Historical Overview

The use of *concejo abierto* dates from centuries ago (for a detailed historical overview, see Salanova Alcalde (2009)). In a primitive form, they appeared in the Christian territories in he Early Middle Ages, when neighbors organized themselves in assemblies to decide on the government of villages. Traditionally, municipalities themselves decided whether to work under direct or representative democracy. A first attempt to introduced a population threshold took place in 1924, when a national law *(Estatuto Municipal)* imposed the use of direct democracy to all municipalities with fewer than 500 inhabitants. However, this provision was never enforced. During the Second Spanish Republic, another attempt was made to extend the use of direct democracy to all municipalities with fewer than 500 inhabitants (1935 Law), but this attempt never materialized either, as the regime lasted only one more year before the onset of the Spanish Civil War in 1936. The situation did not change during Franco's regime. A 1955 law *(Ley de Régimen Local)* required the following of the direct-democracy system only for those municipalities in which that was the traditional form of government. It was not until after the restoration of democracy after Franco's death in 1975 that the rules changed substantially. In 1978, national law *(Ley de Elecciones Locales)* required all municipalities whose population in the election-year was smaller than 25 inhabitants to follow the direct democracy-system. In 1985, a reform extended the requirement to all municipalities with fewer than 100 inhabitants.[39] This regulation was in force until 2011, and is the focus of this paper. In 2011, the law was changed to eliminate the requirement to follow direct democracy for municipalities with fewer than 100 inhabitants. The rationale for this change was that most local politicians preferred to avoid direct democracy (which is consistent with the sorting around the threshold) for the reasons discussed in Section 5. Municipalities can still adopt the direct-democracy system, by following the procedure described in the national law (similar to the one described in footnote 39), but no municipality is now required to follow direct democracy. No official data regarding the use of direct democracy after 2011 exist yet, but partial data indicate that it is low (8.3%), consistent with the purpose of the change in the law.

[39] Municipalities with 100 or more inhabitants could follow a demanding procedure to adopt the direct-democracy system. Specifically, a majority of the citizens of the municipality had to sign a petition, and two-thirds of the members of the council and the regional government had to approve. To the best of my knowledge, no municipality ever used this procedure. This implies that the regression discontinuity component of the estimation is sharp, as the probability of treatment jumps from 0 to 1 at the threshold.

Appendix C: Definition of Variables

In this Appendix, I provide details on the definition of the variables used in the paper.

Budget Variables

Expenditures is defined as the sum of all (non-financial) chapters of the expenditures budget, net of transfers:

$$Expenditures_{myt} = \sum_{k=1}^{7} E_{k,myt} - R_{4,myt} - R_{7,myt},$$

where m denotes municipality, y denotes the year, t denotes the term, E_k is expenditures per capita on Chapter k of the expenditures budget, and R_k denotes revenues per capita from Chapter k of the revenues budget.

Analogously, *Revenues*, *Transfers*, and *Deficit* are defined as follows:

$$\begin{aligned} Revenues_{myt} &= \sum_{k=1}^{7} R_{k,myt} - R_{4,myt} - R_{7,myt}, \\ Transfers_{myt} &= R_{4,myt} + R_{7,myt} - E_{4,myt} - E_{7,myt}, \\ Deficit_{myt} &= Expenditures_{myt} - Revenues_{myt}. \end{aligned}$$

Elections Variables

I use the following variables from national elections:

$$\begin{aligned} Turnout_{mt} &= 100 \frac{VotesCast_{mt}}{ElectoralCensus_{mt}}, \\ PPVoteSh_{mt} &= 100 \frac{PPVotes_{mt}}{ValidVotes_{mt}}, \\ PSOEVoteSh_{mt} &= 100 \frac{PSOEVotes_{mt}}{ValidVotes_{mt}}, \\ IUVoteSh_{mt} &= 100 \frac{IUVotes_{mt}}{ValidVotes_{mt}}, \end{aligned}$$

where

$$VotesCast_{mt} = VotesForCandidates_{mt} + BlankVotes_{mt} + SpoiltVotes_{mt},$$
$$ValidVotes_{mt} = VotesForCandidates_{mt} + BlankVotes_{mt}.$$

That is, *Turnout* measures the proportion of citizens who cast a vote over the set of potential voters (the electoral census). There is no voter registration in Spain: Potential voters are all citizens of Spain and other EU countries as well as countries under Reciprocity Treaties; older than 18, and not disenfranchised by court order.[40] Valid votes include votes for candidates and blank votes but not spoiled votes. I use this denominator because it is relevant for the allocation of seats, as it is used to determine whether parties reach the election threshold to get seats (3% in national elections). Accordingly, it is the one that is normally reported by the media.

Observations for year y are from the Congress elections that take place during the term t to which year y belongs. Because Congress elections have always alternated perfectly with local elections, there is only one Congress election per term. For example, for years 1988-1991, I consider the Congress elections of 1989. As explained in the text, I lag the variables to use them as placebos.

Demographic Variables

The age distribution data are provided by the National Institute for Statistics in intervals of five years, up to "85 or more" until 2010 and up to "100 or more" from 2011. To calculate the average age *Mean Age*, I use the mid-points of those intervals. For ages 85–100 (for which there is no five-year information until 2011), I calculate the average age in 2011 (89.01) and use it for the rest of the years. For ages "100 or more," I assume the average is 102.5 years.

I define the share of young people *Young* as the share of individuals in the first four intervals (19 years old or younger), *Middle-Aged* as the share of individuals in the next nine intervals (ages 20–64), and *Old* as the share of individuals in the final five intervals (ages 65 or older).

[40] Disenfranchisement is mostly for disability reasons. In 2011, the number of disenfranchised individuals was 79,398 (including individuals younger than 18) or approximately 0.18% of the population.

Appendix D: Dicussion of the Sorting Around the Threshold

In Spain, the official population size of a municipality is given by the number of citizens who are registered in the municipal registry *(padrón municipal).* Municipalities keep track of all the variations in the population in the public registry and report periodically the data to the National Statistics Institute (INE). The INE validates the information it receives, checking that there is no fraud—for example, it ensures that, for every registration, there is a corresponding unregistration in another municipality—and, yearly, makes the final population figures public. While this system makes it difficult to imagine that there is direct fraud or manipulation of population figures, sorting around the threshold can appear as mayors (or other local politicians) persuade some individuals to register in the municipality, with the goal of reaching the population threshold and falling into representative democracy. This is facilitated by the fact that individuals who have dwellings in more than one municipality can, in practice, decide in which of them to register: Although individuals are required to register in the municipality in which they spend the most time, this requirement is almost impossible to monitor and is not enforced in practice.

Naturally, the question is why politicians would prefer to be under representative democracy. There are at least three possible reasons. First, representative democracy is easier to operate, as it does not require calling town meetings to adopt decisions.[41] Second, there are five political positions in representative democracy (the five city councilors), but only one in direct democracy (the mayor). Although, in most cases, these positions are not remunerated, people may still derive non-monetary benefits for holding them. Finally, following the logic of the model presented in Section 2, if the mayor is a special interest, he or she will prefer to be under representative democracy as a means to implement his or her preferred policy.[42]

[41] Conversations with mayors and other local government officials make me think that this is probably the most relevant reason.

[42] Of course, if the mayor is a citizen instead of a special interest, he or she will prefer direct democracy. This raises the theoretical possibility of two-way sorting. Although it is not possible to directly test for the existence of two-way sorting, I believe that sorting into direct democracy is negligible. First, as mentioned, mayors and local government officials indicate that the main reason for sorting is the first, i.e., that representative democracy is easier to operate. Second, it is hard to conceive local government officials trying to persuade people to go to register in another municipality. Also, note that the empirical approach and the robustness checks to assess the validity of the strategy, in particular the donut regressions, do not depend on the sorting being one-way.

Appendix E: Robustness Checks

Here I present seven robustness checks to assess the validity of the empirical strategy.

Covariate Balance

Direct- and representative-democracy municipalities should be similar in any predetermined characteristic. To study if this is the case, I estimate Equation (1) with a considerable number of political, demographic, and economic variables as the outcomes. If the empirical strategy is valid, the coefficient on *DirDem* should be zero in these regressions.[43]

The results from these tests are shown Table A5 and represented graphically in Figure A3. All of the variables are balanced the threshold, with no coefficient being significant even at the 10% level. This provides assurance about the validity of the empirical strategy, as it indicates that municipalities in direct- and in representative-democracy are similar in a considerable number of observable characteristics.

Pretrends in the Outcomes of Interest

Here I exploit the longitudinal dimension of the data to test for pretrends in the outcomes of interest. I lag the outcome variables four years (so that they correspond to the outcomes at the same year of the previous term) and perform placebo tests by estimating Equation (1) with these lagged variables as outcomes. Finding a discontinuity in the contemporaneous outcome but not in the lagged one would be a strong piece of evidence supporting the validity of the regression discontinuity design (Lee and Lemieux (2010)). The results of these tests are shown in Table A6 and Figure A4. The coefficients in Table A6 are starkly different to those that show the contemporaneous effect shown in Table 2, and show that there is no effect of direct democracy on the previous term's expenditures, revenues, and deficit.

The results in this section appease the concern that municipalities that switch systems in a given direction were conducting different policies than the rest *before* the switch. In particular, one concern would be that mayors would try to cross the threshold by using taxes or public spending to attract people to the municipality. If this happened, we should see an effect of direct democracy this term on taxes or

[43] To ensure that variables are predetermined, I consider Congress elections held during the previous term, $t-1$.

spending in the previous term. The results rule out this possibility.

Donut Regressions

To address specifically the sorting around the threshold, here I consider donut regressions, following Barreca, Guldi, Lindo, and Waddell (2011) (see Kantorowicz (2017) for a recent application). In this approach, observations in the domain of population sizes that is affected by the sorting are dropped, so that the estimation relies only on "non-sorted" observations. For each outcome variable, I consider six regressions. The first excludes observations within a 1% interval around the threshold (that is, municipalities with 99 and 100 inhabitants), the second excludes those within an interval of 3%, and so on, up to the sixth, which excludes 20%. A visual inspection of Figure 2 suggests that this covers almost all of the "sorted" observations.[44]

Table A7 shows these results. All of the estimates for expenditures and revenues are statistically significant and, in line with the baseline results, imply that direct democracy reduces expenditures and revenues. If anything, these results point to an even larger effect, although the estimates are also more noisy. The estimates for deficits are close to zero and not statistically significant in all but one of the six regressions considered. These results provide assurance that the findings are not driven by the sorting—the effects remain when we exploit only observations that are not affected by it.

Switches into and out of Direct Democracy

Following Pettersson-Lidbom (2012), I estimate Equation (1) separately for those municipalities that switch into direct democracy and for those that switch out of direct democracy. If the results from these two different samples are similar, it would reinforce the credibility of the estimates, as it would be difficult to explain that correlation by some omitted factor. Furthermore, switches from direct to representative democracy could be more problematic, given that the shape of the density of population sizes suggests that municipalities sort this way. It would be reassuring to find that the effects are not driven by this type of switches.

Table A8 shows the results. For expenditures, the preferred specification for switches into direct democracy yields a point estimate of -6.4% (= $100*(\exp(-0.0665)-$

[44]Restricting the sample more than 20% leads to imprecise estimates, but the results still point to a strong reduction in spending and revenues in direct democracy.

1)) (significant at the 5% level), and the one for switches out of direct democracy, -13.1% (significant at the 1% level). The results for revenues also tend to support the main findings. In switches into direct democracy, direct democracy reduces revenues by 4.2% in the preferred specification—the effect is not significant, but it is significant at the 5% level at both 50% and 150% of that bandwidth. In switches out of direct democracy, the effect is -8.4% (significant at 10%). Finally, the results for deficits are mostly insignificant in the two different subsamples.

Placebo Tests at Other Population Thresholds

The donut regressions and the previous test deal with concerns about identification, especially the sorting. Here I address another possible concern, namely, that the model is misspecified and leads to overestimation of the effects. To address this issue, I conduct placebo tests by estimating the effect of crossing population thresholds that are irrelevant (e.g., the effect of having more than 115 inhabitants). Specifically, I create placebo treatments at all other population sizes from 30 to 220 inhabitants, by defining dummies that indicate if the population of a municipality-year is above or below a given population size.[45] I then run equation (1) with every placebo treatment (so I run 190 regressions per outcome variable). If direct democracy has a real causal effect, then the estimate of direct democracy on policy, based on the 100-inhabitant threshold, should be an outlier in the distribution of placebo coefficients.

I show the results from these tests in Figure A5. I show the empirical cumulative distribution function of point estimates and t-statistics for the 190 regressions considered for each variable. I also show the implied p-values, which are the share of placebo regressions in which I obtain a point estimate (or t-statistic) that is larger in absolute value than the one for the true threshold. For expenditures, the implied p-value is 0.036 for point estimates, and 0.01 for t-statistics. For revenues, the p-values are 0.126 and 0.068, respectively and, for deficits, 0.405 and 0.242. These results are consistent with direct democracy reducing expenditures and revenues, but not affecting deficits.

[45]Below 30 (above 220) inhabitants the effects become very imprecisely estimated as there are few observations below (above) that population size—as explained in Section 4, at 250 inhabitants there is another threshold, so I focus on population sizes below that threshold to avoid confounding effects.

Top-Coding Outliers

In Table A9, I study the robustness of the results to top-coding outliers. I winsorize the observations with a dependent-variable value above the top or below the bottom 1% of the observations within the bandwidth. The results are very similar from the main ones, indicating that the main findings are not driven by the outliers.

Missing Observations

If the number of missing observations is unbalanced across the threshold, this might affect the results. In Table A10, I study this question. I estimate Equation 1, with the dependent variable being a dummy that takes the value of 1 if the data are available for a (potential) observation and 0 if the data are missing. The results show no discontinuity at the threshold.

Appendix F: Additional Results

In this Appendix, I provide two additional results that were mentioned in the discussion on the alternative mechanism in Section 7.

Gradual Learning

If the effects of direct democracy on policy were driven by citizens' gradual learning from participating in town meetings, we should observe that the effects grow over time as municipalities spend more time under direct democracy. Here I test whether this is the case, by considering this equation:

$$\begin{aligned} Outcome_{myt} &= \alpha_m + \gamma_y + f(Population_{mt} - 100) + \beta DirDem_{mt} \\ &+ \sum_{j=2}^{4} \Big[N_{j,mt} f_j(Population_{mt} - 100) + \beta_j N_{j,mt} DirDem_{mt} \Big] + u_{myt}, \quad (4) \end{aligned}$$

where $N_{j,mt}$ is a dummy variable that indicates how many terms municipality m at year y has been under direct democracy. For a municipality in its first term, $N_{1,mt} = 1$ and $N_{j,mt} = 0$ for $j \neq 1$ and, analogously, for a municipality in its second or third terms. For a municipality in its fourth or longer term, $N_{4,mt} = 1$ and

$N_{j,mt} = 0$ for all $j \neq 4$.[46] Therefore, β captures the effect of direct democracy in the first term under direct democracy, and the β_j's indicate how the effect varies in subsequent terms.[47]

The results are shown in Table A11. For expenditures, the coefficient on $DirDem_{mt}$ under the optimal bandwidth is -6.3% and significant at the 10% level, indicating that direct democracy reduces expenditures in the first term. The results are also significant for other bandwidths (Columns 2 and 3). For $N_{2,mt}DirDem_{mt}$, $N_{3,mt}DirDem_{mt}$ and $N_{4,mt}DirDem_{mt}$, the coefficients are not significant, suggesting that the effect does not vary over time. It is true, however, that, due to the short length of the panel, the estimates are not very precise, so learning cannot be completely ruled out. A similar pattern appears in the estimation for revenues. For deficits, neither the coefficient on $DirDem_{mt}$ nor the interaction terms are significant.

These results also address one related alternative explanation—that the effects are driven by the transition from one system to the other. For example, if citizens are called to a town meeting for the first time, they might be more hesitant to make large spending commitments or change taxes. Similarly, switching into representative democracy might require additional transition costs. The results in Table A11, however, indicate that the effects are not driven only by the transition year—the difference in policy remains after it.

The Effect of Direct Democracy on Subsequent Elections

Here I test whether being under direct democracy on term t has an effect on national elections after the beginning of term t—during period t itself ($PPVoteSh_{m,t}$, $PSOEVoteSh_{m,t}$, $IUVoteSh_{m,t}$, $Turnout_{m,t}$), and in the subsequent period $t+1$ ($PPVoteSh_{m,t+1}$, $PSOEVoteSh_{m,t+1}$, $IUVoteSh_{m,t+1}$, $Turnout_{m,t+1}$)—by estimating Equation (1) with those variables as outcomes. The results show that there is no

[46] I consider consecutive terms, so a municipality m that switches out of direct democracy and switches back into direct democracy at term t has $N_{1,mt} = 1$ at t. A caveat is that I cannot observe whether a municipality followed direct democracy before the sample period starts, as only municipalities of fewer than 25 (as opposed to 100) inhabitants were required to follow direct democracy before 1987. Larger municipalities could choose between the two systems. My understanding, based on conversations with local government officials, is that very few municipalities opted for direct democracy. Thus, I assume that, for municipalities under direct democracy in the first term $t = 1$, $N_{1,m1} = 1$.

[47] I cannot rule out the possibility that town meetings affect people's preferences in shorter periods of time. For example, attending merely one meeting might make individuals change their minds about policy, thus driving the results.

effect on any of the outcomes (see Table A12). All the coefficients are very close to zero and not significant at conventional levels.

Appendix Tables

Table A1: Number of Municipalities and Switches by Government System

Panel A: Number of Municipalities by Government System

Term	DirDem	RepDem	Total
1987–1991	603	1385	1988
1991–1995	614	1319	1933
1995–1999	695	1387	2082
1999–2003	781	1475	2256
2003–2007	797	1486	2283
2007–2011	827	1510	2337

Panel B: Number of Switches in Government System

Term	RepDem → DirDem	DirDem → RepDem	Total
1st to 2nd	50	13	63
2nd to 3rd	93	21	114
3rd to 4th	94	19	113
4th to 5th	69	38	107
5th to 6th	73	43	116
Total	379	134	513

The data refer to the municipalities used in the estimation—that is, those with a population of 250 or fewer inhabitants, and with non-missing data for some year(s) of the term.

Table A2: Effect of Direct Democracy on Public Finances (No Fixed Effects)

Panel A: Log Expenditures

	(1) Log Exp.	(2) Log Exp.	(3) Log Exp.	(4) Log Exp.	(5) Log Exp.	(6) Log Exp.
DirDem	-0.187***	-0.142***	-0.156***	-0.168***	-0.169***	-0.138***
	(0.0357)	(0.0365)	(0.0288)	(0.0403)	(0.0397)	(0.0445)
Observations	11975	17691	6001	42078	42078	42078
Municipalities	1145	1450	793	2689	2689	2689
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5
Bandwidth	Optimal[a]	1.5 x Opt.	.5 x Opt.	Full	Full	Full

Panel B: Log Revenues

	(1) Log Rev.	(2) Log Rev.	(3) Log Rev.	(4) Log Rev.	(5) Log Rev.	(6) Log Rev.
DirDem	-0.152***	-0.141***	-0.152***	-0.171***	-0.177***	-0.147***
	(0.0320)	(0.0329)	(0.0248)	(0.0379)	(0.0371)	(0.0414)
Observations	10662	16117	5578	42162	42162	42162
Municipalities	1084	1383	765	2689	2689	2689
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5
Bandwidth	Optimal[b]	1.5 x Opt.	.5 x Opt.	Full	Full	Full

Panel C: Deficit (euros)

	(1) Deficit	(2) Deficit	(3) Deficit	(4) Deficit	(5) Deficit	(6) Deficit
DirDem	-1.851	0.266	-1.722	-3.571	12.82	-6.372
	(7.883)	(7.585)	(9.366)	(15.24)	(19.27)	(19.50)
Observations	34617	41811	20694	42162	42162	42162
Municipalities	2286	2670	1603	2689	2689	2689
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5
Bandwidth	Optimal[c]	1.5 x Opt.	.5 x Opt.	Full	Full	Full

Results from estimating Equation (1) without the fixed effects. Each column is a separate regression with a uniform kernel. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 27, [b] Optimal BW = 25, [c] Optimal BW = 98. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Effect of Direct Democracy on Transfers from Upper-Level Governments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.
DirDem	0.0109	-0.0202	0.00613	-0.00143	0.00195	0.0496	-0.0229
	(0.0489)	(0.0435)	(0.0656)	(0.0577)	(0.0662)	(0.0768)	(0.0499)
Observations	12655	19064	6525	40550	40550	40550	11504
Municipalities	1163	1506	803	2630	2630	2630	1096
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5	Linear
Bandwidth	Optimal[a]	1.5 x Opt.	.5 x Opt.	Full	Full	Full	Opt.
Mun. trends							Yes

Results from estimating Equation (1) (columns 1 to 6) and Equation (2) (column 7). Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects, and the last column also includes municipality-specific trends. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 31. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Effect of Direct Democracy on Capital Expenditures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.	Log Tr.
DirDem	0.0183	-0.00323	0.0293	0.0514	0.0499	0.0472	0.0365
	(0.0711)	(0.0660)	(0.0925)	(0.0795)	(0.0930)	(0.105)	(0.0754)
Observations	12101	18236	6234	38929	38929	38929	11003
Municipalities	1140	1490	784	2624	2624	2624	1073
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5	Linear
Bandwidth	Optimal[a]	1.5 x Opt.	.5 x Opt.	Full	Full	Full	Opt.
Mun. trends							Yes

Results from estimating Equation (1) (columns 1 to 6) and Equation (2) (column 7). Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects, and the last column also includes municipality-specific trends. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 31. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Placebo Tests: Covariate Smoothness around the Threshold

Variable	(1) Mean	(2) Optimal BW Value	(3) Opt. BW Results	(4) .5 x Opt. BW Results
PP Vote Sh (%)	51.6	29	-0.0866	0.0524
		[N = 12642]	(0.669)	(0.966)
PSOE Vote Sh (%)	30.6	31	-0.258	-0.382
		[N = 13506]	(0.550)	(0.771)
IU Vote Sh (%)	3.2	28	0.102	0.0510
		[N = 10023]	(0.271)	(0.413)
Votes Difference (%)	31.0	33	0.603	1.284
		[N = 13904]	(1.014)	(1.382)
Votes Winner (%)	59.3	34	0.330	0.374
		[N = 14308]	(0.603)	(0.825)
Turnout (%)	78.3	30	-0.328	-0.144
		[N = 13099]	(0.440)	(0.630)
Mean Age (years)	53.0	34	-0.143	-0.0412
		[N = 9847]	(0.191)	(0.246)
Young (%)	9.2	41	-0.0993	-0.367
		[N = 11428]	(0.243)	(0.290)
Middle-Aged (%)	53.4	37	0.226	0.523
		[N = 10396]	(0.426)	(0.494)
Old (%)	37.3	40	-0.177	-0.225
		[N = 11191]	(0.433)	(0.501)
Immigrants (%)	2.6	39	-0.240	-0.523
		[N = 11185]	(0.324)	(0.397)
EU Immigrants (%)	45.4	29	-0.189	2.452
		[N = 4056]	(3.836)	(4.978)
Unemployment (%)	2.1	34	0.0498	0.0369
		[N = 11212]	(0.0977)	(0.150)

Column 1 shows the mean of the variables. Column 2 shows the optimal bandwidth value and number of observations for a placebo test that estimates the effect of direct democracy on the corresponding variable, according to Equation (1). Columns 3 and 4 show the results for the placebo tests: each column is a separate local linear regression with a uniform kernel. Standard errors, clustered at both municipality and the running variable, are in parentheses. The optimal bandwidth (BW) is based on Imbens and Kalyanaraman (2012). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: Placebo Tests: Lagged Outcomes

Panel A: Log Expenditures (t-1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Exp.	Log Exp.	Log Exp.	Log Exp.	Log Exp.	Log Exp.	Log Exp.
DirDem	-0.0192	0.00203	-0.0124	-0.0148	-0.000786	-0.0176	-0.00907
	(0.0291)	(0.0287)	(0.0420)	(0.0354)	(0.0417)	(0.0493)	(0.0305)
Observations	9433	13679	4738	33486	33486	33486	9433
Municipalities	1003	1294	678	2506	2506	2506	1003
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5	Linear
Bandwidth	Optimal[a]	1.5 x Opt.	.5 x Opt.	Full	Full	Full	Optimal[a]
Mun. trends							Yes

Panel B: Log Revenues (t-1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log Rev.	Log Rev.	Log Rev.	Log Rev.	Log Rev.	Log Rev.	Log Rev.
DirDem	-0.00816	0.00801	-0.00210	-0.00493	0.0175	-0.0126	-0.00908
	(0.0261)	(0.0257)	(0.0376)	(0.0322)	(0.0377)	(0.0446)	(0.0270)
Observations	9455	14027	4753	33553	33553	33553	9455
Municipalities	1003	1311	679	2507	2507	2507	1003
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5	Linear
Bandwidth	Optimal[b]	1.5 x Opt.	.5 x Opt.	Full	Full	Full	Optimal[b]
Mun. trends							Yes

Panel C: Deficit (t-1) (euros)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Deficit	Deficit	Deficit	Deficit	Deficit	Deficit	Deficit
DirDem	-5.832	-2.534	-3.448	-7.204	-17.86	-11.41	-0.517
	(9.638)	(10.88)	(13.62)	(11.91)	(13.97)	(15.42)	(9.515)
Observations	10738	16034	5360	33553	33553	33553	9455
Municipalities	1100	1436	727	2507	2507	2507	1003
Specification	Linear	Linear	Linear	Order 3	Order 4	Order 5	Linear
Bandwidth	Optimal[c]	1.5 x Opt.	.5 x Opt.	Full	Full	Full	Optimal[c]
Mun. trends							Yes

Results from estimating Equation (1) (columns 1 to 6) and Equation (2) (column 7). Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects, and the last column also includes municipality-specific trends. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 27, [b] Optimal BW = 28, [c] Optimal BW = 31. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Donut Regressions

Panel A: Log Expenditures

	(1) Log Exp.	(2) Log Exp.	(3) Log Exp.	(4) Log Exp.	(5) Log Exp.	(6) Log Exp.
DirDem	-0.0956***	-0.0898***	-0.112***	-0.153***	-0.235***	-0.250***
	(0.0239)	(0.0278)	(0.0325)	(0.0414)	(0.0554)	(0.0922)
Observations	41578	40647	39772	37725	35645	33543
Municipalities	2636	2635	2632	2620	2581	2529
Excluded	1%	3%	5%	10%	15%	20%

Panel B: Log Revenues

	(1) Log Rev.	(2) Log Rev.	(3) Log Rev.	(4) Log Rev.	(5) Log Rev.	(6) Log Rev.
DirDem	-0.0959***	-0.0904***	-0.109***	-0.144***	-0.222***	-0.249***
	(0.0226)	(0.0264)	(0.0312)	(0.0401)	(0.0548)	(0.0933)
Observations	41662	40730	39850	37796	35712	33607
Municipalities	2637	2636	2633	2621	2583	2531
Excluded	1%	3%	5%	10%	15%	20%

Panel C: Deficit

	(1) Deficit	(2) Deficit	(3) Deficit	(4) Deficit	(5) Deficit	(6) Deficit
DirDem	-2.407	-0.237	-8.433	-23.20**	-21.56	-10.18
	(9.026)	(10.03)	(9.400)	(10.47)	(13.75)	(22.44)
Observations	41662	40730	39850	37796	35712	33607
Municipalities	2637	2636	2633	2621	2583	2531
Excluded	1%	3%	5%	10%	15%	20%

Results from estimating Equation (1) dropping observations within the window around the threshold indicated in the Excluded row. Each column is a separate linear regression with a uniform kernel. All regressions include municipality and year fixed effects. Standard errors, clustered at both municipality and the running variable, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Effect by Switches into and out of Direct Democracy

Panel A: Log Expenditures

	(1) Log Exp.	(2) Log Exp.	(3) Log Exp.	(4) Log Exp.	(5) Log Exp.	(6) Log Exp.	(7) Log Exp.
DirDem	-0.0799***	-0.0665**	-0.0518*	-0.0786*	-0.141***	-0.0835**	-0.116
	(0.0267)	(0.0289)	(0.0268)	(0.0439)	(0.0452)	(0.0362)	(0.0792)
Observations	11932	11347	16987	5520	10379	16025	4664
Municipalities	1102	1093	1402	736	1090	1398	696
Specification	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Bandwidth	Optimal[a]	Opt.	1.5 x Opt.	.5 x Opt.	Opt.	1.5 x Opt.	.5 x Opt.
Switches	All	Into	Into	Into	Out of	Out of	Out of

Panel B: Log Revenues

	(1) Log Rev.	(2) Log Rev.	(3) Log Rev.	(4) Log Rev.	(5) Log Rev.	(6) Log Rev.	(7) Log Rev.
DirDem	-0.0521**	-0.0422	-0.0551**	-0.0991**	-0.0877*	-0.0968***	-0.0729
	(0.0261)	(0.0283)	(0.0249)	(0.0451)	(0.0467)	(0.0348)	(0.0841)
Observations	10625	10054	15436	5107	9076	14464	4281
Municipalities	1047	1034	1337	709	1029	1333	666
Specification	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Bandwidth	Optimal[b]	Opt.	1.5 x Opt.	.5 x Opt.	Opt.	1.5 x Opt.	.5 x Opt.
Switches	All	Into	Into	Into	Out of	Out of	Out of

Panel C: Deficit (euros)

	(1) Deficit	(2) Deficit	(3) Deficit	(4) Deficit	(5) Deficit	(6) Deficit	(7) Deficit
DirDem	-7.877	-6.187	-3.236	1.468	-17.49*	-13.77	-5.560
	(8.050)	(8.465)	(8.572)	(9.726)	(9.283)	(9.398)	(10.20)
Observations	34570	33680	40796	19965	32726	39842	19006
Municipalities	2239	2233	2614	1555	2232	2613	1554
Specification	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Bandwidth	Optimal[c]	Opt.	1.5 x Opt.	.5 x Opt.	Opt.	1.5 x Opt.	.5 x Opt.
Switches	All	Into	Into	Into	Out of	Out of	Out of

Results from estimating Equation (1). Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects. Columns with switches "Into" ("Out of") exclude from the sample the municipality-years corresponding to terms in which a municipality switched from direct (representative) to representative (direct) democracy. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 27, [b] Optimal BW = 25. [c] Optimal BW = 98. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Robustness to Top-Coding Outliers

Panel A: Log Expenditures

	(1) Log Exp.	(2) Log Exp.	(3) Log Exp.	(4) Log Exp.	(5) Log Exp.	(6) Log Exp.
DirDem	-0.0799***	-0.0815***	-0.0625***	-0.0797***	-0.0394	-0.0793**
	(0.0267)	(0.0255)	(0.0232)	(0.0244)	(0.0278)	(0.0392)
Observations	11932	11932	17646	14763	8914	5964
Municipalities	1102	1102	1405	1263	946	756
Bandwidth	Optimal[a]	Opt.	1.5 x Opt.	1.25 x Opt.	.75 x Opt.	.5 x Opt.
Top-Coding	NO	YES	YES	YES	YES	YES

Panel B: Log Revenues

	(1) Log Rev.	(2) Log Rev.	(3) Log Rev.	(4) Log Rev.	(5) Log Rev.	(6) Log Rev.
DirDem	-0.0521**	-0.0534**	-0.0700***	-0.0742***	-0.0492*	-0.0987**
	(0.0261)	(0.0252)	(0.0217)	(0.0234)	(0.0263)	(0.0394)
Observations	10625	10625	16074	13555	8064	5542
Municipalities	1047	1047	1340	1198	892	729
Bandwidth	Optimal[b]	Opt.	1.5 x Opt.	1.25 x Opt.	.75 x Opt.	.5 x Opt.
Top-Coding	NO	YES	YES	YES	YES	YES

Panel C: Deficit

	(1) Deficit	(2) Deficit	(3) Deficit	(4) Deficit	(5) Deficit	(6) Deficit
DirDem	-7.877	-4.460	-3.072	-3.619	-4.848	-2.212
	(8.050)	(4.990)	(4.848)	(4.902)	(5.211)	(5.935)
Observations	34570	34570	41761	38310	29227	20647
Municipalities	2239	2239	2620	2438	1960	1556
Bandwidth	Optimal[c]	Opt.	1.5 x Opt.	1.25 x Opt.	.75 x Opt.	.5 x Opt.
Top-Coding	NO	YES	YES	YES	YES	YES

Results from estimating Equation (1). The top-coded regressions winsorize the observations with a dependent-variable value above the top or below the bottom 1% of the observations within the bandwidth. Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 27, [b] Optimal BW = 25, [c] Optimal BW = 98. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A10: Number of Observations at the Threshold

	(1)	(2)	(3)	(4)	(5)
	Obs.	Obs.	Obs.	Obs.	Obs.
DirDem	0.0204	0.0168	0.0143	0.00376	0.0462
	(0.0197)	(0.0186)	(0.0190)	(0.0265)	(0.0426)
Observations	18164	27188	22636	9480	4916
Municipalities	1255	1596	1437	901	671
Bandwidth	Optimal[a]	1.5 x Opt.	1.25 x Opt.	.75 x Opt.	.25 x Opt.

Results from estimating Equation (1). The dependent variable is a dummy that takes the value of 1 if data are available for a (potential) observation and 0 if the data are missing. Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a]Optimal BW = 73 inhabitants. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Effects by Number of Terms under Direct Democracy

Panel A: Expenditures, Revenues, and Deficit

	(1) Log Exp.	(2) Log Exp.	(3) Log Exp.	(4) Log Rev.	(5) Log Rev.	(6) Log Rev.	(7) Deficit	(8) Deficit	(9) Deficit
DirDem	-0.0634*	-0.0638**	-0.107**	-0.0436	-0.0661**	-0.106**	-3.098	-2.188	4.623
	(0.0336)	(0.0291)	(0.0483)	(0.0317)	(0.0267)	(0.0451)	(11.27)	(12.00)	(12.25)
DirDem x N_2	-0.0557	-0.0277	0.0673	-0.0587	-0.0272	0.0436	-12.16	-14.92	-25.82*
	(0.0524)	(0.0377)	(0.0958)	(0.0506)	(0.0373)	(0.0879)	(12.60)	(12.01)	(13.72)
DirDem x N_3	-0.106	-0.0491	-0.0176	-0.0566	-0.0941*	-0.0679	-22.93	-28.11	-11.95
	(0.0830)	(0.0578)	(0.133)	(0.0793)	(0.0571)	(0.132)	(21.79)	(21.65)	(23.02)
DirDem x N_4	-0.0683	0.00848	-0.0240	-0.00572	-0.0170	0.00995	-1.626	-6.934	26.38
	(0.104)	(0.0786)	(0.125)	(0.108)	(0.0802)	(0.171)	(17.78)	(18.05)	(26.65)
Observations	11932	17646	5964	10625	16074	5542	34570	41761	20647
Municipalities	1102	1405	756	1047	1340	729	2239	2620	1556
Bandwidth	Optimal[a]	1.5 x Opt.	.5 x Opt.	Optimal[b]	1.5 x Opt.	.5 x Opt.	Optimal[c]	1.5 x Opt.	.5 x Opt.

Results from estimating Equation (4). Each column is a separate regression with a uniform kernel. All regressions include municipality and year fixed effects. Standard errors, clustered at both municipality and the running variable, are in parentheses. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. [a] Optimal BW = 27, [b] Optimal BW = 25, [c] Optimal BW = 98. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Effects of Direct Democracy on Subsequent Elections Behavior

Panel A: Effect of Direct Democracy at t on Election Variables at t

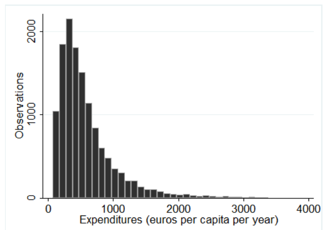
	(1) PP Vote Sh	(2) PP Vote Sh	(3) PSOE Vote Sh	(4) PSOE Vote Sh	(5) IU Vote Sh	(6) IU Vote Sh	(7) Turnout	(8) Turnout
DirDem	-0.649	-0.353	0.212	0.389	0.281	-0.297	-0.217	-0.376
	(0.724)	(1.028)	(0.585)	(0.807)	(0.244)	(0.400)	(0.521)	(0.674)
Observations	11493	5526	12300	6332	10133	4926	11929	5967
Municipalities	1086	727	1125	780	1111	741	1102	754
Bandwidth	Optimal^a	.5 x Opt.	Optimal^a	.5 x Opt.	Optimal^b	.5 x Opt.	Optimal^b	.5 x Opt.

Panel B: Effect of Direct Democracy at t on Election Variables at $t+1$

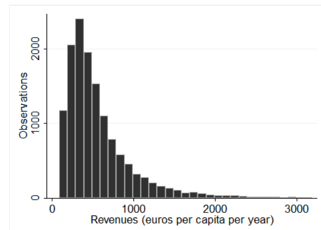
	(1) PP Vote Sh	(2) PP Vote Sh	(3) PSOE Vote Sh	(4) PSOE Vote Sh	(5) IU Vote Sh	(6) IU Vote Sh	(7) Turnout	(8) Turnout
DirDem	-0.686	-1.834	0.300	0.879	0.532*	0.464	0.327	-0.117
	(0.776)	(1.252)	(0.617)	(0.894)	(0.307)	(0.457)	(0.476)	(0.669)
Observations	9263	4390	12845	6414	9753	5039	10917	5231
Municipalities	979	654	1178	797	1090	742	1064	718
Bandwidth	Optimal^c	.5 x Opt.	Optimal^c	.5 x Opt.	Optimal^d	.5 x Opt.	Optimal^d	.5 x Opt.

Each column is a separate local linear regression with a uniform kernel. Standard errors, clustered at both municipality and the running variable, are in parentheses. Fixed effects are municipality and year fixed effects. I calculate the optimal bandwidth following Imbens and Kalyanaraman (2012)'s procedure. ^{a} Optimal BW = 150, ^{b} Optimal BW = 102, ^{c} Optimal BW = 87, ^{d} Optimal BW = 47. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

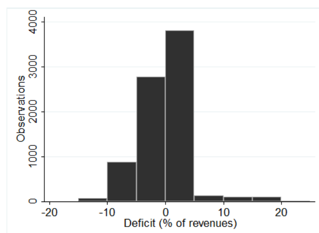
Appendix Figures



(a) Expenditures

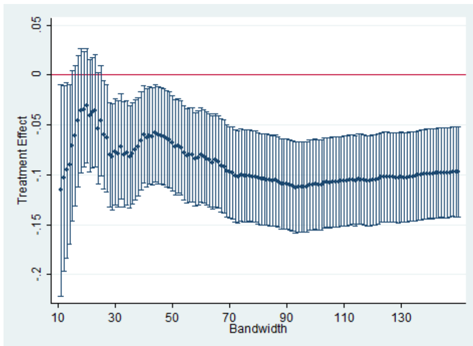


(b) Revenues

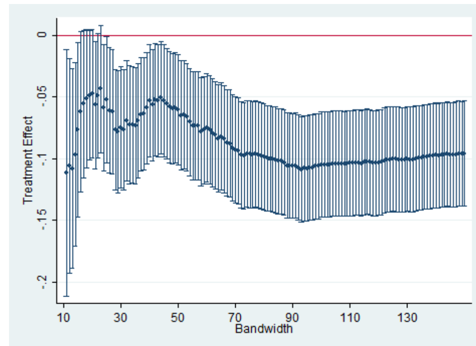


(c) Deficits

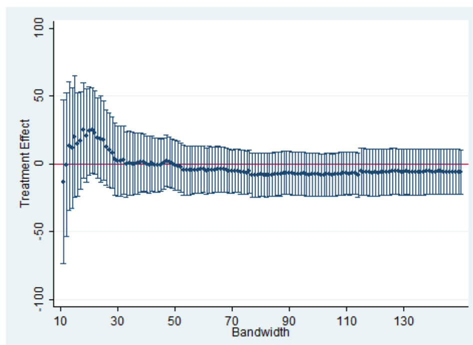
Figure A1: Histograms of expenditures, revenues, and deficits. An observation is a municipality-year. Observations above the 99^{th} and below the 1^{st} percentiles have been excluded to facilitate comprehension (90^{th} and 1^{st} for deficits). Bins are 100-euro wide (5 percentage points for deficit).



(a) Log Expenditures

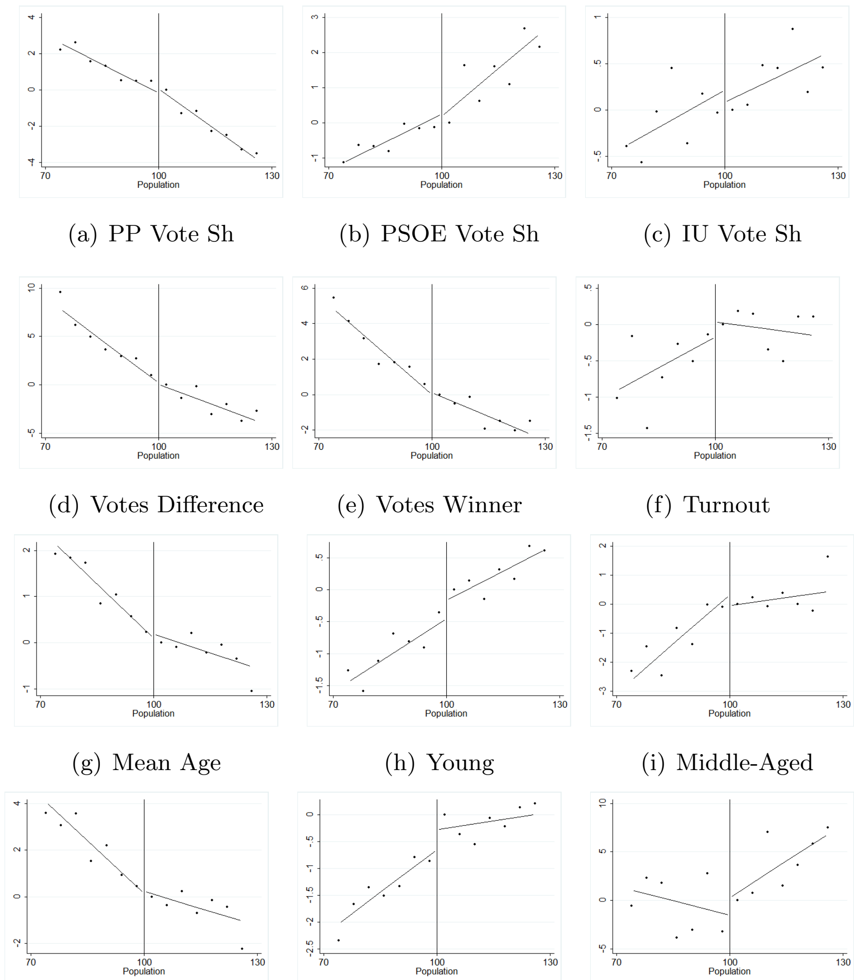


(b) Log Revenues



(c) Deficit

Figure A2: Robustness to Bandwidth Choice. Circles represent the estimated treatment effect, using different bandwidth choices (x-axis). Lines represent the 95% confidence interval (standard errors clustered at the municipality level). I report all possible bandwidths from 10 to 150 inhabitants. (The smaller sample sizes for bandwidths below 10 yield large confidence intervals—larger than .2, or around 20% of expenditures and revenues—for smaller bandwidths.)



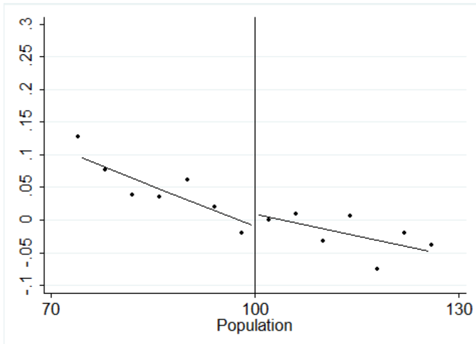
(a) PP Vote Sh (b) PSOE Vote Sh (c) IU Vote Sh

(d) Votes Difference (e) Votes Winner (f) Turnout

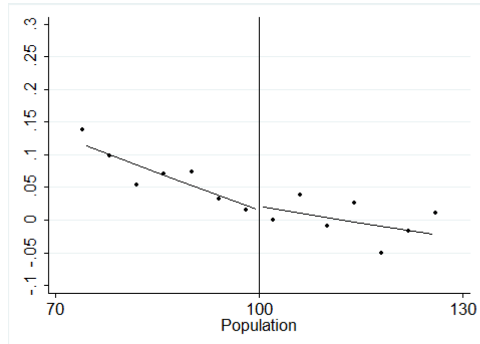
(g) Mean Age (h) Young (i) Middle-Aged

(j) Old (k) Immigrants (l) EU Immigrants

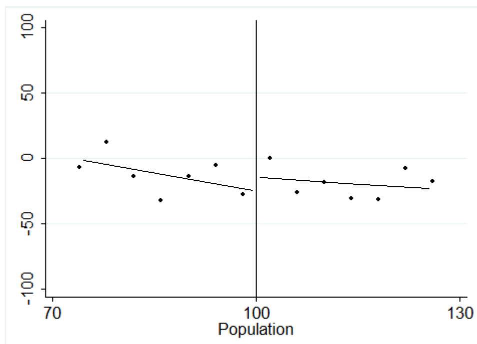
Figure A3: Placebo Tests: Effect of direct democracy on political and demographic variables. I estimate $Outcome_{myt} = \alpha_m + \gamma_y + \sum_{j=100-OBW}^{100+OBW} \delta_j Population_{j,mt} + u_{myt}$, where $Population_{j,mt}$ is a dummy that indicates whether municipality m has population size j at term t. In the y-axis, I plot the estimated coefficients $\hat{\delta}_j$, averaged to 4-inhabitant-wide bins. I normalize the coefficients so that the average bin immediately to the right of the threshold takes the value of zero. The lines are linear fits on $\hat{\delta}_j$, fitted separately for observations above and below the threshold. I use the observations within the optimal bandwidth for $LogExpenditures_{myt}$, so that all graphs show the same range in the x-axis.



(a) Log Expenditures (t-1)

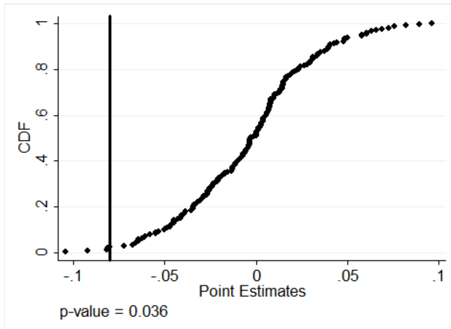


(b) Log Revenues (t-1)

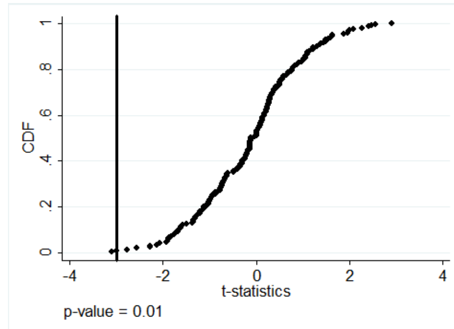


(c) Deficit (t-1)

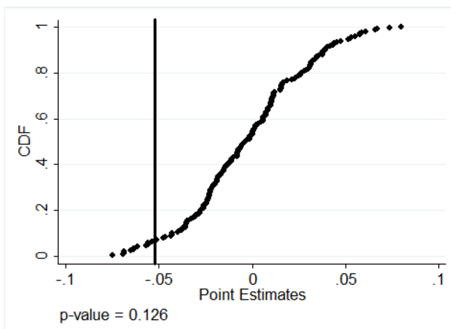
Figure A4: Placebo Tests: Effect of direct democracy on lagged expenditures, revenues, and deficit. I estimate $Outcome_{myt} = \alpha_m + \gamma_y + \sum_{j=100-OBW}^{100+OBW} \delta_j Population_{j,mt} + u_{myt}$, where $Population_{j,mt}$ is a dummy that indicates whether municipality m has population size j at term t. In the y-axis, I plot the estimated coefficients $\hat{\delta}_j$, averaged to 4-inhabitant-wide bins. I normalize the coefficients so that the average bin immediately to the right of the threshold takes the value of zero. The lines are linear fits on $\hat{\delta}_j$, fitted separately for observations above and below the threshold. I use the observations within the optimal bandwidth for $LogExpenditures_{myt}$, so that all graphs show the same range in the x-axis.



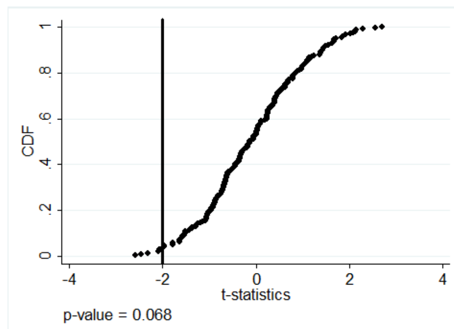
(a) Log Expenditures



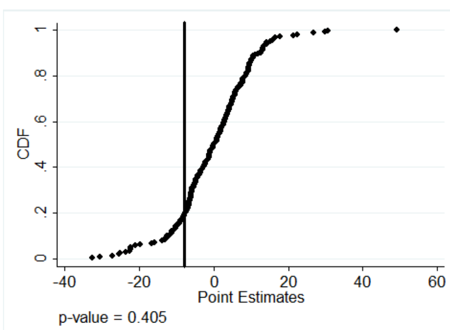
(b) Log Expenditures



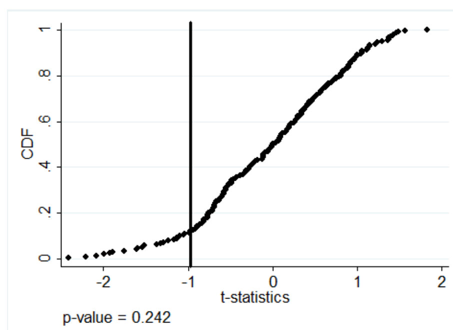
(c) Log Revenues



(d) Log Revenues



(e) Deficit



(f) Deficit

Figure A5: Placebo tests at other population thresholds. I run Equation (1) at all (fake) population thresholds between 30 and 220 inhabitants, by replacing *DD* with a dummy that indicates whether population is larger than the given population threshold. The graphs show the empirical cumulative distribution functions for the resulting point estimates and t-statistics. The vertical lines show the point estimates (or t-statistics) for the effects of direct democracy, obtained at the (true) 100-inhabitant threshold. The p-value below each graph shows the share of point estimates (or t-statistics) that are larger in absolute value than the one for the 100-inhabitant threshold.