HARVARD | BUSINESS | SCHOOL

Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates

Vincent Pons
Clémence Tricaud

Working Paper 17-107

Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates

Vincent Pons
Harvard Business School

Clémence Tricaud
University of Paris-Saclay

Working Paper 17-107

Expressive voting and its cost: Evidence from runoffs with two or three candidates

Vincent Pons* Clémence Tricaud†

May 2017‡

Abstract

In French parliamentary and local elections, candidates ranked first and second in the first round automatically qualify for the second round, while a third candidate qualifies only when selected by more than 12.5 percent of registered citizens. Using a fuzzy RDD around this threshold, we find that the third candidate attracts both "switchers", who would have voted for one of the top two candidates if she were not present, and "loyal" voters, who would have abstained. Switchers vote for the third candidate even when she is very unlikely to win. This disproportionately harms the candidate ideologically closest to her and causes his defeat in one fifth of the races. These results suggest that a large fraction of voters value voting expressively over behaving strategically to ensure the victory of their second-best. We rationalize our findings by a model in which different types of voters trade off expressive and strategic motives.

Keywords: Expressive voting, Strategic voting, Regression discontinuity design, French elections

JEL Codes: D72, K16

*Harvard Business School; BGIE group, Soldiers Field, Boston, MA 02163; vpons@hbs.edu; +1 617 899 7593

†CREST, Ecole Polytechnique, Paris-Saclay University; Route de Saclay, 91128 Palaiseau, France; clemence.tricaud@polytechnique.edu; +33 (0)1 69 33 30 27

‡For suggestions that have improved this article, we are grateful to Daron Acemoglu, Abhijit Banerjee, Pierre Boyer, Bruno Crépon, Torun Dewan, Rafael Di Tella, Esther Duflo, Benjamin Olken, Julio Rotemberg, Kenneth Shepsle, Daniel Smith, James Snyder, Laurent Bouton, Micael Castanheira, seminar participants at MIT, Harvard, the University of Chicago, Bocconi University, the University of Maryland, Ecole Polytechnique, CREST, Sciences Po, and conference participants at APSA and at the RIDGE/LACEA-PEG Workshop on Political Economy. We thank Sebastian Calonico, Matias Cattaneo, Max Farrell and Rocio Titiunik for guiding us through the use of their RDD Stata package (https://sites.google.com/site/rdpackages/rdrobust) and for sharing their upgrades regarding fuzzy design estimations; Nicolas Sauger for his help with the collection of the 1978, 1981 and 1988 French parliamentary election results; Abel François for sharing his data on 1993, 1997 and 2002 candidates' campaign expenditures; and Salomé Drouard and Eric Dubois for their excellent work of data entry and cleaning.

1 Introduction

In an indirect democracy – the form of government prevalent across the West today – representatives rule on behalf of the people. In theory, their representativeness comes from being elected by the people. In practice, it depends on the extent to which voters' choices reflect their true preferences, and on the way in which the voting rule translates vote choices into election outcomes. These two conversions determine who is elected and which policies are enacted.

Under plurality rule, citizens face a difficult tradeoff when they support lower-ranked candidates in elections with more than two candidates: voting for their favorite, or for another candidate with higher chances of winning. By expressing their true preference, voters may split their support over multiple candidates and nominate less-preferred leaders. Hence, the result of the election depends on the extent to which voters are "expressive" – voting based on their preference among candidates only – or "strategic" (or "instrumentally rational") – voting based on likely outcomes of the election.

In his groundbreaking work on strategic behavior, Duverger (1954) posits that voters do not want to waste their vote and will thus exclusively vote for two front-runners in most cases. Models by Palfrey (1989), Myerson and Weber (1993), and Cox (1997) formalize this intuition: under plurality rule, when voters are instrumentally rational, an election with multiple candidates usually boils down to a two-candidate race, and of these two the candidate who is preferred by the majority wins the election.

The division of the American political landscape between the Republicans and the Democrats is a famous illustration of Duverger's law. Yet in other countries, third and lower-ranked candidates frequently receive large fractions of the votes. General elections in the United Kingdom are a case in point: since the Liberal Democrat party emerged in the 1990s, it has regularly split the vote share with the Labour and Conservative parties in many constituencies. In addition, the presence of lower-ranked candidates can have a major impact on the outcome of the election even when they receive only a small fraction of the votes. In the 2000 US Presidential election, the 3 percent of votes going to Ralph Nader in the crucial state of Florida were enough to sway the election in favor of George W. Bush.

To assess the extent to which voters behave strategically, existing studies usually compare people's preferences and vote choices and count the number of voters who vote for a front-runner instead of their favorite. But voters' underlying preferences are difficult to observe, so these studies depend on assumptions regarding the mapping between voters' preferences and vote choices (e.g., Kawai and Watanabe, 2013; Spenkuch, 2015) or on the reliability of survey responses (e.g., Blais, Nadeau, Gidengil, and Nevitte, 2001; Blais, Young, and Turcotte, 2005; Hillygus, 2007).

This paper uses a different approach: instead of estimating preferences, we focus on vote choices. We take electoral results when two candidates are competing, and compare them to cases when three are competing. While the number and types of competing candidates is in general endogenous, French local and parliamentary elections, which are held under a two-round plurality voting rule, provide us with a chance for exogenous variation. In most districts, no single candidate obtains the majority of votes in the first round, and a second round takes place a week later. The top two candidates (the two who obtain the highest vote share in the first round) automatically qualify for the second round. Other candidates also qualify for the second round if they receive a number of votes in the first round higher than 12.5 percent of registered citizens. The candidate who obtains the highest vote share in the second round wins the election.

Our identification strategy exploits the discontinuity generated by the qualification rule for the second round. Using a regression discontinuity design (RDD), we compare second-round results in districts where the third candidate obtains a vote share just above or below the 12.5 percent threshold and, as a result, just passes or misses the qualification requirement. This strategy enables us to estimate the impact of electoral offer on voter behavior and, in particular, examine whether voters adjust expressively or strategically to the presence of a third candidate. The second-round races we study can be thought of as first-past-the-vote elections in which first-round results provided voters with a large amount of information on the chances of the remaining competitors. If anything, this information should stacks the odds in favor of coordination and strategic voting (Cox, 1997).

The threshold is defined as a fraction of registered citizens rather than actual votes: this makes it particularly hard to manipulate and results in a very diverse set of districts close to the threshold. The set includes competitive districts where the third candidate obtained a large vote share in the first round but turnout was relatively low, as well as districts where she obtained a lower vote share but turnout was high. This makes the external validity of our local average treatment effect estimates unusually high and enables us to compare treatment effect size across different settings. The number of elections we consider and the hundreds of districts at each election translate into a large number of observations and secure high statistical power while facilitating treatment effect heterogeneity analysis.

We first disentangle two types of voters who vote for the third candidate in the second round: "loyal" voters who would not have voted for any of the two other candidates if the third candidate were absent; and "switchers", who would have. We find that the presence of the third candidate in the second round increases voter turnout by 3.6 percentage points and reduces the share of blank and null votes by 3.6 percentage points, resulting in an overall increase of the share of people casting a ballot for any of the candidates by 7.2 percentage points (13.0 percent). In addition, the presence of the third candidate decreases the vote share of the top two candidates (expressed as a fraction of registered citizens) by 6.0 percentage points (10.9 percent). Based on these results, we estimate that loyal voters account for 27.3 to 54.5 percent of the voters who vote for the third candidate and switchers for 45.5 to 72.7 percent.

Our key results relate to the impact of switchers' behaviors on the outcome of the elections. We find that switchers' voting choice mainly impacts the vote share of the one candidate of the top two who is closest to the third candidate on the left-right ideological axis, and whom most switchers prefer to the other of top two. The impact remains equally strong and significant when we consider subsets of elections where first-round results indicate that the third candidate is very unlikely to be a "front-runner" (to rank first or second) in the second round. As a consequence, in 19.2 percent of the elections, the presence of the third candidate in the second round causes the victory of the candidate switchers dislike the most when their second-best choice would have won otherwise.[1] The candidate defeated due to switchers' behavior is not only their second-best, but also the Condorcet winner – the candidate who would win in a two-candidate race against each of the other candidates. In sum, switchers' voting choice is not only "costly" for them but also for a majority of voters. These results are difficult to explain through instrumental voting models. Instead, they suggest that a large fraction of voters base their voting decisions on expressive motives.

To characterize the voting decision of voters with a preference for the third candidate and explore the condition under which switchers' behavior causes the defeat of the Condorcet winner, we develop a model in which voters choose between three candidates. Supporters of the third candidate face the following tradeoff: voting for their second-best increases the probability that she wins against the candidate they dislike the most, but induces an ideological cost. We assume that voters are "group rule-utilitarian" as defined in Feddersen and Sandroni (2006a) and Coate and Conlin (2004): they define a cutoff such that voters with a sufficiently low cost of voting for their second-best vote for her and voters with an ideological cost above the threshold vote for their preferred candidate. We model three different types of voters differing in their motives and level of sophistication. *Expressive* (non-strategic) voters only value the expressive benefits of voting and always vote for their favorite candidate. *Strategic-naive* voters follow the cutoff rule but do not take into account the presence of expressive voters when designing it. Finally, *strategic-sophisticated* voters follow the rule and take into account the presence of expressive and strategic-naive voters when choosing the cutoff. The Condorcet winner loses the election when the fraction of sophisticated voters is too low to offset the behavior of expressive and naive voters. We show that our empirical results are consistent with the existence of the three types of voters, and we provide an estimation of their fractions of the electorate.

[1] This estimate is obtained by restricting the analysis to elections where the three candidates are from different political orientations, the third candidate is located either to the right or to the left of both top two candidates (since in these elections the second-best is the same for a large majority of switchers) and where first-round results indicate that the third candidate's chances of becoming a front-runner in the second round are low (so that the impact is not mechanically driven by the third candidate winning the election). More information in Sections 5.2 and 5.3.

1.1 Contribution to the literature

A large literature studies how voting rules shape electoral outcomes and, in turn, how they affect voter behavior. Social choice theory has shown that no electoral system or voting rule is uniformly best under all criteria (Arrow, 1951),[2] and that in any voting system, some voters have an incentive to misrepresent their true preferences in order to affect the outcome of the election (Gibbard, 1973; Satterthwaite, 1975). Building on this result, a normative literature has sought to identify voting rules which deliver outcomes that best represent voters' preferences by most resisting strategic manipulation (e.g., Brams and Fishburn, 2002; Laslier, 2009; Balinski and Laraki, 2011; Posner and Weyl, 2015).

Under existing rules, electing leaders that best correspond to voters' preferences may require that a sufficiently large fraction of them actually engage in such strategic manipulation and misrepresent their preferences when they cast their vote.[3] By contrast, if people vote according to their true preference, a minority candidate may be elected if several majority candidates are campaigning on close platforms. If voters fail to vote strategically, the plurality rule may fail to choose the Condorcet winner, when one exists, thus decreasing the representativeness of the electoral outcome (Nurmi, 1983; Myerson and Weber, 1993).

Instrumental voting models (e.g., Palfrey, 1989; Myerson and Weber, 1993; Cox, 1997; Fey, 1997), posit that voters care only about the winner of the election and thus choose whom to vote for after assessing the relative likelihood that each possible pair of candidates will be in contention for victory. As a result, in most cases, two candidates receive all the votes in equilibrium and the one preferred by the majority of voters wins.[4]

However, predictions alter substantially when we depart from the assumption that voters have common knowledge of candidates' platform and everyone's preferences (which is quite unrealistic in large elections). When instrumentally rational voters only have private information about candidates' support, more than two candidates may receive support in equilibrium (Myatt, 2007). Uncertainty on the level of candidates' support may actually generate an equilibrium where all voters vote for their favorite candidate, which may lead to the victory of the Condorcet loser (Bouton, Llorente-Saguer, and Castanheira, 2015).

A large empirical literature tests these models and examines whether voters behave strate-

[2] Arrow's impossibility theorem states that when there are more than two alternatives, there is no social welfare function that satisfies the Pareto property and the Independence of Irrelevant Alternatives and which is not a dictatorship.

[3] For instance, Demeze, Moyouwou, and Pongou (2016) find that in three-candidate plurality elections the possibility for voters to misrepresent their preferences benefits half to two-thirds of the population.

[4] This result applies to a number of settings. The model of Myerson and Weber (1993) applies to a wide range of single-winner electoral systems such as plurality rule, approval voting, and the Borda system. Cox (1994) extends the model to a multimember context. The case of dual ballot rule is studied in Bouton (2013) and Bouton and Gratton (2015).

gically or not, and how often voting for a non-top-two candidate causes suboptimal outcomes. Small-scale laboratory experiments have provided direct evidence of the existence of strategic behaviors (e.g., Forsythe, Myerson, Rietz, and Weber, 1993; Rietz, 2008; Van der Straeten, Laslier, Sauger, and Blais, 2010; Bouton, Llorente-Saguer, and Castanheira, 2015). Outside of the lab, Cox (1997)'s comprehensive study of strategic coordination across electoral systems identifies patterns consistent with strategic voting. Using RDD on population thresholds, Fujiwara (2011) and Eggers (2015) find that, in general, the top two candidates get more votes under simple plurality than under runoff or proportional elections, in line with Duverger's prediction (but see Bordignon, Nannicini, and Tabellini, 2016). However, the fact that different voting rules may also affect the number and types of competing candidates makes interpreting these findings difficult. Strategic voting often relies on accurately identifying which candidates are the front-runners, which may itself depend upon sufficient information. Consistent with Myatt (2007)'s predictions, Hall and Snyder (2015) find that higher levels of information in US primary elections decrease the number of votes and donations "wasted" on candidates unlikely to win. Anagol and Fujiwara (2016) show that second-place candidates are substantially more likely than close third-place candidates to run in, and win, a subsequent election, suggesting that candidates' rankings in previous elections are an important piece of information used by voters to coordinate their voting decisions.

To determine the actual proportion of strategic voters in the electorate, existing studies compare people's preferences and voting choices, with the hindrance that voters' true preferences are difficult to observe. A first strategy to recover preferences is to rely on surveys and trust people's self-reports (e.g., Niemi, Written, and Franklin, 1992; Alvarez and Nagler, 2000; Blais, Nadeau, Gidengil, and Nevitte, 2001; Blais, Young, and Turcotte, 2005; Hillygus, 2007; Kiewiet, 2013; Eggers and Vivyan, 2016). This method has generated relatively low estimates of the fraction of voters who vote for a candidate other than their preferred one, from 3 percent (Blais, Nadeau, Gidengil, and Nevitte, 2001) to 17 percent (Niemi, Written, and Franklin, 1992). A potential concern is that respondents may misreport their true preference and voting behaviors. For instance, overreporting voting for the winner (Wright, 1990, 1992; Atkeson, 1999; Campbell, 2010) would lead to overestimate strategic behavior. Alternatively, to avoid cognitive dissonance (Festinger, 1962), people may adjust their stated preference to their voting choice, which would lead to underestimating it. Finally, the shortage of survey data at the district level makes it difficult to measure the impact of strategic voting in elections with many districts (Herrmann, Munzert, and Selb, 2016).

Instead of individual surveys, a second strand of the literature relies on aggregate electoral results and studies strategic voting by imposing assumptions on the mapping between voters' preferences and vote choices. Kawai and Watanabe (2013) and Myatt and Fisher (2002) calibrate structural models to estimate the number of voters who did not vote for their preferred candidate and the impact of strategic voting on the number of seats won by a party, respectively. In the context of the German split-ticket voting system, Spenkuch (2014) compares votes cast for party lists under a proportional rule (for which voters have an incentive to follow their true preference) with votes cast for individual candidates under plurality rule (for which they have an incentive to be tactical), and reports that about one third of voters behave strategically. Spenkuch (2015) compares split-ticket voting in a particular district where a party could unusually gain seats by receiving fewer votes to districts where voters did not face such a reverse-incentive and concludes that at least 8 percent of that district's voters did not vote sincerely. Finally, Blais, Dolez, and Laurent (2017) and Kiss (2015) compare electoral outcomes in the first and second rounds of, respectively, French and Hungarian runoff elections, in which more than two candidates can run in the second round. They identify strategic voters as the ones who vote for a non-top-two candidate in the first round but for a front-runner in the second round, relying on the assumption that voters reveal their true preference in the first round of runoff elections (for a discussion of this assumption, see Piketty, 2000; Martinelli, 2002; Dolez and Laurent, 2010; Bouton and Gratton, 2015).

Instead of estimating voter preferences and comparing them with their actual choices, this paper focuses on vote choices only. We compare electoral outcomes when two versus three candidates are competing. This strategy allows us to make two important contributions to the literature. We demonstrate the existence of "switchers" and "loyal" voters and estimate the fraction of the electorate they represent, and we show that in about one in five elections, switchers voting for the third candidate causes the candidate they most dislike to beat the Condorcet winner.

To model the behavior of voters with a preference for the third candidate, we build on "group rule-utilitarian" models of turnout, which assume that voters base their voting decisions on a group rule rather than on pivot probabilities. In Feddersen and Sandroni (2006a), "ethical" voters define a cutoff such that those facing a lower voting cost turn out whereas the others abstain. We extend this framework to model vote choice instead of voter turnout, and study the tradeoff faced by supporters of the third candidate, who have to decide between voting for her or for their second-best choice.

Methodologically, we draw on other studies that exploit vote-share thresholds to estimate causal effects of interest, such as the incumbency effect (Lee, 2008), the effect of having a governor of the same party in office on a presidential candidate's chance of victory in his state (Erikson, Folke, and Snyder, 2015), or the impact of electing a woman on women's future participation in politics (Broockman, 2013).[5]

The remainder of the paper is organized as follows. We describe the data we use in Section 2 and our empirical framework in Section 3. Sections 4 and 5 present our main empirical results. Section 6 develops a model to rationalize our empirical findings. Section 7 concludes.

[5] See de la Cuesta and Imai (2016) for a more comprehensive list of recent studies exploiting vote-share thresholds.

2 Research setting

2.1 French parliamentary and local elections

Our sample includes parliamentary and local elections. Parliamentary elections elect the representatives of the French National Assembly, the lower house of French Parliament. France is divided into 577 constituencies, each of which elects a Member of Parliament every five years. Local elections determine the members of the departmental councils. France is divided into 101 départements, which have authority over education, social assistance, transportation, housing, culture, local development, and tourism.[6] Each département is further divided into small constituencies, the cantons, which elect members of the departmental councils for a length of six years. Until electoral reform in 2013, each canton elected one departmental council member; after that, each elected a ticket composed of a man and a woman. This new rule applied to the 2015 local election, which is included in our sample. We consider a ticket as a single candidate in our analysis, since the two candidates organize a common electoral campaign, run in the election under the same ticket, and get elected or defeated together. [7]

Any French citizen over 18 years old, registered on the voter rolls and not sentenced with ineligibility, can be a candidate at legislative elections. Candidates for departmental councils must also live in the département they mean to represent. Both elections are held under a two-round plurality voting rule. In order to win directly in the first round, a candidate needs to obtain a number of votes greater than 50 percent of the candidate votes and 25 percent of the registered citizens. In the vast majority of districts, no candidate wins in the first round and a second round takes place one week later. In the second round, the election is decided by simple plurality: the candidate who receives the largest vote share in the second round wins the election while the other candidates are left with nothing.

The two candidates who obtain the highest vote share in the first round automatically qualify for the second round. Other candidates qualify only if they obtain a first-round vote share higher than 12.5 percent of the registered citizens. This threshold does not have any other implication than determining which candidates qualify for the second round.

Our sample includes all parliamentary and local elections using the 12.5 percent qualification threshold: the eight parliamentary elections which took place since 1978 as well as the 2011 and

[6] Local government expenditures accounted for 20.6 percent of total government expenditures in 2011 (OECD, 2013).

[7] Before 2013, local elections took place every three years and, in each département, only half of the cantons were electing their council member in a given election. After the 2013 reform, all cantons participated in elections held every six years. The reform further reduced the number of cantons from 4035 to 2054, to leave the total number of council members roughly unchanged. All French territories participate in local elections, except for Paris and Lyon (where the departmental council is elected during municipal elections) and some French territories overseas.

2015 local elections.[8] The fact that this threshold is at a relatively high percentage means at most three candidates qualify for the second round in all but a handful of districts, which is ideal for our study design. In the elections we consider, the third candidate received more than 12.5 percent of the votes in 1,215 districts (16.7 percent of our sample).[9]

All candidates qualified for the second round can decide to drop out of the race between the first and second rounds. Dropouts result from strategic considerations and local and national alliances. For instance, left-wing parties commonly ask their candidates to drop out if they ranked lower than another left candidate. In our sample, when the third candidate qualifies for the second, she decides to drop out in around 50 percent of the cases.

2.2 Data

Our sample includes a total of 7,257 observations: 3,458 (47.7 percent) from local elections and 3,799 (52.3 percent) from parliamentary elections. Official results of local and parliamentary elections were digitalized from printed booklets for the 1978, 1981 and 1988 parliamentary elections and obtained from the French Ministry of the Interior for all others. We exclude districts where only one round took place or with fewer than three candidates in the first round.[10] Table 1 gives the breakdown of the sample data by election type and year.

Table 2 presents some descriptive statistics on our sample. In the average district, 7.78 candidates competed in the first round and turnout was 58.2 percent. On average 56.2 percent of the registered citizens cast a valid vote for one of the candidates. Valid voting entails inserting a ballot pre-printed with the candidate's name in an envelope and putting this envelope in the ballot box. We term these "candidate votes". The difference between turnout and candidate votes arises from voters who cast a blank vote (by putting an empty envelope in the ballot box) or a null vote (by writing something on the ballot or inserting multiple ballots in the envelope). Turnout in the second round was slightly higher than in the first (58.8 percent on average) but the fraction of candidate votes was slightly lower (55.4 percent on average), due to an increased share of blank and null votes. The average number of candidates in the second round was 2.04 and there were three

[8] Each of the 10 elections we consider took place at a different date. Moreover, the local and parliamentary elections we study were never held at the same date as other types of elections such as presidential, mayoral or regional ones.

[9] In local elections, the required vote share was 10 percent of the registered citizens until 2010, when the threshold was increased to 12.5 percent. The lower threshold resulted in more than three candidates qualifying in a large number of constituencies. One exception was made after the change: in the 2011 local elections, the threshold remained at 10 percent in the 9 cantons belonging to Mayotte département (0.6 percent of the 2011 observations). The threshold in parliamentary elections was increased from 5 to 10 percent in 1966 and from 10 to 12.5 percent in 1976. The stated goal of these reforms was to reduce the number of candidates competing in the second round, so that the winning candidate would obtain a larger vote share, increasing her legitimacy (Sénat, 1999).

[10] We also exclude three elections where the second and third candidates in the first round obtained exactly the same number of votes. Here, the 12.5 percent threshold rule did not apply. Both candidates were allowed to move on to the second round, regardless of the number of votes they had obtained in the first.

candidates in the second round in 453 districts (6.2 percent of the sample).

Table 1: Elections in the sample

	Year	Number of observations
Parliamentary elections	1978	423
	1981	333
	1988	455
	1993	496
	1997	565
	2002	519
	2007	467
	2012	541
	Total	3,799
Local elections	2011	1,561
	2015	1,897
	Total	3,458
Total		7,257

Table 2: Summary statistics

	Mean	Sd	Min	Max	Obs.
Panel A. 1st round					
Registered voters	45,964	30,882	883	189,384	7,257
Turnout	0.582	0.124	0.134	0.908	7,257
Candidate votes	0.562	0.122	0.132	0.890	7,257
Blank and Null votes	0.019	0.011	0.001	0.094	7,257
Number of candidates	7.78	4.08	3	29	7,257
Panel B. 2nd round					
Turnout	0.588	0.131	0.128	0.928	7,257
Candidate votes	0.554	0.136	0.124	0.907	7,257
Blank and Null votes	0.035	0.022	0.002	0.278	7,257
Number of candidates	2.04	0.28	1	3	7,257

We further use the political label attributed to the candidates by the French Ministry of the Interior to allocate them to six political orientations: far-right, right, center, left, far-left, and other. In the elections we consider, the candidate who ranked third in the first round was on the right in 19.6 percent of the districts, on the left in 37.1 percent, on the far-right in 36.7 percent, and from another political orientation in the remaining 6.6 percent of the districts.[11]

[11] The Ministry of the Interior attributes political labels based on several indicators: candidates' self-reported polit-

2.3 Vote share of the third candidate

In most cases, candidates who came in third in the first round should be expected to have lower chances of winning the second round or finishing second than the candidates who ranked first and second in the first round, and voters casting a ballot for the third candidate should expect their vote to be "wasted". Strikingly, however, third candidates who qualify and compete in the second round garner more votes than in the first round and get a remarkably high vote share on average: 25.6 percent of the candidate votes, against 33.2 percent for candidates who ranked second and 41.2 percent for candidates who ranked first. This result is not driven by any particular configuration: the vote share obtained by the third candidate in the second round is large when she is on the left (30.6 percent), the right (28.9 percent), and the far-right (21.5 percent).

Voters who vote for the third candidate when she is present may either vote for the top two candidates or instead abstain or vote blank or null when she is absent. We use a regression discontinuity design (RDD) framework to disentangle these two types of voters, which we call "switchers" and "loyal voters" respectively.

3 Empirical strategy

3.1 Identification

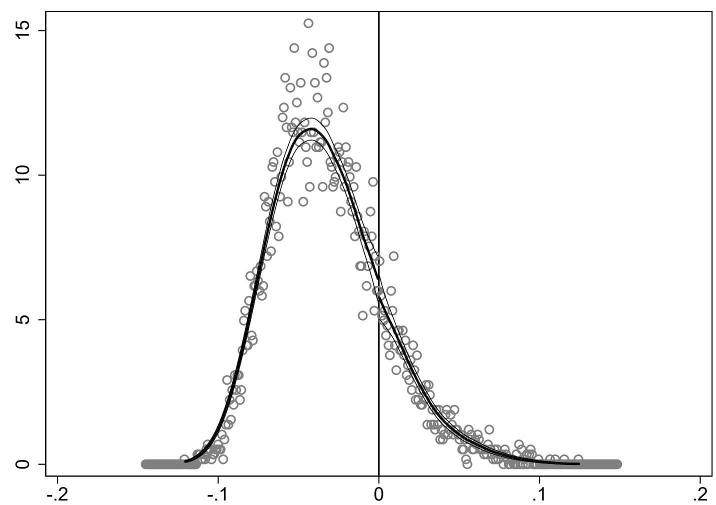
We exploit the 12.5 percent vote share threshold, which determines whether the third-highest-ranking candidate qualifies for the second round, to estimate the impact of her presence on voters' behavior.

Following Rubin (1974)'s potential-outcome framework for causal inference, we define $Y_i(0)$ (respectively $Y_i(1)$) as the potential electoral outcome in district i if the third-highest-ranking candidate in the first round is absent (resp. present) in the second round. Next, we define the running variable X_i as the qualifying margin of this candidate in the first round: X_i is the difference between her vote share, expressed as a fraction of the number of registered citizens, and the 12.5 percent threshold. The third candidate is allowed to compete in the second round if and only if X_i is positive. The assignment variable D_i thus takes two values: $D_i = 0$ if $X_i < 0$ and $D_i = 1$ if $X_i \geq 0$. Finally, T_i indicates the treatment status of district i: $T_i = 1$ if the third candidate is present in the second round, and $T_i = 0$ if she is absent.

ical affiliation, party endorsement, past candidacies, public declarations, local press, etc. Candidates can ask to revise their political label before the final list of first-round candidates is released. We mapped political labels into the six political orientations, mainly based on the allocation chosen by Laurent de Boissieu in his blog "France Politique": http://www.france-politique.fr/. We also used public declarations made by the candidates. Appendix F shows the mapping between labels and political orientations for each election.

The identification assumption is that the distribution of potential confounders changes continuously around the 12.5 percent vote share threshold, so that the only discrete change occurring at this threshold is the shift in treatment status. The growing use of RDD exploiting vote share thresholds spawned a methodological debate on the validity of this source of identification (Caughey and Sekhon, 2011; Hainmueller, Hall, and Snyder, 2015; Eggers, Fowler, Hainmueller, Hall, and Snyder, 2015). As emphasized by de la Cuesta and Imai (2016), sorting of candidates across the threshold only threatens the validity of the RDD if it occurs at the cutoff, with potential losers pushed just above the threshold or potential winners pushed just below. This is unlikely in general, as it requires the ability to predict election outcomes and deploy campaign resources with extreme accuracy. Even if candidates accurately predicted the results of the first round and exerted extra effort to fall above the threshold (or prevent others from doing so) when they expect to be near it, weather conditions on Election Day and other unpredictable events would still make the outcome of the election uncertain.

Figure 1: McCrary test of the density of the running variable



Notes: This Figure tests for a jump in the density of the running variable (the qualifying margin of the third-highest-ranking candidate in the first round) at the threshold. The solid line represents the density of the running variable. Thin lines represent the confidence intervals.

In our setting, manipulation of the threshold is perhaps even more unlikely than in other RDDs using vote share thresholds. First, candidates have very limited information available about voters' intentions in the first round of French parliamentary or local races. District-level polls are very rare during parliamentary elections, and nonexistent during local elections, due to small district size and limited campaign funding. In addition, the threshold is defined as the share of registered citizens. Manipulating it would thus require accurately predicting and manipulating *both* the fraction of registered citizens turning out and the share of candidate votes going to the third candidate.

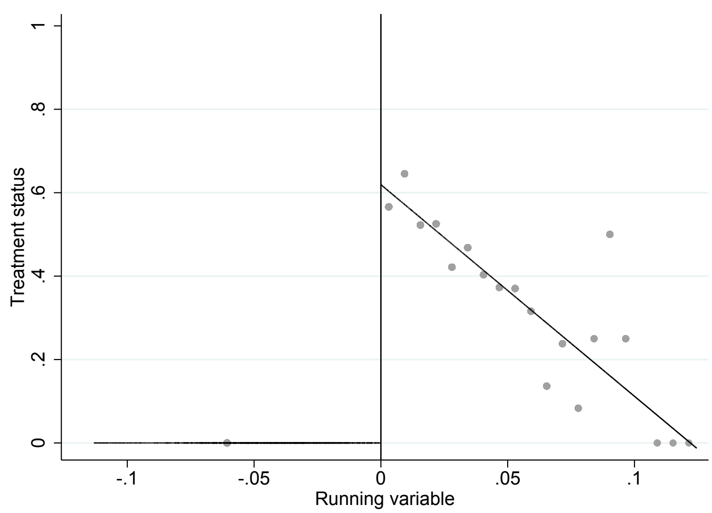
To bring empirical support for the identification assumption, we check if there is a jump in the density of the running variable at the threshold (McCrary, 2008). As Figure 1 shows, we do not observe any. In Section 3.3 we run placebo tests on baseline variables and provide additional evidence supporting the identification assumption.

3.2 Evaluation framework

Candidates are allowed to drop out of the race between the two rounds of the election. As a result of the dropouts of the third candidate, the treatment status is not a deterministic function of the running variable, making our regression discontinuity design fuzzy, as shown in Figure 2. We provide descriptive evidence on the factors affecting the third candidate's decision to drop out in Appendix B.

Following Imbens and Lemieux (2008) and Calonico, Cattaneo, and Tiriunik (2014), our main specification uses a non-parametric approach, which amounts to fitting two linear regressions on districts respectively close to the left, and close to the right of the threshold.

Figure 2: Imperfect Compliance



Notes: Dots represent the local averages of the treatment status (y-axis). Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a linear fit.

To avoid selection bias and deal with imperfect compliance, we instrument the treatment status T with the assignment variable D. We estimate the two following equations:

1st stage:

$$T_i = \alpha_0 + \gamma D_i + \delta_1 X_i + \delta_2 X_i D_i + \varepsilon_i \quad (1)$$

2nd stage:

$$Y_i = \alpha_1 + \tau T_i + \beta_1 X_i + \beta_2 X_i T_i + \mu_i \quad (2)$$

Where T_i in equation [2] is replaced by the predicted value obtained from equation [1]. The coefficients of interest are $\hat{\gamma}$ and $\hat{\tau}$. $\hat{\gamma}$ gives the probability, at the threshold, that a district receives the treatment (the third candidate *is present* in the second round) given that it has been assigned to it (the third candidate *qualified* for the second round). $\hat{\tau}$ estimates the treatment impact on outcome Y at the threshold in the districts where the third candidate qualified and is present in the second round. We test the robustness of our results to a quadratic specification, including X_i^2 and its interaction with D_i as regressors in equation [1] and X_i^2 and its interaction with T_i in equation [2].

Our estimation procedure follows Calonico, Cattaneo, and Tiriunik (2014), which provides robust confidence interval estimators. Our preferred specification uses the MSERD bandwidths developed by Calonico, Cattaneo, Farrell, and Titiunik (2016), which reduce potential bias the most. We also test the robustness of the main results to using the optimal bandwidths computed according to Imbens and Kalyanaraman (2012).

The bandwidths used for the estimations are data-driven and therefore vary depending on the outcomes we consider. Instead, when we provide descriptive statistics on districts "close to the threshold", we will always consider districts in which the vote share of the third candidate was within exactly 2 percentage points from the threshold. Thanks to our large sample size (7,257 districts), the number of districts close to the threshold is higher than 1,800.

Table 3 provides the formal estimates for the first stage. Columns (1) and (2) show the results obtained under the MSRED and IK optimal bandwidths, using a local linear regression. Columns (3) and (4) present the results using a quadratic specification. All four estimates are significant at the 1 percent level. In our preferred specification (column 1), we find that the probability to receive the treatment jumps from 0 to approximately 55.2 percent at the threshold.

Table 3: First stage

	(1)	(2)	(3)	(4)
Outcome	Treatment status			
Assignment status	0.552***	0.611***	0.509***	0.566***
	(0.042)	(0.030)	(0.051)	(0.043)
Observations	1,541	3,579	2,142	3,579
Polynomial order	1	1	2	2
Bandwidth	0.017	0.038	0.023	0.038
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.00	0.00	0.00	0.00

Notes: Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. The outcome is a dummy equal to 1 if the third candidate is present in the second round. The dependent variable is a dummy equal to 1 if the third candidate gathered more than 12.5 percent of the registered votes in the first round. The polynomial order is 1 in columns 1 and 2 and 2 in columns 3 and 4. The bandwidths are derived under the MSERD (columns 1 and 3) and IK (columns 2 and 4) procedures.

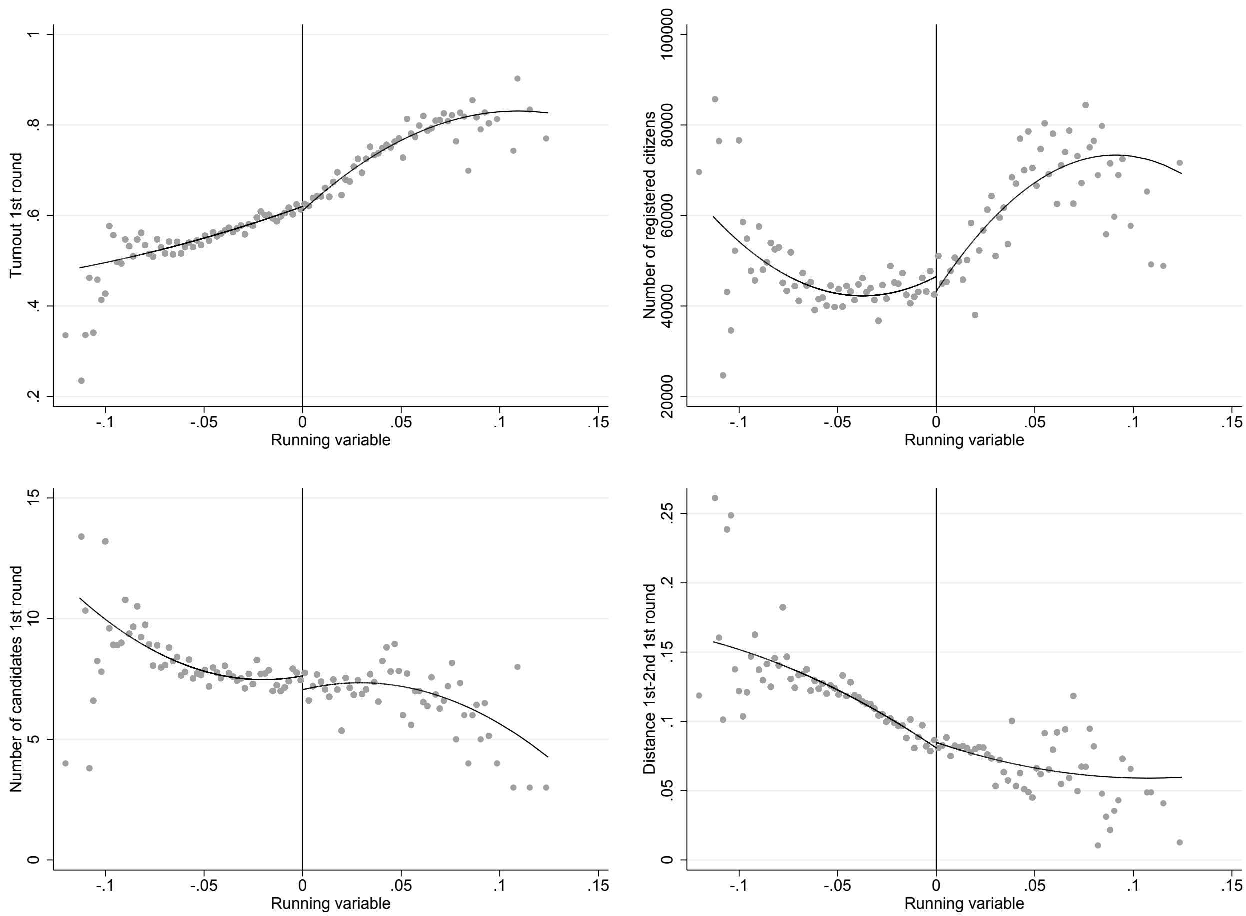
Estimating the impact of the presence of the third candidate on electoral outcomes amounts to comparing voters' behaviors in elections with two versus three candidates modulo two noticeable exceptions. First, in 5.5 percent of the elections near the discontinuity, the candidate ranked second in the first round dropped from the race. Second, in 1.2 percent of the elections, the candidate ranked fourth in the first round also qualified for the second round, and she decided to run in three instances. Appendix D discusses these particular cases at greater length and shows that our results are not driven by them.

3.3 Placebo tests

We perform a series of placebo tests which examine whether there is a discontinuity in any of the following first-round variables at the cutoff: voter turnout, number of registered voters, number of candidates, and closeness (defined as the difference between the vote shares obtained by the two top candidates).

As shown in Figure 3, there is no significant jump at the cutoff for any of these variables. The formal estimation confirms the absence of treatment effect. Columns 1 through 4 of Table A1 in the Appendix present the results obtained for these four outcomes under our preferred specification. None of the estimates is statistically significant at the standard levels. Hence, we cannot reject the null hypothesis that the treatment has no effect on these baseline variables.

Figure 3: Placebo tests on baseline variables



Notes: Dots represent the local averages of the baseline variable (y-axis). Averages are calculated within 0.2 percentage point wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

In addition, we conduct the general following test for imbalance. We regress the assignment variable D on a set of first-round variables including the four aforementioned variables as well as share of candidate votes, vote share of each of the top three candidates, political label and orientation of the three candidates, number of candidates from the left, right, far-right, far-left and center. We then use the coefficients from this regression to predict assignment status, and test whether it jumps at the threshold. As shown in Figure A1, the assignment status predicted by baseline variables increases continuously as a function of the running variable and does not show any discontinuity at the threshold. This suggests that there is no systematic discontinuity in the preexisting observable districts' characteristics at the threshold. The formal estimate in Column 5 of Table A1 confirms this result: the coefficient is small (1.5 percentage points) and non-significant.

If not otherwise specified, in the rest of the analysis, all outcomes, including vote shares, use the number of registered voters as the denominator instead of the number of citizens who vote, as the latter may be affected by the treatment while the former remains unchanged between the two rounds. In addition, our outcomes are defined as a simple difference between the second and first rounds. Using simple differences helps account for a large share of the latent variance of our outcomes.

4 "Loyal voters" and "switchers"

To test whether the third candidate obtains her votes by attracting new voters to the polls ("loyal voters") or by stealing voters away from the two other candidates ("switchers"), we first estimate the impact of the presence of the third candidate on participation and on the votes going to the top two candidates.

4.1 Impact on participation and candidate votes

We consider three outcomes related to participation: turnout, the share of null and blank votes, and the share of candidate votes. We begin with a graphical analysis before presenting formal estimates of treatment effects.

The graphs in Figure 4 depict the impact of the presence of the third candidate on our three outcomes. Each dot represents the average value of the outcome within a given bin of the running variable. To facilitate visualization, a quadratic polynomial is fitted on each side of the 12.5 percent threshold. The graphical evidence shows a clear discontinuity at the cutoff for each outcome: the presence of the third candidate has a large and positive impact on the share of registered citizens who vote and on the share of citizens who vote for one of the competing candidates rather than casting a blank or a null vote.

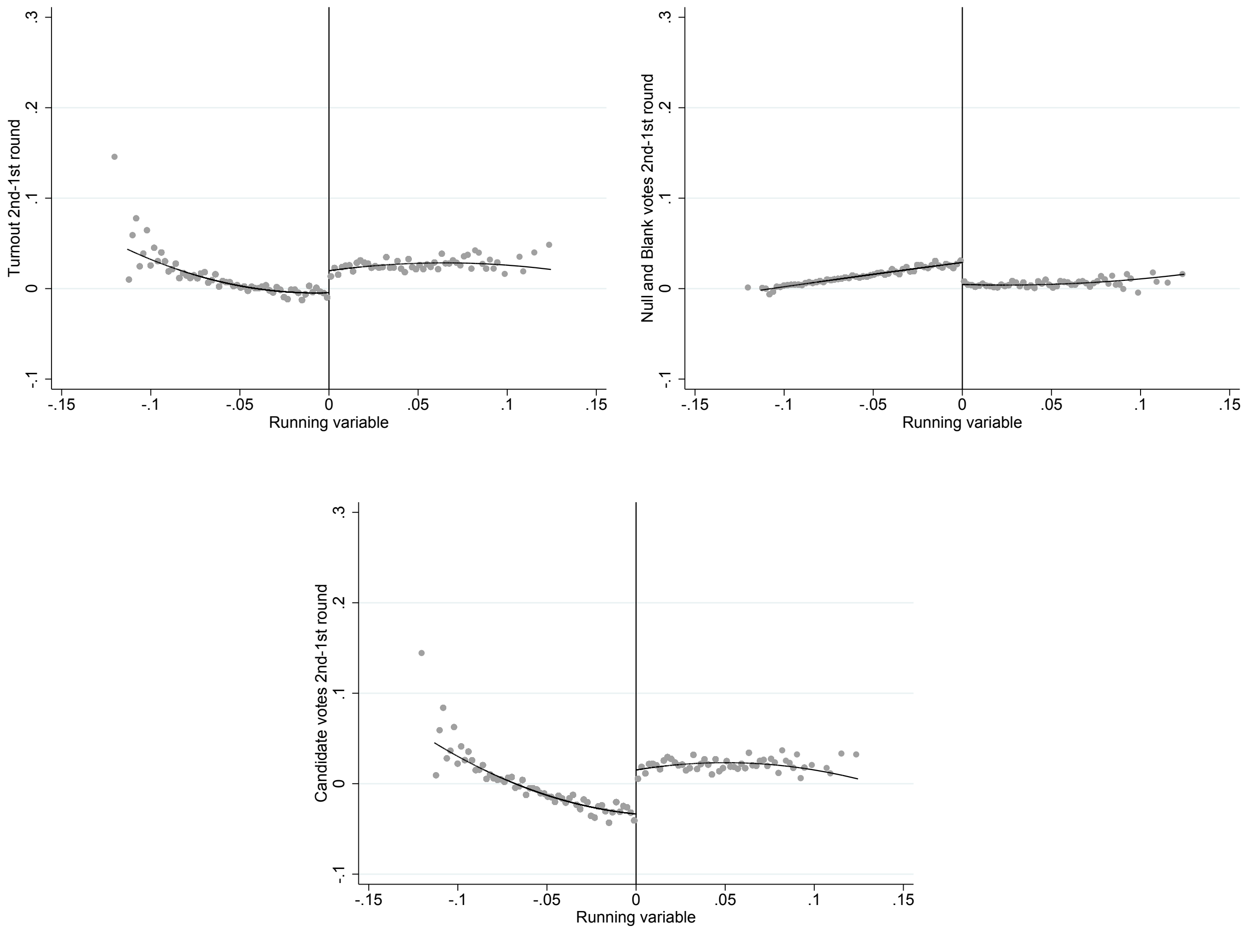
Table 4 provides the formal estimates of the impacts using our preferred specification. On average, the presence of the third candidate in the second round increases turnout by 3.6 percentage points (6.0 percent), reduces the share of null and blank votes by 3.6 percentage points (75.0 percent),[12] and increases the share of candidate votes by 7.2 percentage points (13.0 percent). All effects are significant at the 1 percent level. As should be expected, the impact on the share of candidate votes corresponds to the sum of the absolute value of the impacts on the turnout rate and the share of blank and null votes.

To probe the robustness of the results to specification and bandwidth choices, Table 5 estimates

[12] In the 2015 local elections, the only ones in which blank and null votes were counted separately, the impact on both outcomes was negative (Figure A2 and Table A2 in Appendix).

the treatment effect on the share of candidate votes using four different specifications. Columns (1) and (2) show the results obtained under the MSERD and IK optimal bandwidths, using a local linear regression. Columns (3) and (4) use a quadratic specification. The estimates obtained using these different specifications are all significant at the 1 percent level and very close in magnitude.

Figure 4: Impact on participation and candidate votes



Notes as in Figure 3.

Table 4: Impact on participation and candidate votes

Outcome	(1) Turnout 2nd-1st	(2) Null and Blank votes 2nd-1st	(3) Candidate votes 2nd-1st
3rd present	0.036***	-0.036***	0.072***
	(0.007)	(0.004)	(0.011)
Observations	2,142	2,454	2,103
Polynomial order	1	1	1
Bandwidth	0.023	0.027	0.023
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.600	0.048	0.552

Notes: Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. Each outcome uses the number of registered voters as the denominator and is defined as a simple difference between the second and first rounds. The variable of interest (the presence of the third candidate in the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate in the first round was higher than the threshold). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1 and the optimal bandwidths are derived under the MSERD procedure.

Table 5: Impact on candidate votes

	(1)	(2)	(3)	(4)
Outcome	Candidate votes 2nd-1st rounds			
3rd present	0.072***	0.073***	0.068***	0.071***
	(0.011)	(0.008)	(0.012)	(0.013)
Observations	2,103	3,345	3,985	3,345
Polynomial order	1	1	2	2
Bandwidth	0.023	0.036	0.042	0.036
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.552	0.542	0.539	0.542

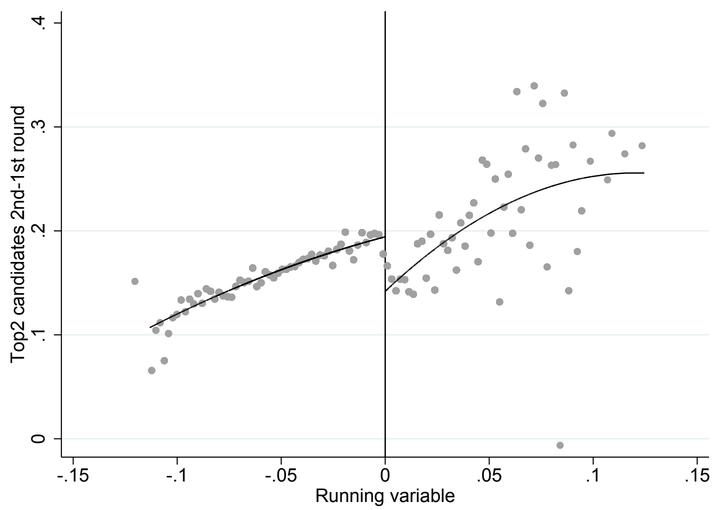
Notes: Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. The outcome uses the number of registered voters as the denominator and is defined as a simple difference between the second and first rounds. The variable of interest (the presence of a third candidate in the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). The polynomial order is 1 in columns 1 and 2 and 2 in columns 3 and 4. The bandwidths are derived under the MSERD (columns 1 and 3) and IK (columns 2 and 4) procedures.

4.2 Impact on votes going to the top two candidates

We now estimate the effect of the presence of the third candidate on the second round vote share of the candidates who placed first and second in the first round. If vote shares were defined using the number of candidate votes as denominator, the total vote share of the top two candidates in the second round would decrease by exactly the fraction of votes going to the third candidate when she is running. Instead, as for all our outcomes, we define vote shares using the number of registered citizens as denominator. As a result, the presence of the third candidate has no mechanical effect: it can either decrease, increase or leave the vote share of the two other candidates unchanged.

Figure 5 plots the vote share of the top two candidates against the running variable. The quadratic polynomial fit indicates a large downward jump at the cutoff.

Figure 5: Impact on votes going to the top two candidates



Notes as in Figure 3.

Consistent with the graphical analysis, the estimates reported in Table 6 indicate a sizable and significant negative impact of the treatment on the vote share of the top two candidates in the second round. Our preferred specification (column 1) shows that the presence of the third candidate decreases the vote share of the top two candidates by 6.0 percentage points (10.9 percent) on average, an effect significant at the 1 percent level. The magnitude and significance of the effect is comparable in other specifications.

Table 6: Impact on votes going to the top two candidates

	(1)	(2)	(3)	(4)
Outcome	Vote share top 2 2nd-1st round			
3rd present	-0.060***	-0.078***	-0.053***	-0.065***
	(0.016)	(0.009)	(0.019)	(0.014)
Observations	1,757	5,108	3,145	5,108
Polynomial order	1	1	2	2
Bandwidth	0.019	0.054	0.034	0.054
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.550	0.534	0.543	0.534

Notes as in Table 5.

We further estimate the impact on the vote share of the first and second candidates separately, and find that both decrease by a similar magnitude when the third candidate is present (Figure A3 and Table A3 in Appendix).

4.3 Interpretation of the results

When the third candidate is present, more citizens vote for a candidate (rather than abstaining or voting blank or null), and fewer voters vote for the first round's top two candidates. In addition to voters' response, the impacts on participation and vote shares could also reflect changes in candidates' strategies to adjust to the presence of an additional competitor. In our setting, however, the fact that there is only one week between the two rounds leaves little time for the top two candidates to do so. While we lack data on candidates' precise political platforms, we collected data on their campaign expenditures for the 2011 and 2015 local elections and for the 1993, 1997, 2002, 2007, and 2012 parliamentary elections (collectively accounting for 77.8 percent of our sample).[13] As shown in Appendix C, we find no impact of the presence of the third candidate on the top two candidates' campaign expenditures or contributions, suggesting that they do not intensify their political campaign when the third candidate is present.

We now use our estimates to compute bounds on the fractions of the two types of voters who vote for the third candidate in the second round: loyal voters – voters who would not vote for any candidate if the third candidate was absent – and switchers – who would vote for one of the top two candidates absent the third.

[13] All data come from the French National Commission on Campaign Accounts and Political Financing (CNCCFP). Data on campaign expenditures for the 1993, 1997, and 2002 parliamentary elections were collected by Abel François and his co-authors for their studies on the impact of electoral expenditures on turnout (Fauvelle-Aymar and François, 2005) and electoral results (Foucault and François, 2005).

At the threshold, the share of registered citizens who vote for the third candidate in the second round, which we use as the denominator, is equal by construction to the impact on candidate votes minus the impact on the vote share of the top two candidates: $7.2-(-6.0) = 13.2$ percent. The impact of the presence of the third candidate on the fraction of null and blank votes (- 3.6 percentage points) is unlikely to be driven by voters turning to the top two candidates. Indeed, voting for them is a choice available regardless of the presence of the third candidate. Hence, this impact provides a lower bound of the fraction of loyal voters among those who vote for the third candidate: $3.6/13.2 = 27.3$ percent.

The impact on turnout may similarly be driven in part by loyal voters who abstain when the choice is limited to the top two candidates but vote when the third candidate is present. Yet, it may also come from an increase in the number of citizens voting for the top two candidates. In particular, as shown in Figure A4 in Appendix, the presence of the third candidate in the second round reduces the winning margin. But more contested elections may drive additional supporters of the top two candidates to the polls.[14] Since the impact on the share of candidate votes (7.2 percentage points) sums the impact on the share of blank and null votes and the impact on turnout, it may itself be driven both by loyal voters supporting the third candidate and by an increased number of votes for the top two candidates. Hence, it provides an upper bound of the fraction of loyal voters: $7.2/13.2 = 54.5$ percent. In sum, a non-negligible fraction of voters prefer not to vote for any candidate if the third candidate is not in the race, suggesting that they only care about voting for their favorite candidate (Davis, Hinich, and Ordeshook, 1970; Cox, 1997; Fiva and Smith, 2017), or that they are strictly indifferent between the top two candidates.

Still, the majority of votes going to the third candidate come from switchers: with loyal voters representing between 27.3 percent and 54.5 percent of those who vote for the third candidate, switchers must account for the remaining 45.5 to 72.7 percent. These voters seem to care about who wins between the top two (since they vote for one of them when the third candidate is not running) and yet they vote for the third candidate when she is present. Section 5 measures the impact of their behavior on the outcome of the election.

[14]The winning margin is computed as the difference between the share of candidate votes obtained by the winner and by the candidate who came in second. Table A4 in the Appendix presents the formal estimate. Using our preferred specification, we find an impact of -5.8 percentage points on average.

5 Impact of switchers' behavior on electoral outcomes

Using sub-sample analysis, we show that switchers' behavior mainly impacts the vote share of the top-two candidate they prefer, that they make this choice even when the third candidate's chances are remote, and that their behavior has important consequences on the results of many elections.

5.1 Impact depending on partisan orientation

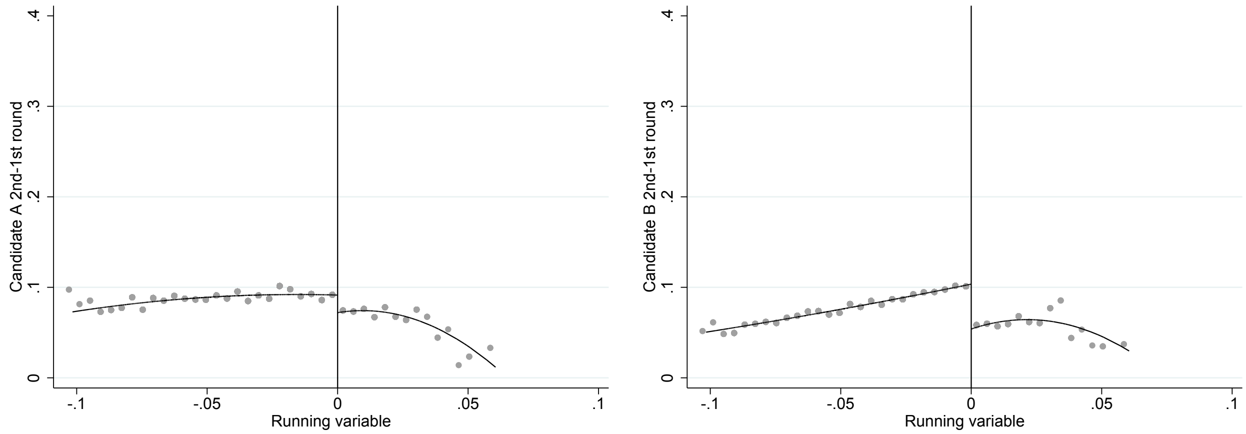
We first examine the extent to which switchers' behavior affects the relative vote share of the top two candidates. Insofar as people make their vote choice based on ideological preferences, the impact of the presence of the third candidate on the vote shares of the top two candidates should depend on their respective ideological positions. To test this hypothesis, we focus on elections in which the political orientations of the three candidates are well-identified and differ from each other, so that they can be ranked on the left-right axis. The resulting sample accounts for 75.7 percent of the districts where the third candidate qualifies and runs in the second round, near the discontinuity. We call A the candidate most to the left, C the candidate most to the right, and B the candidate located between A and C. We study three different settings, each characterized by the ideological position of the third candidate – on the left, right, or in the middle.[15]

In the first setting, the candidate who came in third in the first round is C. She is closer ideologically to B than to A. If C were not present in the second round, we would expect most of the switchers – all of them, if they had single-peaked preferences on the left-right axis – to vote for B, whom we thus call their second-best. Accordingly, we expect B to lose more voters than A from the presence of C. And indeed, as shown in Figure 6, the impact on the vote share of B is much larger than on A.

The regression results are reported in Table 7: while the vote share of candidate B decreases by 4.9 percentage points on average, an estimate significant at the 1 percent level, the vote share of candidate A decreases by only 1.6 percentage points, significant at the 5 percent level.

[15] In 97.3 percent of the elections corresponding to the first setting, the third candidate is on the far-right (C), one of the top two candidates is on the right (B) and the other is on the left (A). In 94.6 percent of the elections corresponding to the second setting, the third candidate is on the left (A), one of the top two candidates is on the right (B) and the other is on the far-right (C). In 62.7 percent of the elections corresponding to the third setting, the third candidate is on the right (B), one of the top two candidates is on the left (A) and the other is on the far-right (C). In 36.4 percent, the third candidate is on the center (B), one of the top two candidates is on the left (A) and the other is on the right (C).

Figure 6: Impact on candidates A and B when the third candidate is C



Notes: Averages are calculated within 0.4 percentage point wide bins of the running variable (x-axis). The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located to the right of both the first and the second candidates. Other notes as in Figure 3.

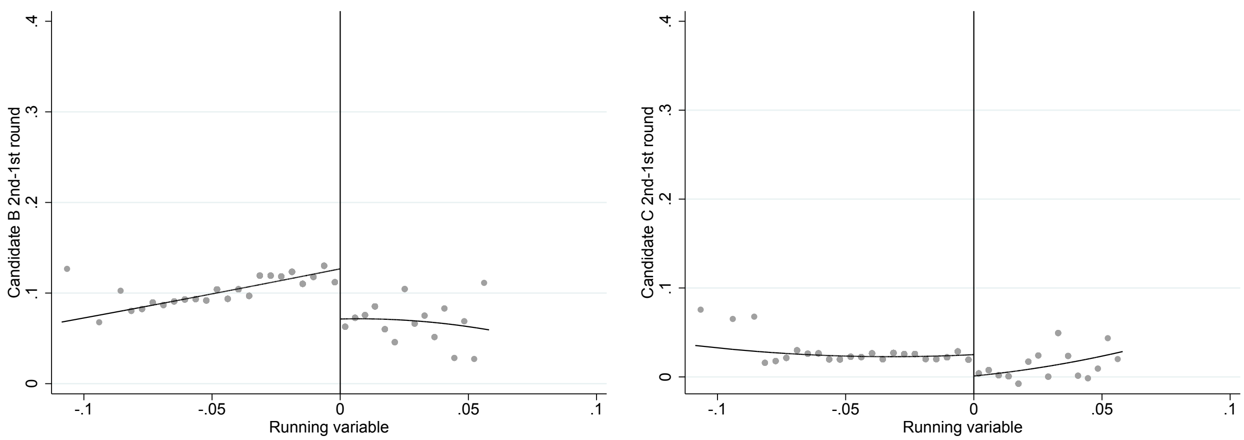
Table 7: Impact on candidates A and B when the third candidate is C

Outcome variable	(1) Top two cand. 2nd-1st round	(2) Cand. A 2nd-1st round	(3) Cand. B 2nd-1st round
3rd present	-0.065***	-0.016**	-0.049***
	(0.012)	(0.007)	(0.006)
Observations	418	431	403
Polynomial order	1	1	1
Bandwidth	0.015	0.015	0.014
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.548	0.264	0.285

Notes: The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located to the right of both the first and the second candidates. Other notes as in Table 4.

In the second setting, the candidate who arrived third in the first round is A. She is closest to B, making B the switchers' second-best again. As expected, we find that the impact of the presence of A on the vote share of B is larger than on C (Figure 7). As shown in Table 8, while the vote share of candidate B decreases by 7.5 percentage points on average, significant at the 1 percent level, the vote share of candidate C decreases by only 2.0 percentage points, significant at the 5 percent level.

Figure 7: Impact on candidates B and C when the third candidate is A



Notes: Averages are calculated within 0.4 percentage point wide bins of the running variable (x-axis). The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located to the left of both the first and the second candidates. Other notes as in Figure 3.

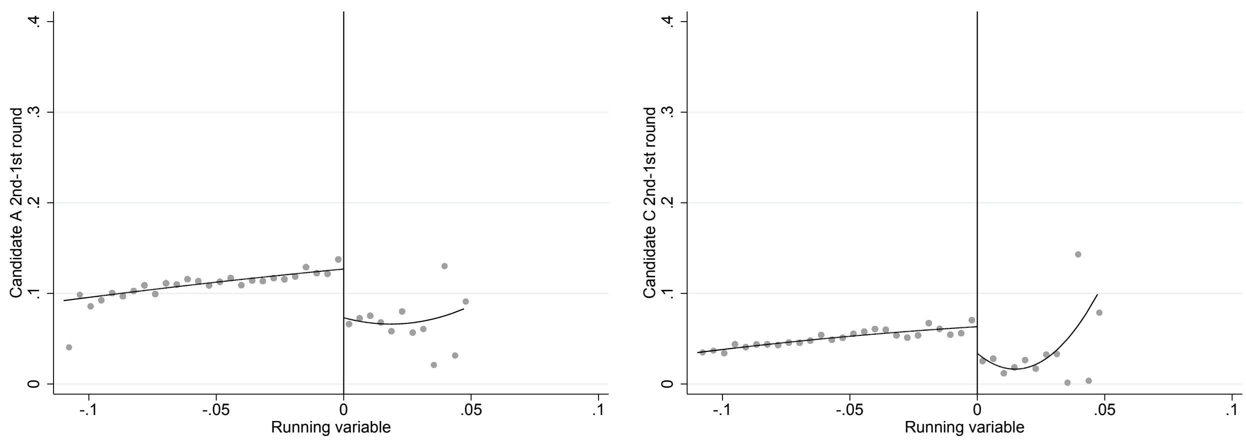
Table 8: Impact on candidates B and C when the third candidate is A

Outcome variable	(1) Top two cand. 2nd-1st round	(2) Cand. B 2nd-1st round	(3) Cand. C 2nd-1st round
3rd present	-0.095***	-0.075***	-0.020**
	(0.014)	(0.016)	(0.008)
Observations	230	135	148
Polynomial order	1	1	1
Bandwidth	0.018	0.011	0.012
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.466	0.305	0.173

Notes: The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located to the left of both the first and the second candidates. Other notes as in Table 4.

Finally, in the third setting, the candidate who arrived third in the first round is B. Since B is located between A and C, we expect sizable numbers of switchers to vote both for A and C in the second round, when B is absent: there is no clear second-best chosen by a large majority of switchers. Accordingly, both A and C should lose a large number of votes from the presence of B. And indeed, we find that B attracts a sizable fraction of voters from both candidates (Figure 8).

Figure 8: Impact on candidates A and C when the third candidate is B



Notes: Averages are calculated within 0.4 percentage point wide bins of the running variable (x-axis). The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located between the first and the second candidates. Other notes as in 3.

Table 9: Impact on candidates A and C when the third candidate is B

Outcome variable	(1) Top two cand. 2nd-1st round	(2) Cand. A 2nd-1st round	(3) Cand. C 2nd-1st round
3rd present	-0.151***	-0.091***	-0.054**
	(0.023)	(0.015)	(0.024)
Observations	141	189	137
Polynomial order	1	1	1
Bandwidth	0.014	0.019	0.014
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.510	0.290	0.218

Notes: The sample includes the elections where the top three candidates are from distinct political orientations and where the third candidate is located between the first and the second candidates. Other notes as in Table 4.

As shown in Table 9, the vote share of candidate A decreases by 9.1 percentage points on average, an estimate significant at the 1 percent level, and the vote share of candidate C decreases by 5.4 percentage points, significant at the 5 percent level.

In sum, the vote shares of the top two candidates are affected in proportion to their ideological proximity with the third candidate. In particular, in the first and second settings, the candidate closer ideologically to the third loses much more from her presence. This suggests that in these two settings, absent the third, a large majority of switchers vote for the same candidate, and that they significantly hurt this second-best by voting for the third candidate, when she is present.

These results also rule out the interpretation that switchers are mainly noisy voters, randomly splitting their votes among competing candidates: such noisy voting should affect the top two candidates equally, which is not what we observe. Finally, the results bring empirical support for modeling assumptions used in Section 6: one-dimensional representation of the political spectrum and single-peaked voter preferences.[16]

5.2 Impact depending on the strength of the third candidate

The rest of this section focuses on the first and second settings, in which a large majority of switchers rank the top two candidates in the same order of preference, so that their behavior reduces the relative vote share of their second-best choice compared to the candidate they dislike the most.

We now examine whether switchers are willing to take this costly choice even when their favorite candidate has low foreseeable chances of being a front-runner in the second round. To do so, we estimate and compare the impact of the presence of the third candidate on the vote shares of the top two candidates in a series of subsamples. In each new subsample, we impose additional restrictions which, arguably, make it less plausible that supporters of the third candidate expect her to pose a challenge on the top two candidates in the second round.

The first subsample combines the first and second settings as defined in Section 5.1, without imposing any further restrictions: it includes all elections in which the three top candidates are from different political orientations, and the third candidate is either on the left or the right of both top two candidates. In this sample, although the third candidate did rank behind the top two candidates in the second round in most cases, she nonetheless won in 12 elections near the threshold and ranked second in 30 (Table 10, column 3, sample 1).

To define the next subsamples, we consider the total voter support that each of the top three candidates may expect to receive in the second round, based on the votes obtained by their political orientations in the first round. A candidate from the Left, for instance, may expect to receive votes not only from her supporters but also from supporters of other left-wing candidates who did not

[16]See also Poole and Rosenthal (1985), (1987) and (2000) for evidence supporting the uni-dimensionality of the political space.

qualify for the second round. We thus define her "strength" as the sum of first round vote shares of all candidates belonging to the Left.

We restrict the second sample to observations from the first sample in which the third candidate's strength is lower than that of each of the top two candidates. For example, if the third candidate is on the left, the second candidate is on the far-right and the first candidate is on the right, we consider only elections where the left candidates gathered fewer votes in total than those on the far-right or on the right in the first round. This restriction makes it arguably less likely that the third candidate could be a front-runner in the second round – and indeed, such candidates never won and ranked second in only 3 cases near the discontinuity in this sample (Table 10, column 3, sample 2).

Candidates' strength computed based on first-round results only provides imperfect information on the level of support that candidates can hope to receive in the second round, not least because not everyone votes sincerely in the first round. The third and fourth samples thus further impose a difference of at least five (resp. ten) percentage points between the strength of the third candidate and the strength of each of the top two candidates. In samples 3 and 4 respectively, the average gap between the strength of the third candidate and of each of the top two candidates is as large as 18.0 and 20.1 percentage points in the first round, close to the discontinuity. Hence, these additional restrictions make it even less plausible that supporters of the third candidate expect her to have reasonable chances to be in contention for victory – and indeed, such a candidate never ranked first or second in those two subsamples (Table 10, column 3, samples 3 and 4).

As shown in Table 10, the impact of switchers' behaviors on the vote share of the top two candidates is robust across the four samples and strikingly close in magnitude: all estimates are significant at the 1 percent level and included between 7.0 and 8.2 percentage points (column 1). Moreover, in all subsamples, switchers' second-best choice loses many more votes from the presence of the third candidate than the candidate switchers dislike the most: while the effects on the second-best range from 4.6 to 5.4 percentage points and are all significant at the 1 percent level (column 4), the effects on the third-best range from 1.9 to 3.1 and are generally less significant (column 5). These results suggest that switchers are equally willing to decrease the vote share of their second-best when the third candidate is unlikely to be among the two front-runners in the second round.

Table 10: Impact on the vote share of the top two candidates in different sub-samples

	(1)	(2)	(3)	(4)	(5)
Impact 3rd present	Top 2 cand. 2nd-1st round	*Bandwith / Observations*	3rd becomes 1st/ 3rd becomes 2nd	Second-best choice 2nd-1st round	Third-best choice 2nd-1st round
Sample 1	-0.075***	*0.014*	12	-0.054***	-0.023**
	(0.012)	*589*	30	(0.006)	(0.009)
Sample 2	-0.070***	*0.017*	0	-0.052***	-0.019**
	(0.012)	*545*	3	(0.006)	(0.007)
Sample 3	-0.082***	*0.012*	0	-0.054***	-0.028***
	(0.016)	*285*	0	(0.008)	(0.010)
Sample 4	-0.077***	*0.008*	0	-0.046***	-0.031*
	(0.027)	*116*	0	(0.013)	(0.017)

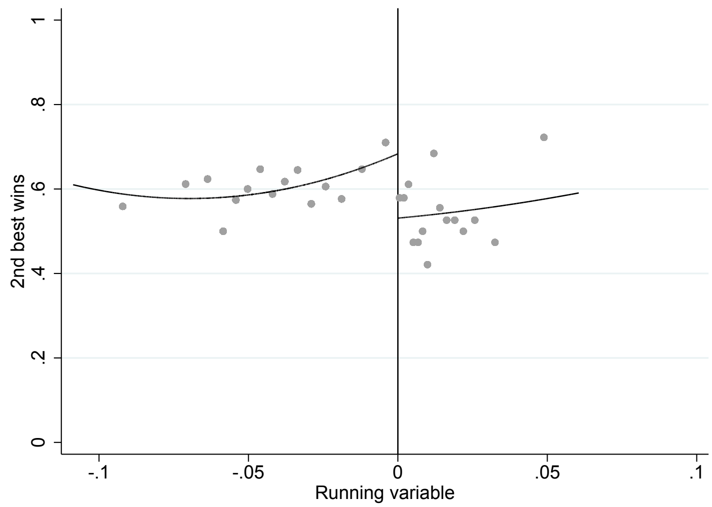
Notes: Sample 1 combines the first and second settings defined in Section 5.1: it includes all elections in which the top three candidates have different political orientations and the third candidate is either on the left or on the right of both top two candidates, making one of them the second-best choice for a large majority of switchers. Sample 2 includes the elections of sample 1 in which the third candidate's strength is lower than that of each of the top two candidates. Sample 3 (resp. 4) includes the elections of sample 2 with a difference of at least five (resp. ten) percentage points between the strength of the third candidate and the strength of each of the top two candidates. Column 2 gives the bandwidths used for the estimation of the impact on the vote share of the top two candidates as well as the number of observations lying in those bandwidths. Column 3 displays the number of cases where the third candidate ranked first or second in the second round in the elections included in the bandwidths defined in column 2. Other notes as in Table 4.

5.3 Impact on winner

We turn to the last and perhaps most important outcome, the winner of the election, and test whether the presence of the third candidate decreases the likelihood that the candidate ideologically closest to her wins the election. Formally, we use a dummy equal to 1 if switchers' second-best wins (and 0 if she loses) as the outcome. We run the test for the elections in sample 2, in which switchers' second-best is clearly identified, and the third candidate never wins (as shown in Table 10), ensuring that the results are not artificially driven by elections in which the outcome takes the value 0 at the right of the threshold due to the victory of the third candidate.

As shown in Figure 9, the presence of the third candidate has a large and negative impact on the probability that her supporters' second-best choice wins. Table 11 provides the formal estimates. In our preferred specification (column 1), we find a negative effect of 19.2 percentage points on average. This estimate is significant at the 5 percent level and robust to other specifications (columns 2 through 4). This result means that, in around one fifth of the elections we consider, the second-best choice of the third candidate's supporters loses as a result of her presence whereas she would have won absent the third, in a two-candidate race against the other top two candidate.

Figure 9: Impact on the probability that the second-best choice wins



Note: Dots represent the local averages of the probability that the second-best choice wins the race in the second round. Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit. Sample includes elections where the top three candidates belong to three distinct political orientations, where the second-best is identified and where the strength of the third candidate is strictly lower than the strengths of each of the top two candidates.

Table 11: Impact on the probability that the second-best choice wins

	(1)	(2)	(3)	(4)
Outcome	Second-best choice wins			
3rd present	-0.192**	-0.178***	-0.101	-0.205**
	(0.089)	(0.069)	(0.140)	(0.094)
Observations	686	1,567	553	1,567
Polynomial order	1	1	2	2
Bandwidth	0.021	0.043	0.017	0.043
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.656	0.622	0.678	0.622

Notes: The outcome is a dummy variable equal to 1 if the second-best choice wins the race in the second round. Sample includes elections where the top three candidates belong to three distinct political orientations, where the third candidate is either on the left or on the right of the two others and where the strength of the third candidate is strictly lower than the strengths of each of the top two candidates. Other notes as in Table 4.

Note, in addition, that switchers' second-best would also win a two-candidate race against the third candidate, making her the Condorcet winner. Indeed, in elections of the first setting, in which the candidate who arrived third in the first round is C and switchers' second-best is B, B should be expected to win a race against C as she obtained more votes in the first round and she would attract relatively more voters of candidate A (based on estimates from Table 8). Similarly, in elections of the second setting, in which the candidate who arrived third in the first round is A and switchers' second-best is B (again), B obtained more votes in the first round than A and she would attract relatively more voters of candidate C, in a race against A (based on estimates of Table 7). In sum, in around one fifth of the elections we consider, the presence of the third candidate causes the defeat of the Condorcet winner, harming the majority of the voters.

5.4 Interpretation of the results

The results obtained in Sections 5.2 and 5.3 show that a large fraction of voters are willing to vote for the third candidate even if she is unlikely to pose a challenge on the top two candidates, and at the cost of causing the defeat of their second-best choice. We interpret these results as evidence that a large fraction of voters value voting expressively for their favorite candidate over behaving strategically to ensure the victory of their second-best choice. We now discuss two competing interpretations and provide additional evidence supporting the interpretation centered on expressive motives.

Our results may alternatively be driven by voters who believe that the third candidate can win or arrive second even when she lags far behind the top two candidates in the first round. Some voters may be prone to wishful thinking and overestimate their favorite candidate's chances of victory (Blais, 2002). Despite the information given by first-round results, others may still feel uncertain about the true level of support that their favorite candidates can hope for and may thus find it instrumentally rational to vote for her (Bouton, Llorente-Saguer, and Castanheira, 2015). The results of the first round are a source of imperfect information on the level of support that candidates who qualify for the second can expect to receive. Not only that, all voters may also not pay equal attention to those results when they choose which candidate to vote for in the second round. Still, if switchers' voting choice was motivated mostly by the expectation that the third candidate may rank first or second, we should expect some fraction of voters to take first-round results into account when forming expectations, and to vote less for the third candidate when her chances are foreseeably lower. Our results do not verify this prediction: as shown in Section 5.2, the impact on the top two candidates is equally strong and remarkably stable when we restrict the sample to elections with a gap of at least 5 or even 10 percentage points between the third candidate's strength and that of each of the top two candidates. This suggests that a large fraction of switchers' behavior cannot be explained by the belief that the third candidate may pose a challenge to the top two candidates.

Another possible interpretation for switchers' behaviors is that these voters are not *short-term* but *long-term* instrumentally rational.[17] It is true that vote shares in the second round only determine who wins the election and are not taken into account for any other purpose such as campaign expenditure reimbursement (which is based on first-round vote shares). Hence, voters cannot expect their candidate or party to benefit directly from a higher vote share in the second round. However, they may choose to vote for the third candidate in order to signal their preferences and affect the policies implemented by the winner, or to influence the opinions and future votes of other voters (Piketty, 2000).

To the extent that dynamic strategic motives are driving voters' behaviors, we should expect them to trade off the impact of their vote on present elections for the impact of their vote on future elections and policies. In particular, we should see fewer people vote for the third candidate when their vote is likely to matter more for the result of the current election, for instance when the second round is expected to be close. Instead, we find that the impact on the vote share of the top two candidates is equally strong in elections where the top two candidates were very close in the first round (see Table A6 in Appendix), suggesting that switchers are willing to decrease the vote share of their second-best whether they expect the race to be close or not.

Overall, it is difficult to fully explain switchers' voting choices by strategic motives. The fact that switchers can cause the victory of the candidate they like least suggests that a large fraction of them vote based on expressive motives, disregarding the impact on election results. This result does not imply, however, that *all* switchers should be considered non-strategic, in races where the winner changes due to the presence of the third candidate: instrumental voters, who mostly care about the outcome of the election, may find it optimal and rational to enjoy the expressive value of voting for the third candidate if they expect expressive voters to swing the election regardless of their own vote choice. Symmetrically, some switchers may vote based on expressive motives even in elections in which the outcome is not affected.

In the section below, we build a model to characterize the behavior of different types of voters with a preference for the third candidate – including those who vote for the top two regardless of her presence – and assess the conditions under which their voting choices cause the defeat of the Condorcet winner.

6 Mechanism: A model with expressive and strategic voters

We consider a voting model with three candidates – two front-runners and a third – and focus on the behavior of voters who prefer the third candidate. They face the following tradeoff: they enjoy expressive value of voting for their preferred candidate but also care about the outcome of the election.

[17] See Cox (1997) for a broader discussion of other types of instrumental motives.

The model serves three purposes. First, we develop a new way of formalizing the tradeoff faced by supporters of the third candidate based on group rule-utilitarian models (Feddersen and Sandroni, 2006a; Coate and Conlin, 2004). A key feature of our model is the distinction of three different types of voters based on the voting rule they adopt: expressive, strategic-naive, and strategic-sophisticated voters.

Second, we assess the reasons that may lead voters of each of these types to vote for the third candidate instead of the front-runners, and show the conditions under which it affects the result of the election.

Finally, we relate the model to our empirical findings and provide an estimation of the fractions of expressive, strategic-naive, and strategic-sophisticated voters. All proofs can be found in Appendix E.

6.1 Set Up

6.1.1 General setting

Consider an election decided by plurality rule. A continuum of voters of mass 1 must choose one of three candidates, I, J, or K. Everybody votes.[18]

Voters have single-peaked preferences and their bliss points are distributed uniformly over the $[0,1]$ left-right axis, where 0 represents the far-left and 1 the far-right. Each candidate is similarly characterized by the location of her political platform on the axis. We assume that candidates' locations are defined ex ante and fixed. Candidate I's platform is closest to 0, candidate K's closest to 1, and candidate J's platform is located between I and K's. The uniform distribution of voters, as well as the candidates' positions, is common knowledge.

The distance between voter i and candidate F is given by:

$$d_i^F = \mid \theta_i - \gamma_F \mid,$$

where θ_i is voter i's bliss point and γ_F is candidate F's position.

We define $\mathscr{G}_F$ with $F \in \{I,J,K\}$ as the set of voters which are closest to candidate F. We have:

$$\mathscr{G}_I = \left\{i \mid d_i^I < d_i^J, d_i^I < d_i^K\right\}; \mathscr{G}_J = \left\{i \mid d_i^J < d_i^I, d_i^J < d_i^K\right\}; \mathscr{G}_K = \left\{i \mid d_i^K < d_i^I, d_i^K < d_i^J\right\}.$$

Define x (resp. y) as the bliss point of voters that are exactly indifferent between I and J (resp. J and K). The fractions of voters in $\mathscr{G}_I$, $\mathscr{G}_J$, and $\mathscr{G}_K$ are then respectively x, $y-x$, and $1-y$.

[18] This model does not include turnout, as we are interested in explaining why voters choose the third candidate, conditional on voting.

We make the two following assumptions on x and y:

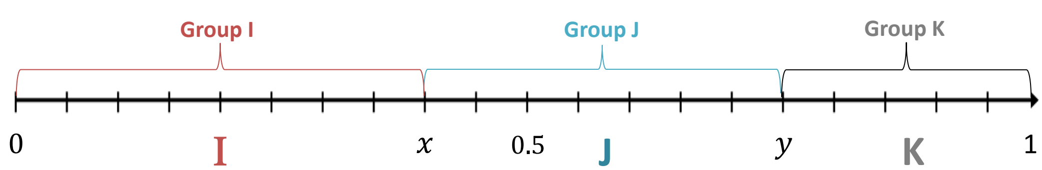
$$x > y - x > 1 - y, \quad \text{(A1)}$$

and

$$x < 1/2. \quad \text{(A2)}$$

Assumption (A1) states that $\mathscr{G}_I$ is larger than $\mathscr{G}_J$ which is itself larger than $\mathscr{G}_K$. If all voters were voting for the candidate they are closest to, K would arrive third, J second, and I would win the election. However, the Condorcet winner is J: if opposed only to K, she would receive y votes, which is higher than $1/2$ under assumption (A1); and, if opposed only to I, $1-x$ votes, which is also higher than $1/2$, under assumption (A2). Figure 10 shows candidates' positions which satisfy both assumptions.

Figure 10: Example of candidates' positions



Since voters' and candidates' positions are common knowledge, so is the fraction of voters in each group. We assume that voters use this common knowledge as a coordination device and identify the two candidates with the largest support, I and J, as the two front-runners.[19] In this context, the dominant strategy for voters in $\mathscr{G}_I$ and $\mathscr{G}_J$ is to vote for I and J respectively. Instead, voters in $\mathscr{G}_K$ need to choose between voting for K, who is their favorite candidate but has no chance of winning, or for J, who is their second-preferred candidate and is in contention for victory. Under assumptions (A1) and (A2), J can win only if a sufficiently large fraction of voters in $\mathscr{G}_K$ vote for her. The tie-breaking rule is such that J wins if she gets a fraction of votes higher than or equal to I's.

6.1.2 Preferences of voters in $\mathscr{G}_K$

Voters in $\mathscr{G}_K$ have preferences over the outcome of the election (who wins) as well as over which candidate they personally vote for. We assume that the utility voter i derives from the victory of candidate F, and the ideological cost he incurs when voting for her are both linear functions of d_i^F, the distance between i and F.

[19] In the two-round elections we consider in the empirical analysis, voters have common knowledge on first-round results, which convey some information about candidates' support. They can coordinate on the two candidates who received the largest number of votes in the first round, as in Duverger (1954) where voters coordinate on the two parties which received the largest number of votes in the preceding elections.

Therefore, without loss of generality, the additional utility voter i of group $\mathscr{G}_K$ derives from the victory of candidate J over I, and the additional cost he incurs when voting for J instead of K, can be written respectively as

$$w_i^{J,I} = b(d_i^I - d_i^J),$$

$$c_i^{J,K} = d_i^J - d_i^K,$$

where b represents the relative importance that voters in $\mathscr{G}_K$ give to the identity of the winner, compared to the identity of the candidate they personally vote for.

Lemma 1. Preferences of voters in $\mathscr{G}_K$.

i) $\forall i \in \mathscr{G}_K$, $w_i^{J,I} = w$, *where* w *is a constant*

ii) $c_i^{J,K} = \begin{cases} 2\theta_i - \gamma_J - \gamma_K & if\, \theta_i \leq \gamma_K, \\ \gamma_K - \gamma_J & if\, \theta_i \geq \gamma_K. \end{cases}$

Voters in $\mathscr{G}_K$ differ in their ideological cost of voting for J instead of their favorite candidate K, but they all enjoy the same benefit if J wins over I.

6.1.3 Group rule-utilitarian voters

We assume that voters are group rule-utilitarian. A single vote is never pivotal and voters are aware that their individual vote cannot change the result of the election. Yet, each voter has an action he understands he should take and feels morally obligated to do so. To determine the action he should follow, a voter considers different voting rules and ranks them according to the outcome they would produce for voters in $\mathscr{G}_K$ if followed by everyone in their group.

We now solve this problem and derive agents' voting decisions.[20]

6.2 Voting decision of voters in $\mathscr{G}_K$

We distinguish three types of voters in $\mathscr{G}_K$, which differ in their motives and level of sophistication: expressive, strategic-naive, and strategic-sophisticated voters. We derive the voting behavior each of these types adopt based on their preferences and beliefs.

[20]The results we show hold whether candidate K is located on the right of both candidates I and J (as in Figure 10) or on the left of both of them, the two settings which were also the focus of our empirical analysis in Sections 5.2 and 5.3. In these settings, all supporters of K have the same second-best and they all face an identical tradeoff, which justifies modeling them as a group defining a common voting rule.

6.2.1 Expressive voters

As in Feddersen and Sandroni (2006a), some voters receive zero payoff for following the rule, either because they do not feel bound by it or because they are too focused on the immediate expressive value of their vote. As a result, they base their voting decision solely on minimizing ideological cost and they all vote for K. We call these voters “expressive”. We define the fraction of expressive voters in $\mathcal{G}_K$ as α and assume that they are uniformly distributed over $\mathcal{G}_K$.

6.2.2 Strategic voters

Other voters, called “strategic”, appreciate the expressive dimension of the vote, but they also take into account the possible effects of their group’s voting choices on the outcome of the election. Formally, they receive a payoff D for following the rule that maximizes the aggregate utility of voters in $\mathcal{G}_K$ given by

$$U = pW - C, \quad (1)$$

where p is the probability that J beats I, W is the aggregate benefit that voters in $\mathcal{G}_K$ derive from the victory of J, and C is the aggregate ideological cost of voters in $\mathcal{G}_K$.

We assume $D \geq c_i^{J,K}$ for all i so that any strategic voter in $\mathcal{G}_K$ will vote for J if the rule asks her to do so.

Proposition 1. Optimal cutoff rule.

The aggregate utility of voters in $\mathcal{G}_K$ is maximized by a rule e^* which is defined by a cutoff point $\sigma^* \in [y;1]$ such that:

$$e^*(\theta_i) = \begin{cases} vote\, for\, J & if\, \theta_i \in [y;\sigma^*], \\ vote\, for\, K & if\, \theta_i \in (\sigma^*;1]. \end{cases}$$

The underlying intuition is that the least costly way to achieve a given level of votes for J, and thus a given probability p that J wins the election, is to have voters in $\mathcal{G}_K$ with low ideological cost (those located close to J) vote for J and voters with high ideological cost (those located far from J) vote for K.

As a result, strategic voters’ problem can be rewritten as

$$\underset{\sigma\in[y;1]}{Max}\, U(\sigma) = p(\sigma)W - C(\sigma),$$

where σ defines a cutoff.

The aggregate benefit, W, is independent on the cutoff and given by

$$W = \int_y^1 w d_i = (1-y)w.$$

Instead, $p(\sigma)$ and $C(\sigma)$ are both increasing functions of σ: a higher cutoff increases the number of votes for J and thus the likelihood that she wins the election, but also the aggregate ideological cost as more voters are required to vote for J instead of K. Strategic voters need to trade off both aspects when choosing the optimal cutoff. We now formalize this tradeoff for two different types of strategic voters, according to their beliefs about how other voters in $\mathscr{G}_K$ behave.

Strategic-naive voters

A fraction β of $\mathscr{G}_K$'s voters are strategic but naive, uniformly distributed over $\mathscr{G}_K$'s support. These "strategic-naive" voters do not take into account the existence of expressive voters when defining the cutoff rule: they wrongly believe that everybody in $\mathscr{G}_K$ is strategic (i.e., that $\alpha = 0$). As a result, they also have wrong beliefs on $p(\sigma)$ and $C(\sigma)$. We write their beliefs as $p_n(\sigma_n)$ and $C_n(\sigma_n)$, and their maximization problem as

$$\underset{\sigma_n \in [y;1]}{Max} U_n(\sigma_n) = p_n(\sigma_n)W - C_n(\sigma_n).$$

Lemma 2. Strategic-naive voters' beliefs.

i) $p_n(\sigma_n) = \begin{cases} 1 & if\, \sigma_n \geq 2x, \\ 0 & otherwise. \end{cases}$

ii) $C_n(\sigma_n) = \int_y^{\sigma_n} c_i^{J,K} d_i.$

Strategic-naive voters believe that J will win as long as they set a cutoff higher or equal to $2x$. Indeed, if the rule is followed by everyone in $\mathscr{G}_K$, this cutoff gives J just enough votes to beat I: while I receives the votes from voters in $\mathscr{G}_I$ (a fraction x), J receives both the votes from voters in $\mathscr{G}_J$ (a fraction $y-x$) and from voters in $\mathscr{G}_K$ with a bliss point lower than the cutoff (a fraction $2x-y$), for a total vote share of $(y-x)+(2x-y) = x$. When computing the total cost for voters in $\mathscr{G}_K$, strategic-naive voters consider once again that all agents in $\mathscr{G}_K$ will follow the rule so that all those located to the left of the cutoff will vote for J and bear the corresponding ideological cost.

Lemma 3. Voting rule of strategic-naive voters.

For b sufficiently large, strategic-naive voters in $\mathscr{G}_K$ follow the rule e_n^* defined by the cutoff $\sigma_n^* = 2x$ such that:

$$e_n^*(\theta_i) = \begin{cases} vote\, for\, J & if\, \theta_i \in [y;\sigma_n^*], \\ vote\, for\, K & if\, \theta_i \in (\sigma_n^*;1]. \end{cases}$$

Strategic-naive voters consider two types of rules: rules that let I win or rules that make J win the election. Of all rules leading to the victory of I (resp. J), the rule which minimizes ideological costs gives J no vote (resp. just enough votes to beat I). Under strategic-naive voters' beliefs, the second option involves choosing the cutoff $2x$, which is always possible since $2x$ is lower than 1 by assumption (A2).

When b, the relative weight of winner identity in $\mathscr{G}_K$ voters' utility is sufficiently large, the benefit they derive from the victory of J overcomes the corresponding ideological cost, and the second option is chosen.

But strategic-naive voters fail to take into account that expressive voters in $\mathscr{G}_K$ do *not* follow their rule: they always vote for K, including when their bliss point is lower than the cutoff. This makes the rule ill-defined to ensure the victory of J.

Strategic-sophisticated voters.

By contrast, there is a share $1-\alpha-\beta$ of voters in $\mathscr{G}_K$ called "strategic-sophisticated" who take into account the existence of both expressive and strategic-naive voters when choosing their cutoff. As a result, they have correct beliefs on $p(\sigma)$ and $C(\sigma)$. We write their beliefs as $p_s(\sigma_s)$ and $C_s(\sigma_s)$, and their maximization problem as

$$\underset{\sigma_s \in [y;1]}{Max}\, U_s(\sigma_s) = p_s(\sigma_s)W - C_s(\sigma_s).$$

Lemma 4. Strategic-sophisticated voters' beliefs.

i) $p_s(\sigma_s) = \begin{cases} 1 & if\, \sigma_s \geq y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y), \\ 0 & otherwise. \end{cases}$

ii) $C_s(\sigma_s) = \beta C_n(\sigma_n^*) + (1-\alpha-\beta)\int_y^{\sigma_s} c_i^{J,K} d_i.$

Given the behavior of expressive and strategic-naive voters, strategic-sophisticated voters know that J can win only if they set a cutoff higher or equal to $y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y)$. The votes of strategic-sophisticated voters located to the left of this cutoff, added to the votes of naive voters located to the left of σ_n^* and to the votes of voters in $\mathscr{G}_J$ then add up to x, which is exactly equal to the votes from voters in $\mathscr{G}_I$ received by I. When computing the total cost for voters in $\mathscr{G}_K$, strategic-sophisticated voters take into account both the costs of strategic-naive voters voting for J, which does not depend on σ_s, and the costs that strategic-sophisticated voters located to the left of their cutoff will have to bear.

Lemma 5. Voting rule of strategic-sophisticated voters.

Define g as the difference between the fractions of voters in $\mathcal{G}_I$ and $\mathcal{G}_J$: $g = 2x - y$ and $g^* = \left(1 - \frac{\alpha}{1-\beta}\right)(1-y)$. For b sufficiently large:
i) if $g \leq g^*$, strategic-sophisticated voters in $\mathcal{G}_K$ follow the rule e^*_s defined by the the cutoff $\sigma^*_s = y + \frac{(1-\beta)}{(1-\alpha-\beta)} g \geq \sigma^*_n$ such that:

$$e^*_s(\theta_i) = \begin{cases} vote\, for\, J & if\, \theta_i \in [y; \sigma^*_s], \\ vote\, for\, K & if\, \theta_i \in (\sigma^*_s; 1]. \end{cases}$$

ii) if $g > g^*$, strategic-sophisticated voters in $\mathcal{G}_K$ follow the rule e^*_s such that:

$$e^*_s(\theta_i) = vote\, for\, K\, for\, all\, \theta_i \in [y; 1].$$

For b (and thus W) large enough, strategic-sophisticated agents seek to define a rule that leads to J's victory, when possible. Contrary to the naive voters, they realize that making J win is not always possible. This depends on two conditions. The gap g between the shares of voters supporting candidates I and J needs to be sufficiently narrow and the fractions of expressive and strategic-naive voters among supporters of candidate K sufficiently small, reducing g^*. If $g \leq g^*$, strategic-sophisticated voters are able to ensure the victory of J and they define their cutoff such that J receives just enough votes to win against I. Their cutoff is higher than the cutoff chosen by the strategic-naive voters, as they need to compensate for expressive and strategic-naive voters' behaviors. Instead, if $g > g^*$, strategic-sophisticated voters realize J cannot win and they all vote for K.

Overall, candidate K receives votes from the three following groups of voters in $\mathcal{G}_K$: expressive voters; strategic-naive voters with a sufficiently large ideological cost; and part or all of the strategic-sophisticated voters, depending on their ability to ensure the victory of J.

We now derive equilibrium vote shares and election results and examine the conditions under which the presence of the third candidate affects the outcome of the election.

6.3 Equilibrium

Proposition 2. Outcome of the election.

For b sufficiently large, the unique equilibrium is such that:
i) if $g \leq g^*$, J wins the election with a vote share of x, I ranks second with a vote share of x and K comes third with a vote share of $1 - 2x$.
ii) if $g > g^*$, I wins the election with a vote share of x, J ranks second with a vote share of $x - (2x - y)(1-\beta)$, and K comes third with a vote share of $1 - y - \beta(2x - y)$.

If there are enough strategic-sophisticated to compensate for expressive and naive voters, J obtains a vote share just high enough to ensure her victory and the outcome of the election is the same as in the absence of K: the Condorcet winner wins the election. On the contrary, if the shares of expressive and naive voters are too high and the gap between the supports of I and J is too large, J loses the election and the only voters in $\mathscr{G}_K$ who vote for J are strategic-naive voters with a sufficiently low ideological cost. In this case, the presence of candidate K changes the outcome of the election: the Condorcet winner loses the election.

6.4 Connecting the Theory to Empirical Results

Bringing our model to the data, we first provide evidence of the existence of the three types of voters described in the model: expressive, strategic-naive and strategic-sophisticated voters. We then use the output of the model to provide estimates of their relative shares among voters who have a preference for the third candidate but vote for one of the top two candidates if she is not in the race.

Candidates K, J, and I of the model correspond, in our data, to the candidate ranked third in the first round, the top-two candidate closest to her on the left-right axis, and the candidate her supporters dislike the most. We proxy their support and thus the fraction of voters in $\mathscr{G}_I$, $\mathscr{G}_J$, and $\mathscr{G}_K$, by the strength of the corresponding candidates, which again is defined as the sum of first-round vote shares of candidates belonging to the same political orientation.

Our tests and estimates should be interpreted with caution as we use a restrictive subsample and proxy support with first-round results which are an imperfect measure of voters' preferences.

Existence of expressive, strategic-naive and strategic-sophisticated voters

If all voters in $\mathscr{G}_K$ were strategic, then strategic-naive and strategic-sophisticated voters would define the same cutoff and the Condorcet winner would always win the election (proof in Appendix E). Hence, in the elections we consider in Section 5.3, if all voters were strategic, the presence of the third candidate should let the result of the election unchanged. Instead, we find that the presence of the third candidate changes who wins in a large number of elections and causes the defeat of the Condorcet winner. We conclude that the fraction of expressive voters is not null.

If all voters in $\mathscr{G}_K$ were expressive, candidate I, whom voters in $\mathscr{G}_K$ dislike the most, would always win the election (proof in Appendix E). We test this prediction in elections satisfying both assumptions (A1) and (A2): elections where, first, the strength of the third candidate is lower than that of her supporters' second-best, which is itself lower than that of the candidate they dislike the most; and, second, the combined strengths of the third candidate and her supporters' second-best is higher than that of the candidate they dislike the most. In this sample, we find that the second-best wins in 26.9 percent of the cases close to the threshold, contradicting the prediction. We conclude that the fraction of strategic voters is not null. In one quarter of the elections, their voting behavior actually enables their second-best choice to win.

Moreover, in these cases, the victory margin of this candidate is only 2.9 percentage points on average, close to the discontinuity, which is twice as low as her average victory margin when the third candidate is not present in the second round (5.9 percentage points). This is consistent with the model predicting that, when strategic-sophisticated voters can make J win, they define an optimal voting rule which gives J just enough votes to win against I.

Finally, if all strategic voters in $\mathscr{G}_K$ were sophisticated (i.e., no strategic-naive agent), in the cases where the share of expressive voters is too large to make J win, all voters in $\mathscr{G}_K$ would vote for K. We test this prediction using the same sample as defined above and we focus on the cases where the second-best choice of the third candidate's supporters loses the election. In these elections, close to the threshold, the vote share of the third candidate in the second round is on average 3.7 percentage points lower than the total vote share of his political orientation in the first round. This suggests that some voters with a preference for the third candidate vote for their second-best even when she loses the election, a behavior that can only be attributed to strategic-naive voters.

In line with the model, the empirical evidence supports the existence of three types of voters who prefer the third candidate but vote for one of the top two candidates when she is absent.

Shares of expressive, strategic-naive and strategic-sophisticated voters

To provide estimates of the shares of expressive, strategic-naive and strategic-sophisticated voters, we solve a system of two equations in two unknowns, α and β:

$$\begin{cases} g^* = \left(1 - \frac{\alpha}{1-\beta}\right)(1-y) & (4) \\ f(g) = 1 - y - \beta g \text{ for } g > g^* & (5) \end{cases}$$

where $f(g)$ is the vote share of candidate K. Equation [4] comes from Lemma 5 and Equation [5] from Proposition 2 (ii).

To solve the system, we first estimate g^*, $1-y$, g and $f(g)$ in elections close to the threshold in which assumptions (A1) and (A2) are satisfied.

g^* is the difference between the strength of I and J below which J wins the election and above which I does. We regress the difference in vote shares between I and J in the second round on the difference between their strengths and estimate g^* as the smallest value of the regressor for which I is predicted to win: $g^* = 0.041$.

We obtain $(1-y)$ by estimating the average strength of the third candidate (minus the fraction of loyal voters) in the entire sample (for equation [4]) and when $g > 0.041$ (for equation [5]). We estimate g as the average difference between the strengths of candidates I and J when $g > 0.041$. Finally, we estimate $f(g)$ as the average vote share of the third candidate in the second round (minus the fraction of loyal voters) when $g > 0.041$. Replacing these values in equations [4] and [5], we obtain $\alpha = 0.510$ and $\beta = 0.348$.

Hence, on average, expressive voters represent 51.0 percent, strategic-naive 34.8 percent, and strategic-sophisticated voters , the remaining 14.2 percent of supporters of the third candidate who cast a valid vote whether she is present or not (details of the computation are in Appendix E).[21]

7 Conclusion

This paper highlights the motivations and consequences of citizens voting for lower-ranked candidates in elections held under plurality rule. We use a fuzzy regression discontinuity design in the context of French local and parliamentary elections. In these elections, the third-highest-ranking candidate can compete in the second round if and only if in the first round she gets a vote count at least as high as 12.5 percent of registered citizens. Exploiting this threshold, we compare electoral outcomes when voters have to choose between two or three candidates.

The presence of the third candidate increases the share of registered citizens who vote for any of the candidates by 7.2 percentage points and reduces the vote share of the top two candidates by 6.0 percentage points. Based on these results, we infer that 27.3 to 54.5 percent of the votes going to the third candidate come from "loyal" voters who would not vote for the top two candidates if their favorite candidate was not running, and 45.5 to 72.7 percent come from "switchers" who would, and yet vote for the third candidate when she is present.

Switchers' voting choices disproportionately affect the vote share of the one candidate of the top two who is ideologically closest to the third candidate. This effect is large enough to impact election results: focusing on cases where the second-best is the same for a large majority of switchers and where the third candidate does not have any chance of winning, we find that her presence in the second round changes who wins one fifth of the time. This suggests that a large fraction of voters are what we term "expressive" and vote for their favorite candidate at the cost of causing the defeat of their second-best choice. This behavior is also costly for the majority of voters, as

[21] In our model, (1-y) represents the share of voters who have a preference for the third candidate and who vote whether she is present or not (since we do not account for abstention). To estimate (1-y) in our data, we thus need to subtract the share of loyal voters from the strength of the third candidate. Similarly, f(g) represents the share of voters who vote for the third candidate in the second round and would have voted anyway. We thus need to substract the share of loyal voters once again. We use the lower bound of the share of loyal voters as derived in Section 4.3. The estimates are very close in magnitude when we use instead the upper bound. In that case we obtain that expressive, strategic-naive and strategic-sophisticated represent 45.0 percent, 34.8 percent, and 20.2 percent of supporters of the third candidate who cast a valid vote whether she is present or not.

switchers' second-best is the Condorcet winner in these elections.

In our model, the Condorcet winner will win the election if the behavior of expressive voters (who always vote for their favorite candidate) and strategic-naive voters (who fail to take into account the presence of expressive voters when making their voting decisions) is offset by a sufficient number of strategic-sophisticated voters (those who *do* account for the presence of expressive voters). Where expressive voting is too prevalent, the victory of the Condorcet winner hinges on the third candidate dropping out of the race. The RDD strategy we use to estimate voters' response to the presence of a third candidate does not enable us to provide causal evidence on the factors affecting this candidate's decision to stay in the second round when qualified. We can highlight a few striking regularities, however. Candidates who ranked third in the first round drop out in 11.8 percent of the races when they belong to a different political orientation than both top two candidates (14.8 percent close to the threshold). The higher their first-round vote share, the less likely they are to drop out. By contrast, third candidates drop out in 96.4 percent of the races (91.1 percent, close to the threshold) when they belong to the same political orientation as one of the top two candidates. In that case, the higher the vote share of the third candidate – and thus the more she threatens the victory of the top two belonging to the same orientation – the more likely she is to drop out (see Figure B1). These patterns suggest that ideological proximity – which may also proxy for trust and habit to govern together – facilitates candidates' coordination. This coordination often takes the form of national agreements, such as in the 2012 parliamentary elections, when left-wing parties asked all their candidates to drop out in the second round if another candidate on the left was better ranked. Instead, ideological distance decreases candidates' willingness to adjust their decision to compete or not in the second round to the risk of the least-preferred candidate winning the election.

Further research is needed to better understand how and why such agreements succeed or fail and to disentangle the part played by national parties and local candidates in them. Together with voters' expressive behavior, candidates' lack of coordination is responsible for the large number of defeats of Condorcet winners. Ultimately, these repeated failures of Duverger's law call into question the widespread use of the plurality rule to aggregate voter' preferences.

References

ALVAREZ, R. M., AND J. NAGLER (2000): "A New Approach for Modeling Strategic Voting in Multiparty Elections," *British Journal of Political Science*, 30(01), 57–75.

ANAGOL, S., AND T. FUJIWARA (2016): "The Runner-Up Effect," *Journal of Political Economy*, 124(4), 927–991.

ARROW, K. J. (1951): *Social Choice and Individual Values*. Wiley, New York.

ATKESON, L. R. (1999): "Sure, I Voted for the Winner! Overreport of the Primary Vote for the Party Nominee in the National Election Studies," *Political Behavior*, 21(3), 197–215.

BALINSKI, M., AND R. LARAKI (2011): *Majority judgment: measuring, ranking, and electing*. MIT press.

BLAIS, A. (2002): "Why is there so little strategic voting in Canadian plurality rule elections?," *Political Studies*, 50(3), 445–454.

BLAIS, A., B. DOLEZ, AND A. LAURENT (2017): "Strategic voting in the second round of a two-round system: The 2014 French municipal elections," *French Politics*, 15(1), 27–42.

BLAIS, A., R. NADEAU, E. GIDENGIL, AND N. NEVITTE (2001): "Measuring strategic voting in multiparty plurality elections," *Electoral Studies*, 20(3), 343–352.

BLAIS, A., R. YOUNG, AND M. TURCOTTE (2005): "Direct or indirect? Assessing two approaches to the measurement of strategic voting," *Electoral Studies*, 24(2), 163–176.

BORDIGNON, M., T. NANNICINI, AND G. TABELLINI (2016): "Moderating Political Extremism: Single Round versus Runoff Elections under Plurality Rule," *American Economic Review*, 106(8), 2349–2370.

BOUTON, L. (2013): "A theory of strategic voting in runoff elections," *American Economic Review*, 103(4), 1248–1288.

BOUTON, L., AND G. GRATTON (2015): "Majority Runoff Elections: Strategic Voting and Duverger's Hypothesis," *Theoretical Economics*, 10(2), 283–314.

BOUTON, L., A. LLORENTE-SAGUER, AND M. CASTANHEIRA (2015): "Multicandidate Elections: Aggregate Uncertainty in the Laboratory," *Games and Economic Behavior*.

BRAMS, S. J., AND P. C. FISHBURN (2002): "Voting procedures," *Handbook of Social Choice and Welfare*, 1, 173–236.

BROOCKMAN, D. E. (2013): "Do female politicians empower women to vote or run for office? A regression discontinuity approach," *Electoral Studies*, 34, 190–204.

CALONICO, S., M. D. CATTANEO, M. H. FARRELL, AND R. TITIUNIK (2016): "Regression Discontinuity Designs Using Covariates," *Working Paper, University of Michigan*.

CALONICO, S., M. D. CATTANEO, AND R. TIRIUNIK (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82(6), 2295–2326.

CAMPBELL, J. E. (2010): "Explaining Politics, Not Polls: Reexamining Macropartisanship with Recalibrated NES Data," *Public Opinion Quarterly*, 74(4), 616–642.

CAUGHEY, D., AND J. S. SEKHON (2011): "Elections and the regression discontinuity design: Lessons from close U.S. House Races, 1942-2008," *Political Analysis*, 19(4), 385–408.

COATE, S., AND M. CONLIN (2004): "A Group-Rule Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 94(5), 1476–1504.

COX, G. W. (1994): "Strategic Voting Equilibria Under the Single Nontransferable Vote," *American Political Science Review*, 88(3), 608–621.

——— (1997): *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge University Press, Cambridge.

DAVIS, O. A., M. J. HINICH, AND P. C. ORDESHOOK (1970): "An Expository Development of a Mathematical Model of the Electoral Process," *American Policital Science Review*, 64(2), 426–448.

DE LA CUESTA, B., AND K. IMAI (2016): "Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections," *Annual Review of Political Science*, 19, 375–396.

DEMEZE, H., I. MOYOUWOU, AND R. PONGOU (2016): "The Welfare Economics of Tactical Voting in Democracies: A Partial Identification Equilibrium Analysis," *Working Paper*.

DOLEZ, B., AND A. LAURENT (2010): "Strategic voting in a semi-presidential system with a two-ballot electoral system. The 2007 French legislative election," *French Politics*, 8(1), 1–20.

DUVERGER, M. (1954): *Political Parties*. Wiley, New York.

EGGERS, A. C. (2015): "Proportionality and Turnout: Evidence From French Municipalities," *Comparative Political Studies*, 48(2), 135–167.

EGGERS, A. C., A. FOWLER, J. HAINMUELLER, A. B. HALL, AND J. M. SNYDER (2015): "On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races," *American Journal of Political Science*, 59(1), 259–274.

EGGERS, A. C., AND N. VIVYAN (2016): "Who votes strategically?," *Working Paper*.

ERIKSON, R. S., O. FOLKE, AND J. M. SNYDER (2015): "A Gubernatorial Helping Hand? How Governors Affect Presidential Elections," *Journal of Politics*, 77(2), 491–504.

FAUVELLE-AYMAR, C., AND A. FRANÇOIS (2005): "Campaigns, Political Preferences and Turnout: An Empirical Study of the 1997 French Legislative Elections," *French Politics*, 3(1), 49–72.

FEDDERSEN, T., AND A. SANDRONI (2006a): "A theory of participation in elections," *American Economic Review*, 96(4), 1271–1282.

——— (2006b): "The calculus of ethical voting," *International Journal of Game Theory*, 35(1), 1–25.

FESTINGER, L. (1962): "Cognitive dissonance," *Scientific American*, 207(4), 93–107.

FEY, M. (1997): "Stability and Coordination in Duverger's Law : A Formal Model of Preelection Polls and Strategic Voting," *American Political Science Review*, 91(1), 135–147.

FIVA, J. H., AND D. M. SMITH (2017): "Local Candidates and Voter Mobilization : Evidence from Historical Two-Round Elections in Norway," *Electoral Studies*, 45, 130–140.

FORSYTHE, R., R. B. MYERSON, T. A. RIETZ, AND R. J. WEBER (1993): "An experiment on coordination in multi-candidate elections: The importance of polls and election histories," *Social Choice and Welfare*, 10(3), 223–247.

FOUCAULT, M., AND A. FRANÇOIS (2005): "Le rendement des dépenses électorales en France," *Revue Economique*, 56(5), 1125–1143.

FUJIWARA, T. (2011): "A Regression Discontinuity Test of Strategic Voting and Duverger's Law," *Quarterly Journal of Political Science*, 6(3-4), 197–233.

GIBBARD, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41(4), 587–601.

HAINMUELLER, J., A. B. HALL, AND J. M. SNYDER (2015): "Assessing the External Validity of Election RD Estimates: An Investigation of the Incumbency Advantage," *Journal of Politics*, 77(3), 707–720.

HALL, A. B., AND J. M. SNYDER (2015): "Information and Wasted Votes: A Study of U.S. Primary Elections," *Working Paper*.

HERRMANN, M., S. MUNZERT, AND P. SELB (2016): "Determining the effect of strategic voting on election results," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 179(2), 583–605.

HILLYGUS, D. S. (2007): "The Dynamics of Voter Decision Making Among Minor-Party Supporters: The 2000 Presidential Election in the United States," *British Journal of Political Science*, 37(02), 225–244.

IMBENS, G., AND K. KALYANARAMAN (2012): "Optimal bandwidth choice for the regression discontinuity estimator," *Review of Economic Studies*, 79(3), 933–959.

IMBENS, G., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635.

KAWAI, K., AND Y. WATANABE (2013): "Inferring strategic voting," *American Economic Review*, 103(2), 624–662.

KIEWIET, D. R. (2013): "The Ecology of Tactical Voting in Britain," *Journal of Elections, Public Opinion and Parties*, 23(1), 86–110.

KISS, Á. (2015): "Identifying strategic voting in two-round elections," *Electoral Studies*, 40, 127–135.

LASLIER, J.-F. (2009): "The Leader Rule: A Model of Strategic Approval Voting in a Large Electorate," *Journal of Theoretical Politics*, 21(1), 113–136.

LEE, D. S. (2008): "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 142(2), 675–697.

MARTINELLI, C. (2002): "Simple plurality versus plurality runoff with privately informed voters," *Social Choice and Welfare*, 19(4), 901–919.

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

MYATT, D. P. (2007): "On the Theory of Strategic Voting," *Review of Economic Studies*, 74(1), 255–281.

MYATT, D. P., AND S. D. FISHER (2002): "Tactical Coordination in Plurality Electoral Systems," *Oxford Review of Economic Policy*, 18(4), 504–522.

MYERSON, R. B., AND R. J. WEBER (1993): "A Theory of Voting Equilibria," *American Political Science Review*, 87(1), 102–114.

NIEMI, R. G., G. WRITTEN, AND M. N. FRANKLIN (1992): "Constituency characteristics, individual characteristics and tactical voting in the 1987 British general election," *British Journal of Political Science*, 22(02), 229–240.

NURMI, H. (1983): "Voting Procedures : A Summary Analysis," *British Journal of Political Science*, 13(2), 181–208.

OECD (2013): *Government at a Glance 2013*.

PALFREY, T. R. (1989): "A mathematical proof of Duverger's law," in *Models of Strategic Choice in Politics*, ed. by P. C. Ordeshook, pp. 69–92. University of Michigan Press, Ann Arbor.

PIKETTY, T. (2000): "Voting as Communicating," *Review of Economic Studies*, 67, 169–191.

POOLE, K. T., AND H. ROSENTHAL (1985): "A Spatial Model for Legislative Roll Call Analysis," *American Journal of Political Science*, 29(2), 357–384.

——— (1987): "Analysis of Congressional Coalition Patterns: A Unidimensional Spatial Model," *Legislative Studies Quarterly*, 12(1), 55–75.

——— (2000): *Congress: A political-economic history of roll call voting*. Oxford University Press on Demand.

POSNER, E. A., AND E. G. WEYL (2015): "Voting squared: Quadratic voting in democratic politics," *Vand. L. Rev.*, 68, 441.

RIETZ, T. (2008): "Three-way experimental election results: strategic voting, coordinated outcomes and Duverger's law," *Handbook of Experimental Economics Results*, 1, 889–897.

RUBIN, D. B. (1974): "Estimating causal effects of treatments in randomized and nonrandomized studies.," *Journal of Educational Psuchylogy*, 66, 688–701.

SATTERTHWAITE, M. A. (1975): "Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory*, 10(2), 187–217.

SÉNAT (1999): "Propositions de loi concernant les dispositions relatives aux élections municipales, cantonales et législatives," *RAPPORT 62*.

SPENKUCH, J. L. (2014): "(Ir)rational Voters ?," *Working paper*.

——— (2015): "Please Don't Vote for Me : Voting in a Natural Experiment with Perverse Incentives," *Economic Journal*, 125(585), 1025–1052.

VAN DER STRAETEN, K., J.-F. LASLIER, N. SAUGER, AND A. BLAIS (2010): "Strategic, sincere, and heuristic voting under four election rules: An experimental study," *Social Choice and Welfare*, 35(3), 435–472.

WRIGHT, G. C. (1990): "Misreports of vote choice in the 1988 NES senate election study," *Legislative Studies Quarterly*, pp. 543–563.

——— (1992): "Reported versus actual vote: there is a difference and it matters," *Legislative Studies Quarterly*, pp. 131–142.

Appendix (for Online Publication only)

Appendix A. Tables and figures

Table A1: Placebo checks

Outcome	(1) Nb reg. citizens	(2) Nb cand. 1st round	(3) Turnout 1st round	(4) Dist. 1-2 1st round	(5) Predicted Assignment
3rd present	5,601	-0.14	0.003	0.003	0.015
	(5,639)	(0.6)	(0.015)	(0.011)	(0.017)
Observations	1,879	2,428	2,189	2,018	1,930
Polynomial order	1	1	1	1	1
Bandwidth	0.021	0.026	0.024	0.022	0.021
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD
Mean, left of the threshold	43,982	7.5	0.604	0.092	0.301

Notes: Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. The variable of interest (the presence of a third candidate at the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). The polynomial order is 1 and the optimal bandwidths are derived under the MSERD procedure. The outcomes are: the number of registered citizens (column 1), the number of candidates running in the first round (column 2), the share of registered citizens turning out in the first round (column 3), the difference between the share of candidate votes obtained by the candidate ranked first and by the candidate who came in second in the first round (column 4) and the predicted value of the assignment status (column 5). The last outcome is obtained in two steps. We first regress the assignment variable D on first-round variables, including the four aforementioned outcomes as well as share of candidate votes, vote share of each of the top three candidates, political label and orientation of the three candidates, number of candidates from the left, right, far-right, far-left and center. We then use the coefficients from this regression to predict assignment status.

Table A2: Impact on blank and null votes separately for the 2015 local elections

Outcome	(1) Null and Blank votes 2nd-1st	(2) Blank votes 2nd-1st	(3) Null votes 2nd-1st
3rd present	-0.027***	-0.016***	-0.011***
	(0.002)	(0.002)	(0.001)
Observations	481	416	599
Polynomial order	1	1	1
Bandwidth	0.014	0.012	0.017
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.053	0.036	0.017

Notes: Sample includes only the 2015 local elections. Each outcome uses the number of registered voters as the denominator and is defined as a simple difference between the second and first rounds. Other notes as in Table A1.

Table A3: Impact on the vote shares of the first and second candidates separately

Outcome	(1) Vote share top 2 2nd-1st	(2) Vote share 1st 2nd-1st	(3) Vote share 2nd 2nd-1st
3rd present	-0.060***	-0.031***	-0.027**
	(0.016)	(0.010)	(0.011)
Observations	1,757	1,585	1,547
Polynomial order	1	1	1
Bandwidth	0.019	0.018	0.018
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.550	0.311	0.250

Notes: Each outcome uses the number of registered voters as the denominator and is defined as a simple difference between the second and first rounds. Other notes as in Table A1.

Table A4: Impact on the winning margin

Outcome	(1) Distance winner - 2nd candidate 2nd round
3rd present	-0.058*** (0.014)
Observations	2,677
Polynomial order	1
Bandwidth	0.030
Band. method	MSERD
Mean, left of the threshold	0.155

Notes: The outcome variable is the vote share of the winner minus the vote share of the second candidate in the second round, as fractions of candidate votes. Other notes as in Table A1.

Table A5: Impact on the top two candidates depending on the closeness of the race

We estimate the impact of the presence of the third candidate on the vote share of the top two candidates depending on the closeness of the race in the first round. Closeness is defined as the difference in vote share (as fraction of candidate votes) between the first and second candidates.

As defined in Section 5.2, sample 1 includes all elections in which the top three candidates have different political orientations and the third candidate is either on the left or on the right of both top two candidates, making one of them the second-best choice for a large majority of switchers. We then consider two subsamples: one in which the distance between the top two candidates in the first round is smaller than 5 percentage points and one in which the distance is smaller than 2.5 percentage points. In those subsamples, the distance between the first and second candidates is on average equal to 2.3 and 1.1 percentage points, respectively, close to the discontinuity.

As shown in Table A5, the effects of switchers' behaviors on the vote share of the top two candidates are robust across the three samples and strikingly close in magnitude: all estimates are significant at the 1 percent level and included between 7.5 and 8.2 percentage points (column 1). Moreover, in all subsamples, switchers' second-best choice loses much more votes from the presence of the third candidate than the candidate switchers dislike the most: while the effects on the second-best range from 5.4 to 5.7 percentage points and are all significant at the 1 percent level (column 3), the effects on the third best range from 2.3 to 3.2 percentage points and are less significant (column 4).

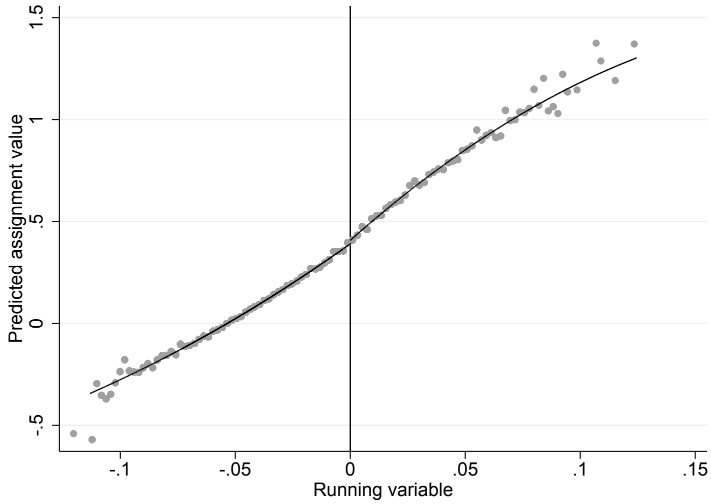
	(1)	(2)	(3)	(4)
Impact 3rd present	Top 2 cand. 2nd-1st round	*Bandwith /* *Observations*	Second-best choice 2nd-1st round	Third-best choice 2nd-1st round
Sample 1	-0.075*** (0.012)	*0.014* *589*	-0.054*** (0.006)	-0.023** (0.009)
Sample 1 + distance12≤5pp	-0.076*** (0.018)	*0.015* *243*	-0.057*** (0.006)	-0.023* (0.014)
Sample 1 + distance12≤2.5pp	-0.082*** (0.023)	*0.015* *131*	-0.056*** (0.009)	-0.032* (0.018)

Notes: Sample 1 combines the first and second settings defined in Section 5.1: it includes all elections in which the top three candidates have different political orientations and the third candidate is either on the left or on the right of both top two candidates, making one of them the second-best choice for a large majority of switchers. Sample 2 includes the elections of sample 1 in which the distance between the first and second candidates in the first round is smaller than 5 percentage points. Sample 3 includes the elections of sample 1 in which the distance between the first and second candidates in the first round is smaller than 2.5 percentage points. Column 2 gives the bandwidths used for the estimation of the impact on the vote share of the top two candidates as well as the number of observations lying in those bandwidths. Other notes as in Table A1.

These results suggest that switchers are equally willing to vote for the third candidate and thus decrease the vote share of their second-best when the race is close in the first round.

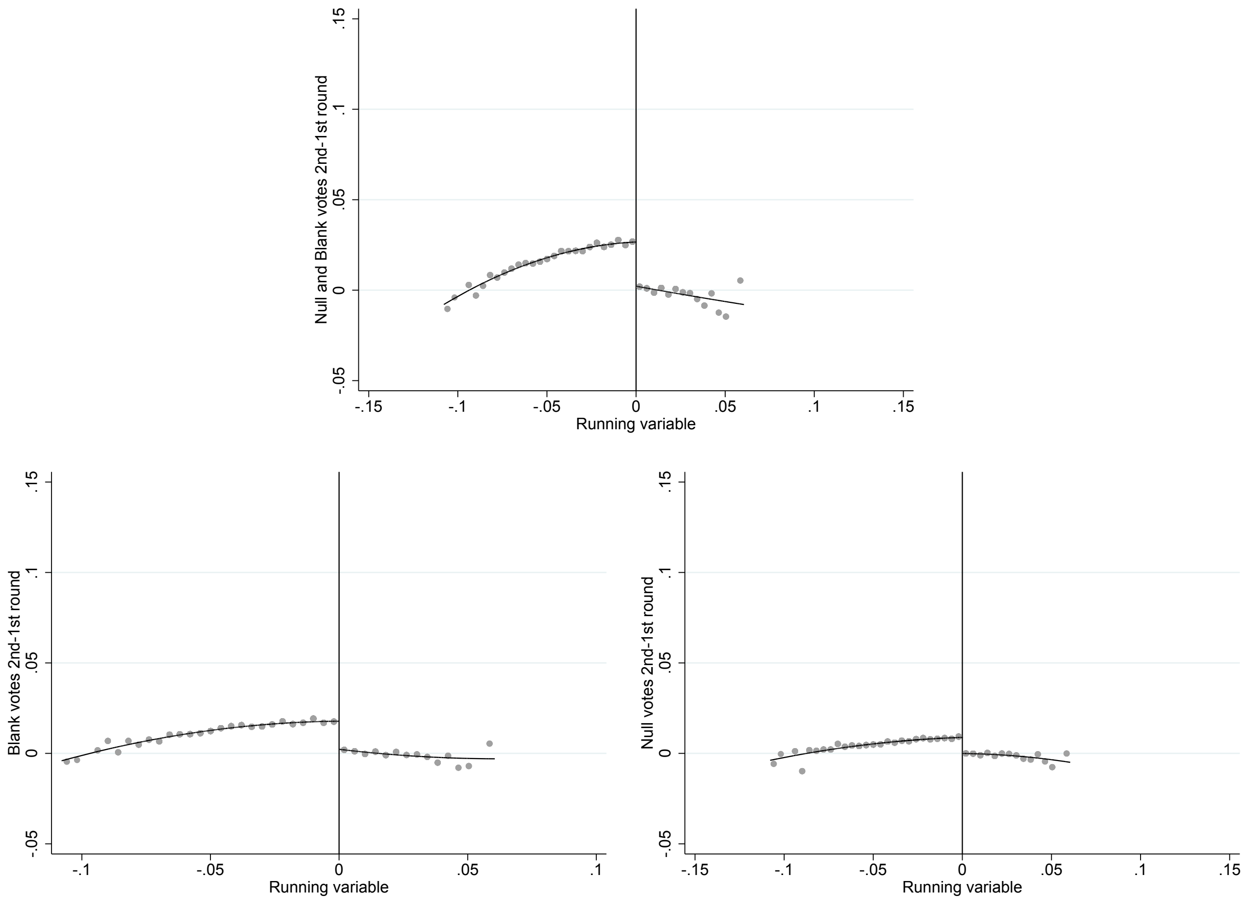
Our results are robust to considering sample 2 as defined in Section 5.2 instead of sample 1 and they remain unchanged if we compute closeness as the distance between the top two candidates' strengths rather than between candidates' vote shares (results available upon request).

Figure A1: General balance check



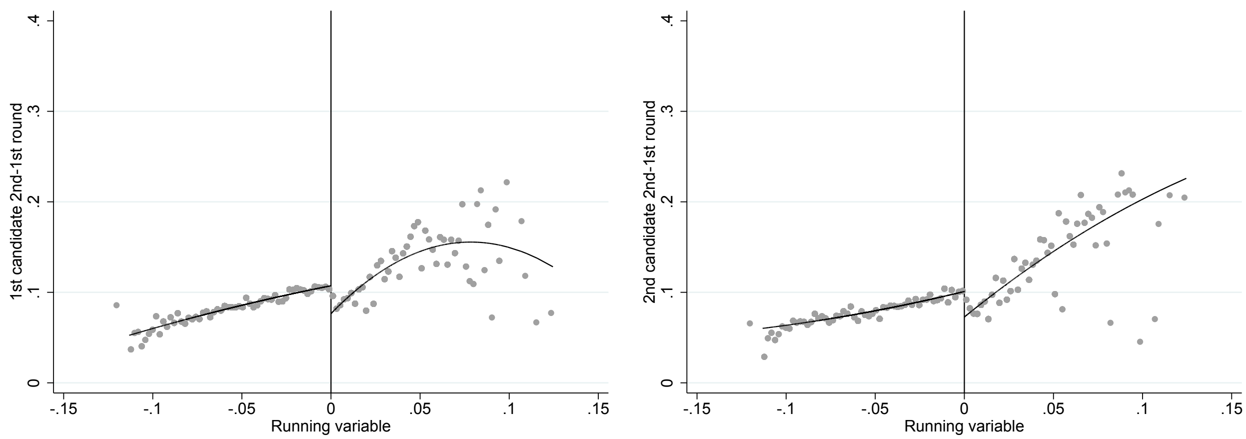
Notes: Dots represent the local averages of the predicted assignment status (y-axis). Averages are calculated within 0.2 percentage point wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit. Outcome defined as in Table A1.

Figure A2: Impact on blank and null votes separately in the 2015 local elections



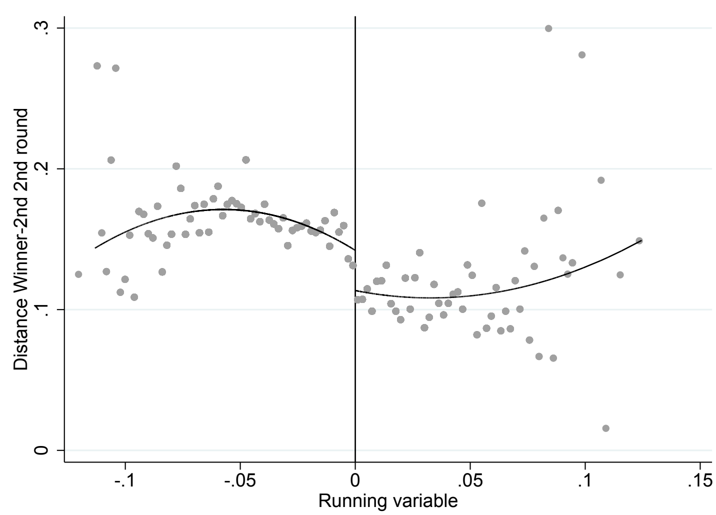
Notes: Sample includes only the 2015 local elections. Dots represent the local averages of the outcome variable (y-axis). Averages are calculated within 0.4 percentage point wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured in percentage points. Continuous lines are a quadratic fit.

Figure A3: Impact on the vote shares of the first and second candidates separately



Notes: Dots represent the local averages of the vote share of the first candidate (resp. second) in the second round (y-axis). Vote shares are computed using the number of registered citizens as the denominator. Other notes as in Figure A1.

Figure A4: Impact on the distance between the winner and the second candidate in the 2nd round



Notes: Dots represent the local averages of the difference between the share of candidate votes obtained by the winner and by the candidate who came in second. Other notes as in Figure A1.

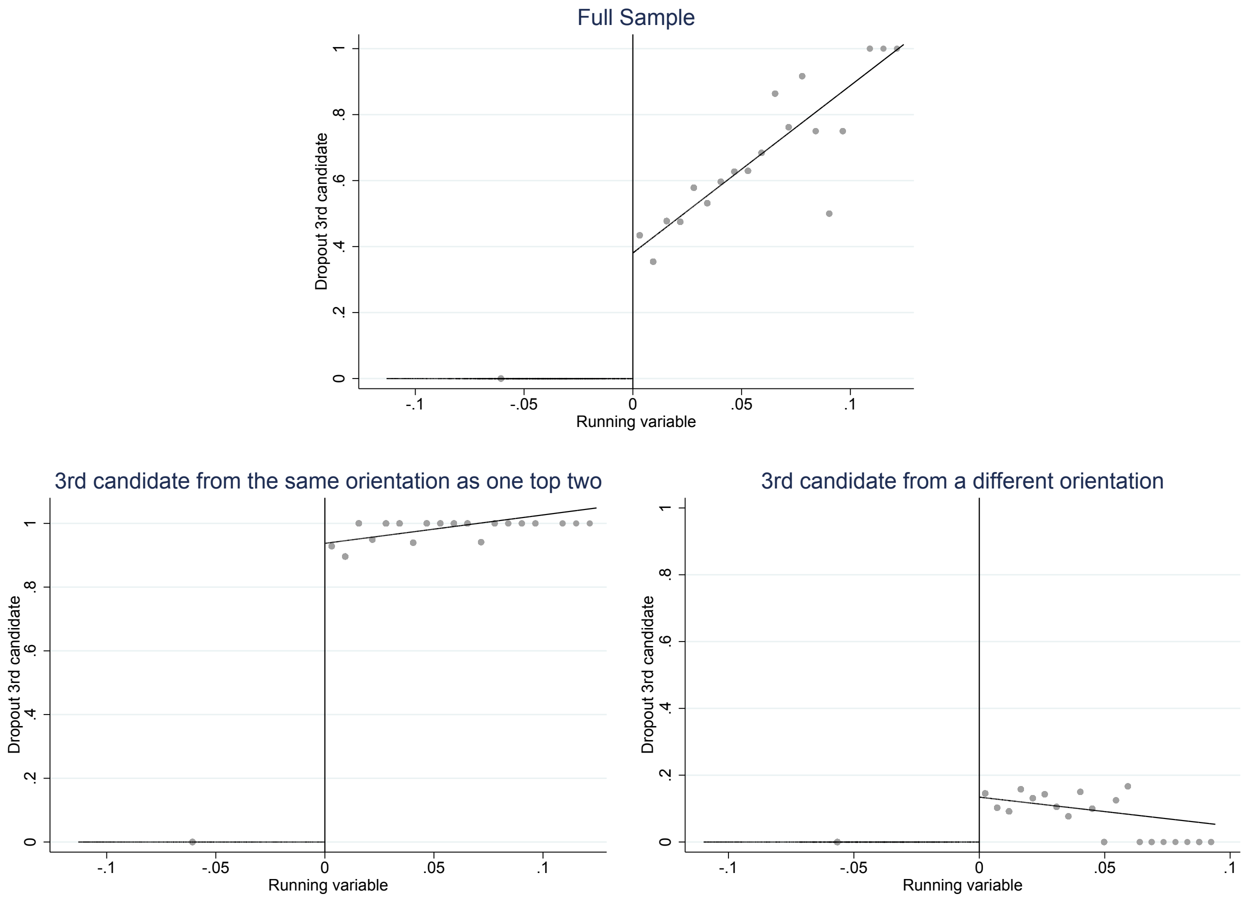
Appendix B. Third candidate dropouts

In this section, we provide additional graphical evidence to characterize the cases in which the third candidate decides to drop out of the race when she qualifies for the second round. In the graphs below, the outcome is a dummy equal to 1 if the third candidate decides to drop out of the race. By definition, it always takes value 0 at the left of the threshold.

The first graph plots the probability to drop out against the running variable in the whole sample. Note that it exactly mimicks the first-stage figure shown in Section 3.2. We then differentiate elections where the third candidate belongs to the same political orientation as one of the top two candidates from elections where she does not. On average, the third candidate is much more likely do drop out when of the same political orientation as one of the top two candidates (96.4 percent of the races, 91.1 percent close to the threshold) than when she belongs to a different orientation than both of them (11.8 percent of the races, 14.8 percent close to the threshold).

In addition, the linear fit on the right hand side of the graph shows that the third candidate is more likely to drop out the higher her vote share in the first round, when she belongs to the same political orientation as one of the top two candidates. Instead, when the third candidate belongs to a different orientation than both top two candidates, we observe the opposite trend: the higher her vote share in the first round, the less likely she is to drop out.

Figure B1: Probability that the third candidate drops out depending on her political orientation



Notes: The outcome is a dummy equal to 1 if the third candidate drops out of the race in the second round. Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a linear fit.

Appendix C. Campaign expenditures

In French local and parliamentary elections, candidates who receive at least 1 percent of candidate votes in the first round must submit their campaign accounts to the French National Commission on Campaign Accounts and Political Financing (CNCCFP). The CNCCFP then examines the accounts, checks whether candidates respected the maximal amount they were authorized to spend in their district and assesses whether they are eligible to be reimbursed by the French State.

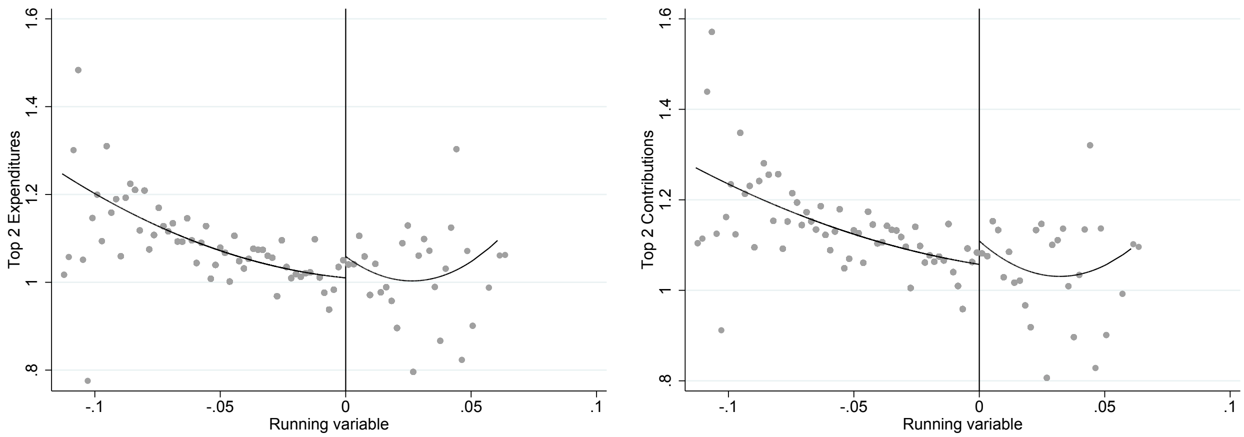
Data on campaign expenditures are publicly made available by the CNCCFP. The CNCCFP was created in 1990. Hence, data for elections held before 1990 are not available. Official accounts for the most recent elections - 2011 and 2015 local elections as well as 2007 and 2012 parliamentary elections - are available online on the CNCCFP website (http://www.cnccfp.fr/index.php?art=584). Official accounts for the 1993, 1997 and 2002 parliamentary elections were digitalized from printed booklets by Abel François and his co-authors for their studies on the impact of electoral expenditures on turnout (Fauvelle-Aymar and François, 2005) and electoral results (Foucault and François, 2005). In total, we have been able to gather data corresponding to 77.8 percent of our sample.[22]

For each election, district and candidate, we observe the total amount spent by the candidate (summing up expenditures incurred before the first round and between the first and second rounds), the total amount of contributions she received and the amount of each different type of contribution (contributions received from the candidate's political party, personal funds, donations, natural advantages and other sources), as well as the decision of the CNCCFP to accept, modify or reject the proposed account.

These data enable us to test whether the top two candidates increase their campaign expenditures in response to the presence of the third candidate. As we can see in Figure C1, the presence of the third candidate does not seem to affect the campaign expenditures of the top two candidates, or the contributions they receive to finance their campaign. Table C1 provides the formal estimates. Neither the effect on top two candidates' total expenditures nor the estimate on total contributions is statistically significant. The estimate on contributions received from candidates' political parties is significant at the 5 percent level, and positive. Nevertheless, the estimate on total contributions is small, not significant, and actually negative: the top two candidates do not receive significantly more money overall when the third candidate is present.

[22]Note that for the 2011 local elections, data are available only for district exceeding 9,000 inhabitants. We thus have data on campaign expenditures for 74.4 percent of that election's observations. Also, for 1993, data for two French territories overseas ("outre-mer") are missing. Finally, data are missing for some candidates, either because they receive less than 1 percent of the candidate votes, or because they did not release their campaign account on time: we have 1 observation missing for the first candidate, 8 for the second candidate and 34 for the third candidate.

Figure C1: Campaign expenditures of the top two candidates



Notes: Sample includes 2011 and 2015 local elections and 1993, 1997, 2002, 2007 and 2012 parliamentary elections. One outlier has been removed to make the graph clearer (the district "Saint-Pierre-et-Miquelon" in 1997 parliamentary elections). Dots represent the local averages of the outcome variable (y-axis). Each outcome uses the number of registered voters as the denominator. Averages are calculated within 0.2 percentage point wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

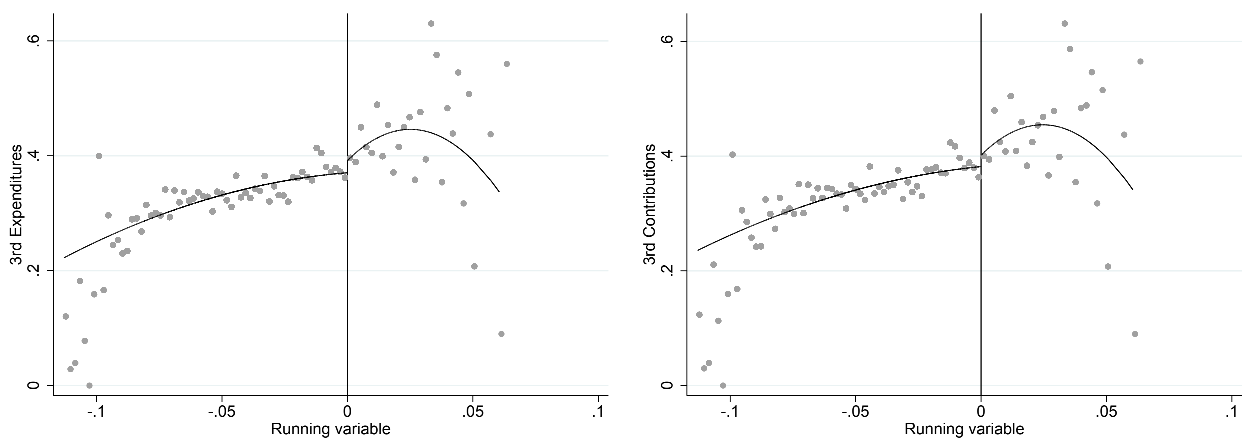
Table C1: Campaign expenditures of the top two candidates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Top two candidates							
	Total Expenditures	Total Contributions	Personal	Party's Contributions	Other	Natural Advantages	Donations	Balance
3rd present	0.013	-0.024	-0.078	0.080**	0.012	0.003	0.003	-0.010
	(0.079)	(0.092)	(0.068)	(0.039)	(0.007)	(0.012)	(0.061)	(0.021)
Observations	890	786	800	1,236	1,009	1,437	954	774
Polyn. order	1	1	1	1	1	1	1	1
Bandwidth	0.013	0.012	0.012	0.018	0.016	0.023	0.014	0.011
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	1.017	1.045	0.688	0.113	0.012	0.037	0.198	0.044

Notes: Sample includes 2015 and 2011 local elections and 1993, 1997, 2002, 2007 and 2012 parliamentary elections. Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. Each outcome uses the number of registered voters as the denominator. The variable of interest (the presence of a third candidate at the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). The polynomial order is 1 and the optimal bandwidths are derived under the MSERD procedure.

We now turn to the impact of the presence of the third candidate on her own overall campaign expenditures (which again sum up expenditures incurred before the first round and between the first and second rounds). We do not find any significant impact on the third candidate's total expenditures or on the total contributions she received (Figure C2 and Table C2). When we disentangle between the different sources of contributions, only one coefficient out of six is significant at the 10 percent level, and negative.

Figure C2: Campaign expenditures of the third candidate



Notes as in Figure C1.

Table C2: Campaign expenditures of the third candidate

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Third candidate							
	Total Expenditures	Total Contributions	Personal	Party's Contributions	Other	Natural Advantages	Donations	Balance
3rd present	0.035	0.038	0.054	-0.031*	-0.001	-0.000	-0.004	-0.000
	(0.037)	(0.040)	(0.035)	(0.018)	(0.005)	(0.007)	(0.024)	(0.011)
Observations	929	842	831	679	835	1,034	678	701
Polyn. order	1	1	1	1	1	1	1	1
Bandwidth	0.014	0.013	0.012	0.010	0.014	0.017	0.010	0.010
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	0.382	0.397	0.307	0.023	0.005	0.015	0.043	0.010

Notes as in Table C1.

Appendix D. Robustness of the results to two special cases

Case 1: Second candidate dropouts

While the first candidate never drops out of the race in our sample, the second candidate drops out between the two rounds in 5.5 percent of the elections near the discontinuity. When the second candidate drops out of the race on the left of the discontinuity, the second round is uncontested and the only candidate remaining in the race wins the election. When the second candidate drops out from the race on the right of the discontinuity, in all but one election the third candidate remains in the race and the second round takes place between the candidates arrived first and third in the first round.

As shown by Table D1, the likelihood of the second candidate dropping out is not significantly affected by the presence of the third candidate. Moreover, we derive our main results restricting the sample to configurations where all three candidates belong to distinct political orientations (see Section 5) and where, as a result, the second candidate almost never drops out (she does so in only 4 elections near the discontinuity or 0.4 percent of the cases).

In sum, our results are not driven by second candidate dropouts.

Table D1: Second candidate dropouts

	(1)	(2)	(3)	(4)
Outcome	Second candidate drops out			
3rd present	0.025	0.025	0.024	0.036
	(0.045)	(0.032)	(0.058)	(0.053)
Observations	1,966	3,430	3,092	3,430
Polynomial order	1	1	2	2
Bandwidth	0.021	0.037	0.033	0.037
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.038	0.034	0.037	0.034

Notes: Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. The outcome is a dummy equal to 1 if the second candidate drops out of the race in the second round. The variable of interest (the presence of a third candidate in the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). The polynomial order is 1 in columns 1 and 2 and 2 in columns 3 and 4. The bandwidths are derived under the MSERD (columns 1 and 3) and IK (columns 2 and 4) procedures.

Case 2: Fourth candidate qualifies and runs in the second round

In 22 races or 1.2 percent of the elections near the discontinuity, the candidate ranked fourth in the first round also obtained a number of votes higher than 12.5 percent of the registered citizens and qualified for the second round. She decided to run in the second round in only 3 races close to the discontinuity, and in these cases the third candidate always dropped out of the race.

When turnout is low, it is difficult for more than three candidates to reach the 12.5 threshold. Restricting our sample to elections where turnout in the first round is lower than 58 percent enables us to get a subsample of elections where the fourth candidate never qualifies. As shown in Table D2, the impacts on our three main outcomes are robust in size and significance in this sample: using our preferred specification, we find that the presence of the third candidate raises the share of candidate votes by 6.6 percentage points on average (compared with 7.2 for the whole sample), decreases the vote share of the top two candidates by 8.2 percentage points on average (compared with 6.0 for the whole sample) and decreases the probability that the second-best wins by 32.3 percentage points on average (compared with 19.2 for the whole sample). As on the whole sample, the first two coefficients are significant at the 1 percent level and the third coefficient at the 5 percent level.

In sum, our results are not driven by elections where the fourth candidate qualifies and runs in the second round.

Table D2: Main results in elections where the 4th candidate never qualifies

Outcome	(1) Candidate votes 2nd-1st	(2) Vote share top 2 2nd-1st	(3) Second-best choice wins
3rd present	0.066*** (0.011)	-0.082*** (0.012)	-0.323** (0.149)
Observations	339	413	191
Polynomial order	1	1	1
Bandwidth	0.010	0.012	0.012
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.469	0.469	0.802

Notes: Sample includes all elections with first round turnout lower than 58%. In column (3), the sample is further restricted to elections where the three candidates are from distinct political orientations, the second-best is identified and the strength of the third candidate is lower than that of each of the top two candidates (see Section 5.3). In Columns (1) and (2), each outcome uses the number of registered voters as the denominator and is defined as a simple difference between the second and first rounds. In Column (3) the outcome is a dummy equal to 1 if the second-best wins the election. Standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Each column reports the results from a separate local polynomial regression. The variable of interest (the presence of the third candidate in the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate in the first round was higher than the threshold). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1 and the optimal bandwidths are derived under the MSERD procedure.

Appendix E. Model

Proofs of Lemmas and Propositions

Proof of Lemma 1.

We show that all voters in $\mathscr{G}_K$ enjoy the same benefit if J wins over I but that they differ in their ideological cost of voting for J instead of their favorite candidate.
i) For voter i in $\mathscr{G}_K$, the benefit of J winning over I is given by

$$w_i^{J,I} = b(d_i^I - d_i^J) = b(\mid \theta_i - \gamma_I \mid - \mid \theta_i - \gamma_J \mid).$$

As $\theta_i > \gamma_J > \gamma_I$ for all $i \in \mathscr{G}_K$, we get

$$w_i^{J,I} = b(\theta_i - \gamma_I - \theta_i + \gamma_J) = b(\gamma_J - \gamma_I) = w.$$

ii) For voter i in $\mathscr{G}_K$, the cost of voting for J instead of K is given by

$$\begin{aligned} c_i^{J,K} &= d_i^J - d_i^K \\ &= \mid \theta_i - \gamma_J \mid - \mid \theta_i - \gamma_K \mid \\ &= \begin{cases} \theta_i - \gamma_J + \theta_i - \gamma_K & if\, \theta_i \leq \gamma_K \\ \theta_i - \gamma_J - \theta_i + \gamma_K & if\, \theta_i \geq \gamma_K \end{cases} \\ &= \begin{cases} 2\theta_i - \gamma_J - \gamma_K & if\, \theta_i \leq \gamma_K \\ \gamma_K - \gamma_J & if\, \theta_i \geq \gamma_K. \end{cases} \end{aligned}$$

Proof of Proposition 1.

We show that there is no loss of generality in assuming that an optimal rule can be defined as a cutoff point.

Considering all voters in $\mathscr{G}_K$, note first that $c_i^{J,K}$ is weakly increasing in θ_i over $[y,1]$: $c_i^{J,K} = 2\theta_i - \gamma_J - \gamma_K$ strictly increases in θ_i for $\theta_i \in [y, \gamma_K]$, reaches a maximum at γ_K, and then remains constant and equal to $c_i^{J,K} = \gamma_K - \gamma_J$ for $\theta_i \in [\gamma_K, 1]$.

We next show by contradiction that an optimal rule defined as a cutoff on voters' bliss points weakly dominates any other voting rule. The following proof is adapted from Feddersen and Sandroni (2006b).

We define e^* as an optimal rule for voters in $\mathscr{G}_K$ and $e^*(\theta)$ denotes the action that a voter with a bliss point θ should take. Let θ_i^* be the largest bliss point in $[y; \gamma_K]$ such that for every bliss point $\theta \leq \theta_i^*$, $e^*(\theta) = \{$vote for $J\}$.

Suppose that $\theta_i^* \geq \gamma_K$. In this case the proof is complete: as $c_i^{J,K}$ remains constant for $\theta_i \in [\gamma_K, 1]$, any rule requiring the same positive measure of voters in $[\gamma_K, 1]$ to vote for J is equivalently costly and gives the same chance for J to win. Hence, a cutoff rule weakly dominates any other voting rule in this case.

Suppose that $\theta_i^* \in [y; \gamma_K)$. The proof would also be complete if $e^*(\theta) = \{\text{vote for } K\}$ for all bliss points $\theta_i > \theta_i^*$. So, consider a strictly positive measure set of bliss points $\overline{\Theta} \subseteq (\gamma_K; 1]$ such that $e^*(\theta) = \{\text{vote for } J\}$ if $\theta \in \overline{\Theta}$. By the definition of θ_i^* there is a strictly positive measure set of bliss points $\underline{\Theta}$ in $(\theta_i^*; \gamma_K]$ such that $e^*(\theta) = \{\text{vote for } K\}$ if $\theta \in \underline{\Theta}$. Pick a subset of $\underline{\Theta}$ and $\overline{\Theta}$ with the same strictly positive measure. Consider an alternative rule in which the required actions for the bliss points in $\underline{\Theta}$ and $\overline{\Theta}$ are exchanged, but that is otherwise identical to e^*. As $c_i^{J,K}$ is weakly increasing over $[y, 1]$ and strictly increasing over $[y; \gamma_K]$, the aggregate ideological cost for the group $\mathscr{G}_K$ would be reduced while the chance that J wins the election would remain the same. Thus, this rule would have a higher ranking than e^*, which contradicts the definition of e^*.

Hence, for $\theta_i^* \in [y; \gamma_K)$ (i.e., as long as there is a positive measure of voters to the left of K who vote for K), the optimal rule must be defined such that there exists no positive measure $\overline{\Theta}$ in $(\gamma_K; 1]$ such that $e^*(\theta) = \{\text{vote for } J\}$ if $\theta \in \overline{\Theta}$.

We now apply the same reasoning as above to show that the optimal rule to determine which voters should vote for J in $[y; \gamma_K]$ can be defined as a cutoff point in $[y; \gamma_K]$. Again, the proof would be complete if $e^*(\theta) = \{\text{vote for } K\}$ for all bliss points $\theta_i > \theta_i^*$. Hence, we consider a strictly positive measure $\widehat{\Theta} \subseteq (\theta_i^*; \gamma_K]$ such that $e^*(\theta) = \{\text{vote for } J\}$ if $\theta \in \widehat{\Theta}$. There exists $\bar{\theta}_i > \theta_i^*$ such that $e^*(\theta) = \{\text{vote for } J\}$ if $\theta \in \widehat{\Theta} \cap (\bar{\theta}_i; \gamma_K]$. By the definition of θ_i^*, there is a strictly positive measure set of bliss points $\widetilde{\Theta}$ in $(\theta_i^*; \bar{\theta}_i]$ such that $e^*(\theta) = \{\text{vote for } K\}$ if $\theta \in \widetilde{\Theta}$. Pick a subset of $\widetilde{\Theta}$ and $\widehat{\Theta} \cap (\bar{\theta}_i; \gamma_K]$ with the same strictly positive measure. Consider an alternative rule in which the required actions for the bliss points in $\widetilde{\Theta}$ and $\widehat{\Theta} \cap (\bar{\theta}_i; \gamma_K]$ are exchanged, but that is otherwise identical to e^*. As $c_i^{J,K}$ is strictly increasing over $[y; \gamma_K]$, the aggregate ideological cost for the group $\mathscr{G}_K$ would be reduced while the chance that J wins the election would remain the same. This rule would have a higher ranking than e^*, which contradicts the definition of e^*.

Hence a cutoff rule on voters' bliss points weakly dominates any other voting rule and there is no loss in generality in defining the optimal rule e^* by a cutoff point σ^*.

As the cutoff rule only weakly dominates other rules, the rule e^* defined in Proposition 1 is not necessarily unique. To achieve uniqueness, we need $c_i^{J,K}$ to be *strictly* increasing in θ_i over the entire support of $\mathscr{G}_K$: $[y; 1]$. This can be obtained by re-defining the ideological cost of voter i as $c_i^{J,K} = \mid \theta_i - \gamma_J \mid^{(1+\varepsilon)} - \mid \theta_i - \gamma_K \mid^{(1+\varepsilon)}$ with $\varepsilon > 0$. Our results would remain the same with this definition of the costs.

Proof of Lemma 2.

Strategic-naive voters believe that everyone in $\mathscr{G}_K$ is strategic (that $\alpha = 0$). Suppose that strategic-naive agents follow the rule $\sigma_n \in [y;1]$.

Candidate J is elected if she receives more votes than I. Strategic-naive voters believe that this occurs if

$$(y-x)+(\sigma_n-y) \geq x \iff \sigma_n \geq 2x,$$

where $y-x$ is the number of voters in $\mathscr{G}_J$ and $\sigma_n - y$ the number of voters in $\mathscr{G}_K$ located to the left of the cutoff.

So, they believe that candidate J is elected with probability

$$p_n(\sigma_n) = \begin{cases} 1 & if\, \sigma_n \geq 2x, \\ 0 & otherwise. \end{cases}$$

As strategic-naive voters believe that all voters in $\mathscr{G}_K$ and to the left of their cutoff will follow the rule and vote for J, they anticipate the following aggregate ideological cost for group $\mathscr{G}_K$:

$$C_n(\sigma_n) = \int_y^{\sigma_n} c_i^{J,K} d_i.$$

Proof of Lemma 3.

Given (1) and Lemma 2, if strategic-naive voters act according to the rule σ_n, they consider the following payoffs for voters in $\mathscr{G}_K$:

$$U_n(\sigma_n) = p_n(\sigma_n)W - C_n(\sigma_n) = \begin{cases} W - C_n(\sigma_n) & if\, \sigma_n \geq 2x, \\ -C_n(\sigma_n) & otherwise. \end{cases} \tag{2}$$

Assume that b is large enough so that $W > C_n(1)$ (or b higher than $\frac{\gamma_K^2 - y^2 - (\gamma_K - y)(\gamma_J + \gamma_K) + (1-\gamma_K)(\gamma_K - \gamma_J)}{(1-y)(\gamma_J - \gamma_I)}$). Following (2), strategic-naive voters will define a rule that ensures J's victory, if possible.

Given assumption (A2), we have

$$x \leq 1/2 \iff 2x \leq 1.$$

Hence, strategic-naive voters believe that they can set a threshold high enough to ensure the victory of J and their maximization program becomes

$$\underset{\sigma_n \in [y;1]}{Min}\; C_n(\sigma_n),$$

subject to $p_n(\sigma_n) = 1$.

Equivalent to

$$\underset{\sigma_n \in [y;1]}{Min} \int_y^{\sigma_n} c_i^{J,K} d_i,$$

subject to $\sigma_n \geq 2x$.

Equivalent to

$$\underset{\sigma_n \in [2x;1]}{Min} \int_y^{2x} c_i^{J,K} d_i + \int_{2x}^{\sigma_n} c_i^{J,K} d_i.$$

As $c_i^{J,K} > 0$ for all $i \in [2x;1]$, the cost is a monotonically increasing function of σ_n and we get

$$\sigma_n^* = 2x.$$

All strategic-naive voters follow the rule since the corresponding payoff D is higher than their cost of voting for J instead of K. Hence, a fraction $\sigma_n^* - y$ of them vote for J while a fraction $1 - \sigma_n^*$ vote for K.

Proof of Lemma 4.

Strategic-sophisticated voters are aware that expressive voters in $\mathscr{G}_K$ always vote for K and that the rule designed by the strategic-naive voters fails to take this into account. Hence, they are aware that $\alpha + \beta(1-2x)$ voters in $\mathscr{G}_K$ vote for K while $\beta(2x-y)$ of them vote for J.

Suppose that strategic-sophisticated agents follow the rule $\sigma_s \in [y;1]$.

Candidate J is elected if she receives more votes than I. This occurs if

$$(y-x) + \beta(2x-y) + (1-\alpha-\beta)(\sigma_s - y) \geq x \Longleftrightarrow \sigma_s \geq y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y),$$

where $(y-x)$ is the number of voters in $\mathscr{G}_J$, $\beta(2x-y)$ the number of strategic-naive voters in $\mathscr{G}_K$ located to the left of σ_n^* and $(1-\alpha-\beta)(\sigma_s - y)$ the number of strategic-sophisticated voters in $\mathscr{G}_K$ located to the left of their cutoff.

So, candidate J is elected with probability

$$p_s(\sigma_s) = \begin{cases} 1 & if\ \sigma_s \geq y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y), \\ 0 & otherwise. \end{cases}$$

Given the behavior of expressive and naive voters, strategic-sophisticated voters anticipate the following aggregate cost for the group $\mathscr{G}_K$

$$C_s(\sigma_s) = \beta C_n(\sigma_n^*) + (1-\alpha-\beta)\int_y^{\sigma_s} c_i^{J,K} d_i,$$

where $\beta C_n(\sigma_n^*)$ is the aggregate cost borne by strategic-naive voters who vote for J, which is independent from σ_s.

Proof of Lemma 5.

Given (1) and Lemma 3.1, it follows that if strategic-sophisticated voters act according to the rule σ_s, then the induced payoffs for voters in $\mathscr{G}_K$ are

$$U_s(\sigma_s) = p_s(\sigma_s)W - C_s(\sigma_s) = \begin{cases} W - C_s(\sigma_s) & if\ \sigma_s \geq y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y), \\ -C_s(\sigma_s) & otherwise. \end{cases} \tag{3}$$

We assume that b is large enough such that $W > C_s(1)$, (or b higher than $\frac{(1-\alpha)[\gamma_K^2 - y^2 - (\gamma_K - y)(\gamma_J + \gamma_K)] + (1-\alpha-\beta)(1-\gamma}{(1-y)(\gamma_J - \gamma_I)}$
Given (3), strategic-sophisticated voters will define a rule that ensures J's victory if possible. Define g as the difference between the fractions of voters in $\mathscr{G}_I$ and $\mathscr{G}_J$: $g = x - (y-x) = 2x - y$ and $g^* = \left(1 - \frac{\alpha}{1-\beta}\right)(1-y)$. We have

$$y + \frac{(1-\beta)}{(1-\alpha-\beta)}(2x-y) \leq 1 \Longleftrightarrow g \leq g^*.$$

Hence, strategic-sophisticated are able to ensure the victory of J if and only if $g \leq g^*$. Note that g^* is decreasing in α and β: $\frac{\partial g^*}{\partial \alpha} = -\frac{1-y}{1-\beta} < 0$ and $\frac{\partial g^*}{\partial \beta} = -\frac{\alpha(1-y)}{(1-\beta)^2} < 0$. The higher the share of expressive and naive voters, the more binding the condition is. The optimal cutoff chosen by strategic-sophisticated depends on whether they can ensure the victory of J or not.

i) if $g \leq g^*$, strategic-sophisticated voters know they can make J win. Their maximization problem becomes

$$\underset{\sigma_s \in [y;1]}{Min}\ C_s(\sigma_s),$$

subject to $p_s(\sigma_s) = 1$.

Equivalent to

$$\underset{\sigma_s \in [y;1]}{Min}\ \int_y^{\sigma_s} c_i^{J,K} d_i,$$

subject to $\sigma_s \geq y + \frac{(1-\beta)}{(1-\alpha-\beta)} g$.

Equivalent to

$$\underset{\sigma_s \in \left[y+\frac{(1-\beta)}{(1-\alpha-\beta)}g;1\right]}{Min} \int_y^{y+\frac{(1-\beta)}{(1-\alpha-\beta)}g} c_i^{J,K} d_i + \int_{y+\frac{(1-\beta)}{(1-\alpha-\beta)}g}^{\sigma_s} c_i^{J,K} d_i.$$

As $c_i^{J,K} > 0$ for all $i \in [y + \frac{(1-\beta)}{(1-\alpha-\beta)}g; 1]$, the cost is a monotonically increasing function of σ_s and we get

$$\sigma_s^* = y + \frac{(1-\beta)}{(1-\alpha-\beta)}g.$$

Note that σ_s^* is larger than σ_n^*: $\sigma_s^* = y + \frac{(1-\beta)}{(1-\alpha-\beta)}g = 2x + \frac{\alpha}{(1-\alpha-\beta)}g \geq 2x = \sigma_n^*$.

All strategic-sophisticated voters follow the rule as the corresponding payoff D is higher than their cost of voting for J instead of the third candidate. Hence, if $g \leq g^*$, a fraction $\sigma_s^* - y$ of them vote for J while a fraction $1 - \sigma_s^*$ vote for K.

ii) If $g > g^*$, strategic-sophisticated voters know they cannot make J win. Hence, they only seek to minimize costs and their maximization problem becomes

$$\underset{\sigma_s \in [y;1]}{Min} C_s(\sigma_s).$$

Equivalent to

$$\underset{\sigma_s \in [y;1]}{Min} \int_y^{\sigma_s} c_i^{J,K} d_i,$$

As $c_i^{J,K} \geq 0$ for all $i \in [y; 1]$, the cost is a monotonically increasing function of σ_s and we get

$$\sigma_s^* = y.$$

Hence, if $g > g^*$, all strategic-sophisticated voters vote for K.

Proof of Proposition 2.

The proof of Proposition 2 directly follows from Lemma 3 and Lemma 5.

Predictions of the model in two special cases

Predictions of the model with only strategic voters.

Consider the case in which there are only strategic voters ($\alpha = 0$).

Strategic-naive voters believe everyone is strategic. Therefore, their behavior is unaffected by α and, according to Lemma 3, they define the cutoff $\sigma_n^* = 2x.$ as

Strategic-sophisticated voters have accurate beliefs and thus they know everyone is strategic. For $\alpha = 0$, we have: $g^* = 1 - y$. By Assumption (A1),

$$x \leq 1/2 \iff 2x \leq 1 \iff 2x - y \leq 1 - y \Rightarrow g \leq g^*$$

Hence, strategic-sophisticated voters realize they can always make J win. Following Lemma 5, they define the optimal cutoff

$$\sigma_s^* = y + \frac{(1-\beta)}{(1-\alpha-\beta)} g.$$

Replacing $\alpha = 0$, we have:

$$\sigma_s^* = y + g = 2x.$$

Hence both strategic-naive and strategic-sophisticated voters define the same cutoff. As they get a payoff D higher than their cost if they follow the rule, all strategic agents (and thus all voters in $\mathscr{G}_K$) with a bliss point between y and $2x$ vote for J.

As a result, J gets $(y-x)+(2x-y) = x$ votes and wins over I.

Predictions of the model with only expressive voters.

Consider the case in which there are only expressive voters ($\alpha = 1$).

In this case, all voters in $\mathscr{G}_K$ vote for K.

As a result, J gets $(y-x)$ votes. Given A2, $(y-x) < x$. Hence, J loses and I wins the election.

Shares of expressive, strategic-naive and strategic-sophisticated voters.

We solve a system of two equations in two unknowns, α and β:

$$\begin{cases} g^* = \left(1 - \frac{\alpha}{1-\beta}\right)(1-y) & (4) \\ f(g) = 1 - y - \beta g \text{ for } g > g^* & (5) \end{cases}$$

where $f(g)$ is the vote share of candidate K. Equation [4] comes from Lemma 5 and Equation [5] from Proposition 2 (ii).

We estimate each quantity on elections close to the threshold in which assumptions (A1) and (A2) are satisfied.

We begin by solving equation [5] to obtain β. Equation [5] holds only for observations for which $g > g^*$. Hence, we first need to estimate g^*: the difference between the strengths of I and J below which J wins the election and above which I does. We regress the difference in vote shares between I and J in the second round on the difference between their strengths and estimate g^* as the smallest value of the regressor for which I is predicted to win: $g^* = 0.041$. We then estimate g, $(1-y)$ and $f(g)$ in elections for which $g > 0.041$. We compute g as the average gap between the strength of I and J in elections where this gap is larger than 0.041. We get: $g_{g>g^*} = 0.112$.

In the model, we consider only voters who would vote whether the third candidate is present or not. Hence, to estimate $(1-y)$ and $f(g)$, we need to subtract the fraction of loyal voters from her support and from her second round vote share respectively. We know from section 4.3 that at least 27.3 percent and up to 54.5 percent of voters who vote for the third candidate in the second round are loyal. We choose to focus on the lower bound. Hence, the fraction of loyal voters in elections for which $g > 0.041$ is equal to $0.209 \times 0.273 = 0.057$, where 20.9 percent is the share of voters voting for the third candidate in the second round in elections for which $g > 0.041$.

We estimate $(1-y)$ by averaging the strength of the third candidate in elections for which $g > 0.041$, minus the fraction of loyal voters: $(1-y)_{g>g^*} = 0.248 - 0.057 = 0.191$.

We estimate $f(g)$ as the average vote share of the third candidate in the second round in elections where $g > 0.041$, minus the fraction of loyal voters. We get: $f(g)_{g>g^*} = 0.209 - 0.057 = 0.152$.

From equation [5] we have
$\beta = \frac{1-y-f(g)}{g}$ for $g > g^*$.
Replacing by the values computed above, we obtain
$\beta = \frac{0.191-0.152}{0.112} = 0.348$.

We next turn to equation [4] to obtain α. We first compute the fraction of loyal voters in our sample by multiplying the share of voters voting for the third candidate in the second round by the lower bound of the share of loyal voters. We get $0.207 \times 0.273 = 0.057$. We estimate $(1-y)$ by averaging the strength of the third candidate, minus the fraction of loyal voters: $(1-y) = 0.247 - 0.057 = 0.190$.

From equation [5] we have
$\alpha = (1 - \frac{g^*}{1-y})(1-\beta)$
Replacing by the values computed above along with the estimated β, we obtain
$\alpha = (1 - \frac{0.041}{0.190})(1-0.348) = 0.511$.

Note that we obtained these estimates using the lower bound of the share of loyal voters among voters who voted for the third candidate in the second round. We can alternatively use the upper bound by replacing 0.273 with 0.545 in the above computations. This leads to similar results: we obtain the same beta ($\beta = 0.348$) and a slightly lower alpha ($\alpha = 0.450$).

Appendix F. Political orientations

Political labels are attributed by the French Ministry of Interior. Tables below show how we allocate each political label to one of our six political orientations for each election and year. The 1978 and 1981 parliamentary elections are shown together as the political parties competing in both elections were identical.

1978 and 1981 parliamentary elections	
Political label	**Political orientation**
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Parti Communiste Français	Left
Parti Socialiste	Left
Rassemblement Pour la République	Right
Union pour la Démocratie Française	Right
Non Classés	Other
Indépendants	Other

1988 parliamentary elections	
Political label	**Political orientation**
Communiste	Left
Divers Droite	Right
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Majorité Présidentielle	Left
Radical de Gauche	Left
Régionalistes	Other
Rassemblement Pour la République	Right
Parti Socialiste	Left
Union pour la Démocratie Française	Right

1993 parliamentary elections	
Political label	**Political orientation**
Communiste	Left
Divers	Other
Divers Droite	Right
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Gestion Ecologie	Other
Majorité Présidentielle	Left
Radical de Gauche	Left
Régionalistes	Other
Rassemblement Pour la République	Right
Parti Socialiste	Left
Union pour la Démocratie Française	Right
Europe Ecologie les Verts	Left

1997 parliamentary elections	
Political label	**Political orientation**
Communiste	Left
Divers	Other
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Parti Radical Socialiste	Left
Rassemblement Pour la République	Right
Parti Socialiste	Left
Union pour la Démocratie Française	Right

2002 parliamentary elections	
Political label	**Political orientation**
Communiste	Left
Chasse, Pêche, Nature et Traditions	Right
Divers	Other
Démocratie Libérale	Right
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Ligue Communiste Révolutionnaire	Far-left
Lutte Ouvrière	Far-left
Mouvement National Républicain	Far-right
Mouvement Pour la France	Right
Pôle Républicain	Left
Radical de Gauche	Left
Régionalistes	Other
Rassemblement Pour la France	Right
Parti Socialiste	Left
Union pour la Démocratie Française	Center
Union pour un Mouvement Populaire	Right
Europe Ecologie les Verts	Left

2007 parliamentary elections	
Political label	**Political orientation**
Communiste	Left
Chasse, Pêche, Nature et Traditions	Right
Divers	Other
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Majorité présidentielle	Right
Mouvement Pour la France	Right
Radical de Gauche	Left
Régionalistes	Other
Parti Socialiste	Left
Union pour la Démocratie Française -Mouvement Démocrate	Center
Union pour un Mouvement Populaire	Right
Europe Ecologie les Verts	Left

2011 local elections	
Political label	**Political orientation**
Autres	Other
Communiste	Left
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Majorité présidentielle	Right
Nouveau Centre	Right
Modem	Center
Parti de Gauche	Left
Radical de Gauche	Left
Régionalistes	Other
Parti Socialiste	Left
Union pour un Mouvement Populaire	Right
Europe Ecologie les Verts	Left

2012 parliamentary elections	
Political label	**Political orientation**
Alliance Centriste	Center
Autres	Other
Centre pour la France	Center
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Other
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front de Gauche	Left
Front National	Far-right
Nouveau Centre	Right
Parti Radical	Right
Radical de Gauche	Left
Régionalistes	Other
Parti Socialiste	Left
Union pour un Mouvement Populaire	Right
Europe Ecologie les Verts	Left

2015 local elections	
Political label	**Political orientation**
Communiste	Left
Divers	Other
Debout la France	Right
Divers Droite	Right
Divers Gauche	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front de Gauche	Left
Front National	Far-right
Modem	Center
Parti de Gauche	Left
Radical de Gauche	Left
Parti Socialiste	Left
Union Centriste	Center
Union pour la Démocratie	Right
Union des Démocrates et Indépendants	Right
Union de Gauche	Left
Union pour un Mouvement Populaire	Right
Europe Ecologie les Verts	Left