I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15589

Is Demonstrating against the Far Right Worth It? Evidence from French Presidential Elections

Nicolas Lagios
Pierre-Guillaume Méon
Ilan Tojerow

SEPTEMBER 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15589

Is Demonstrating against the Far Right Worth It? Evidence from French Presidential Elections

Nicolas Lagios
ULB, CEBRIG and DULBEA

Pierre-Guillaume Méon
ULB, CEBRIG and DULBEA

Ilan Tojerow
ULB, CEBRIG, DULBEA and IZA

SEPTEMBER 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Is Demonstrating against the Far Right Worth It? Evidence from French Presidential Elections*

We study the electoral impact of protesting against the far right by investigating the demonstrations held during the 2002 French presidential elections against far-right candidate Jean-Marie Le Pen. Instrumenting rally attendance with rainfall while factoring in that some municipalities never host protests, we find that larger protests reduced both the number of votes for Le Pen and the number of abstentions, while increasing the number of votes for Chirac. Regarding the mechanisms behind these results, we show that protests reduced the social desirability of voting for Le Pen, the support for his policies, and generated spatial spillovers through local media.

JEL Classification: D72

Keywords: protest, election, demonstration, far right, populism

Corresponding author:
Ilan Tojerow
Université Libre de Bruxelles
Avenue F.D. Roosevelt, 50
1050 Brussels
Belgium
E-mail: ilan.tojerow@ulb.be

* We are grateful to Roland Hodler, Laurent Simula, Bruno Pessoa Carvalho, Laura Duthilleul, Nanna Lauritz Schönhage, Yiqing Xu, Sébastien Fontenay, Abel François, Andreas Madestam, Rikhil Bhavnani, Apurav Yash Bhatiya, Stephan Schneider, Gabriel Leon-Ablan, Dominic Rohner, Oliver Feltham, and the participants of the Economics and Politics Workshop in Brussels, the Silvaplana Workshop in Political Economy in Pontresina, the Workshop on Empirical Public and Political Economics in Leuven, the Annual Meeting of the European Public Choice Society, the Annual ISNIE/SIOE Conference, the Beyond Basic Questions workshop in Nuremberg, the Workshop on the Politicial Economy of Protests and Demonstrations in Brussels, the Quantitative Political Economy Workshop at King's College London in London, and the Institutional and Organizational Economics Academy in Cargèse for useful comments. We also thank Simon Legendre for his research assistance. This research benefitted from a "fonds exceptionnel de recherche" grant from Université libre de Bruxelles (ULB). Remaining errors and approximations are the sole responsibility of the authors.

1 Introduction

Far-right parties have seen their influence grow in countries as diverse as Austria, Brazil, the Netherlands, France, Hungary, Germany, the United Kingdom, and the United States. Although each has unique characteristics, these parties — generally described as far right, radical right, or right-wing populist — share common traits: anti-immigration policies, nationalist or even xenophobic or racist positions, an anti-elite discourse, charismatic leaders, and a defense of traditional values (Guriev and Papaioannou, Forthcoming). Another common characteristic is that they have prompted protests by their opponents not only after their election but also preemptively, with the aim to reduce electoral support for the party's candidate. Protests against Donald Trump started during his 2016 campaign and marked his term in office. In Brazil, the Ele Não movement, which means "not him", was created during the 2018 presidential campaign to protest Jair Bolsonaro and his campaign. In Italy, the Sardines movement was set up to oppose far-right politician Mateo Salvini and his party Lega Nord. In Hungary, Viktor Orban's 2018 reelection was also met with protests. Can these protests reach their stated objective to reduce the influence of far-right parties and, if so, how? These are the questions we address in this paper.

We study the 2002 French presidential election that saw far-right candidate Jean-Marie Le Pen reach the second round of voting. In the first round, on April 21, Jean-Marie Le Pen had to the surprise of all obtained the second largest number of votes behind Jacques Chirac, the right-of-center incumbent. Jean-Marie Le Pen thus became the first far-right candidate to make it to the second round of presidential elections in the history of the French Fifth republic, potentially challenging the social stigma attached to the far-right party and its ideas, as observed by Bursztyn, Egorov, and Fiorin (2020) in the case of Donald Trump's election. The second round took place on May 5. Four days before, on May 1, around 300 peaceful protests against Jean-Marie Le Pen and

his party, Front National, gathered more than one and half million participants across the country.

The 2002 French election provides an ideal natural experiment to study the effect of protests against a far-right candidate for at least three reasons. First, it is a case where a far-right candidate obtained an unexpected electoral success potentially challenging the prevailing social norm but also prompting protests on a single day. Second, France's presidential election is a two-round system, and the protests took place between the two rounds of the election. This means that the first round can be used as a measure of the initial performance of the candidates in each municipality to finely evaluate the movement of votes between the two rounds caused by the protests that took place in between. Moreover, as the protests took place only four days before the second round, this enables us to rule out certain mechanisms and threats to identification. Third, the weather on the day of the protests varied from beautiful to rainy across municipalities. Under the assumption that rainfall is uncorrelated with determinants of electoral outcomes other than protests, differences in rainfall across municipalities on the day of protests can be used as exogenous source of variation in rally attendance, which allows us to build upon Madestam et al.'s (2013) identification strategy to develop an innovative two-part model (Belotti et al., 2015, Cameron and Trivedi, 2010).

In this two-part model, we construct from two different data-generating processes a rain-based synthetic instrument that factors in the probability of a protest occurring and the number of participants conditional on that probability. The key advantage of the model is that it accounts for the fact that some municipalities are unlikely to host a protest. In particular, because of a municipality's size, residents wishing to protest may typically join a protest in a larger nearby municipality instead.

We also factor in the fact that the effect of rain on the day of the protests is conditional on the average rain in a municipality on the day of the protests, since residents living in a municipality that does not typically get rain on that

day are more likely to be deterred by it from attending a protest. We do so by interacting weather conditions on the day of the protest with historical rain frequency on the same day, which is another innovation.

Our results show that protests were effective in impacting the outcomes of the second round of the election. Specifically, we observe that larger protests in a municipality resulted in a lower score for far-right candidate Jean-Marie Le Pen and a higher score for right-of-center candidate Jacques Chirac in that municipality. Larger protests also resulted in fewer voters deciding to abstain or cast a blank ballot. These findings stand up to a series of robustness checks addressing spatial correlation, outliers, the way we code rally attendance, how we construct the rain-based instrument, and the way we define a rainy protest. Our results are also robust to relaxing the exclusion restriction, thus making them immune to Mellon's (2021) criticism that many weather-based instrumental variables fail to meet the exclusion restriction.

Most of all, we investigate the different mechanisms that extant theory has put forward to explain the effect of protests and that may explain our baseline results. We first focus on how information generated by protests is channeled by the local press raising public awareness around the issue at stake (Lohmann, 1994, Battaglini, 2017, Wasow, 2020). To do so, we condition the effect of the number of participants on the reach, or dissemination of news, of the local press. Our results show that the effect of protests was larger in municipalities with a larger diffusion of the local press, suggesting that the local media played a role in conveying information about the protests, in line with García-Jimeno et al. (2022). We then test whether the effects of protests spilled over to other municipalities using a spatial model. We find that they did. Taken together, those results suggest that protests have an informational content that is channeled at least partly by local media, which is in line with the logic of Lohmann (1994), Battaglini (2017), and Wasow (2020). Our results also suggest that the effect of protests can go beyond the local networks observed by Madestam et al. (2013).

We then leverage individual survey data to understand how protests affected the behavior of individual voters. We observe that larger protests increased the probability of both self-declared left- and right-wing voters voting for Jacques Chirac in the second round. However, those voters faced different trade-offs. Specifically, larger protests reduced the probability of right-wing voters voting for Jean-Marie Le Pen but had no statistically significant effect on their probability of casting a blank or invalid ballot. By contrast, larger protests reduced the probably of left-wing voters casting a blank or invalid ballot but had no statistically significant effect on their probability of voting for Jean-Marie Le Pen. These findings suggest that some right-wing voters swung between extreme and moderate right-wing candidates whereas some left-wing voters swung between casting a blank or invalid ballot and voting for the candidate they viewed as the lesser of two evils.

We observe, moreover, that respondents living in municipalities with larger protests showed less support for the views and policies supported by Jean-Marie Le Pen and his party. In line with the models of Lohmann (1994) and Battaglini (2017), we can interpret these results as implying that protests served as a signal to voters of the potential negative consequences of voting for the far right.

Most of all, we find suggestive evidence that the number of participants in the protests emphasized a social norm that voting for the far right was socially undesirable. Bursztyn, Egorov, and Fiorin (2020) argue that election results can signal the prevalence of a social norm in the population, hence changing individuals' behaviors. We argue that protests can play the same role and even compensate the signal sent by an unexpected electoral outcome. A larger number of participants in protests against a candidate can signal that they consider voting for that candidate inadequate. In the 2002 election, the unexpected outcome of the first round could have challenged the norm of not voting for a far-right candidate. By contrast, the success of the May 1 protests could reinforce the original social norm.

In line with that presumption, we observe that larger protests reduced the probability of respondents declaring that they had voted for Jean-Marie Le Pen in the *first* round of the election. As the survey was carried out after the *second* round and protests took place ten days after the first round, they could by construction not affect votes in the first round. Conversely, we find no effect of protests on reporting a vote for Jacques Chirac or a blank or invalid ballot. We can therefore conclude that protests affected respondents' willingness to reveal that they had voted for the far-right candidate. This finding echoes the evidence that the election of Donald Trump affected social norms surrounding sexism and racism (Bursztyn, Egorov, and Fiorin, 2020, Giani and Méon, 2021). Although the finding pertains to votes reported in a survey as opposed to actual votes, it is suggestive that protests can sway voters by reinforcing the social norm that voting for the far right is undesirable. To our knowledge, our paper is the first to suggest and empirically illustrate this mechanism.

Overall, the present paper contributes to several strands of literature. The first is the literature on populism that has grown by documenting a series of determinants of support for populist parties or politicians (Guriev and Papaioannou, Forthcoming). However, this strand of literature has so far overlooked the role of the public reaction opposing those parties. We show that the 2002 protests reduced support for the populist candidate.

The second strand of literature to which we contribute is research on the consequences of protests. A primary question of this literature is whether protests can achieve their stated goals. Evidence from the US on the effect of protests against racial and gender discrimination suggest that while protests can increase the perception of discrimination and the support to policies to reduce it, the protests simultaneously can result in polarization (Mazumder, 2018, Reny and Newman, 2018, Wasow, 2020, Larreboure and González, 2021, Klein Teeselink and Melios, 2022). Our findings show that the French protests affected the outcome of the election as intended. We also complement the empirical contribution of Madestam et al. (2013) and Snyder and Yousaf

(2020), who studied the effect of rallies *in favor* of a party or a candidate by studying the effect of a protest *against* a party and its candidate. To our knowledge, our paper is the first to directly study the effect of a protest *against* a candidate on his electoral performance.

We complement the empirical literature on the effects of protests by documenting mechanisms through which they operate. In line with the theoretical contributions of Lohmann (1994) and Battaglini (2017), we report evidence that protests effectively conveyed information to the public. Further, our finding that the effect was larger in areas where more local newspapers were available is consistent with Wasow's (2020) agenda seeding theory.[1]

Most of all, we suggest a new channel of transmission of the effects of protests: the signaling of a social norm. We thereby contribute to the rising literature on changes in social norms (Bursztyn, González, and Yanagizawa-Drott, 2020, Bursztyn, Egorov, and Fiorin, 2020, Giani and Méon, 2021). The role of social norms has been documented for abstention using field experiments by Gerber et al. (2008) and Della Vigna et al. (2016) and for election outcomes by Bursztyn, Egorov, and Fiorin (2020) and Giani and Méon (2021). We report evidence suggesting that the protests in France had a similar effect. In so doing, our paper contributes to the literature on social norms and social desirability by showing that protests can reinforce a norm that may have been challenged by an electoral outcome.

The paper also makes a methodological contribution to the literature on protests by showing how a two-part model à la Belotti et al. (2015) and Cameron and Trivedi (2010) can be used to account for the fact that protests do not occur randomly across municipalities. That model allows us to leverage exogenous

[1] Another mechanism posits that protests change votes by creating or strengthening networks of activists, in line with the model of Murphy and Shleifer (2004) and the findings of Pons (2018) and Pons and Liegey (2019) on canvassing. In accordance with this theory, Madestam et al. (2013) have shown that the demonstrations of the "Tax Day" that launched the Tea Party in the US allowed it to develop its network of activists and tilted voters in favor of the Republican Party. Since only four days elapsed between the protests and the election, and campaigning in the last two days before an election is banned in France, networks of activists likely played a minor role in our case.

variations in weather conditions affecting the number of participants, while accounting for the fact that some municipalities are unlikely to ever host a protest.

The remainder of the paper is organized as follows. The next section describes the historical and institutional context of the 2002 French presidential election. The third section introduces the data, and the fourth section presents the empirical methodology. The fifth section reports and discusses the baseline results. The sixth section investigates the role of the local press in spreading the effect of protests and their spatial spillovers. The seventh section uses survey data to better understand how individual voters reacted to the protests and suggests a series of transmission channels. The eighth section concludes.

2 Historical context

Since 1965, the president of the French republic has been elected in a two-round direct election. Unless a candidate garners an absolute majority in the first round, which to date has never happened, the second round pits the two candidates who received the most votes against one another in an election that takes place two weeks after the first round. Since 2002, the president's mandate is for five years.

As the requirements for running for president are not stringent, the number of candidates in the first round can be large. In 2002, there were 16 total candidates in the first round. Two candidates were far-right (Jean-Marie Le Pen and Bruno Mégret), four were right (Jacques Chirac, Jean Saint-Josse, Christine Boutin, and Alain Madelin), two were centrist (Corinne Lepage and François Bayrou), five were left (Christiane Taubira, Noel Mamere, Lionel Jospin, Robert Hue, and Jean-Pierre Chevènement), and three were far-left (Daniel Gluckstein, Arlette Laguiller, and Olivier Besancenot). The two candidates who were considered as the most likely contenders for the second round were incumbent president Jacques Chirac and the incumbent prime minister, Lionel Jospin (Lewis-Beck, 2004). They had shared the executive since 1997 following

an electoral defeat of the right-wing coalition supporting Jacques Chirac's party in a legislative election. The executive was therefore split between a right-wing president and a left-wing prime minister belonging to the socialist party.

On April 21, 2002, contrary to expectations, Jean-Marie Le Pen, leader of far-right xenophobic Front National beat Lionel Jospin in the first round with 16.86 percent of votes against Jospin's 16.18 percent of votes. As a result, the second round pitted Jean-Marie Le Pen against Jacques Chirac, who had gotten the most votes in the first round.[2]

Whereas moderate right-wing voters had a clear candidate around whom to rally, left-wing voters faced a moral dilemma: They had to vote for a candidate they disapproved of to prevent the election of a candidate they disapproved of even more. The matter was made worse by the fact that Jacques Chirac was suspected of corruption while he was the mayor of Paris and perceived as dishonest and corrupt. Hence, many left-wing voters may have been tempted to abstain or cast a blank or invalid ballot. The dilemma faced by left-wing voters is summarized in the title of an article published by left-wing daily newspaper Libération: "Vote for the crook, not the fascist!" (Libération, April 23, 2002).

Nonetheless, the main leaders of the left quickly called for citizens to vote for Jacques Chirac. Importantly, political parties and trade unions called for Labor Day protests to be devoted to expressing an opposition to far-right candidate Jean-Marie Le Pen. On Wednesday, May 1, 2002, around 300 protests were held across France, attracting around one and half million participants. Those protests were peaceful and festive. No violent incident was reported, and the press underlined their peaceful nature. For Instance, France 2, the main

[2] Unlike in legislative elections, third-place candidates in the first round cannot run in the second round of the presidential election, even if they have garnered more than 12.5 percent of votes. Pons and Tricaud (2018) have found that some voters use those third-place candidates to vote expressively. As there is no third-place candidate in presidential elections, voters had to choose between voting for Jacques Chirac, voting for Jean-Marie Le Pen, abstaining or casting a blank ballot.

public TV channel, described the protests as “good natured” on the 8:00 pm news.

The official political campaign ended two days later on Friday, May 3 at midnight, after which candidates were no longer allowed to campaign, and the media were forbidden from publishing surveys of voters. On Sunday, May 5, 2002, Jacques Chirac was re-elected in a landslide victory with 82.21 percent of votes against 17.79 for Jean-Marie Le Pen.

3 The data

Our dataset merges information on electoral outcomes, rally attendance, rainfall, and sociodemographic characteristics for all municipalities in mainland France.

Voting outcomes

The official results of the first and second rounds of the elections reported by the Ministry of the Interior were collected from the public data portal of the French government. For each municipality and each round, we observe the number of voters registered for the election, the vote share of each competing candidate, and the share of abstentions and blank and invalid ballots.

All figures are scaled down by the number of registered voters in the first round of the election. This allows us to interpret the variations in the scores of the candidates between the two rounds as changes in the absolute number of voters.

*** Insert Figure 1 around here ***

Figure 1 reports the election results of the two rounds averaged over all the municipalities in the sample. Figure 1 shows that turnout increased in the second round. Likewise, the shares of the two candidates increased, implying that both candidates attracted voters who had not voted for them in the first round while the share of voters who abstained or cast a blank or invalid ballot

decreased. Figure A1 in the appendix documents the geographical distribution of the vote share.

Rally attendance

Information on protests and rally attendance were drawn from national and local newspapers.[3] We hand-collected all news articles that were published between May 1 and May 18 and contained the word "manifestants", "manifestations", or "Front National". We then extracted information regarding the location of the protest and the number of attendees. Rally attendance for municipalities for which we found no news articles was set to zero. Figure 2 describes the size and the location of the protests identified in the sample.

*** Insert Figure 2 around here ***

Unsurprisingly, different news articles may report different turnouts for the same protest, as sources of information across newspapers can differ (e.g., some rely on police figures, and others use information from the organizers or from the Ministry of the Interior). As a result, figures may vary since some sources tend to exaggerate, while others tend to minimize rally attendance. We therefore use the average reported number of participants as our main measure of attendance.[4]

Rainfall

We obtained detailed data on rainfall from the public data portal of Météo France, the French national meteorological service. Our database reports information on rainfall (in millimeters) from 1,500 to 4,000 weather stations spread throughout mainland France from 1983 to 2002.[5] To be able to study the

[3] We contacted the Ministry of the Interior to obtain official figures, to no avail. In any case, the Ministry of the Interior would have likely focused on the protests taking place in the largest municipalities, while using the local press allows us to obtain information on protests that occurred in smaller municipalities.

[4] In the robustness checks, we show that the findings are robust to the uncertainty surrounding the number of participants.

[5] Rainfall is the amount of rain collected between 06H00 UTC on day D and 06H00 UTC the next day (D+1). The number of weather stations varies in function of the measurement date. Appendix Figure A2shows the localization of the weather stations used.

municipality level, we matched each municipality with the meteorological information of the closest weather station. For each municipality, we therefore observe the intensity of rainfall on the day of the protests — that is, on May 1, 2002 — as well as the intensity of rainfall that occurred the same day from 1983 to 2001. In our baseline specifications, we use a binary measurement of rainfall: According to the official definition of Météo France, a day (and, therefore, a protest) is said to be rainy in Metropolitan France if it rained more than 1 millimeter.[6]

Sociodemographic data

We complemented the dataset with sociodemographic data on each municipality, including: population size, average age, landmass, variation in the unemployment rate between 1995 and 2002, and the proportion of individuals aged 60 years or older. We also created a dummy $Parisian\ agglomeration$ that takes the value 1 if the municipality is part of the Parisian agglomeration and 0 otherwise, as well as created a dummy $Special\ administrative\ status$ that takes the value 1 if the municipality is a canton capital, a prefecture, or a sub-prefecture.

Finally, we created four dummies which correspond to the four types of municipalities defined by France's National Institute of Statistics and Economic Studies (INSEE). Thus, a municipality can either be a city center, a suburb, an isolated municipality, or a rural area. It is a city center if the municipality represents more than 50% of the population of the urban unit. Suburbs are urban municipalities that are not city centers. When an urban unit is constituted by a single municipality, it is classified as an isolated municipality. A rural municipality is a municipality that does not belong to an urban unit.

[6] In the robustness checks, we show that the results are robust to alternative definitions of a rainy protest.

4 Empirical framework

4.1 Identification strategy

To assess the impact of the protests against Jean-Marie Le Pen on the voting outcome of the second round, we estimate variants of

$$y_{2,m} = \zeta_0 + \zeta_1 y_{1,m} + \zeta_2 \ln(Participants_m) + \xi_m + \epsilon_m, \quad (1)$$

where

- $y_{2,m}$ is the relevant post-rally voting outcome in the second round in municipality m. We use the vote share of Jean-Marie Le Pen, the vote share of Jacques Chirac, and the sum of abstentions and blank and invalid ballots. All figures are scaled down by the number of registered voters in the first round of the election and therefore range from 0 to 100.
- $y_{1,m}$ is the relevant voting outcome in the first round in municipality m;
- $Participants_m$ is a measure of the number of participants who rallied against Jean-Marie Le Pen on May 1, 2002, in municipality m;
- ξ_m is a vector of dummy variables coding municipality type (city center, isolated municipality, rural area or suburb); and
- ϵ_m is the error term.

Controlling for the initial performance of the candidates allows us to finely evaluate the change in votes between the two rounds that can be attributed to the between-round protests. In our baseline estimations, we use the logarithmic transformation of the number of participants as it allows us to conveniently interpret the effect of the prize as a semi-elasticity.[7]

Hence, the identification strategy is to compare voting outcomes across municipalities that hosted protests of different sizes. Doing so overlooks the nationwide effect of protests, which may for instance be driven by the coverage of protests by national media outlets. As the nationwide effect is likely

[7] We show in the robustness checks that the results are robust to alternative functional forms.

qualitatively similar to the cross-municipality effect, the estimates must be interpreted as lower bounds on the total effect of protests on nationwide voting outcomes.

The standard errors are clustered by department, which are the administrative division between municipalities and regions.[8] Clustering at that level allows for arbitrary dependance between municipalities of the same department.[9]

It may be argued that the same variables — for example, political preferences — drive both the score of the candidates and the number of participants in the protests. Similarly, the score of the far-right candidates may have affected the number of participants. To address potential endogeneity, we estimate Equation (1) using two-stage least squares (2SLS) with $\ln(Participants_m)$ instrumented by an exogenous instrument based on rainfall on the day of protests. In line with Madestam et al. (2013) and Wasow (2020), we assume that inclement weather conditions deter some potential participants from joining a protest. However, we depart from their approach by accounting for the possibility that some municipalities are unlikely to host a protest because their relatively small size means that their residents will typically join a protest in a larger nearby municipality.

Specifically, we construct a rain-based synthetic instrument by estimating a two-part model (Belotti et al., 2015, Cameron and Trivedi, 2010) where we allow the occurrence of a protest in a municipality on May 1, 2002, and the number of participants to be generated by two different processes. By factoring in the fact that the intensity of rainfall is irrelevant for predicting rally attendance in a municipality that never experience a protest, regardless of weather conditions, we avoid a weak instrument. If rainfall on May 1, 2002, is

[8] In mainland France, there are 94 departments. The average population is 619,697 (s.e. = 48,043), and the average landmass is 5,741 km^2 (s.e. = 202).

[9] Clustering at the department level assumes zero spatial correlation between departments. In Section 5.2, we investigate the sensitivity of the results when we depart from this assumption by computing standard errors that account for spatial dependance. In Section 6.2, we formally model the spatial dimension of the effect of protests.

uncorrelated with determinants of electoral outcomes other than protests, then the resulting 2SLS estimator is consistent and identifies the causal impact of the number of participants on electoral outcomes.

The first-stage equation thus reads:

$$\ln(Participants_m) = \pi_0 + \pi_1 y_{1,m} + \pi_2 z_m + \xi_m + \eta_m. \quad (2)$$

where z_m is our rain-based synthetic instrument described in the next section and η_m the error term.

The exclusion restriction underlying our approach rests on two assumptions. The first is that rainfall on the day of protests should be conditionally independent of individual or municipal characteristics affecting electoral behavior. We believe this assumption to be plausible as rainfall is uncorrelated with pre-rally electoral outcomes and municipality characteristics, as shown by Table B1 and Table B2 in Appendix B. In addition, leveraging survey data that we present in Section 7, we provide evidence that rain did not affect the qualitative composition of protests or cause different types of voters to mobilize in the second round (Table B3).[10] Finally, in Table B4, we observe that rain on the day of protests (May 1) is uncorrelated with rain on the day of the second round (May 5).

The second assumption is that rainfall should affect voting outcomes only through rally attendance. This assumption is reasonable because protests took place just four days before the vote, leaving little time or room for the weather to affect voting outcomes through a channel other than protests. In addition, following Mellon's (2021) advice, we surveyed the literature and

[10] We document this by showing that the sociodemographic characteristics of individuals in rainy municipalities who reported attending a demonstration and voting in the second round are similar to those in non-rainy municipalities. We consider the following characteristics: respondents' political orientation (left, right, or neither), gender (dummy equal to one if female), level of education (dummy equal to one if high school diploma or higher), employment status (dummy equal to one if currently working), income (dummy equal to one if earning more than sample median income), interest in politics (dummy equal to one if some or a lot of interest), whether they are a member of a political association, and religiosity (dummy equal to one if attend religious services more than once a month).

found no result suggesting that the weather four days before an election could affect its outcome.[11] Possibly, the weather *on* the day of the election may affect the mood of voters, as Meier et al. (2019) report. In Section 5.2, we show that controlling for the weather on the day of the vote does not affect our results. Realistically, however, we cannot rule out with certainty all violations of the exclusion restriction. In Section 5.2, we therefore also provide a more general discussion of the plausibility of the exclusion restriction. In particular, we show that the reduced-form estimates of electoral outcomes on the instrument, which is correctly identified even when the exclusion restriction is violated, leads to consistent results. Using Conley et al.'s (2012) method, we also show that our results remain similar even when the exclusion restriction is violated.

4.2 A rain-based synthetic instrument

In the two-part model, the probability of a protest occurring and the number of participants conditional on a protest occurring come from different data-generating processes. Specifically, in the first part, we let the characteristics of municipalities determine the probability of a protest taking place. In the second part, the variation in rainfall across municipalities provides an exogenous source of variation in rally attendance. Doing so increases the precision of the estimates and the strength of the instrument because it factors in the fact that weather conditions are irrelevant in municipalities that never host a protest.

The first part models the probability that municipality m has a protest — that is, a strictly positive number of participants. To model this, we use a binomial regression with a complementary log-log link function:

$$\Pr(Participants > 0 \mid X_1) = F\left(X_1'\alpha\right) = 1 - \exp\left\{-\exp\left(X_1'\alpha\right)\right\}, \qquad (3)$$

where X_1 is a vector of dummies classifying the municipality type (city center, isolated municipality, rural area, or suburb), and α the corresponding vector of

[11] Mellon (2021) surveys the literature using weather as an instrument and calls for more attention to the exclusion restriction.

parameters to be estimated. The subscript m denoting the municipality is omitted for convenience.

We expect city centers and isolated municipalities to be more likely than suburbs to experience a protest as city centers are generally larger, more populated, and more important areas, while the isolated municipalities are geographically further away from other large municipalities. Conversely, we expect rural areas to be unlikely to host protests as individuals will tend to rally in larger nearby municipalities.

Because we cannot rule out that the type of municipality also influences voting behavior, we control for municipality type at each stage of our estimations to restrict the analysis to variations within each type of municipality, as the instrument only needs to be exogenous conditional on covariates (Angrist and Pischke, 2008).

The second part of the model estimates the number of participants attending the protest as a function of rainfall, conditional on the protest taking place in the municipality. Specifically, we take the natural logarithm of $Participants$ and estimate the following semi-log linear regression[12]:

$$\begin{aligned} &\mathbb{E}(\ln(\text{Participants}) \mid \text{Participants} > 0, \mathbf{X}, \beta_1, \beta_2) \\ &\quad = \beta_0 + \beta_1 Rainy\ Protest + \beta_2 Historical\ Rain \\ &\quad + \beta_3 Rainy\ Protest \times Historical\ Rain + X_1' \gamma \\ &\quad + X_2' \delta + \epsilon, \end{aligned} \tag{4}$$

where $Rainy\ Protest$ is a dummy variable equal to 1 if it rained more than 1 millimeter in municipality m on the day of the rally. $Historical\ Rain$ is the probability that the rainfall in municipality m on that day was greater or equal to 1 millimeter, computed by counting the number of times it rained 1 millimeter

[12] We estimate a semi-log linear regression to increase the quality of the prediction as the number of participants is right-skewed (Cameron and Trivedi, 2010).

or more on May 1 between 1983 and 2001. Q is a vector of the rain variables, and β_1, β_2, and β_3 are parameters to be estimated.

X_1 is the vector of dummies classifying the municipality type also included in the first part, X_2 is a vector of pre-determined municipality characteristics (population size, average age, landmass, variation in the unemployment rate between 1995 and 2002, the proportion of individuals aged 60 years or older, whether the municipality belongs to the Parisian agglomeration, and whether the municipality has a special administrative status), and γ and δ are vectors of parameters to be estimated.

The interaction term between $Rainy\ Protest$ and $Historical\ Rain$ captures the fact that the effect on participation of rain on the day of protests depends on how common rain is in the municipality. Specifically, rain should have a more detrimental effect on rally attendance in municipalities that do not typically get rain on that day.

We define the instrument z to be used in Equation (2) as the number of log participants predicted by the model. That is, $z = (\hat{p}|X_1) \times \left(\ln(\widehat{Participants}) \,\middle|\, Participant > 0, Q, X_1, X_2\right)$, where $(\hat{p}|X_1)$ is the predicted probability that municipality m has a protest, estimated in Equation (3), and $\left(\ln(\widehat{Participants}) \,\middle|\, Participants > 0, Q, X_1, X_2\right)$ is the predicted number of participants in municipality m given that it experiences a protest, which is estimated in Equation (4). We provide more details on the construction and estimation of the synthetic instrument in Appendix B2.

4.3 Estimation of the two-part model

Figure 3 shows the results of the two-part model. The left-hand side reports the estimates of the first part, which models the probability of a protest occurring in a municipality by using the dummy variables classifying the municipality type. All dummies exhibit the expected sign. For example, the average marginal effect of the variable coding city centers is equal to 0.16, implying that city centers were 16% more likely to host a protest than a suburban municipality.

Overall, the findings are in line with the notion that the residents of small rural municipalities tend to join protests in neighboring large municipalities. Conversely, the residents of a municipality are more likely to protest in that municipality if they live in an isolated municipality.

*** Insert Figure 3 around here ***

The right-hand side of Figure 3 plots the effect of rain on the number of participants in municipalities that experienced a protest (i.e., the second part of the two-part model). In particular, we report the estimated marginal effect of rain on May 1, 2002, and its 90-percent confidence interval against historical rain frequency, together with the distribution of historical rain frequency.[13] It shows that the marginal effect of rain on rally attendance was negative and statistically significant in the vast majority of municipalities. This result is in line with our assumption that rain had a detrimental effect on attendance by dissuading some potential participants from joining a protest. For example, for the average municipality in terms of historical rain frequency, rainfall decreased rally attendance by approximatively 50% (0.46 log points), which is consistent with the estimates reported by Madestam et al. (2013). The graph further shows that the marginal effect of rain on attendance decreases as historical rain frequency increases and is statistically insignificant when the frequency is large enough. Figure 3 confirms our presumption that the effect of rain on May 1 was lower in municipalities more accustomed to rain and validates our instrument.

5 Baseline results

5.1 The effect of protests on electoral outcomes

We can now turn to the key results that pertain to the effect of protests on electoral outcomes. They are based on parsimonious specifications controlling only for the relevant first-round outcome and the municipality type. If the

[13] The raw coefficients are reported alongside the coefficients of the full set of variables used in the estimation of the two-part model in Table B5 of Appendix B3.

instrument is valid then the 2SLS estimator is consistent and the inclusion of additional control variables is not necessary to consistently identify the causal impact of protests on electoral outcomes.[14] The first stage is reported in Table C2 of Appendix C1 and shows that our rain-based synthetic instrument — the number of participants predicted by the two-part model — strongly correlates with the effective number of participants.

The results of the second stage are reported in Table 1. Each column is devoted to a different electoral outcome: specifically, the vote share of Jean-Marie Le Pen, the vote share of Jacques Chirac, and the share of abstentions and blank and invalid ballots. In all specifications, the F statistic for the excluded instrument is well above the rule-of-thumb of 10, meaning that the synthetic instrument is a strong instrument for rally attendance.

In all specifications, too, the coefficient of the relevant first-round voting outcome is positive and statistically significant at the one-percent level. Accordingly, voting outcomes unsurprisingly displayed persistence between the two rounds of the election.

The parameter of interest is, however, the coefficient of the number of participants during the protests. Column (1) of Table 1. Baseline results – The impact of protests on voting outcomes reports the effect of those protests on Jean-Marie Le Pen's vote share. The coefficient of the number of participants bears a negative sign that is significant at the one-percent level. Accordingly, larger protests reduced the far-right candidate's vote share.

As the number of participants is a log value in the model, the coefficient can be interpreted as a semi-elasticity. Accordingly, a one-percent increase in the number of participants in a protest in a municipality resulted in a 0.399

[14] Additionally, since z_i is constructed based on these covariates, including them in the 2SLS estimates will lead to severe collinearity issues and, hence, to unstable p-values and coefficients (James et al. 2013, Sheather, 2009). We would then not be able to fully trust the estimates. Table C1 in Appendix C1 shows that including all the socio-demographic covariates yields a variance inflation factor (VIF) of 21 for $\ln(Participants)$. A VIF higher than five generally indicates a problematic level of collinearity (James et al., 2013, Sheather, 2009). In light of the few benefits and high costs associated with the inclusion of those covariates in the 2SLS estimates, we do not include them.

percentage point lower vote share for Jean-Marie Le Pen in that municipality. The magnitude of the effect is therefore substantial without being implausibly large.

In Column (2), the dependent variable is the vote share for Jacques Chirac. Here, the number of participants exhibits a positive coefficient that is statistically significant at the one-percent level in all specifications. Accordingly, a larger number of participants in the May 1 protests increased the share and the number of voters who cast a ballot for the right of center candidate. The magnitude of the effect is again substantial without being implausibly large: A one-percent increase in the number of participants in a municipality increased Jacques Chirac's vote share by 0.818 percentage points.

Finally, Column (3) suggests that protests also affected the share of abstentions and blank and invalid ballots. The number of participants exhibits a negative coefficient statistically significant at the one-percent level, implying that a one-percent increase in the number of participants reduced the share of abstentions and blank and invalid ballots by 0.304 percentage points.

*** Insert Table 1 around here ***

As vote shares are scaled down by the number of registered voters in the first round of the election, which is by definition constant, changes in vote shares can also be interpreted in absolute terms. For instance, the negative marginal effect of the number of participants on the vote share of Jean-Marie Le Pen also means that larger protests reduced the absolute number of voters who cast a ballot for that candidate.

The results of Table 1, therefore, sketch a consistent picture: A larger number of participants in the May 1 protests did affect the outcome of the second round of the election in the way that participants hoped. Specifically, it reduced the share and the number of votes for the candidate they demonstrated against, Jean-Marie Le Pen, and increased the number of votes for his opponent, Jacques Chirac. In addition, it decreased the number of abstentions and blank and invalid ballots.

To better understand the quantitative meaning of the results, we did a back-of-the-envelope calculation to simulate the electoral results in the absence of protests. We found that Jean-Marie Le Pen would have gained 0.9 to 2.8 points in the second round, Jacques Chirac would have lost 2.9 to 5.3 points, and the number of abstentions and blank and invalid ballots would have been higher by 1 to 5.9 points. The lower-bound estimates are based on a specification that overlooked the spatial dimension of protests, as is the case in this section, while the upper-bound estimates come from a specification that models spatial spillovers. We introduce this specification in Section 6.2.

5.2 Robustness checks

We now consider and discuss a series of robustness tests to investigate the strength of our baseline findings. Those additional analyses are reported in Appendix C2.

Specification

Theory gives little guidance as to the functional form of the relationship between rain, municipal characteristics, and rally attendance. To make sure that the baseline estimates are not driven by a misspecification of the two-part model, we estimate several alternative specifications that allow for more flexibility. First, we replace all continuous variables by quintile dummies to allow for non-linearities (see Table C3 in Appendix C). Second, we estimate the second part of the two-part model using a quantile regression for the 0.25, 0.5, and 0.75 quantiles (Table C4). We also run an additional regression where we only included demographic covariates: specifically, population size, average age, and the proportion of individuals aged 60 years or older (Table C5). Regardless of the specification, the baseline results remain unchanged.

Similarly, as the true functional form of the relationship between rally attendance and voting outcomes is unknown, we present results where we scale down the number of participants by the municipality's population (Table C6).

Regardless of the voting outcome considered, the results are in line with the baseline.

In our baseline estimates, we rely on a binary definition of rainfall whereby a protest is said to be rainy if rainfall exceeds one millimeter. To assess the sensitivity of our results to the way rainfall is defined, we consider three alternative rain variables. First, we use two alternative thresholds by defining a protest as rainy if rainfall exceeded two (Panel A of Table C7) or three millimeters (Panel B). Second, we define a protest as rainy if rain on the day of the protests was above the historical average measured between 1983 and 2001 (Table C8). Finally, we use a continuous measurement of rainfall alongside flexible measurements of average rainfall by quintile dummies (Table C9). In all cases, rain is negatively correlated with the number of participants, and the estimates of the effect of protests remain in line with baseline results.

Our baseline findings also rely on the use of the average number of reported participants as the dependent variable, as different newspapers sometimes reported different estimates of the number of participants in the same protest. We therefore present results using the maximum (Panel A of Table C10) and the minimum (Panel B) of the reported number of participants. We also assess the sensitivity of the results to dropping from the sample the five municipalities with the highest differential between the maximum and minimum number of reported participants (Table C11). The estimates are significant at conventional levels.

Finally, we restrict the sample in four ways. First, we trim the sample at the 95th percentile of the participants variable to show that our findings are not driven by outliers (Table C12). Second, as the size of French municipalities is right skewed, with a large number of very small municipalities, we exclude the municipalities that have a population lower than the 95 percentile to make sure that the skewed sample does not bias the results (Table C13). Third, we exclude Paris as it has a special role in French politics and geography (Table C14). Fourth, we drop from de sample the municipalities matched with a weather

station located more than 15 and 20 kilometers away (Table C15). In all cases, the results are insensitive to such changes.

Spatial correlation

By clustering standard errors at the department level, we have so far assumed that there was no correlation across departments. Yet, if protests and electoral outcomes in adjacent municipalities are correlated, this assumption may be violated, and the standard errors may be underestimated (Colella et al., 2019). To address this concern, we check the robustness of our findings by computing spatially corrected standard errors (Conley, 1999). This approach assumes that the error of each observation is correlated with the errors of other observations located within a given radius and uncorrelated with the errors of observations located beyond it. We test a series of radii: 50, 100, 250, 500, and 750 km (see Table C16). Regardless of the size of the radius, the magnitude of the spatially corrected standard errors remains similar to that of the clustered standard errors, which suggests that spatial correlation does not bias our baseline results.

The exclusion restriction

The key identification assumption of our approach is that, conditional on the municipality type, rain is not correlated with drivers of electoral behavior other than protests. As mentioned in Section 4.1, weather conditions can affect the mood of voters and thus have a direct effect on electoral outcomes, as observed by Meier et al. (2019). As a result, if weather conditions persisted from the day of the protests, May 1, to the day of the second round of the election, May 5, the exclusion restriction may be violated, as Mellon (2021) underlines. In Section 4.1, we showed that weather conditions on the day of protests are not correlated with those on the day of the second round, which limits the risk that the exclusion restriction is violated. As an additional test, we directly control for weather conditions on the day of the second round in the 2SLS estimation (Table C17). Doing so results in estimates that are qualitatively and quantitatively very similar to our baseline estimates.

Realistically, however, we cannot prove with certainty that our rain-based instrument affected electoral outcomes only through rally attendance. To mitigate this, we provide two additional results. First, we provide reduced-form estimates of electoral outcomes on the instrument. The main advantage of this method is it allows for correct identification even when the exclusion restriction is violated (Table C18). Second, we investigate how our estimates change if we relax the exclusion restriction by allowing the instrument to have a direct impact on electoral outcomes (Conley et al., 2012).[15] In all cases and for all outcomes, the interpretations remain unchanged.

6 How information spread

Most voters did not directly witness or participate in protests. Instead, they learned about them in the media or by directly or indirectly interacting with participants or witnesses of the protests. In this section, we investigate how the effect of protests spread to municipalities that did not host them. We start by investigating the role of the local press and then gauge spatial spillovers.

6.1 The role of the press

To gauge the role of the press, we leverage the variation across departments in the distribution of local newspapers.[16] Specifically, we condition the effect of protests on access to the local press by adding an interaction term between the number of participants in a protest in a municipality and the average distribution of local newspapers sold in the department of that municipality, scaled down by the number of inhabitants in the department. We interpret that measure of

[15] See Appendix D for a description of Conley et al.'s (2012) approach and its results.

[16] Distribution is measured by the average number of copies sold per month in a department. Local newspapers in France are typically distributed through one department or region, the exception being "Le Parisien", which has a local version for the Ile-de-France region around Paris as well as a national edition. We obtained data on newspapers' distribution in 2006 from the "Office de justification de la diffusion", which is a non-profit organization that certifies the distribution of newspapers and periodicals in France. Table E1 Appendix E1 provides descriptive statistics on newspapers.

distribution as a measure of the ease with which the inhabitants of a department can access information.

Since raw coefficients in interaction models are uninterpretable, Figure 4 plots the marginal effect of the number of participants as a function of the distribution of local newspapers available in the department of their municipality.[17]

*** Insert Figure 4 around here ***

The first and second panels of Figure 4 plot the marginal effect of protests on the vote shares of Jean-Marie Le Pen and Jacques Chirac. In both cases, the absolute size of the effect increases with the level of distribution of local newspapers. In addition, the effect is statistically insignificant in departments with very few local newspapers. Those results suggest that local newspapers were instrumental in conveying and amplifying information about protests.

By contrast, Panel C of Figure 4 suggests that interacting the level of distribution of local newspapers with protests does not improve the fit of the estimation when the dependent variable is the number of blank and invalid ballots. The result is even more striking when we allow for more flexibility in the interaction specification, as shown in Figure E1 in Appendix E2.

6.2 The geography of the effect

We have so far implicitly assumed that the effects of protests on electoral outcomes were confined to the municipalities in which they took place, which is in line with the findings of Madestam et al. (2013) on the effects of the Tea Party rallies. We now relax this assumption to determine whether the effect of protests on electoral outcomes spilled over to other municipalities.

[17] The outcome of those regressions is reported in Table E2 of Appendix E1. As noted by Hainmueller et al. (2019), multiplicative interaction models may be biased if the linear interaction effect assumption does not hold. To show that our results are robust to such a concern, in Appendix E2 we estimate a flexible 2SLS interaction model based on the binning estimator proposed by Hainmueller et al. (2019). The outcomes are summarized in Figure E1 and show a similar trend to our main results.

To do so, we estimate an extension of Equation (1) where the spatial lag of the number of participants, $\sum_{k=1}^{n} w_{mk} \ln(Participants_k)$, is introduced. Equation (1) thus becomes:

$$y_{2,m} = \alpha_0 + \alpha_1 y_{1,m} + \beta \ln(Participants_m) + \theta \sum_{k=1}^{n} w_{mk} \ln(Participants_k) + \xi_m + \epsilon_m, \quad (5)$$

where w_{mk} are elements of the $N \times N$ spatial weights matrix W defined as

$$W = \begin{pmatrix} 0 & w_{12} & \cdots & w_{1N} \\ w_{21} & 0 & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & 0 \end{pmatrix} \quad (6)$$

and

$$w_{mk} = \frac{1}{d_{mk}^{\gamma}}, \quad (7)$$

following Vega and Elhorst (2015). W quantifies the spatial interdependencies between municipalities[18], N is the number of municipalities, d_{mk} is the great circle distance between municipalities m and k's centroids, and γ is the distance decay parameter and is set to 0.5.[19] By introducing $\sum_{k=1}^{n} w_{mk} \ln(Participants_k)$, we allow the voting outcomes of each

[18] Following Kelejian and Prucha (2010), we normalize W by dividing each of its elements by the scalar $m = \min\left[\max_m \sum_{k=1}^{N} w_{mk}, \max_j \sum_{j=1}^{N} w_{mk}\right]$.

[19] The distance decay parameter γ indicates how the influence of a protest declines with distance. Here, we assume that this decline is less than proportional to distance. Figure F2 in Appendix F2 shows the sensitivity of the spatial spillover estimates to γ. For high values of γ, the coefficients explode in size, and the magnitude of the effect becomes implausible. This may result from the fact that high values of γ put too much weight on the role distance plays in the spatial relationship between municipalities.

municipality to be affected by the number of participants in protests in all other municipalities.[20]

Table 2 reports the estimates of Equation (5), and the first stage is reported in Table F5 of Appendix F2. We first focus on the direct impact of protests — that is, the effect of a change in the number of participants in municipality m on the voting outcomes of municipality m. In all specifications, the magnitude and the significance of the effect are very similar to the baseline.

We can now turn to the spillover effects, or the effect of a change in the number of participants in all municipalities *except* m on the voting outcomes of municipality m. The coefficients of the spatial lags are all statistically significant at the one-percent level, showing that the effects of protests were not confined to the municipalities that hosted a protest. Moreover, in line with previous findings, spillovers are negative for Jean-Marie Le Pen, negative for abstentions and blank ballots, and positive for Jacques Chirac. The effect of protests on neighboring municipalities was therefore qualitatively similar to their effect on the municipalities where they took place.

To test the robustness of those results to the way spillovers are modelled, we estimate in Table F6 of Appendix F2 a variant of the model where the elements of the distance matrix are defined like in (7) but are set to zero if the distance between two municipalities exceeds the mean of the distance matrix. Although smaller than in the previous series of estimations, the spillovers remain statistically significant at conventional levels. The finding is not specific to one definition of the distance power matrix.

[20] To generate an instrument for the spatial lag of the number of participants, we estimate the same two-part model as in the baseline but instead use the spatial lag of the rain and control variables in the second part of the model. Specifically, we regress the spatial lag of the number of participants on the spatial lag of rainfall (measured on a continuous scale as the spatial lag of a binary variable is meaningless), controlling for the spatial lag of average rainfall (quintile dummies) and an unweighted linear combination of the spatial lag of municipality's continuous characteristics (population, landmass, average age, variation in the unemployment rate between 1995 and 2002, and the proportion of individuals aged 60 years or older). Table F4 of Appendix F2 shows that the spatial lag of rain is negatively and significantly correlated with the spatial lag of the number of participants, in line with our assumption that bad weather conditions deter some individuals from joining a protest.

*** Insert Table 2 around here ***

In Appendix F, we complement our spatial model with a more intuitive approach used in a different context by Mamo et al. (2019) that consists in estimating the baseline model on observations aggregated at a higher geographic level, first by including all votes and next by dropping the votes of municipalities that hosted a protest. The difference between the two sets of estimates provides an estimate of the magnitude of spillovers. The results are in line with those of the spatial model we report in this section.

The upshot of this sub-section is that the effects of protests spread out of the municipalities that hosted them. This finding contrasts with Madestam et al. (2013) who observed that the effect of Tea Party rallies was only local. This suggests that the mechanisms at work in the protests that we study differed from those of the Tea Party rallies. The next section specifically tests those mechanisms.

7 Individual behaviors

To provide some evidence about the mechanisms that enabled protests to sway voters, we now use individual survey data from the 2002 wave of *Panel électoral français*, a survey carried out by three French research centers in the days following the second round of the election.[21] The survey features questions about political preferences, policy views, and votes in the two rounds. More than 4,000 respondents were interviewed. Their municipality of residence can be identified, which allows us to match them with protests and apply the same estimation strategy as in our baseline regressions to individual answers in the survey. Table 3 presents descriptive statistics for the voting variable.

In the first sub-section, we use the survey data to assess the effect of protests on reported individual votes in the second round and compare those

[21] Specifically, the survey is a joint effort of the Centre de recherches politiques de Sciences Po (CEVIPOF), Centre d'informatisation des données socio-politiques (CIDSP), and Centre d'études sur l'opinion publique (CECOP). It is available to researchers for free for research projects approved by the consortium.

results to baseline results to validate the survey. We then test whether the reactions to protests of voters was a function of their identification as left- or right-wing. In the second sub-section, we test the band-wagon effect and look at whether protests affected respondents' perception of the policies that Jean-Marie Le Pen advocated. We finally test whether the protests affected the social desirability of reporting having voted for the far-right candidate.

*** Insert Table 3 around here ***

7.1 The specific trade-offs of left- and right-wing voters

One may wonder if the effect was driven by left- or right-wing voters. In the 2002 election, left- and right-wing voters likely faced different moral dilemmas. Although voting for Jacques Chirac was the only way to reduce the vote share of Jean-Marie Le Pen and prevent him from winning the election, Jacques Chirac was unpalatable to left-wing voters because he was not only right-wing but had also been involved in various corruption scandals, particularly while he was the mayor of Paris. A popular slogan in the May 1 protests was "vote for the crook, not the fascist", and the media reported anecdotes of voters casting their ballot wearing plastic gloves or pinching their noses with a clothes peg (Libération, 2002). Left-wing voters could therefore be torn between voting for Jacques Chirac and abstaining or casting a blank ballot.

Conversely, some right-wing voters may have hesitated between voting for Jean-Marie Le Pen and Jacques Chirac. While still the leader of the main right-wing party UMP, Jacques Chirac had addressed the issue of immigration in a way that was likely directed to voters who could be tempted to vote for Jean-Marie Le Pen. The speech he had made in Orléans on June 19, 1991, in which he referred to the "noise and smell" of foreigners had particularly left a mark. Jean-Marie Le Pen had responded to it two days later by stating that "French voters would prefer the original to the copy" (Libération, 2019).

As a result, one may contend that the 2002 protests likely swayed left- and right-wing voters in different ways. We test for that possibility by separately

estimating a multinomial probit model for left-wing, center, and right-wing voters where the dependent variable could take three values: voting for Jean-Marie Le Pen, voting for Jacques Chirac, and casting a blank or invalid ballot. The number of participants in the respondent's municipality is instrumented by the logarithm of the number of participants predicted by the two-part model.[22] We controlled for the respondent's individual characteristics, their political orientation, and the type of municipality they live in.[23]

Figure 5 reports the average marginal effect of the number of participants on the probability of respondents to vote for Jean-Marie Le Pen, Jacques Chirac, or cast a blank ballot. It starts by reporting the results of an estimation pooling all respondents together, regardless of their political preferences. The results of that regression are in line with those of baseline estimations using true voting outcomes at the municipality level, lending credence to the new results obtained with the survey.

If we now focus on effects by group of voters, the figure shows that protests increased the propensity of both left- and right-wing respondents to vote for Jacques Chirac. However, the protests decreased the probability of left-wing respondents only of casting a blank or invalid ballot but did not affect their probability of voting for Jean-Marie Le Pen. Conversely, protests decreased the probability of right-wing respondents voting for Jean-Marie Le Pen. Protests did not affect the behavior of undeclared voters.

*** Insert Figure 5 around here ***

Those findings confirm that left- and right-wing voters faced different trade-offs and that protests prompted some left-wing voters to vote for Jacques

[22] Specifically, as the second stage is nonlinear, the first and second stages are jointly estimated in a conditional mixed-process framework (Roodman, 2011).

[23] Individual controls include gender (dummy equal to one if female), level of education (dummy equal to one if high school diploma or higher), employment status (dummy equal to one if currently working), income (dummy equal to one if earning more than the sample median income), interest in politics (dummy equal to one if some or a lot of interest), whether they are a member of a political association, religiosity (dummy equal to one if attend religious services more than once a month), and latitude and longitude of the municipality where they live. The results are robust to the exclusion of these control variables, as shown in Appendix G.

Chirac rather than abstain and some right-wing voters to vote for Jacques Chirac rather than for Jean-Marie Le Pen.

7.2 Changes in the perception of policies

The theoretical contributions of Lohmann (1994) and Battaglini (2017) suggest that the number of participants in a protest signals the importance of the cause. In the context of the 2002 presidential election, these protests may have informed voters of the negative consequences of the policies sponsored by Jean-Marie Le Pen. If that is true, we should expect survey respondents to be less positive about those policies where protests attracted more participants.[24]

To test that possibility, we regress on the number of participants in a protest the opinion of individual respondents about policies championed by Jean-Marie Le Pen expressed on a one-to-four scale using our baseline identification strategy. Specifically, we estimate with 2SLS and the rain-based synthetic instrument a variant of Equation (1) using individuals' answers as dependent variables.[25]

*** Insert Figure 6 around here ***

Figure 6 reports the estimated marginal effect of the number of participants on the agreement of respondents with the statement that the following issues, sponsored by Jean-Marie Le Pen, are important: immigration, security, traditional values, criticism of the political class, the abolition of the income tax, and an exit of France from the EU. The marginal effect is always negative and statistically significant at standard levels.

We also looked at two more specific measures of the position of respondents on immigration, which was the theme on which Jean-Marie Le Pen

[24] One may argue that opinions about those policies were subject to a social desirability bias, insofar as respondents may have been reluctant to reveal their true opinion instead of having been convinced to change it by the protests. We cannot rule out that protests triggered a social desirability bias, especially as we show in the next subsection that it affected who respondents said they voted for. However, while some policies — like those related to immigration — were clearly associated with Jean-Marie Le Pen, others — such as an exit from the EU or a critique of the political class — were sponsored by other candidates unrelated to Jean-Marie Le Pen.
[25] The exact wording of the questions is reported in Appendix G.

was the most salient. Specifically, respondents were asked their level of agreement with the following two statements: “there are too many immigrants” and “immigrants enrich a culture”. Here, Jean-Marie Le Pen was not explicitly mentioned. We observe that respondents in municipalities that experienced larger protests were less likely to agree strongly with the statement that there are too many immigrants and more likely to strongly agree with the statement that immigrants enrich a culture.

This series of findings is consistent with the models of Lohmann (1994) and Battaglini (2017) that argue that protests signal the importance of an issue. Larger protests against Jean-Marie Le Pen reduced the importance that respondents assigned to the issues that he championed, thereby reducing the incentive to vote for him.

7.3 Social desirability

Bursztyn, Egorov, and Fiorin (2020) and Giani and Méon (2021) argue and report evidence that election results provide information on the preferences of the population and can therefore change the ongoing status quo and reinforce or delegitimize social norms stigmatizing sexist or racist behaviors. The same reasoning can be adjusted to protests. If a large number of people participate in protests opposing a far-right candidate and condemning his positions as racist, some voters who contemplated voting for that candidate may decide against it, while others who contemplated abstaining will cast a ballot against him to avoid social stigma. In a nutshell, protests can reduce the social desirability of voting for the far-right candidate, prompting some voters who would have considered voting for him to cast a vote for a different candidate. What makes that mechanism likely in the 2002 French protests is that the selection of a far-right candidate for the second round was unprecedented and came as a complete surprise. It could thus change the status quo that voting for that type of candidate was unacceptable in line with Bursztyn, Egorov, and Fiorin’s (2020) model. By demonstrating against the far-right candidate, protesters reinforced the pre-

existing norm by showing that, despite the surprising result, many people were appalled and still considered it inadequate to vote for such a candidate.

A challenge to this argument is that ballots are secret. Someone could therefore vote for the far-right candidate without incurring a social stigma. However, as DellaVigna et al. (2016) point out, it is common for acquaintances and family to ask whether we voted. One may add that it is also common to ask for whom people voted. In line with that presumption, using survey data, Gerber et al. (2013) observe that 80 percent of the American population believed that someone will know for whom they voted, either because they will reveal it themselves or because they do not believe in ballot secrecy. 87 percent report that they are asked for whom they voted at least sometimes. Moreover, many people would feel uncomfortable lying about their vote. Refusing to disclose their vote may also effectively reveal it. In Gerber et al.'s (2013) survey, 84.4 percent of respondents declared that they truthfully reveal their vote to a close friend almost all the time and 9.8 more percent reported they do so most of the time. As a result, social desirability may influence both the decision to vote and the choice of the candidate if people care about what others think and are uncomfortable lying (Gerber et al., 2013).

Although survey data does not allow us to distinguish between a true and a misreported vote, we can nonetheless report direct evidence of an effect of protests on the social desirability bias affecting the far-right candidate. To do so, we leverage the timing of the election and the survey. Recall that the protests took place after the first round and that the survey was carried out after the second round. At the time of both the protests and the survey, the actual first-round vote had already been cast and could therefore not be affected by protests. If we observe a causal effect of protests on the vote that respondents *report* in the survey, we can unambiguously attribute it to a reduction by protests of the social desirability of reporting a given vote.

*** Insert Table 4 around here ***

Table 4 reports the marginal effect of the number of participants on the reported first-round vote. While the effect on reporting a vote for Jacques Chirac or having cast a blank ballot is statistically insignificant, the effect on reporting a vote for Jean-Marie Le Pen in the first round is negative and statistically significant. As first-round votes were pre-determined, this is evidence that protests decreased the social desirability of reporting a vote for Jean-Marie Le Pen. In a nutshell, more respondents lied about their vote in the first round in municipalities that hosted larger protests. Although we cannot report evidence that the effect extended to actual votes in the second round, this finding is suggestive that social desirability was at play.

The findings of this section are consistent with a model where protests signal that voting for the far-right candidate was not socially desirable. In this vein, the protests affected differently left- and right-wing voters in different ways because they were perceiving different trade-offs. Protests pushed right-wing voters who could have considered voting for Jean-Marie Le Pen to vote for Jacques Chirac instead. For left-wing voters, the protests signaled to voters that it was more acceptable to vote for the moderate right-wing candidate than to abstain. As a result, larger protests reduced the number of votes cast for Jean-Marie Le Pen among right-wing voters and reduced abstentions and blank ballots among left-wing voters. Consequently, the number of votes for Jacques Chirac increased among both left- and right-wing voters.

8 Conclusion

Protests against far-right parties have become common, but whether these protests achieve their stated goals and how they do is seldom discussed. Using data on a historical mobilization against a far-right candidate who had passed the first round of the 2002 French presidential election and a new variant of an instrumental strategy exploiting weather conditions, we find that the protests that we study did achieve their goal. Specifically, we show that the May 1, 2002, peaceful protests reduced the vote share and the absolute number of voters for

far-right candidate Jean-Marie Le Pen and reduced the share of abstentions and blank or invalid ballots. By the same token, they increased the vote share and the absolute number of votes for moderate candidate Jacques Chirac. In a nutshell, those protests were effective. That finding holds up to a series of robustness checks.

Moreover, we report suggestive evidence that the local press contributed to spreading the effect of protests. We can also document spillovers from municipalities that hosted a protest to municipalities that did not. Not only do those spillovers exist but they are of the same sign as the direct effects of the protests. In other words, protests in a municipality reduced both the number of votes for the far-right candidate and the number of abstentions and blank or invalid ballots, and the protests increased the number of votes for the moderate candidate in neighboring municipalities.

Finally, using individual survey data, we explore how individual voters reacted to the protests and suggest a series of transmission channels. We show that the effect of protests on voters depended on their political preferences. Left-wing voters were less likely to cast a blank or invalid ballot and right-wing voters less likely to vote for the far-right candidate. However, they prompted both left- and right-wing voters to vote more for the moderate candidate. We find no statistically significant effect on voters identifying with the center. These findings suggest that some right-wing voters swung between the extreme and moderate right-wing candidates, whereas some left-wing voters swung between casting a blank or invalid ballot and voting for the candidate they viewed as the lesser of two evils. Moreover, in line with the view that protests signal the importance of policy issues, we observe that protests decreased public support for the policies championed by the far-right candidate.

Most of all, we provide suggestive evidence that those effects were the result of a change in the social desirability of voting for the far-right candidate, as we document that voters who had voted for the far-right candidate in the first round were more likely to misreport their vote in municipalities that had hosted

larger protests, despite the protests having taken place a week and a half after the first round. In other words, protests increased their propensity to lie about their vote. To our knowledge, this is the first evidence of an effect of protests on the social desirability of a voting behavior.

We do not claim that the present paper provides an exhaustive list of mechanisms through which the May 1, 2002, protests affected voting outcomes on May 5. In particular, we do not rule out that they could help build political movements like in Madestam et al. (2013), although the proximity between the two days probably limited that possibility. That mechanism may, however, have affected later elections and possibly later protests. Investigating the long-term effects of the May 1, 2002, protests is therefore food for future research.

References

Angrist, J. D., & Pischke, J. S. (2008). *Mostly Harmless Econometrics*. Princeton University Press.

Battaglini, M. (2017). Public Protests and Policy Making. *Quarterly Journal of Economics*, 132(1), 485-549.

Belotti, F., Deb, P., Manning, W. G., & Norton, E. C. (2015). Twopm: Two-Part Models. *The Stata Journal*, 15(1), 3-20.

Bursztyn, L., González, A.L., & Yanagizawa-Drott, D., 2020. Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review*, 110(10), 2997-3029.

Bursztyn, L., Egorov, G., & Fiorin, S. (2020). From Extreme to Mainstream: The Erosion of Social Norms, *American Economic Review*, 110(11), 3522-48.

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics Using Stata* (Vol. 2). College Station, TX: Stata press.

Colella, F., Lalive, R., Sakalli, S. O., & Thoenig, M. (2019). Inference with Arbitrary Clustering. IZA Discussion Paper No. 12584.

Conley, T. G. (1999). GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, *92*(1), 1-45.

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly Exogenous. *Review of Economics and Statistics*, *94*(1), 260-272.

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2016). Voting to Tell Others. *Review of Economic Studies*, 84(1), 143-181.

García-Jimeno, C., Iglesias, A. & Yildirim, P., 2022. Information Networks and Collective Action: Evidence from the Women's Temperance Crusade. *American Economic Review*, 112(1), 41-80.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment. *American Political Science Review*, 102(1), 33-48.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2013). Is there a Secret Ballot? Ballot Secrecy Perceptions and their Implications for Voting Behaviour. *British Journal of Political Science*, 43(1), 77-102.

Giani, M., & Méon, P. G. (2021). Global Racist Contagion Following Donald Trump's Election. *British Journal of Political Science*, 51(3), 1332-1339.

Guriev, S., & Papaioannou, E. (forthcoming). The Political Economy of Populism, *Journal of Economic Literature*.

Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice. *Political Analysis*, 27(2), 163-192.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 112, p. 18). New York: Springer.

Kelejian, H. H., & Prucha, I. R. (2010). Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. *Journal of Econometrics*, 157(1), 53-67.

Klein Teeselink, B., & Melios, G. (2022). Weather to Protest: The Effect of Black Lives Matter Protests on the 2020 Presidential Election, CEBRIG Working Paper N°22-007.

Lewis-Beck, M. S. (2004). French Election Theories and the 2002 Results: An Introduction. In *The French Voter* (pp. 1-11). Palgrave Macmillan, London.

Larreboure, M., & González, F. (2021). The impact of the Women's March on the US House Election. Documento de Trabajo IE-PUC, (560).

Libération, "Votez escroc, pas facho !", 23 April 2002.

Libération, "Retour sur Quand Jacques Chirac évoquait « le bruit et l'odeur »... ", 26 September 2019.

Lohmann, S. (1994). Information Aggregation Through Costly Political Action. *American Economic Review*, 84(3), 518-530.

Madestam, A., Shoag, D., Veuger, S., & Yanagizawa-Drott, D. (2013). Do Political Protests Matter? Evidence from the Tea Party Movement. *Quarterly Journal of Economics*, 128(4), 1633-1685.

Mamo, N., Bhattacharyya, S., & Moradi, A. (2019). Intensive and Extensive Margins of Mining and Development: Evidence from Sub-Saharan Africa. *Journal of Development Economics*, *139*, 28-49.

Mazumder, S. (2018). The Persistent Effect of US Civil Rights Protests on Political Attitudes. *American Journal of Political Science*, 62(4), 922-935.

Meier, A. N., Schmid, L., & Stutzer, A. (2019). Rain, Emotions and Voting for the Status Quo. *European Economic Review*, 119, 434-451.

Mellon, J. (2021). Rain, Rain, Go Away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable. Available at SSRN 3715610.

Murphy, K. M., & Shleifer, A. (2004). Persuasion in Politics. *American Economic Review*, 94(2), 435-439.

Pons, V. (2018). Will a Five-minute Discussion Change your Mind? A Countrywide Experiment on Voter Choice in France. *American Economic Review*, 108(6), 1322-63.

Pons, V., & Liegey, G. (2019). Increasing the Electoral Participation of Immigrants: Experimental Evidence from France. *Economic Journal*, 129(617), 481-508.

Pons, V., & Tricaud, C. (2018). Expressive Voting and its Cost: Evidence from Runoffs with Two or Three Candidates. *Econometrica*, 86(5), 1621-1649.

Reny, T. T., & Newman, B. J. (2018). Protecting the Right to Discriminate: the Second Great Migration and Racial Threat in the American West. *American Political Science Review*, 112(4), 1104-1110.

Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with cmp. *The Stata Journal*, 11(2), 159-206.

Sheather, S. (2009). *A Modern Approach to Regression with R*. Springer Science & Business Media.

Snyder Jr, J. M., & Yousaf, H. (2020). Making Rallies Great Again: The Effects of Presidential Campaign Rallies on Voter Behavior, 2008-2016, National Bureau of Economic Research, working paper No. 28043.

Wasow, O. (2020). Agenda seeding: How 1960s black protests moved elites, public opinion and voting. *American Political Science Review*, 114(3), 638-659.

Tables

Table 1. Baseline results – The impact of protests on voting outcomes

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants (ln)	-0.399	0.818	-0.304
	(0.124)	(0.273)	(0.0997)
First-round outcome	0.875	0.478	0.524
	(0.0221)	(0.0341)	(0.0142)
F statistics	127.3	127.5	127.5
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. The instrument is the rain-based synthetic instrument estimated in Section 4.2. The first stage is reported in Table C2 of Appendix C1. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table 2. The impact of protests on voting outcomes – Spatial spillovers

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Direct effect of the number of participants (ln)	-0.378	0.775	-0.256
	(0.0787)	(0.158)	(0.0696)
Spatial spillover effect of the number of participants (ln)	-0.287	0.520	-0.649
	(0.0666)	(0.196)	(0.0954)
F statistics	1624	1572	1621
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (5). The dependent variable of each specification is reported at the top of each panel. The instrument for the direct effect is the rain-based synthetic instrument estimated in Section 4.2 and the instrument for the spillover effect is the spatial version of synthetic instrument. The first stage is reported in Table F5 of Appendix F2. In each specification, we control for municipality type. Robust standard errors are reported in parentheses.

Table 3. Survey data: Descriptive statistics

	First-round vote				Second-round vote			
	Mean	SD	Min	Max	Mean	SD	Min	Max
J.-M. Le Pen	0.0995	0.299	0	1	0.0759	0.265	0	1
J. Chirac	0.205	0.404	0	1	0.862	0.345	0	1
Blank/invalid	.0213	0.144	0	1	0.0625	0.242	0	1
Observations	3,246				3,519			

Notes. The unit of analysis is a survey respondent.

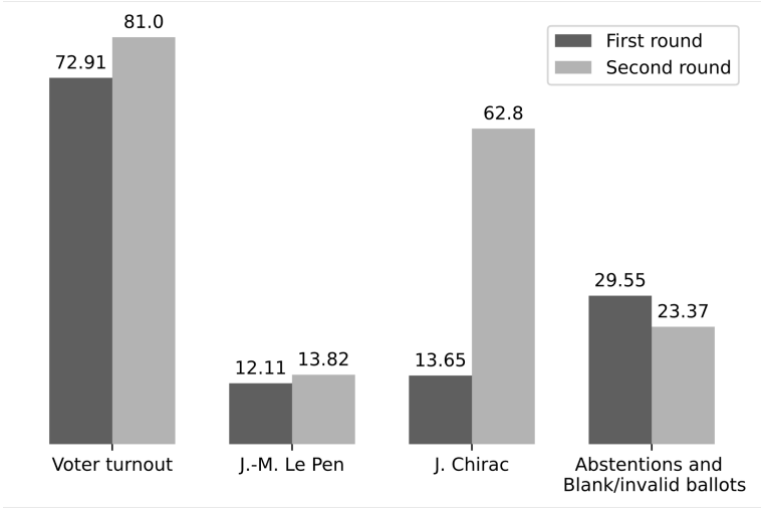
Table 4. Average marginal effect of the number of participants on the probability of declaring a vote for Jean-Marie Le Pen, Jacques Chirac, or a blank or invalid ballot

	(1)	(2)	(3)
	J.-M. Le Pen	J. Chirac	Blank or
First-round declared vote			invalid ballot
Number of participants (ln)	-0.0106	-0.00201	-0.00192
	(0.00369)	(0.00284)	(0.00179)
F statistics	155.79	155.79	155.79
Observations	3,241	3,241	3,241

Notes. The unit of analysis is a survey respondent. The linear first stage and the second stage multinomial probit model are jointly estimated in a conditional mixed-process framework (Roodman, 2011). The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for respondents' political orientation, gender (dummy equal to one if female), level of education (dummy equal to one if high school diploma or higher), employment status (dummy equal to one if currently working), income (dummy equal to one if earning more than the sample median income), interest in politics (dummy equal to one if some or a lot of interest), whether they are a member of a political association, religiosity (dummy equal to one if attend religious services more than once a month), and the type, latitude, and longitude of the municipality where they live. Standard errors clustered at the department level are reported in parentheses. The exact wording of the questions is reported in Appendix G.

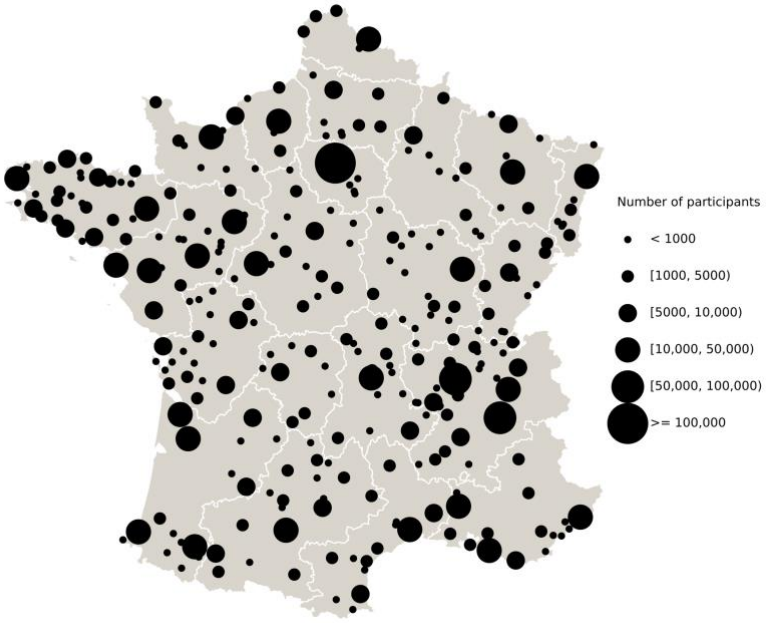
Figures

Figure 1: Outcomes of the two rounds of the election



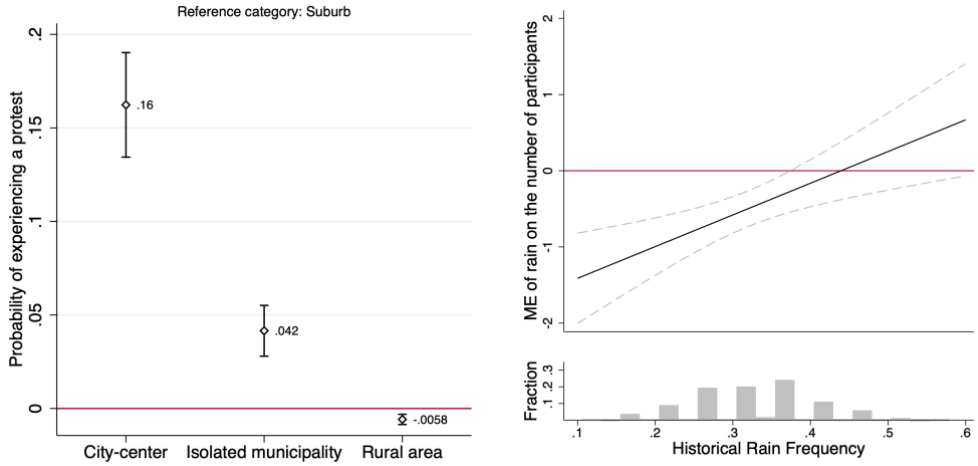
Note: Voting outcomes are reported as shares of the number of registered voters, which is the same for the two rounds. The figure reports the outcome's mean over municipalities in mainland France.

Figure 2. Location of protests on May 1, 2002



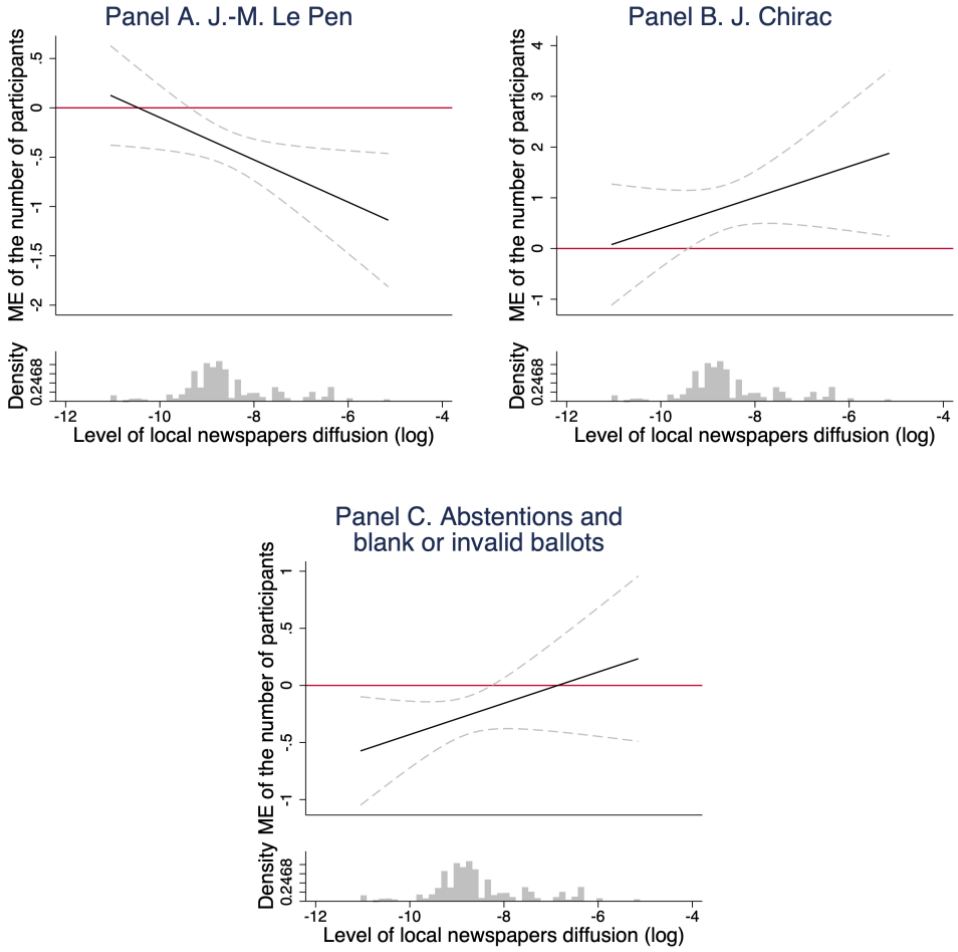
Note: The figure reports the location and the size of the protests held on May 1, 2002, against Jean-Marie Le Pen.

Figure 3. Synthetic instrument: Two-part model



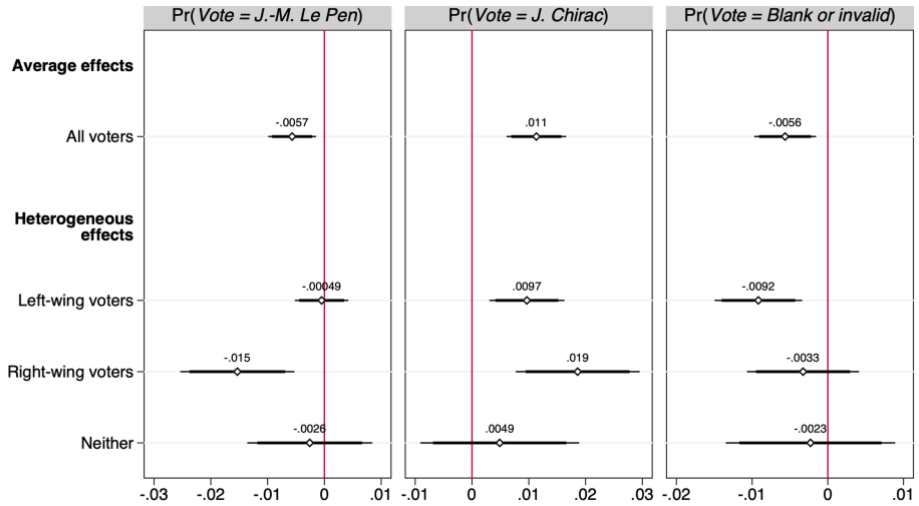
Note: The unit of analysis is a municipality. The figure reports the results of the two-part model estimated in Section 4.2 to generate our rain-based synthetic instrument. The left-hand reports the estimates of the first part, which models the probability of a protest occurring in a municipality by using the dummy variables classifying the municipality type. The model specification follows Equation (3). The black lines indicate 90% confidence intervals based on standard errors clustered at the department level. The right-hand side plots the marginal effect of rain on the number of participants against historical rain frequency in municipalities that experienced a protest (the second part of the two-part model). The model specification follows Equation (4). The

Figure 4. Marginal effect of the number of participants (ln) on voting outcomes as a function of the level of local newspaper distribution



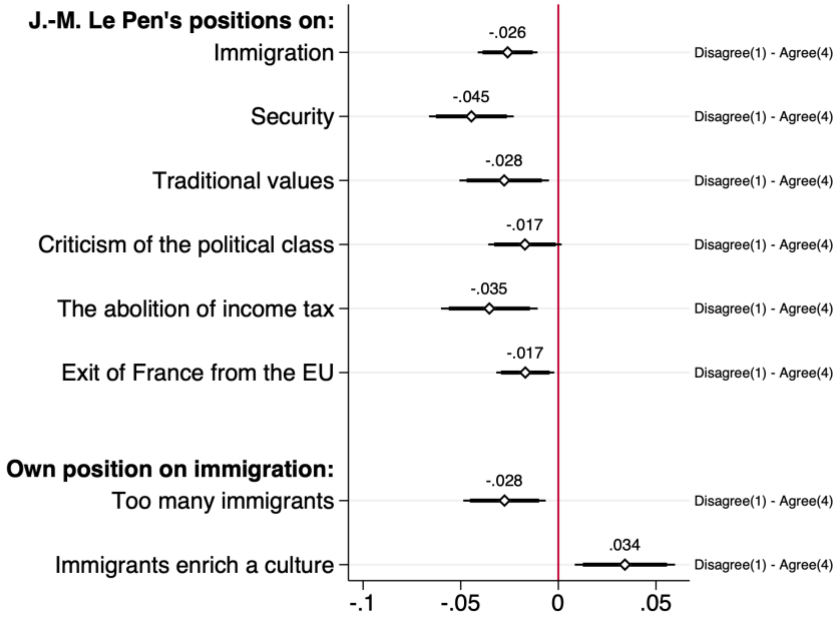
Note: 2SLS estimates. The unit of analysis is a municipality. The dotted lines indicate 90% confidence intervals based on standard errors clustered at the department level. In each specification, we control for municipality type.

Figure 5. Average marginal effect of the number of participants (ln) on the probability of voting for Jean-Marie Le Pen, voting for Jacques Chirac, or casting a blank or invalid ballot



Notes. The unit of analysis is a survey respondent. The linear first stage and the second stage multinomial probit model are jointly estimated in a conditional mixed-process framework (Roodman, 2011). The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for respondents' gender (dummy equal to one if female), level of education (dummy equal to one if high school diploma or higher), employment status (dummy equal to one if currently working), income (dummy equal to one if earning more than the sample median income), interest in politics (dummy equal to one if some or a lot of interest), whether they are a member of a political association, religiosity (dummy equal to one if attend religious services more than once a month), and the type, latitude, and longitude of the municipality where they live. The specification of the average effect further controls for respondents' political orientation. The thick black line indicates 90% confidence intervals, while the thin black line indicates 95% confidence intervals. The confidence intervals are based on standard errors clustered at the department level. The F statistics of the linear first stage range from 96 to 153 across specifications. The exact wording of the questions is reported in Appendix G.

Figure 6. Marginal effect of the number of participants (ln) on the support of policies championed by Jean-Marie Le Pen



Notes. The unit of analysis is a survey respondent. 2SLS estimates. The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for respondents' political orientation, gender (dummy equal to one if female), level of education (dummy equal to one if high school diploma or higher), employment status (dummy equal to one if currently working), income (dummy equal to one if earning more than the sample median income), interest in politics (dummy equal to one if some or a lot of interest), whether they are a member of a political association, religiosity (dummy equal to one if attend religious services more than once a month), and the type, latitude, and longitude of the municipality where they live. The thick black line indicates 90% confidence intervals, while the thin black line indicates 95% confidence intervals. Confidence intervals are based on standard errors clustered at the department level. The F statistics range from 125 to 161 across specifications. The exact wording of the questions is reported in Appendix G.

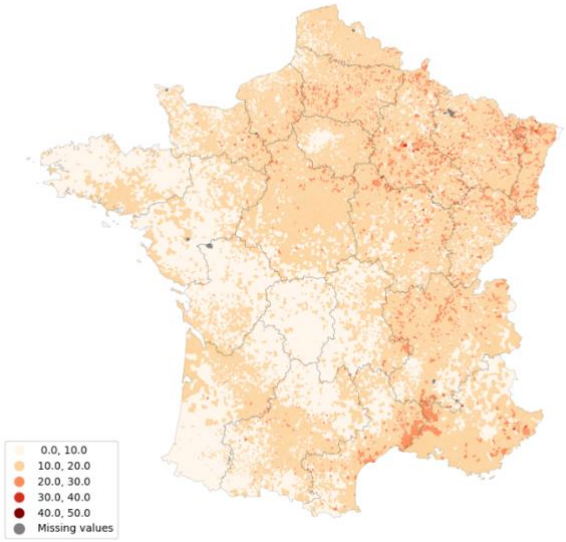
For Online Publication

Is Demonstrating Against the Far Right Worth it? Evidence From French Presidential Elections

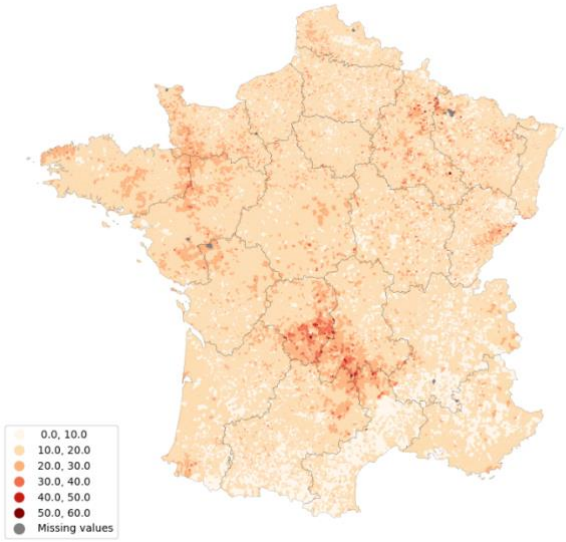
Nicolas Lagios, Pierre-Guillaume Méon, Ilan Tojerow

Appendix A. Descriptive statistics

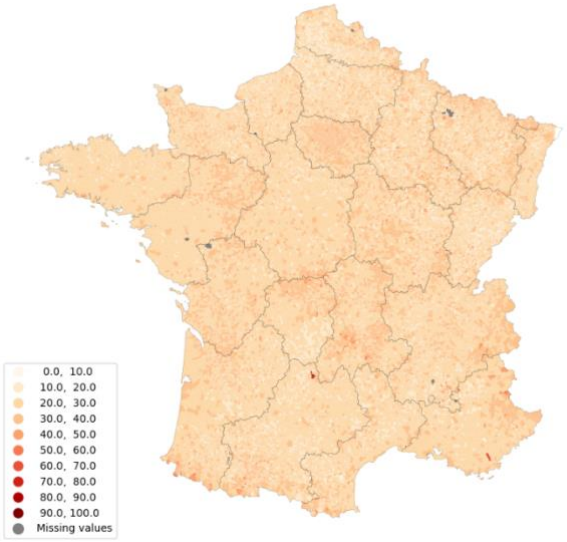
Figure A1. First and second round electoral outcomes at the municipal level



a) J.-M. Le Pen vote share in the 1st round

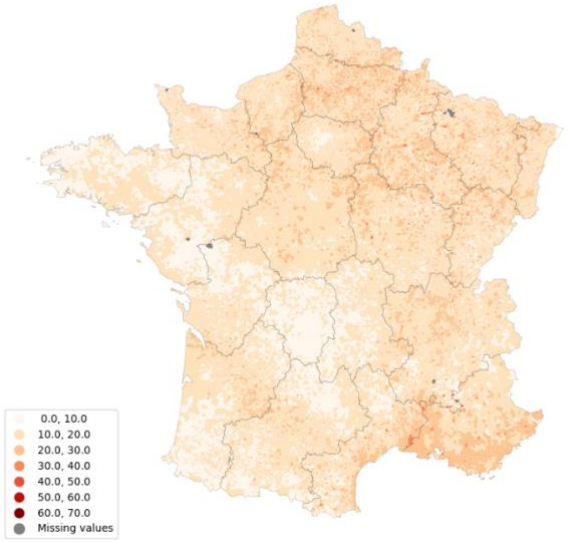


b) J. Chirac vote share in the 1st round

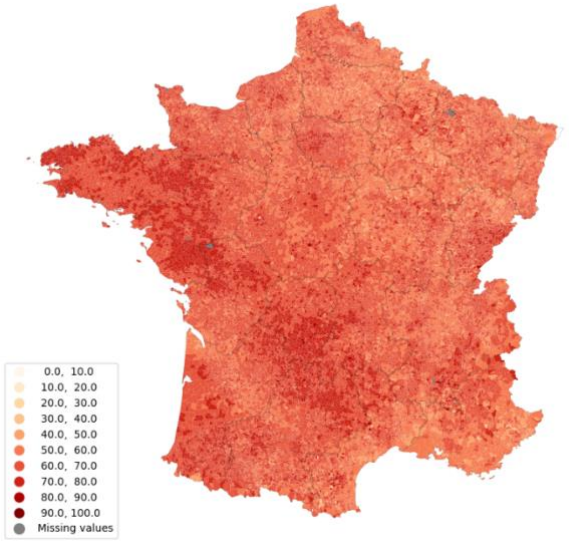


c) Blank/invalid and abstention share in the 1st round

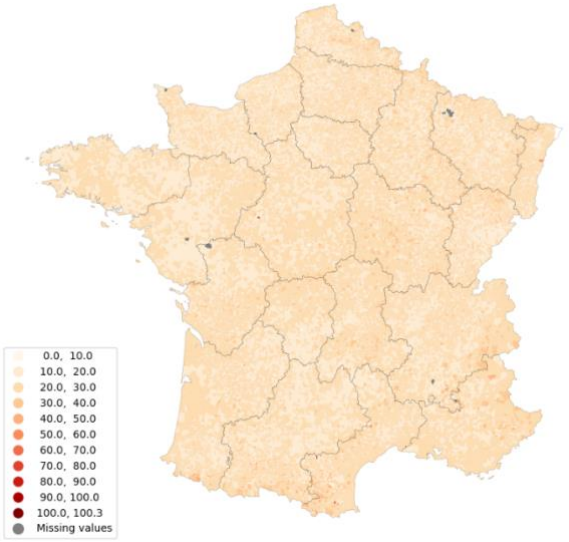
(continued)



d) J.-M. Le Pen vote share in the 2nd round

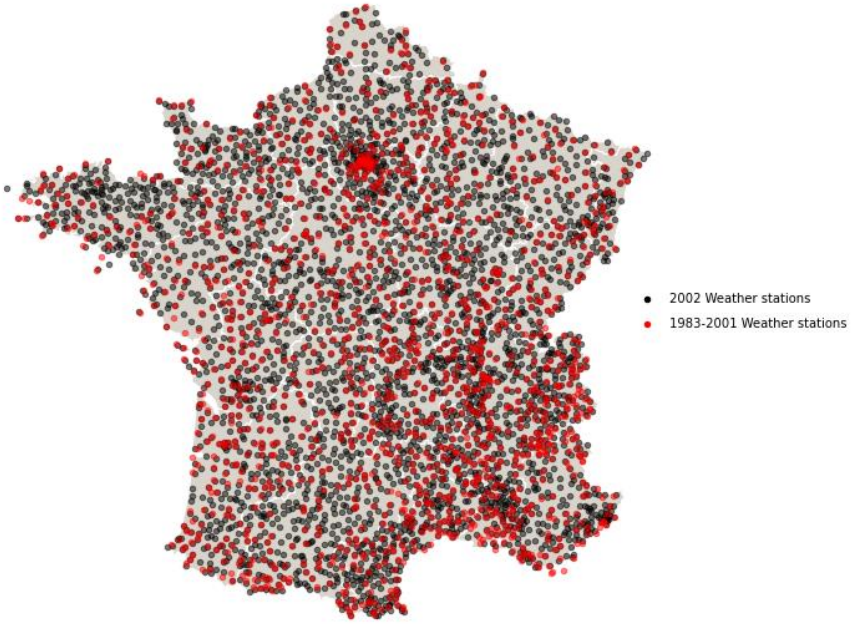


e) J. Chirac vote share in the 2nd round



f) Blank/invalid and abstention share in the 2nd round

Figure A2. Location of weather stations



Note: The figure reports the location and the size of weather stations used to assess the level of precipitations.

Appendix B. The synthetic instrument

Appendix B1. Exogeneity checks

In this subsection, we provide a series of placebo and balance tests to assess our identifying assumption that conditional on the municipality type, rain does not correlate with drivers of electoral behavior other than protests.

Table B1. The effect of rain on the protest day on pre-rally electoral outcomes

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blanks
Rainy protest	0.382	-0.581	-0.150
	(0.681)	(0.534)	(0.304)
Observations	36,153	36,153	36,153

Notes. OLS estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day), and 0 otherwise. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table B2. Municipal characteristics' balance

	(1) Change in unemployment rate |b| 95-01	(2) Special administrative status	(3) Landmass	(4) Population Size	(5) Average age	(6) % of ind. aged 60 and +
Rainy protest	-0.00655	-0.0149	1.483	-253.6	0.440	0.00651
	(0.00762)	(0.0442)	(1.108)	(224.5)	(0.474)	(0.00884)
Model	Linear	Probit	Linear	Linear	Linear	Linear
Observations	36,153	36,153	36,153	36,153	36,153	36,153

Notes. OLS estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day), and 0 otherwise. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table B3. Individual characteristics' balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Political orientation (base outcome = neither)								
	Left	Right	Gender	Education	Employment status	Income	Interest in politics	Member of a political association	Religiosity
Panel A. Qualitative composition of protests									
Rainy protest	0.0587	-0.0235	-0.105	-0.109	0.172	0.0898	0.117	-0.0591	-0.327
	(0.392)	(0.360)	(0.177)	(0.217)	(0.182)	(0.157)	(0.178)	(0.121)	(0.462)
Observations	480		482	482	482	482	482	482	468
Panel B. Qualitative composition of voter mobilization in the second round									
Rainy vote	0.0987	0.0391	-0.0408	0.0730	0.00332	0.0670	-0.0479	-0.0160	-0.0878
	(0.0696)	(0.0632)	(0.0442)	(0.0746)	(0.0409)	(0.0823)	(0.0617)	(0.0446)	(0.0596)
Model	Multinomial probit		Probit	Probit	Probit	Probit	Probit	Probit	Probit
Observations	3,595		3,655	3,652	3,655	3,655	3,655	3,655	3,655

Notes. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day), and 0 otherwise. $Rainy\ vote$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 5, 2002 (the second round), and 0 otherwise. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table B4. Correlation between rain on the protest day (May 1) and rain on the second-round election day (May 5)

	(1) Rainy vote
Rainy protest	0.0693
	(0.0960)
Observations	36,153

Notes. OLS estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day). $Rainy\ vote$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 5, 2002 (the day of the second round). In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Appendix B2. Construction of the synthetic instrument

In this subsection, we provide more details on the construction of the synthetic instrument.

Given (3) and (4) in the main text, the overall expected value of $\ln(Participants)$ is given by

$$\begin{aligned} &\mathbb{E}(\ln(\text{[illegible]}) \mid \text{[illegible]}, \text{[illegible]}_1, \text{[illegible]}_2) \\ &\quad = \Pr(Participants > 0 \mid X_1) \\ &\quad \times \mathbb{E}(\ln(Participants) \mid Participants > 0, Q, X_1, X_2). \end{aligned} \tag{B1}$$

Then, following (B1), we define the instrument z_m to be used in Equation (2) as the prediction of $\ln(Participants)$, that is $\left(\ln(\widehat{Participants}) \mid Q, X_1, X_2\right)$:

$$\begin{aligned} z &= \left(\ln(\widehat{Participants}) \mid Q, X_1, X_2\right) \\ &\quad = (\hat{p} \mid X_1) \\ &\quad \times \left(\ln(\widehat{Participants}) \mid Participant > 0, Q, X_1, X_2\right), \end{aligned} \tag{B2}$$

where $(\hat{p} \mid X_1)$ is the predicted probability that municipality m experiences a protest, estimated in Equation (3), and $\left(\ln(\widehat{Participants}) \mid Participants > 0, Q, X_1, X_2\right)$ is the predicted number of

participants in municipality m given that it experiences a protest, which is estimated in Equation (4).[1]

The model is estimated by maximum likelihood. Let d be a binary indicator equal to 1 when $Participants > 0$ and 0 otherwise. Then, the density for observation m is given by

$$\begin{aligned}&\phi(\ln(Participants) \mid Q, X_1, X_2)\\&\quad = \left(1 - F\left(X_1'\alpha\right)\right)^{1-d} \times \left\{F\left(X_1'\alpha\right)h\left(Q'\beta, X_1'\gamma, X_2'\delta\right)\right\}^{d},\end{aligned} \tag{B3}$$

where $h(\cdot)$ is the probability density function of the normal distribution for $\ln(Participants) \mid Participants > 0$. Following (B3), the log-likelihood function can be written as

$$\begin{aligned}\mathcal{L}(\alpha, \beta, \gamma, \delta) &= \mathcal{L}(\alpha) + \mathcal{L}(\beta, \gamma, \delta)\\&= \sum_{m=1}^{n} \Big((1 - d_m) \ln\left(1 - F\left(X_{1,m}'\alpha\right)\right)\\&\quad + d_m \ln\left(F\left(X_{1,m}'\alpha\right)\right)\Big)\\&\quad + \sum_{m=1}^{n} d_m \ln\left(h\left(Q_m'\beta, X_{1,m}'\gamma, X_{2,m}'\delta\right)\right).\end{aligned} \tag{B4}$$

Because the two parts of the model are assumed to be independent, the joint likelihood function can be maximized by separately maximizing the likelihood functions for the model for the zeros and for the model for the positives, thus meaning that α and (β, γ, δ) can be estimated separately.

[1] Note that we are using as instrument the prediction of $\mathbb{E}(\ln(Participants) \mid Q, X_1, X_2)$ rather than the prediction of $\mathbb{E}(Participants \mid Q, X_1, X_2)$. For our purpose of predicting the number of participants, this makes conceptually no difference as ln is monotonically increasing. However, the former has the advantage of being much easier to predict as there is no need to deal with the retransformation problem that is non-trivial in non-homoscedastic settings (Belotti et al., 2015, Mullahy, 1998).

Appendix B3. Estimation of the two-part model

Table B5 reports the full set of results of the two-part model estimated in the main text to construct the synthetic instrument.

Table B5. Synthetic instrument: Two-part model

	(1) Coef.	(2) St. errors
Panel A. First part – Probability of experiencing a protest		
Municipality type		
Suburb	(reference category)	
Rural	-2.057	0.325
Isolated municipality	1.994	0.301
City center	3.314	0.258
Panel B. Second part – Number of participants conditional on experiencing a protest		
Rain variables		
Rainy protest	-1.828	0.502
Historical rain frequency	-1.467	1.217
Rainy protest*Historical rain frequency	4.159	1.509
Municipality type		
Suburb	(reference category)	
Rural	-0.262	0.272
Isolated municipality	0.647	0.243
City center	1.661	0.227
Municipal characteristics		
Parisian agglomeration	-0.0976	0.617
Change in unemployment rate |b| 1995-2001	-0.831	1.609
Special administrative status	0.640	0.152
Landmass (in square km)	0.00386	0.00202
Population Size	4.41e-06	1.35e-06
Average age	0.0308	0.0829
% of ind. aged 60 and +	-6.930	4.458
First part		*Second part*	
Observations	36,153	Observations	302
Log pseudolikelihood	-1097.529	Log pseudolikelihood	-465.238
Zero outcomes	35,851	R-squared	0.506

Notes. The unit of analysis is a municipality. $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day), and 0 otherwise. $Historical\ rain\ frequency$ is the share of May 1 days where rainfall on that day was higher or equal to 1 millimeter between 1983 and 2001. Column (2) reports standard errors clustered at the department level.

Appendix C. Baseline results

Appendix C1. Additional results

Table C1. Variance inflation factor (VIF) of Participants when all covariates are included

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
VIF	21.55	21.41	21.40

Notes. The unit of analysis is a municipality. The table shows variance inflation factor (VIF) when we include all covariates in the 2SLS specifications. Covariates include population size, average age, landmass, variation in the unemployment rate between 1995 and 2002, the proportion of individuals aged 60 years or older, whether the municipality belongs to the Parisian agglomeration, whether it has a special administrative status, and the municipality type.

Table C2. Baseline model – First-stage estimates

ln(number of participants)	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
$\ln(\widehat{Participants})$	9.777	9.795	9.775
	(0.866)	(0.867)	(0.866)
First-round outcome	-0.004	0.001	0.002
	(0.001)	(0.000)	(0.001)
Observations	36,153	36,153	36,153

Notes. The unit of analysis is a municipality. The dependent variable is the number of participants (ln). $\ln(\widehat{Participants})$ is the number of participants predicted by the two-part model in Section 4.2. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Appendix C2. Robustness checks of the baseline results

C2.1. Specification

Table C3. The impact of protests on voting outcomes – Flexible specification

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.307	0.682	-0.370
	(0.148)	(0.354)	(0.172)
First-round outcome	0.875	0.478	0.524
	(0.0219)	(0.0341)	(0.0142)
F statistics	114.4	114.4	113.8
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C4. The impact of protests on voting outcomes – Quantile regression

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Quantile = 0.25			
Number of participants	-0.383	1.072	-0.408
	(0.130)	(0.270)	(0.0935)
Quantile = 0.5			
Number of participants	-0.222	0.550	-0.232
	(0.0786)	(0.168)	(0.0817)
Quantile = 0.75			
Number of participants	-0.233	0.525	-0.192
	(0.0714)	(0.161)	(0.0831)
Observations	36,153	36,153	36,153

Notes. Quantile regressions. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C5. The impact of protests on voting outcomes – Only demographic covariables

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.518	1.140	-0.404
	(0.169)	(0.318)	(0.109)
First-round outcome	0.875	0.478	0.524
	(0.0223)	(0.0341)	(0.0142)
F statistics	52.91	52.81	52.89
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C6. The impact of protests on voting outcomes – Participant variable: in % of population

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-32.62	66.81	-24.76
	(10.63)	(21.08)	(7.493)
First-round outcome	0.874	0.480	0.524
	(0.0222)	(0.0341)	(0.0142)
F statistics	88.13	87.94	88.20
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C7. The impact of protests on voting outcomes – A protest is said to be rainy if rain on the day of the protests exceeded 2 and 3 millimeters

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
A. A protest is said to be rainy if rain on the day of the protests exceeded 2 millimeters			
Number of participants	-0.366	0.732	-0.274
	(0.119)	(0.249)	(0.0988)
First-round outcome	0.875	0.478	0.524
	(0.0220)	(0.0341)	(0.0142)
F statistics	122.9	123.2	123.2
B. A protest is said to be rainy if rain on the day of the protests exceeded 3 millimeters			
Number of participants	-0.277	0.522	-0.248
	(0.117)	(0.238)	(0.0908)
First-round outcome	0.876	0.478	0.523
	(0.0220)	(0.0341)	(0.0142)
F statistics	104.6	104.8	105
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C8. The impact of protests on voting outcomes – A protest is said to be rainy if rain on the day of the protests was above the historical average

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.191	0.555	-0.348
	(0.109)	(0.259)	(0.0881)
First-round outcome	0.876	0.478	0.524
	(0.0220)	(0.0341)	(0.0142)
F statistics	108.9	108.9	109

Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C9. The impact of protests on voting outcomes – Continuous measurement of rainfall

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.204	0.817	-0.326
	(0.117)	(0.278)	(0.0898)
First-round outcome	0.876	0.478	0.524
	(0.0220)	(0.0341)	(0.0142)
F statistics	121.7	121.8	121.9
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C10. The impact of protests on voting outcomes - Maximum and minimum number of participants

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
A. Maximum number of participants			
Number of participants	-0.384	0.769	-0.275
	(0.117)	(0.254)	(0.0936)
First-round outcome	0.875	0.478	0.524
	(0.0221)	(0.0341)	(0.0142)
F statistics	128.5	128.6	128.6
B. Minimum number of participants			
Number of participants	-0.388	0.843	-0.352
	(0.133)	(0.294)	(0.106)
First-round outcome	0.875	0.478	0.524
	(0.0220)	(0.0341)	(0.0142)
F statistics	122.5	122.6	122.5
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C11. The impact of protests on voting outcomes – Without the five municipalities with the highest difference between the maximum and the minimum number of participants

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.306	0.577	-0.110
	(0.0567)	(0.147)	(0.0593)
First-round outcome	0.876	0.478	0.523
	(0.0220)	(0.0341)	(0.0142)
F statistics	225.4	225	224.9
Observations	36,148	36,148	36,148

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C12. The impact of protests on voting outcomes – Excluding outliers

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.673	1.078	-0.392
	(0.203)	(0.496)	(0.185)
First-round outcome	0.874	0.478	0.524
	(0.0221)	(0.0341)	(0.0142)
F statistics	57.44	57.60	57.16
Observations	36,138	36,138	36,138

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C13. The impact of protests on voting outcomes – Excluding small municipalities

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.156	0.841	-0.236
	(0.0910)	(0.231)	(0.0963)
First-round outcome	1.108	1.002	0.676
	(0.0574)	(0.133)	(0.0509)
F statistics	38.81	40.79	42.54
Observations	1,808	1,808	1,808

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C14. The impact of protests on voting outcomes – Excluding Paris

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.350	0.694	-0.200
	(0.0904)	(0.204)	(0.0696)
First-round outcome	0.875	0.478	0.523
	(0.0221)	(0.0341)	(0.0142)
F statistics	135.3	134.7	134.8
Observations	36,152	36,152	36,152

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Table C15. The impact of protests on voting outcomes – Only municipalities within 15km and 20km of a weather station

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
A. Distance: 15km			
Number of participants	-0.344	0.674	-0.227
	(0.111)	(0.257)	(0.0957)
First-round outcome	0.875	0.470	0.518
	(0.0211)	(0.0346)	(0.0158)
F statistics	110.6	110.9	111
Observations	30,520	30,520	30,520
B. Distance: 20km			
Number of participants	-0.380	0.776	-0.276
	(0.120)	(0.273)	(0.0978)
First-round outcome	0.874	0.476	0.523
	(0.0219)	(0.0343)	(0.0147)
F statistics	116.3	116.4	116.4
Observations	34,599	34,599	34,599

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

C2.2. Spatial correlation

Table C16. The impact of protests on voting outcomes – Spatially corrected standard errors

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Estimates	-0.399	0.818	-0.304
Cluster robust s.e.	(0.124)	(0.273)	(0.0997)
Spatial robust s.e.			
Radius: 50km	(0.100)	(0.231)	(0.0993)
Radius: 100km	(0.115)	(0.257)	(0.104)
Radius: 250km	(0.103)	(0.243)	(0.115)
Radius: 500km	(0.0700)	(0.193)	(0.0904)
Radius: 750km	(0.0724)	(0.188)	(0.0788)
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The model specification follows Equation (1). The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors corrected for spatial correlation are reported in parentheses.

C2.3. The exclusion restriction

Table C17. The impact of protests on voting outcomes – Controlling for the weather on the day of the second round

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-0.423	0.893	-0.315
	(0.122)	(0.261)	(0.0972)
First-round outcome	0.840	0.476	0.526
	(0.0201)	(0.0293)	(0.0144)
F statistics	128.2	128.6	128.5
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type and the weather on the day of the second round. Standard errors clustered at the department level are reported in parentheses.

Table C18. Reduced-form estimates

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
$\ln(\widehat{Participants})$	-.0846***	0.174***	-0.0643***
Observations	36,148	36,148	36,148

Notes. Reduced-form estimates of the effect of a 0.1 standard deviation increase in $\ln(\widehat{Participants})$ on electoral outcomes. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. ***Significant at the 1% level.

Appendix D. Conley's et al. (2012) plausibly exogenous approach

Table B4 in Appendix B3 shows that rain on the day of protests (May 1) is uncorrelated with rain on the day of the second round (May 5). This suggests that weather conditions were not persistent over time. Table C17 in Appendix C2 shows that even when controlling for weather on the day of the second-round election, the estimates are quantitatively and qualitatively very similar to the baseline 2SLS estimates. Taken together, these results suggest that rain on the day of protests likely had no direct impact on the second-round electoral outcomes, lending credence to the exclusion restriction. However, as we cannot rule out with certainty all violations of the exclusion restriction, we explore the robustness of the 2SLS estimates to departures from it to provide further evidence of the validity of our instrumental variable (IV) framework.

To do so, we implement Conley's et al. (2012) plausibly exogenous approach that allows for a direct impact of the instrument on the outcome. Specifically, denoting by Y, X, and Z the vectors of outcomes, endogenous variables, and instruments and disregarding control variables, the IV equations defined in (1) and (2) in the main text are a special case of

$$Y = X\zeta + Z\gamma + \epsilon, \tag{D1}$$

$$X = Z\Pi + V, \tag{D2}$$

where we impose that $\gamma = 0$. That is, we assume that the exclusion restriction holds. In other words, we assume that the instrument has no direct effect on the outcome. To relax this assumption, we follow the Local to Zero Approximation method considered in Conley et al. (2012), which consists of hypothesizing that the prior on γ follows a normal distribution with mean μ_γ and variance Ω_γ, and where the uncertainty regarding γ decreases with the sample size. The plausibly exogenous estimator is thus defined as

$$\hat{\zeta} \sim N\left(\zeta_{2SLS} + A\mu_\gamma, W_{2SLS} + A\Omega_\gamma A'\right), \quad \text{(D3)}$$

where $A = (X'Z(Z'Z)^{-1}Z'X)^{-1}(X'Z)$, ζ_{2SLS} is the 2SLS point estimate, W_{2SLS} is the 2SLS variance-covariance matrix, and μ_γ and Ω_γ are inputs to be specified.

To find a plausible estimate of the direct effect of the instrument on electoral outcomes in the full sample, γ, we follow the guidance of van Kippersluis and Rietveld (2018) and use the zero-first-stage test. Consider the following reduced form equation where Equation (D2) is substituted into Equation (D1):

$$Y = Z(\gamma + \zeta\Pi) + (\epsilon + \zeta V). \quad \text{(D4)}$$

If one can find a sub-sample for which the instrument has no effect on rally attendance — that is, a sub-sample for which $\Pi = 0$ in the first stage — then the reduced-form coefficient of the instrument will be given by γ. We refer to this sub-sample as the zero-first-stage group. Thus, by estimating Equation (D4) only in the zero-first-stage group, one obtains the estimator $\hat{\gamma}$ which can then be used as a plausible estimate for γ in the full sample. The authors therefore recommend specifying $\mu_\gamma = \hat{\gamma}$ in Equation (D3).

To take into account uncertainty regarding the direct effect of the instrument on electoral outcomes, Ω_γ, we again follow van Kippersluis and Rietveld (2018) and set $\Omega_\gamma = \left(0.125\sqrt{S_0^2 + S_{-0}^2}\right)^2$, where S_0^2 and S_{-0}^2 are the standard errors of $\hat{\gamma}$ in the zero-first-stage group and in the remainder of the sample.

Municipalities in the Parisian agglomeration provide a quasi-natural zero-first-stage group as they experienced nearly no protests, their residents being more likely to go and protest in Paris. Table D1 shows that the instrument has indeed no predictive power in this sub-sample, which makes it a suitable zero-first-stage group. In addition, the instrument has no direct effect on

electoral outcomes. The outcome of the plausibly exogenous approach is summarized in Table D2. The results indicate that the 2SLS estimates are robust to violations of the exclusion restriction.

Table D1. Zero-First-Stage Group

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
First stage	8.548	1.164	-3.743
	(18.296)	(19.503)	(18.368)
Direct effect	5.536	-44.903	27.395
	(16.392)	(58.152)	(32.369)
Observations	396	396	396

Notes. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. Standard errors clustered at the department level are reported in parentheses.

Table D2. The impact of protests on voting outcomes – Relaxing the exclusion restriction

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
2SLS	-0.399	0.818	-0.304
	(0.124)	(0.273)	(0.0997)
Plausibly Exogenous	-0.966	5.402	-3.106
(with uncertainty)	(0.244)	(0.792)	(0.426)
Observations	36,153	36,153	36,153

Notes. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. Plausibly Exogenous is the estimates obtained by implementing Conley's et al. (2012) plausibly exogenous approach that allows for a direct impact of the instrument on the outcome (see Appendix D for more detail). Standard errors clustered at the department level are reported in parentheses.

Appendix E. The role of the press

Appendix E1. Additional results

Table E1. Local newspapers

	Newspaper copies sold			
	Mean	SD	Min	Max
Log(National Newspapers per capita)	-8.543	0.998	-11.048	-5.111

Table E2. Interaction between the number of participants and press diffusion – Raw coefficients

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Number of participants	-2.242	3.448	0.939
	(0.997)	(2.398)	(1.043)
Press diffusion per capita	0.471	-0.627	0.233
	(0.186)	(0.429)	(0.113)
Number of participants*Press diffusion per capita	-0.214	0.305	0.137
	(0.116)	(0.277)	(0.119)
F statistics	73.33	73.26	73.25
Observations	36,153	36,153	36,153

Notes. 2SLS estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Standard errors clustered at the department level are reported in parentheses.

Appendix E2. Flexible interaction model

To assess the moderating role of the press on voting outcomes, we estimate in Section 6.1 the following multiplicative interaction model:

$$y_{2,m}^{j} = \zeta_0 + \zeta_1 y_{1,m}^{j} + \zeta_2 \ln(Participants_m) + \zeta_3 PD_d + \zeta_4 \ln(Participants_m) \times PD_d + \xi_m + \epsilon_m^{j}, \quad \text{(E1)}$$

where $y_{2,m}^{j}$ is the relevant voting outcome in the second round in municipality m, $y_{1,m}^{j}$ is the relevant voting outcome in the first round, $Participants_m$ is the number of participants who rally the protests against J.M. Le Pen, PD_d is the press diffusion per capita in department d, and ξ_m is a vector of municipality classifications.

As noted by Hainmueller et al. (2019), multiplicative interaction models may be biased if the linear interaction effect assumption does not hold. To show that our baseline results are robust to violation of this assumption, we estimate a flexible interaction model that allows for non-linear interaction effects. The starting point of the model follows the steps of Hainmueller et al. (2019).

Step 1. Discretize the moderator variable PD_d into three groups and create a dummy variable corresponding to each group.

$$G_1 = \begin{cases} 1, PD_d < \delta_{1/3} \\ 0, \text{otherwise} \end{cases} \quad G_2 = \begin{cases} 1, PD_d \in [\delta_{1/2}, \delta_{2/3}) \\ 0 \text{ otherwise} \end{cases} \quad G_3 = \begin{cases} 1, PD_d \geq \delta_{2/3} \\ 0 \text{ otherwise} \end{cases} \quad \text{(E2)}$$

where $\delta_{1/3}$ and $\delta_{2/3}$ correspond to the first and second terciles of the moderator, respectively.

Step 2. Within each group, define the evaluation point x_j as being the median of the group. These three evaluation points — which correspond to low, intermediate, and high values of the moderator — are the values of the moderator where the conditional marginal effects of the number of participants on the voting outcome will be computed. These conditional marginal effects are assumed to be linear within each group but are allowed to freely vary across each group.

Step 3. The flexible interaction model reads:

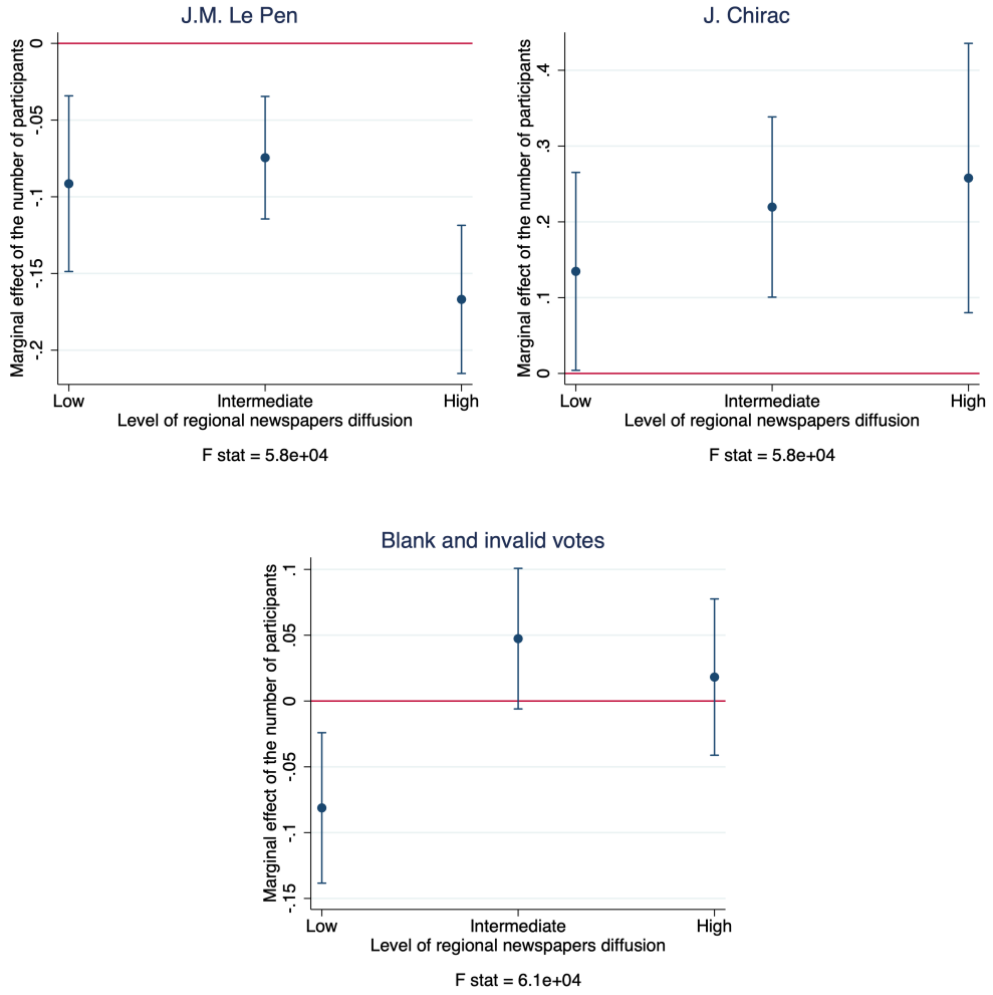
$$y_{2,m}^{j} = \sum_{j=1}^{3} \{\mu_j + \alpha_j \ln(Participants_m) + \eta_j (PD_d - x_j) + \beta_j \ln(Participants_m) \times (PD_d - x_j)\} \times G_j + y_{1,m}^{j} + \xi_m + \epsilon_m^j. \quad \text{(E3)}$$

Equation (E3) has six endogenous variables: $\ln(Participants_m) \times G_j$ and $\ln(Participants_m) \times (PD_d - x_j) \times G_j$ $(j = 1,2,3)$. A way to address endogeneity is to use in each local linear regression a 2SLS estimator instead of an OLS estimator. $\ln(Participants_m) \times G_j$ will thus be instrumented by $z_m \times G_j$ while $\ln(Participants_m) \times (PD_d - x_j) \times G_j$ will be instrumented by $z_m \times (PD_d - x_j) \times G_j$, where z_m is the rain-based synthetic instrument generated in Section 4.2 of the paper. We thus have six instruments.

However, these six instruments have poor predictive power, resulting in weak-instruments issues (F stat < 10). To tackle this problem, we augment these six (external) instruments with generated (internal) instruments, following Lewbel (2012)[1]. Intuitively, Lewbel's (2012) approach uses the presence of heteroscedasticity in the error term of the first stage to generate a set of instruments from a set of covariates.

[1] To implement Lewbel's (2012) approach, we use the Stata command developed by Baum and Schaffer (2019).

Figure E1. Marginal effect of the number of participants on the voting outcome conditional of press diffusion



Note: The unit of analysis is a municipality. The dotted lines indicate 90% confidence intervals based on standard errors clustered at the department level. In each specification, we control for municipality type.

Figure E1 documents the marginal effect of the number of participants on the voting outcome conditional of press diffusion, as estimated in Equation (E3). For each specification, the F statistic is well beyond 10, which rules out any weak instrument concerns. The results show similar trends to the baseline, with the caveat that the estimates are likely to be less precise than the conventional interaction IV model used in the paper, since Lewbel (2012) relies upon higher order moments to identify the parameter of interest.

Appendix F. The geography of the effect

Appendix F1. Arrondissements excluding municipalities with protests

In this subsection, we provide another way of computing spatial spillovers. Following Mamo et al. (2019), we aggregate observations at a higher geographical level, in our case *arrondissements,* and re-estimate the baseline specification.[2,3] First, we aggregate participants and votes in arrondissements using all the municipalities of each arrondissement. The regression is therefore similar to the baseline but performed at an aggregated level. Second, we aggregate participants using all municipalities but aggregate votes using only the municipalities that did not experience any protest. This regression only captures spillovers. A positive coefficient would mean that within each arrondissement, the number of participants of a municipality had a positive impact on the voting outcomes of the other municipalities and vice versa.

Panel A of Table F1 presents the outcome of regressions where voting outcomes are aggregated for all municipalities. The estimates, although smaller, remain qualitatively similar to the baseline despite the higher level of aggregation.

In Panel B, we report the outcome of regressions where voting outcomes are only aggregated for municipalities that hosted no protest. Those estimations therefore report pure spillover effects. The first noteworthy finding is that the

[2] A*rrondissements* are a supra-municipal administrative division. There are 324 arrondissements in mainland France. Their average population is 183,257, with a standard deviation of 957, and their average landmass is 1,981 km^2, with a standard deviation of 5.

[3] As more than one third of arrondissements did not have a protest, we followed the same steps as in the baseline estimates and generated an instrument using a two-part model based on weather conditions aggregated at the arrondissement level. We observe that arrondissements with a higher proportion of city-center and isolated municipalities had a higher probability of hosting a protest, as shown in Table F2. In addition, similar to the baseline, we observe that rain had a detrimental effect on the number of participants at the arrondissement level (Figure F1). The results of the first-stage estimates are reported in Table F3. Because the municipalities of Paris, Strasbourg, and Metz each constitute an arrondissement on their own, we drop them from the sample. Their inclusion does not, however, affect the results.

number of participants is statistically significant at conventional levels in all specifications. Accordingly, the effects of protests on voting outcomes were not only limited to the municipalities in which they took place but also extended to other municipalities within the arrondissement. Those spillovers were, as expected, negative for J.M. Le Pen and abstentions and blank ballots, but positive for J. Chirac.

Table F1. Spatial spillovers – Aggregating observations at the level of arrondissements

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Panel A. Votes aggregated over all municipalities in each arrondissement			
Number of participants (ln)	-0.137	0.689	-0.173
	(0.0518)	(0.168)	(0.0479)
First-round outcome	0.991	0.904	0.449
	(0.0476)	(0.112)	(0.0816)
F statistic	70.01	77.86	78.07
Panel B. Votes aggregated over all municipalities in each arrondissement – Pure spillovers			
Number of participants (ln)	-0.124	0.651	-0.195
	(0.0464)	(0.138)	(0.0395)
First-round outcome	0.997	0.894	0.369
	(0.0480)	(0.102)	(0.0769)
F statistic	85.48	87.88	85.33
Observations	321	321	321

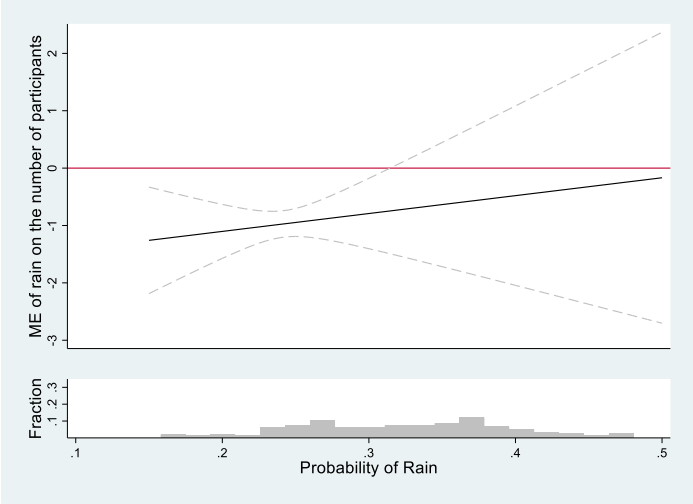
Notes. 2SLS estimates. The unit of analysis is an arrondissement. The dependent variable of each specification is reported at the top of each panel. The first stage is reported in Table F3. In each specification, we control for the proportion of city center and isolated municipalities. Standard errors clustered at the department level are reported in parentheses.

Table F2. Two-part model – Arrondissement level

			(1) Coef.	(2) St. errors
Panel A. First part – Probability of experiencing a protest				
Municipality type				
Proportion of city center and isolated municipalities			2.033	0.895
Panel B. Second part – Number of participants conditional on experiencing a protest				
Rain variables				
Rainy protest			-1.726	1.445
Historical rain frequency			-0.293	5.874
Rainy protest*Historical rain frequency			3.117	5.947
Municipality type				
Proportion of city center and isolated municipalities			1.798	0.949
Municipal characteristics				
Change in unemployment rate |b| 1995-2001			0.00556	1.336
Landmass			0.000476	9.16e-05
Population Size			4.06e-06	5.81e-07
Average age			0.195	0.148
% of ind. aged 60 and +			-12.65	8.767
First part		*Second part*		
Observations	321	Observations		204
Log pseudolikelihood	-207.673	Log pseudolikelihood		-285.996
Zero outcomes	117	R-squared		0.493

Notes. The unit of analysis is an arrondissement (supra-municipal administrative division). $Rainy\ protest$ is a binary variable equal to 1 if it rained more than 1 millimeter on May 1, 2002 (the protest day), and 0 otherwise. $Historical\ rain\ frequency$ is the share of May 1 days where rainfall on that day was higher or equal to 1 millimeter between 1983 and 2001. Column (2) reports standard errors clustered at the department level. ***Significant at 1% level; **Significant at 5% level; *Significant at 10% level.

Figure F1. Marginal effect of rain on the number of participants as a function of historical rain frequency on May 1 – Arrondissement level



Notes. The unit of analysis is an arrondissement. The graph plots the marginal effect of rain on the number of participants against historical rain frequency in municipalities that experienced a protest (the second part of the two-part model). The dotted lines indicate 90% confidence intervals based on standard errors clustered at the department level. In each specification, we control for municipality type.

Table F3. Arrondissement level – First-stage estimates

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Panel A. Arrondissements with all municipalities – No spillovers			
$\ln(\widehat{Participants})$	2.618	2.816	2.841
	(0.313)	(0.319)	(0.322)
First-round outcome	-0.171	0.146	-0.272
	(0.067)	(0.065)	(0.089)
Panel B. Arrondissements excluding municipalities with protests – Spillovers			
$\ln(\widehat{Participants})$	3.153	3.247	3.087
	(0.341)	(0.346)	(0.334)
First-round outcome	-0.071	0.145	-0.465
	(0.067)	(0.057)	(0.068)
Observations	321	321	321

Notes. The unit of analysis is an arrondissement. The dependent variable of each specification is reported at the top of each panel. In each specification, we control for the proportion of city center and isolated municipalities.

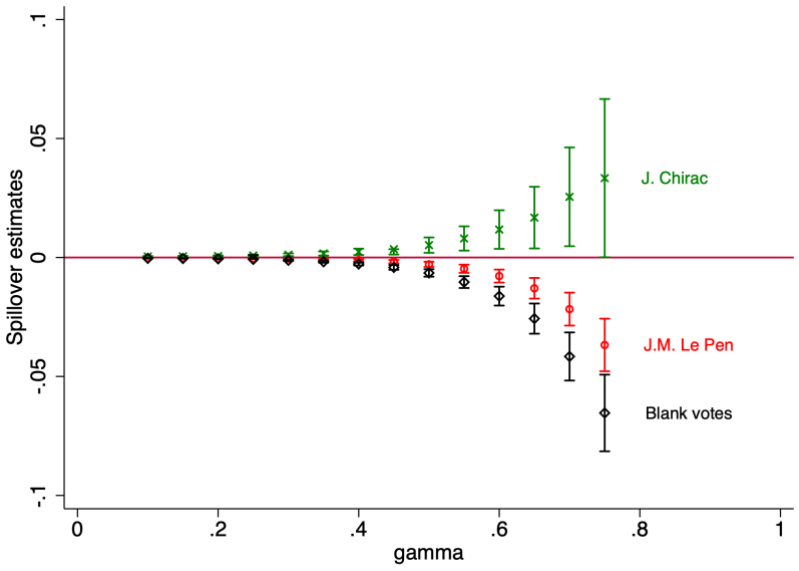
Appendix F2. Spatial model

Table F4. SLX Spatial Model – Effect of the spatial lag of rain on the spatial lag of the number of participants (ln)

	(1) Coef.	(2) St. errors
Rain (continuous)	-0.000240	1.14e-05
Observations	36,153	

Notes. The unit of analysis is a municipality. *Rain* refers to the spatial lag of the intensity of precipitation on the day of protests (May 1, 2002). The specification controls for the spatial lag of average rainfall (quintile dummies) and an unweighted linear combination of the spatial lag of continuous municipal characteristics (population, landmass, average age, variation in the unemployment rate between 1995 and 2002, and the proportion of individuals aged 60 years or older). Column (2) reports standard errors clustered at the department level.

Figure F2. Gamma sensitivity



Notes. The figure indicates how the spillover effects of protests vary with the distance decay parameter γ.

Table F5. SLX Spatial Model – First stage

ln(number of participants)	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Panel A. Outcome: ln($Participants$)			
ln($\widehat{Participants}$)	9.842	9.860	9.842
	(0.954)	(0.955)	(0.953)
$\sum_{k=1}^{n} w_{mk}$ ln($\widehat{Participants}$)	-0.564	-0.569	-0.584
	(0.794)	(0.795)	(0.795)
Panel B. Outcome: $\sum_{k=1}^{n} w_{mk}$ ln($Participants$)			
ln($\widehat{Participants}$)	-0.237	-0.162	-0.251
	(0.158)	(0.165)	(0.151)
$\sum_{k=1}^{n} w_{mk}$ ln($\widehat{Participants}$)	8.464	8.468	8.403
	(0.143)	(0.143)	(0.145)
Observations	36,153	36,153	36,153

Notes. The unit of analysis is a municipality. The dependent variable is the number of participants (ln). In each specification, we control for municipality type. Robust standard errors are reported in parentheses.

Table F6. Spatial spillovers with a threshold

Second-round outcome	(1) J.-M. Le Pen	(2) J. Chirac	(3) Abstentions and blank/invalid ballots
Direct effect	-0.386	0.778	-0.272
	(0.0806)	(0.160)	(0.0708)
Spatial spillover effect	-0.0986	0.258	-0.243
	(0.0301)	(0.0871)	(0.0431)
F statistics	554.4	514.7	583.9
Observations	36,153	36,153	36,153

Notes. 2SLS Estimates. The unit of analysis is a municipality. The dependent variable of each specification is reported at the top of each panel. In each specification, we control for municipality type. Robust standard errors are reported in parentheses.

Appendix G. Individual behaviors

Wording of the question used in Figure 5

If the respondent voted in the second round of the 2002 presidential election, which candidate did you vote for? (Si l'interviewé a voté au second tour de l'élection présidentielle 2002. Pour quel candidat avez-vous voté ?).

Wording of the questions used in Figure 6

Immigration: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: immigrants (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - Les immigrés);

Security: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: security (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - La sécurité);

Traditional values: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: the defense of traditional values (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - La défense des valeurs traditionnelles);

Criticism against the political class: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: the criticism of the political class (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - Les critiques contre la classe politique);

The abolition of income tax: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: the abolition of the income tax (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - La suppression de l'impôt sur le revenu);

Exit of France from the EU: More specifically, do you approve or disapprove of Jean-Marie Le Pen's positions on: France's exit from the EU (Plus précisément, approuvez-vous/ou désapprouvez-vous les prises de position de Jean-Marie Le Pen sur : - La sortie de la France de l'Union Européenne);

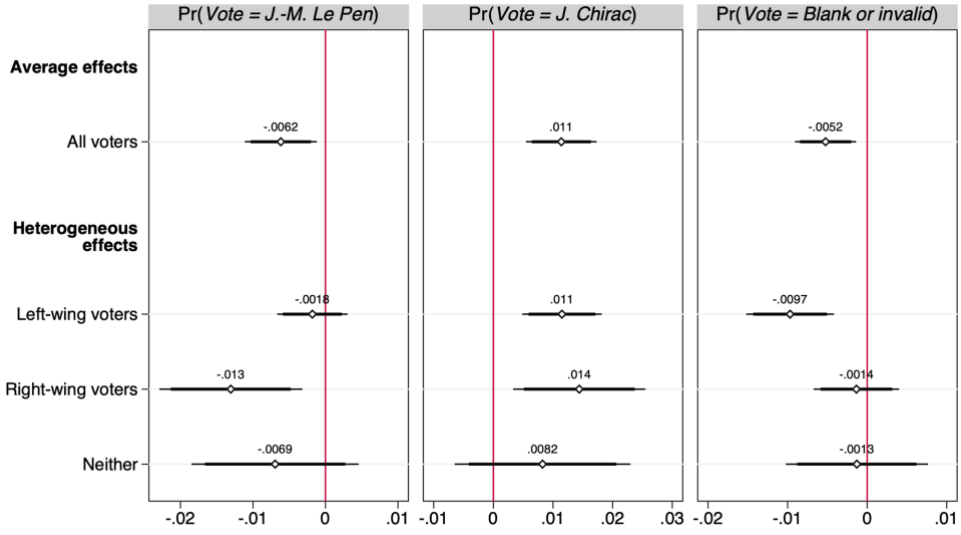
Too many immigrants: Here is a list of sentences. For each of them, can you tell me if you completely agree, somewhat agree, somewhat disagree, or completely disagree? There are too many immigrants in France (Voici maintenant une liste de phrases. Pour chacune d'elles, pouvez-vous me dire si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord ou pas d'accord du tout ? - Il y a trop d'immigrés en France);

Imigrants enrich a culture: Here is a list of sentences. For each of them, can you tell me if you completely agree, somewhat agree, somewhat disagree, or completely disagree? The presence of immigrants in France is a source of cultural enrichment (Voici maintenant une liste de phrases. Pour chacune d'elles, pouvez-vous me dire si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord ou pas d'accord du tout ? - La présence d'immigrés en France est une source d'enrichissement culturel).

Wording of the question used in Table 4

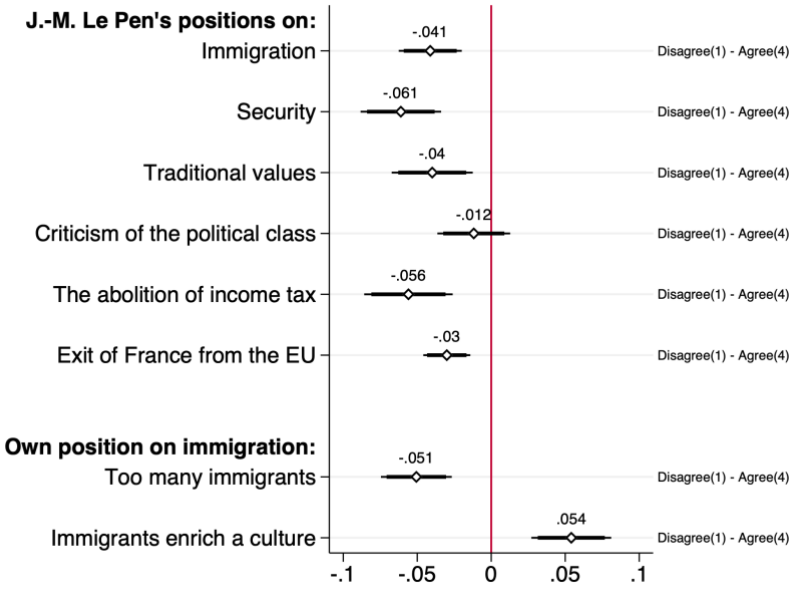
If the respondent voted in the first round of the 2002 presidential election, which candidate did you vote for? (Si l'interviewé a voté au premier tour de l'élection présidentielle 2002. Pour quel candidat avez-vous voté ?).

Figure G1. Average marginal effect of the number of participants (ln) on the probability of voting for Jean-Marie Le Pen, voting for Jacques Chirac, or casting a blank or invalid ballot – No controls



Notes. The unit of analysis is a survey respondent. The linear first stage and the second stage multinomial probit model are jointly estimated in a conditional mixed-process framework (Roodman, 2011). The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for type of municipality where respondents live. The specification of the average effect further controls for respondents' political orientation. The thick black line indicates 90% confidence intervals, while the thin black line indicates 95% confidence intervals. The confidence intervals are based on standard errors clustered at the department level. The F statistics of the linear first stage range from 96 to 153 across specifications. The exact wording of the questions is reported in Appendix G.

Figure G2. Marginal effect of the number of participants (ln) on the support of policies championed by Jean-Marie Le Pen – No controls



Notes. The unit of analysis is a survey respondent. 2SLS estimates. The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for respondents' political orientation and the type of municipality where they live. The thick black line indicates 90% confidence intervals while the thin black line indicates 95% confidence intervals. Confidence intervals are based on standard errors clustered at the department level. The F statistics range from 125 to 161 across specifications. The exact wording of the questions is reported in Appendix G.

Table G1. Average marginal effect of the number of participants on the probability of declaring a vote for Jean-Marie Le Pen, Jacques Chirac, or a blank or invalid ballot – No controls

	(1)	(2)	(3)
Second-round outcome	J.-M. Le Pen	J. Chirac	Blank or invalid ballot
Number of participants (ln)	-0.0105	-0.00401	-0.00300
	(0.00428)	(0.00294)	(0.00210)
F statistics	124.08	124.08	124.08
Observations	3,244	3,244	3,244

Notes. The unit of analysis is a survey respondent. The linear first stage and the second stage multinomial probit model are jointly estimated in a conditional mixed-process framework (Roodman, 2011). The number of participants is instrumented by the logarithm of the synthetic instrument defined in Section 4.2. In each specification, we control for respondents' political orientation and the type of municipality where they live. Standard errors clustered at the department level are reported in parentheses. The exact wording of the questions is reported in Appendix G.

References

Baum, C., & Schaffer, M. (2019). IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments.

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly Exogenous. *Review of Economics and Statistics*, *94*(1), 260-272.

Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How Much should we trust estimates from multiplicative interaction models? Simple Tools to Improve Empirical Practice. *Political Analysis*, 27(2), 163-192.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67-80.

Mamo, N., Bhattacharyya, S., & Moradi, A. (2019). Intensive and extensive margins of mining and development: Evidence from Sub-Saharan Africa. *Journal of Development Economics*, *139*, 28-49.

Mullahy, J. (1998). Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics*, *17*(3), 247-281.

Nevo, A., & Rosen, A. M. (2012). Identification With Imperfect Instruments. *Review of Economics and Statistics*, *94*(3), 659-671.

Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with cmp. *The Stata Journal*, *11*(2), 159-206.

van Kippersluis, H., & Rietveld, C. A. (2018). Beyond plausibly exogenous. *Econometrics Journal*, 21(3), 316-331.