



HAL
open science

Essays in political economy and public economics

Clémence Tricaud

► **To cite this version:**

Clémence Tricaud. Essays in political economy and public economics. Economics and Finance. Institut Polytechnique de Paris, 2020. English. NNT : 2020IPPAX019 . tel-02865258

HAL Id: tel-02865258

<https://theses.hal.science/tel-02865258>

Submitted on 11 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Essais en Économie Politique et Publique

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École Polytechnique

École doctorale n°626 : Ecole Doctorale de l'Institut Polytechnique de
Paris (ED IP Paris)

Spécialité de doctorat: Sciences économiques

Thèse présentée et soutenue à Paris, le 04/06/2020, par

Clémence Tricaud

Composition du Jury :

Ekaterina ZHURAVSKAYA Professeure, École d'Économie de Paris	Président
Julia CAGÉ Professeure Assistante, Sciences Po	Rapporteur
Pierre-Philippe COMBES Professeur CNRS, GATE-LSE, Lyon-ENS	Rapporteur
Alberto ALESINA Professeur, Harvard University	Examineur (Absent)
Edward GLAESER Professeur, Harvard University	Examineur
Vincent PONS Professeur Associé, Harvard Business School	Examineur
Pierre BOYER Professeur, CREST-École Polytechnique	Directeur de thèse
Bruno CRÉPON Professeur, CREST-École Polytechnique	Co-Directeur de thèse

Acknowledgments - Remerciements

This dissertation is inspired and shaped by my interactions with many great people, and has been made possible thanks to their support.

I would like first to thank my two mentors, Pierre Boyer and Vincent Pons.

Vincent Pons has provided me with consistent guidance for more than five years, and I am extremely grateful for his support. He guided my first steps in research during my pre-doctoral internship, supervised my work during my PhD, and shared his experience when working on our joint projects. Working with him was—and still is—the best training I could hope for. His intellectual honesty will always inspire me.

I met Pierre Boyer in the classroom, during my first political economy course. I was impressed by the quality and depth of this teaching, and thrilled when he agreed to become my PhD advisor. I discovered what advising means: thanks to him, I always felt supported and I always knew I had someone watching my back. On top of providing key and detailed advice on my research, he kept paying attention to me, aware of my weaknesses and fears, and caring for my well-being. I will never thank him enough for his attention to all the details and his ability to plan ahead the major steps of my thesis.

I also benefited from the insightful guidance of my co-advisor, Bruno Crépon. He introduced me to the tools of policy evaluation during my undergraduate studies, and provided me with crucial advice on the empirical methods used in this dissertation.

I would like to thank Alberto Alesina and Edward Glaeser who played a key role during my visiting periods at Harvard University. The time spent in Harvard has been determinant for the completion of my thesis and this experience would not have been possible without their support. Alberto Alesina welcomed me in the Economics Department and made me part of the amazing political economy group. He followed my work since the beginning and his deep and powerful insights greatly improved my thesis. Edward Glaeser detected the urban economist in me, and our fascinating discussions shaped the last part of this dissertation.

I am also honored to have Julia Cagé, Pierre-Philippe Combes, and Ekaterina Zhuravskaya in my jury. I thank Julia Cagé and Pierre-Philippe Combes for agreeing to be reviewers. The time spent with Julia Cagé has always been a powerful source of motivation, and her involvement in the public debate a great source of inspiration. Pierre-Philippe Combes shared with me his vast knowledge of urban economics. I learned a lot in his class, and I thank him for

introducing me to his world.

I benefited from the advice of many great researchers in France and abroad. I would like to thank them for their support: Philippe Aghion, Laurent Bouton, Micael Castanheira, Guillaume Chapelle, Xavier D'Haultfoeuille, Andrew Eggers, Francis Kramarz, Benjamin Marx, Rida Laraki, Yukio Koriyama, Alessandro Riboni, Benoit Schmutz, and Stefanie Stantcheva. I thank Allan Drazen for sharing his perspective on the political economy literature, and for the delightful dinners we shared. A particular thanks to Isabelle Méjean who played a key role in making my time in Harvard possible, who provided me with extremely valuable feedbacks, and who has been a fierce tennis partner. Finally, I would like to thank Clémentine Van Effenterre who guided me through all the difficult steps of the PhD and always knew when I needed her most.

I also benefited a lot from helpful and stimulating discussions with the members of the Political Economy seminar at the Harvard Government Department. The organizers gave me the great opportunity to present my work regularly in their workshop. In particular, thanks to Jame Snyder, Jeffrey Frieden, Pia Raffler, and Ken Shepsle.

I would like to thank the PhD students I met in my home institution and abroad. Discussing my research with them and sharing our PhD experience was a huge help. Special thanks to Antoine Ferey, with whom I shared my office: he was always ready to brainstorm and helped me think deeper about the theoretical foundations of my work. Thanks also to: Alicia Marguerie who welcomed me in CREST, Audrey Rain, Yvan Ousse, Daphné Scandalis, Antoine Bertheau, Clémence Lenoir, Bérangère Patault, Yannick Guyonvarch, Malka Guillot, Anasuya Raj, Reda Aboutajdine, Pierre-Edouard Collignon, Thomas Delemotte, Germain Gauthier, Elio Nimier-David, Emilie Sartre, and Jérôme Trinh. Thanks to the PhD students who made my time far from home easier: Anaïs Goburdhun, Martin Fernandez Sanchez, Lydia Assouad, Gloria Gennaro, Gemma Dipoppa, Ellora Derenoncourt, Elisa Macchi, Cathrin Mohr, Louise Paul-Delvaux, Itzhak Tzachi Raz, and Moya Chin.

De nombreuses personnes en France m'ont aidé à obtenir les données empiriques essentielles à ce travail. Je remercie Nicolas Sauger et Abel François pour leur aide dans la collecte des données électorales françaises; l'équipe de Médiamétrie pour leur partage des données d'audience radio et télévisée en France, en particulier Philippe Tassi, David Bernier, Jean-Pierre Panzani et Anne-Marie Dussaix; Brice Fabre pour les données sur les comptes des communes et son aide précieuse pour les déchiffrer; l'équipe en charge des données sur les permis de

construire au Ministère du Développement Durable; Bernard Pelamourgues pour le partage des données de la CAF sur les crèches municipales; Cécile Queffelec pour le partage des enquêtes du Ministère de la Culture sur les bibliothèques municipales; Laurent Chevereau du CEREMA pour m'avoir transmis les données sur les transports urbains et avoir passé de longs moments à me les expliquer.

J'aimerais aussi remercier celles et ceux, au sein du CREST-Ecole Polytechnique, qui m'ont permis de réaliser ma thèse dans les meilleures conditions: Pascale Deniau, Murielle Jules, Weronika Leduc, Eliane Madelaine, Lyza Racon, Arnaud Richet, Fanda Traore et Edith Verger. Merci à Francis Kramarz, directeur du CREST pendant ma thèse, pour son soutien et pour nos discussions sur l'avenir des communes françaises.

Pour finir, j'aimerais remercier mes proches. Mon père Christophe pour son amour de la concision et de la logique; ma mère Brigitte pour son écoute toujours bienveillante et ses conseils toujours pertinents; mon beau-père Wassim pour m'avoir montré que l'important est avant tout le chemin. Merci à mes deux petites sœurs, Lucie et Emma, qui me rappellent à chaque instant l'importance d'être heureuse. Merci à mes amis les plus proches, Aurélia, Mathilde, Thomas et Georges pour leur soutien indéfectible dans les moments éprouvants, et leur présence toute aussi indéfectible dans les grands moments de joie. Enfin et surtout, merci à mon mari et co-équipier, Jean-Charles. Lui qui a toujours su où se situait l'essentiel, qui m'a toujours comprise et soutenue, et qui se lance avec moi dans une nouvelle aventure.

To Emma and Lucie,

To Jean-Charles,

Summary

This dissertation lies at the intersection of Political Economy and Public Economics. It combines quasi-experimental designs with administrative data to understand the determinants and consequences of citizen and policymaker behavior.

The first two chapters, in collaboration with Vincent Pons, study the behavior of voters in elections. In Chapter 1, we assess the extent to which voters behave expressively or strategically and the consequences it has on electoral results. Using a regression discontinuity design (RDD) around the qualification threshold for the runoff of two-round elections in France, we compare electoral results when two or three candidates compete. We find that a third candidate's presence substantially increases participation and reduces the vote share of the top two candidates. Crucially, in about 1/5 of the elections, the presence of a third candidate causes the loss of the candidate among the top two that is ideologically closest to her. Hence, we show that the plurality rule often leads to suboptimal outcomes, because many voters value voting expressively over voting strategically.

Chapter 2 investigates which type of information affects voter and candidate behavior, focusing on one specific piece of information: candidate rankings. Building on the data collected for the first project and digitizing electoral results for older elections, we isolate the impact of first-round rankings on second-round outcomes. We use an RDD in French local and parliamentary elections, where up to three or four candidates can qualify for the second round. Results show that arriving marginally first, second, or third in the first round has large effects on a candidate's likelihood of running and winning in the second round. These results remain strong even when only two candidates qualify (and thus there is no need for coordination), suggesting that the desire to vote for the winner is an important driver of voter behavior.

Chapter 3 studies the behavior of policymakers and seeks to understand mayors' reluctance to cooperate over local public policies. Exploiting a 2010 reform in France that forced non-integrated municipalities to join an intermunicipal community, I assess the causal impact of integration on municipalities that were reluctant to cooperate. I provide new evidence that resistance is driven by local costs of integration: urban municipalities resisted integration to prevent further housing construction, while rural municipalities resisted to avoid losing local public services. Using a difference-in-differences strategy, I first find that municipalities forced to enter a community experienced a large increase in the number of building permits delivered

per year. Consistent with NIMBYsm (Not In My BackYard) explaining urban municipalities' resistance, this effect is driven by high-demand and densely built municipalities, where a rise in construction is the most likely to create congestion. Second, exploiting new data on daycare facilities and public libraries, I show that rural municipalities forced to enter a community experienced a decrease in the number of facilities available in their territory. Finally, I find that resisting municipalities benefited from a greater access to public transport and higher fiscal revenues after integration, but that these benefits were not high enough to compensate for the costs of integration.

Field: Economics

Key words: Political Economy; Public Economics; Elections; Voting; Local public policies; Intermunicipal cooperation; Regression-discontinuity-design; Difference-in-differences

Résumé

Cette thèse s'inscrit dans les domaines de l'Économie Politique et Publique. À travers l'utilisation de méthodes quasi-expérimentales et de données administratives, son objectif est d'atteindre une meilleure compréhension du comportement des citoyens et des élus.

Les deux premiers chapitres, en collaboration avec Vincent Pons, étudient le comportement des électeurs. Le premier chapitre évalue dans quelle mesure les électeurs votent de manière expressive ou stratégique et les conséquences sur les résultats électoraux. Nous exploitons le seuil qui détermine la qualification des candidats pour le second tour des élections législatives et cantonales françaises. En utilisant la méthode de la régression sur discontinuités (*Regression Discontinuity Design* RDD), nous comparons les résultats électoraux selon que deux ou trois candidats se présentent. Les résultats montrent que la présence d'un troisième candidat augmente significativement la participation électorale et réduit le nombre de voix obtenue par les deux premiers candidats. Surtout, dans environ 1/5^{ème} des élections, la présence d'un.e troisième candidat.e provoque la perte du candidat qui, parmi les deux premiers, était le plus proche idéologiquement. Nous montrons ainsi qu'une large fraction des électeurs préfère voter expressivement plutôt que stratégiquement et que, dans ce contexte, la règle majoritaire conduit souvent à des résultats sous-optimaux.

Le deuxième chapitre explore le type d'information susceptible d'affecter le comportement des électeurs et des candidats, en évaluant l'impact d'une information spécifique: le classement des candidats. En utilisant une RDD, nous isolons l'impact du classement des candidats au premier tour sur les résultats du second tour, au sein d'élections françaises où jusqu'à trois ou quatre candidats peuvent se qualifier pour le second tour. Nous montrons qu'être classé.e marginalement premier.ère, deuxième ou troisième au premier tour augmente significativement la probabilité qu'un.e candidat.e maintienne sa candidature au second tour et gagne. L'impact reste significatif même lorsque deux candidats seulement se qualifient (et qu'il n'y a donc pas de place pour le vote stratégique), ce qui suggère que le désir de voter pour le.a vainqueur.e joue un rôle clé dans la décision de vote des électeurs.

Le troisième chapitre étudie le comportement des élus et en particulier la réticence des maires à coopérer en matière de politiques publiques. J'exploite la loi de 2010, qui en France, a forcé les communes non intégrées à rejoindre une intercommunalité. J'étudie l'impact causal de leur intégration en utilisant la méthode des doubles différences. Mes résultats montrent que les

communes résistaient afin d'échapper aux conséquences économiques locales de l'intégration: les communes urbaines résistaient afin d'éviter la construction de nouveaux logements, tandis que les communes rurales résistaient afin d'éviter la perte de services publics locaux. Je constate premièrement que les communes contraintes d'entrer dans une intercommunalité ont connu une augmentation significative du nombre de permis de construire délivrés par an. Cet effet se manifeste principalement dans les communes densément construites et qui font face à une forte demande immobilière: des communes où davantage de constructions accentuerait la congestion, et dans lesquelles les résidents ont donc de forte chance d'être sujets au syndrome NIMBY (*Not In My BackYard*). Deuxièmement, je montre, à travers l'exploitation de données sur les crèches et bibliothèques municipales, que les communes rurales ont subi une diminution du nombre de services publics de proximité. Enfin, les résultats révèlent que les communes forcées de s'intégrer ont bénéficié de revenus fiscaux plus élevés et d'un meilleur accès aux transports publics, mais que ces bénéfices n'étaient pas suffisants pour compenser les effets négatifs de l'intégration.

Discipline: Sciences économiques

Mots-clés: Économie Politique; Économie Publique; Élections; Comportements de vote; Politiques publiques locales; Coopération intercommunale; Régression sur discontinuités; Doubles différences

Contents

Acknowledgements - Remerciements	i
Summary	v
Résumé	vii
General Introduction	1
Introduction Générale	13
1. Expressive Voting and its Cost: Evidence from Runoffs with Two or Three Candidates	28
1 Introduction	29
1.1 Contribution to the literature	32
2 Research setting	34
2.1 French parliamentary and local elections	34
2.2 Data	35
2.3 Vote share of the third candidate	36
3 Empirical strategy	37
3.1 Evaluation framework	37
3.2 Identification	38
4 Empirical results	41
4.1 Impact on participation and candidate votes	41
4.2 Impact on votes going to the top two candidates	43
4.3 Impact on the top two candidates depending on political orientation	45
4.4 Impact on the top two candidates depending on the strength of the third	47
4.5 Impact on winner identity	48
5 Interpretation of the results and discussion	50
5.1 Voters' behavior	50
5.1.1 Loyals and switchers	50
5.1.2 Theoretical implications	52
5.2 Candidates' decision to drop out	55

6 Conclusion	58
Appendix	60
2. Coordination and Bandwagon Effects of Candidate Rankings: Evidence from Runoff Elections	87
1 Introduction	88
2 Empirical strategy	92
2.1 Setting	92
2.2 Data	93
2.3 Evaluation framework	96
2.4 Identification assumption	97
3 Empirical results	102
3.1 Impact on winning	102
3.2 Impact on running	106
3.3 Impact on winning and vote shares conditional on running	108
4 Mechanisms	113
4.1 Impact depending on political orientation	113
4.2 The role of coordination	117
4.3 Party norms and bandwagon effect	119
4.4 Discussion of alternative mechanisms	121
5 Conclusion	125
3. Better Alone? Evidence on the Costs of Intermunicipal Cooperation	128
1 Introduction	129
2 Institutional background and data	134
2.1 Intermunicipal cooperation in France	134
2.2 The 2010 law	136
2.3 Data	137
3 Empirical Strategy	139
3.1 Treatment and control groups	139
3.2 Specification and identification	142

4	Main results: the costs of integration	144
4.1	Impact on housing supply	144
4.1.1	Impact on municipalities forced to integrate	144
4.1.2	Impact on municipalities that voluntarily integrated	147
4.1.3	Further evidence on the costs linked to housing supply	149
4.2	Impact on local public services	156
5	Evidence on the benefits of integration	160
5.1	Impact on public transport	160
5.2	Impact on fiscal revenues	163
6	Discussion	166
6.1	Interpretation	166
6.2	External validity	167
7	Conclusion	169
	Appendix	171
	References	209

General Introduction

The public economics literature studies the economics of the public sector. It analyzes how government policies affect the economy, and how they should be designed to optimally reach their goals. The political economy literature studies the role of political institutions. It analyzes how political constraints shape the decisions of political agents (e.g., voters, policymakers), and to which extent public policies are determined by their preferences and interests.

This dissertation lies at the intersection between the two fields and aims at better understanding both the determinants and consequences of citizen and policymaker behavior. The first two chapters study the determinants of voters' behavior at the electoral stage, and the consequences of their voting choices on political selection; the third chapter studies policymakers' behavior at the post-electoral stage, and the consequences of their policy choices on the economy.

In the remainder of the introduction, I first outline the theoretical motivations for studying voter behavior, and then for studying policymaker behavior. The last section provides an overview of each chapter and details the contribution of this dissertation to the literature.

Voter Behavior

A large theoretical literature has sought to model and predict citizens' voting choices.

Canonical models of electoral participation focus on two-candidate races and formalize the decision to vote as a cost-benefit calculation. Downs (1957a) posits that voters face a cost of voting and receive instrumental benefits depending on their likelihood to be pivotal—the probability that their vote changes the result of the election—and on the differential utility associated to the victory of their favorite candidate over the other. Given the very low probability of being pivotal in most elections, it only takes the costs of voting to be slightly positive for the net benefit of voting to be negative. Hence, while quite intuitive, this model cannot explain the high turnout rates we observe in reality. To solve this issue, game theory models endogenize the probability of being pivotal as the result of strategic interactions (Palfrey and Rosenthal, 1983; 1985). These models generate multiple equilibria, some of which involving high turnout rates, but at the cost of assuming that citizens have perfect information about the voting preferences

of others.

Departing from the assumption that citizens decide to vote based on pivot probabilities, recent models assume that participation is driven by ethical motives. Rule-utilitarian models consider an electorate divided into two groups, each one supporting one candidate, and advance that voters want to “do their part” to help their group win (Coate and Conlin, 2004; Feddersen and Sandroni, 2006). As in the previous models, utility is characterized by the tradeoff between the benefits derived from the victory of the favorite candidate, and the costs of voting. The difference lies in the fact that agents do not maximize their individual utility, but care about the group’s aggregate one. As a result, they vote according to the rule that, if followed by everyone else in their group, would maximize their group’s aggregate utility. In each group, the voting rule is defined by a cut-off point such that agents with voting costs below this threshold will vote for their favorite candidate, and others will abstain. Group rule utilitarian models generate predictions that are more in line with what we observe: they predict high turnout rates under imperfect information as well as realistic comparative statics, such as the fact that participation tends to be larger in close races.

Models of electoral participation only consider two-candidate races and thus do not address the question of voting choices in elections with more than two candidates. When more than two candidates are running, citizens who support lower-ranked candidates may face a trade-off: voting for their favorite, or for another candidate with higher chances of winning. In his groundbreaking work on strategic behavior, Duverger (1954) posits that voters do not want to “waste” their vote by voting for a candidate with a very low chance of winning, and thus vote only for the two front-runners. Myerson and Weber (1993) formalize this intuition under the classical game theory framework. They show that, when voters are instrumentally rational, an election with multiple candidates usually boils down to a two-candidate race, and of these two, the candidate who is preferred by the majority wins the election (see also Palfrey, 1989; Fey, 1997; Cox, 1997). Interestingly, rule-utilitarian models also predict that most votes will go to the front runners in multi-candidate elections, as the groups’ aggregate utility is only affected by who wins the election (Bouton and Ogden, 2017). All in all, voting models suggest that there is actually no loss of generality in considering two-candidate races when modeling elections.

However, do voters really consider their vote as “wasted” if not cast for a candidate with a chance of winning the election? One could think that they also care about expressing their true preferences and that voters derive some benefits from voting for their favorite candidate,

independently from her probability of winning. Under such assumption, we would predict that more than two candidates receive votes in equilibrium, as it is the case in real elections. Indeed, while the division of the American political landscape between Republicans and Democrats is a famous illustration of Duverger's law, in other countries third- and lower-ranked candidates frequently receive votes (see UK general elections for instance).

Assuming that voters are instrumentally rational or instead expressive plays a key role when predicting the number of candidates receiving votes in equilibrium. It is also crucial when assessing which electoral rule best aggregate preferences. In particular, under the plurality rule, a sufficiently large fraction of voters may need to engage in strategic manipulation in order to elect leaders who best correspond to citizens' preferences. If too many people vote according to their true preference instead of behaving strategically, then plurality rule may fail to choose the Condorcet winner, when one exists, thus decreasing the representativeness of the electoral outcome (Nurmi, 1983; Myerson and Weber, 1993).

Yet, there is still little systematic evidence about the actual drivers of voter behavior, and about the role played by their behavior in the aggregation of preferences. This is the focus of the first two chapters of this dissertation. Chapter 1 assesses the extent to which voters behave strategically or expressively, and the consequences it has on who gets elected. Chapter 2 focuses on the role of information and investigates the impact of past electoral results on electoral behaviors and outcomes.

Policymaker Behavior

Public policies are not only affected by the way preferences are aggregated, but also by how policymakers behave once in office.

First, as stressed by the literature on legislative bargaining, in practice, many public policies are negotiated by groups of individuals (e.g., legislatures, monetary policy boards) and are thus shaped by the bargaining process. For instance, Baron and Ferejohn (1989) study the decision making process of a committee where agents are perfectly informed and have to decide over the allocation of resources. They show that, in equilibrium, a minimum coalition is formed and resources are split unequally, with the agenda setter securing a higher share for herself. Under legislative bargaining, the equilibrium policies depend upon several factors: how impatient the legislators are, the voting rule inside the committee, as well as the characteristics of the status

quo and whether it is assumed to be endogenous or exogenous (Baron, 1996; Bowen et al., 2014).

Second, policies may be shaped by policymakers' electoral incentives. While elections can be used as a disciplining device to make officeholders accountable (Barro, 1973; Ferejohn, 1986), they can also generate opportunistic behaviors. When incumbents are self-interested and voters are myopic and imperfectly informed, Nordhaus (1975) shows that incumbents have the incentive to manipulate economic time paths to their electoral benefits. In particular, they will tend to stimulate the economy right before the election to reduce unemployment, ignoring the detrimental effect it will have on inflation after the election. Political business cycles can also occur with rational voters, when two ideological parties with different preferred policies compete for election (Alesina, 1987; Hibbs Jr, 1994). In this case, the cycle is driven by partisan forces and comes from the alternation of the two political parties in power.

In a context where policies can be influenced by legislative bargaining and electoral incentives, the scale of the decision making is key. As shown by Battaglini and Coate (2008), equilibrium policies differ depending on whether the decision is made by a single social planner or by a legislature where each legislator seeks to attract resources for her own district. In the latter case, policies are affected by the bargaining among legislators and by their desire to be re-elected at the local level, resulting in suboptimal tax rate, public good provision and debt levels. Similarly, Rodden (2002; 2006) stresses the risk coming from fiscal federalism: if the power is given to state governments, they are likely to spend and borrow excessively, attempting to shift their burdens onto the central government and on one another.

Both the political economy and public economics literature have long thought about what should be the optimal jurisdiction size at which policy decisions are taken. This question was addressed early on by Tocqueville (1835), Tiebout (1956) and Oates (1972), and more recently formalized by Alesina and Spolaore (1997; 2003). They argue that countries' sizes are the results of tradeoffs between the benefits and the costs of size. On one side, large countries benefit from economies of scale, reducing the per capita costs of public good provision; on the other side, large countries are more heterogeneous in terms of preferences and more individuals or regions end up dissatisfied with central policies. They also show that, when political agents decide over jurisdictions' sizes, we end up with too many jurisdictions compared to what a social planner would choose. This is due to the fact that political agents do not internalize the aggregate benefits of size in terms of public goods provision. As a result, political agents with

different preferences over public policies prefer to split and create their own jurisdictions.

Urban planning is one example of policies that involves tensions between aggregate benefits and local preferences is urban planning. As argued by the urban literature, more housing construction in high-productivity areas can be beneficial by making housing affordable, fostering worker mobility and avoiding urban sprawl (Glaeser and Gyourko, 2018; Hsieh and Moretti, 2019; Glaeser and Maré, 2001; Chetty et al., 2016; Glaeser and Kahn, 2010; Jones and Kammen, 2013). At the same time, residents living in those areas are likely to oppose new construction to preserve their quality of life and avoid having strangers settling in their backyard (see Gyourko and Molloy, 2015 for a review on NIMBYsm). Hence, ignoring the aggregate benefits of construction, local policymakers will prefer to keep control over zoning and housing policies rather than have them centralized.

Still, in practice, little is known about which public policies policymakers are reluctant to share. Over the last century, many central governments across the developed world have faced the opposition of mayors when trying to enforce intermunicipal cooperation. More empirical evidence is needed to understand why: are mayors reluctant to cooperate because they want to protect the interests of their constituents? Or are they only self-interested and do not want to lose their own political power? This is the focus of the last chapter. Chapter 3 assesses the impact of integration on municipalities that were forced to cooperate, to better understand the reasons why mayors were reluctant to share their public policies in the first place. A special attention is given to urban planning policies to assess the role played by NIMBYsm in explaining municipalities' resistance against cooperation.

This Dissertation

This dissertation builds upon these theoretical insights and presents three empirical investigations of political agent behavior. Using administrative data and quasi-experimental designs, my work aims at providing new causal evidence on the determinants and consequences of political agent behavior. We will first challenge voting models' assumption by showing that voters do not care only about who wins the election when deciding for whom to vote. We will provide evidence that they also care about voting expressively for their favorite candidate, and that it matters for how preferences are aggregated under the plurality rule. We will also explore which type of information affects candidate and voter behavior, and show that past electoral results

impact their choices, by enhancing coordination as well as through behavioral channels. Finally, we will study the behavior of policymakers, and provide evidence that the local impacts of consolidation help explain mayors' reluctance to share their public policies.

Chapter 1 - Expressive Voting and its Costs: Evidence from Runoffs with Two or Three Candidates (with Vincent Pons)

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

To assess the extent to which voters behave strategically, existing studies usually compare people's preferences and vote choices and count the number of voters who vote for a front-runner instead of their favorite. But voters' underlying preferences are difficult to observe, so these studies depend on the reliability of survey responses (e.g., Blais et al., 2001; Hillygus, 2007) or on assumptions regarding the mapping between voters' preferences and vote choices (e.g., Kawai and Watanabe, 2013; Spenkuch, 2015). This chapter uses a different approach: instead of estimating preferences, we focus on vote choices. We exploit the qualification threshold for the runoff of two-round elections in France, and compare electoral results when two or three candidates compete in the second round.

We find that a third candidate's presence substantially increases participation and reduces the vote share of the top two candidates. Crucially, in about 1/5 of the elections, the presence of a third candidate causes the loss of the candidate among the top two that is ideologically closest to her. Hence, we show that the plurality rule often leads to suboptimal outcomes, because many voters value voting expressively over voting strategically. Third candidates could prevent this effect by dropping out of the race between the first and second rounds, but those who have different political orientations from the top two usually refuse to do so.

Contribution. A large empirical literature examines whether voters behave strategically or ex-

pressively. To determine the actual proportion of voters voting for a candidate other than their preferred one, existing studies compare people's preferences and voting choices. Estimates based on surveys are typically low, below 20 percent (e.g., Alvarez and Nagler, 2000; Blais et al., 2001; Hillygus, 2007; Kiewiet, 2013), but potentially biased by misreporting. For instance, overreporting voting for the winner (Wright, 1993; Atkeson, 1999) may lead to overestimate strategic behavior. Alternatively, to avoid cognitive dissonance (Festinger, 1962), people may adjust their stated preference to their voting choice, which would lead to underestimating it.

A second strand of the literature relies on aggregate electoral results and studies strategic voting by imposing assumptions on the mapping between voters' preferences and vote choices. Kawai and Watanabe (2013) and Myatt and Fisher (2002) calibrate structural models to estimate the number of voters who did not vote for their preferred candidate and the impact of strategic voting on the number of seats won by a party, respectively. In the context of the German split-ticket voting system, Spenkuch (2018) compares votes cast for party lists under a proportional rule with votes cast for individual candidates under plurality rule, and reports that about one third of voters behave strategically (also see Spenkuch, 2015).

Instead of estimating voter preferences and comparing them with their actual choices, this paper focuses on vote choices only. We compare electoral outcomes when two versus three candidates are competing. Our strategy allows us to make three important contributions to the literature. First, we estimate the impact of the presence of the third candidate on both participation and vote shares and demonstrate that the third candidate obtains votes both from voters who would have voted for the top two had she been absent and from supporters who would have abstained or voted blank or null. Second, we can precisely estimate the fraction of races whose final outcome changes as a result of these behaviors. Third, in addition to studying voters' behavior, we highlight the main factors affecting candidates' own decision to stay in the race or drop out.

We further contribute to the theoretical literature on strategic voting by exploring the extent to which our results can be explained by rational voter models (e.g., Myerson and Weber, 1993), and group-rule utilitarian models (Coate and Conlin, 2004; Feddersen and Sandroni, 2006; Bouton and Ogden, 2017), in which the individual or group's utility is only affected by who wins the election.

Methodology. Our identification strategy exploits the discontinuity generated by the qualifica-

tion rule for the second round in French local and parliamentary elections. In those elections, the top two candidates in the second round automatically qualify for the second round; other candidates also qualify if they receive a number of votes higher than 12.5 percent of registered citizens. Using a regression discontinuity design (RDD), we compare second-round results in districts where the third candidate obtains a vote share just above or below the 12.5 percent threshold and, as a result, just passes or misses the qualification requirement. This strategy enables us to compare electoral outcomes under two versus three candidates, and to assess the impact of the presence of a third candidate. The identification assumption is that the distribution of potential confounders changes continuously around the 12.5 percent vote share threshold, so that the only discrete change occurring at this threshold is the shift in qualification.

Chapter 2 - Coordination and Bandwagon Effects of Candidate Rankings: Evidence from Runoff Elections (with Vincent Pons)

To form their expectations about others' preferences and strategies and make their own decisions, voters and candidates can rely on information provided by polls and past election results. Despite large evidence that political actors' overall level of information matters (e.g., Hall and Snyder, 2015; Le Pennec and Pons, 2019), little is known about which exact pieces of information they actually base their decisions on.

This chapter focuses on one specific piece of information, candidate rankings: the ordering of candidates in polls, in previous elections, or in the previous round (in elections with multiple rounds). While past and predicted vote shares provide detailed information on the distribution of preferences, rough-hewn rankings can serve as a coordination device in and of themselves. When more than two candidates are in the running, their past rankings can be used by strategic voters as a focal point to coordinate on the same subset of candidates, and by sister parties to determine which of their candidates should drop out. These coordination mechanisms can be reinforced by behavioral motives such as a bandwagon effect. Voters who desire to vote for the winner might decide to "jump on the bandwagon" and rally behind higher-ranked candidates.

In this chapter, we estimate the impact of rankings on voters and candidates' decisions and examine whether this impact reflects strategic coordination or bandwagon effects. To do so, we consider French two-round elections where up to three or four candidates can qualify for the

second round, and we isolate the impact of first-round rankings on second-round outcomes.

Results show that arriving marginally first, second, or third in the first round has large effects on a candidate's likelihood of running and winning in the second round. We find that the effects are much larger when the higher- and lower-ranked candidates have the same political orientation, and we also show comparative statics consistent with rankings being used as a coordination device both by candidates and voters. However, coordination alone does not suffice to explain the results: the impacts of rankings remain strong even when only two candidates qualify (and thus there is no need for coordination), suggesting that the desire to vote for the winner is an important driver of voter behavior.

Contribution. Our analysis builds on previous studies estimating the impact of candidate rankings across elections. Following Lee (2008), a large literature has examined the impact of ranking first (instead of second) on future elections and shown that close winners generally benefit from an incumbency advantage when they run again (e.g., Eggers et al., 2015). These studies have been unable to distinguish the effect of holding office from the pure effect of being labeled first, which we can instead isolate by studying two rounds of the same election. Complementing this work on the effect of ranking first, Anagol and Fujiwara (2016) show that arriving second (instead of third) in past elections also increases candidates' likelihood to run in the next elections and win them, and they attribute these effects to strategic coordination by voters. Our paper completes these studies in four important ways.

First, we can estimate the effect of arriving second or third, as well as the effect of arriving first (independently from incumbency advantage). Second, the short time span (one week) between the first and second rounds helps us isolating the direct effect of rankings from reinforcing mechanisms which are more likely to matter when considering elections separated by several years, such as increased notoriety of the higher-ranked candidates. Third, we observe the full set of candidates eligible to compete in the runoff, whether or not they actually stay in the race. This enables us to cleanly estimate and characterize the contribution of party coordination to the effects of rankings. Fourth, we can isolate the contribution of the bandwagon effect to the impact of rankings by focusing on second rounds in which only two candidates qualified and there is thus no need or even room for strategic coordination.

The positive effects we find on vote shares and winning are in line with the predictions of models assuming that voters gain utility from voting for the winner (Hinich, 1981; Callander, 2007). The bandwagon effect of candidate rankings is akin to the effects measured in other

contexts of asset rankings on trading behavior (Hartzmark, 2015), hospitals' rankings on their number of patients and revenues (Pope, 2009), and employees' rankings on their sales performance (Barankay, 2018).

Methodology. To isolate the effect of rankings from the effect of past vote shares, we use a regression discontinuity design (RDD) and compare the likelihood of running, the likelihood of winning, and the vote share obtained by candidates who received close-to-identical numbers of votes in the first round, but ranked just below or just above one another. Our analysis is conducted at the candidate level. We use two observations per district, corresponding to the higher-ranked and lower-ranked candidates and run the RDD on the vote share difference between the two candidates. The identification assumption is that the distribution of candidate characteristics changes continuously around the threshold, so that the only discrete change occurring at this threshold is the shift in candidates' first-round ranking.

Chapter 3 - Better Alone? Evidence on the Costs of Intermunicipal Cooperation

Over the last century, central governments in developed countries have sought to consolidate their local jurisdictions in order to achieve economies of scale in the provision of public goods. In particular, they tend to set reforms to encourage intermunicipal cooperation. However, municipalities are often reluctant to cooperate, slowing down or even blocking the consolidation process. On the one hand, cooperating municipalities benefit from efficiency gains and economies of scale. On the other, the loss of autonomy over public policies can be costly, most notably when integration implies cooperating with other municipalities that have different preferences (Tiebout, 1956; Alesina and Spolaore, 1997; Alesina et al., 2004).

This chapter studies the factors explaining municipalities' resistance. In particular, it seeks to assess whether resistance can be explained by actual local costs of integration, beyond ideological or political considerations.

Exploiting a 2010 reform in France that forced non-integrated municipalities to join an intermunicipal community (IC), I assess the causal impact of integration on municipalities that were reluctant to cooperate. Doing so, I provide new evidence that resistance is driven by local costs of integration: urban municipalities resisted integration to prevent further housing

construction, while rural municipalities resisted to avoid losing local public services. Using a difference-in-differences strategy, I first find that municipalities forced to enter a community experienced a large increase in the number of building permits delivered per year. Consistent with NIMBYsm (Not In My BackYard) explaining urban municipalities' resistance, this effect is driven by high-demand and densely built municipalities, where a rise in construction is the most likely to create congestion. Second, exploiting new data on daycare facilities and public libraries, I show that rural municipalities forced to enter a community experienced a decrease in the number of facilities available in their territory. Finally, in looking at public transport and fiscal revenues, I find that resisting municipalities benefited from a greater access to public transport and higher revenues after integration in the same proportion as municipalities that voluntarily integrated. This suggests that they did not oppose integration because they anticipated getting a lower share of the benefits, but rather that these benefits were not high enough to compensate for the costs of integration.

Contribution. This paper relates to several strands of literature. The first analyzes patterns of integration to study the factors explaining jurisdictions' choice to cooperate. Using structural or spatial econometric models, papers have emphasized the role of expected change in fiscal revenues, expected distance to public services, and municipalities' socio-demographic characteristics, as well as the role of neighbors' characteristics and thus heterogeneity (Gordon and Knight, 2009; Saarimaa and Tukiainen, 2014; Weese, 2015; Bel and Warner, 2016; Di Porto and Paty, 2018). In contrast, this paper provides new evidence on the factors explaining resistance by directly assessing the causal impact of integration on resisting municipalities.

Closer to my empirical strategy, a second strand of literature seeks to identify the causal impact of merger laws on cost reduction in the newly formed jurisdictions. The results are mixed (Bel and Warner, 2015): Although Reingewertz (2012) finds evidence of overall efficiency gains in Israel, Blom-Hansen et al. (2016) find no effect of recent merger laws in Denmark. The vast majority of these papers focus on the aggregate effects of consolidation. Two recent papers use geocoded data to study the distributional impact of mergers on local public sector jobs (Harjunen et al., 2019) and night-light intensity (Egger et al., 2017). In contrast, this paper measures the impact of forced collaboration (not mergers) on resisting municipalities and provides evidence on both the local costs and benefits of integration by looking at the effects on the housing market, local public services, public transport, and fiscal revenues.

Third, I contribute to the literature on housing restrictions. Recent empirical papers show

that regulations in high-demand places are responsible for lowering productivity, preventing households from moving to opportunity and fostering urban sprawl (Glaeser and Gyourko, 2018; Hsieh and Moretti, 2019; Glaeser and Maré, 2001; Chetty et al., 2016; Glaeser and Kahn, 2010; Jones and Kammen, 2013). Despite these consequences, housing regulations are unlikely to lessen if decisions are made by residents at the local level, as most residents consider a new construction project nothing but a bother (Glaeser, 2014 and see Gyourko and Molloy, 2015 for a review of the literature on NIMBYsm). In line with this argument, results show that municipalities are reluctant to lose control over urban planning. They also suggest that transferring urban planning to a higher level, and thus allowing outsiders and potential future residents to participate in the decision making, helps overcome local housing restrictions. These results stress the tension between local preferences and aggregate effects and highlight the risk of the local trap, where decisions made at the local scale ignore macro consequences (Purcell, 2006; Hankinson, 2018).

Methodology. This chapter uses a difference-in-differences strategy in order to assess the impact of integration on municipalities that resisted integration and were forced to enter an intermunicipal community. More precisely, I compare, before and after the 2010 law, municipalities that were not part of an intermunicipal community in 2010 and thus forced to integrate (treatment group) with municipalities that were already part of a community in 2010 and thus not affected by the law (control group). The identification assumption is that absent the law, municipalities would have evolved the same way in the control and in the treatment groups.

Introduction Générale

L'Économie Publique étudie le rôle du secteur public dans l'économie. Cette littérature analyse l'impact économique des politiques mises en place par les gouvernements et étudie la façon dont les politiques publiques devraient être conçues pour atteindre leurs objectifs. L'Économie Politique étudie le rôle des institutions politiques. Cette littérature analyse la manière dont les contraintes politiques façonnent les décisions des agents politiques (électeurs, élus) et dans quelle mesure les politiques publiques sont influencées par leurs préférences et intérêts.

Cette thèse se situe à l'intersection de ces deux domaines et vise à mieux comprendre les déterminants et les conséquences du comportement des citoyens et des élus. Les deux premiers chapitres étudient les déterminants du comportement de vote des électeurs et leurs conséquences sur les résultats électoraux; le troisième chapitre étudie le comportement des élus et les conséquences de leurs politiques publiques sur l'économie.

La suite de l'introduction est structurée de la façon suivante: je décris les fondations théoriques qui motivent l'étude des comportements de vote, puis les fondations théoriques qui motivent l'étude du comportement des élus. La dernière partie de l'introduction donne un aperçu de chaque chapitre et détaille la contribution de cette thèse à la littérature.

Comportement des Électeurs

Une vaste littérature théorique s'emploie à modéliser et prédire les choix de vote des citoyens.

Les modèles canoniques de participation électorale présentent la décision de voter comme un calcul coûts-bénéfices. Dans le cadre d'élections à deux candidats, Downs (1957a) postule que voter entraîne un coût pour les électeurs (incluant par exemple les coûts associés à l'inscription sur les listes et le coût de se rendre aux urnes), et génère d'autre part des bénéfices dits instrumentaux. Ces bénéfices dépendent de la probabilité que l'électeur.ice soit pivot - la probabilité que son vote modifie le résultat de l'élection - ainsi que de l'utilité différentielle associée à la victoire de son.a candidat.e préféré.e. Étant donnée la très faible probabilité d'être pivot dans la plupart des élections, il suffit que le coût associé au vote soit légèrement positif pour que l'utilité nette du vote soit négative. Ce modèle, bien que très intuitif, ne permet donc pas d'expliquer les taux de participation élevés que nous observons. Pour résoudre ce problème, les modèles de théorie des jeux rendent la probabilité d'être pivot endogène en la modélisant

comme le résultat d'interactions stratégiques entre agents (Palfrey and Rosenthal, 1983; 1985). Ces modèles génèrent des équilibres avec un taux de participation élevé, mais uniquement sous l'hypothèse que les citoyens disposent d'informations parfaites sur les préférences de vote des autres.

Des modèles plus récents considèrent que la participation électorale s'appuie sur des motivations éthiques plutôt que sur la probabilité d'être pivot. Les modèles appelés "*groupe rule utilitarian*" présentent l'électorat comme étant divisé en deux groupes, chacun soutenant un.e candidat.e, et supposent que les électeurs veulent "faire leur part" pour contribuer à la victoire de leur groupe (Coate and Conlin, 2004; Feddersen and Sandroni, 2006). Comme dans les modèles précédents, l'utilité du vote se caractérise par l'arbitrage entre les coûts de vote et les bénéfices générés par la victoire du ou de la candidat.e préféré.e. La différence entre ces modèles réside dans le fait que les agents ne maximisent pas leur utilité individuelle mais se soucient de l'utilité agrégée du groupe. En conséquence, ils votent selon la règle qui, si elle était suivie par tous les membres de leur groupe, maximiserait l'utilité globale de leur groupe. Au sein de chaque groupe, la règle de vote est définie par un seuil tel que les électeurs dont les coûts de vote sont inférieurs à ce seuil votent pour leur candidat.e préféré.e, et les autres s'abstiennent. Ces modèles génèrent des prédictions fidèles à la réalité: ils prédisent un taux de participation élevé dans un univers d'information imparfaite, ainsi que des statistiques comparatives réalistes, comme le fait que la participation a tendance à être plus élevée lors d'élections serrées.

Les modèles de participation électorale s'intéressent exclusivement aux élections à deux candidats et par conséquent n'abordent pas la question des choix de vote lors d'élections à plus de deux candidats. Lorsque plus de deux candidats sont en lice, les citoyens qui soutiennent les candidats les moins bien placés peuvent être confrontés à un choix difficile: voter pour leur.e favori.e même si ses chances de victoire sont faibles ou voter pour un.e autre candidat.e qui a plus de chances de gagner. Duverger (1954), pionnier de la réflexion sur les comportements stratégiques, postule que les électeurs ne veulent pas "gaspiller" leur voix en l'attribuant à un.e candidat.e qui a une faible chance de gagner, et qu'ils votent donc uniquement pour les deux premiers candidats. Myerson and Weber (1993) formalisent cette intuition dans le cadre des modèles classiques de théorie des jeux. Ils montrent que, lorsque les électeurs sont motivés par des bénéfices instrumentaux, une élection avec plusieurs candidats se résume généralement à une compétition entre deux candidats, et parmi eux, le.a candidat.e préféré.e par la majorité

remporte l'élection (voir aussi Palfrey, 1989; Fey, 1997; Cox, 1997). La même prédiction émerge dans le cadre des modèles de groupe: dans la mesure où l'utilité des groupes n'est affectée que par l'identité du ou de la vainqueur.e de l'élection, les deux premiers candidats reçoivent l'essentiel des voix dans la majorité des élections (Bouton and Ogden, 2017). Dans l'ensemble, les modèles de vote suggèrent que les modèles électoraux peuvent se restreindre aux élections à deux candidats sans que cela ne réduise la portée de leurs analyses et prédictions.

Néanmoins, une question se pose : les électeurs considèrent-ils vraiment que leur vote est "gaspillé" s'il n'est pas attribué à un.e candidat.e ayant une chance de gagner l'élection? On pourrait envisager que les électeurs attachent de l'importance à exprimer pleinement leurs préférences, et que voter pour leur.e candidat.e préféré.e génère de l'utilité, indépendamment de la probabilité qu'il ou elle gagne l'élection. Si cette intuition est correcte, à l'équilibre, les voix devraient se répartir entre davantage de candidats, comme c'est le cas lors de nombreuses élections. En effet, alors que la loi de Duverger est parfaitement illustrée par la division du paysage politique américain entre Républicains et Démocrates, dans d'autres pays, plus de deux candidats se partagent régulièrement les voix (lors des élections générales au Royaume-Uni par exemple).

Faire l'hypothèse que les électeurs votent expressivement ou stratégiquement change entièrement les prédictions sur le nombre de candidats recevant des votes à l'équilibre. Cette hypothèse est également cruciale lorsqu'il s'agit d'évaluer la capacité d'une règle électorale à agréger efficacement les préférences individuelles. En particulier, sous la règle majoritaire, il est parfois nécessaire qu'une fraction importante de l'électorat vote stratégiquement afin d'élire un.e dirigeant.e qui correspond aux préférences de la majorité des citoyens. À l'inverse, si une proportion trop importante d'électeurs vote expressivement, la règle majoritaire peut élire un.e candidat.e qui n'est pas le.a vainqueur.e de Condorcet, s'il en existe un.e, diminuant ainsi la représentativité de l'élu.e (Nurmi, 1983; Myerson and Weber, 1993).

Néanmoins, il existe encore peu d'études ayant mené une analyse systématique des véritables moteurs du vote et du rôle joué par les comportements de vote dans l'agrégation des préférences. Les deux premiers chapitres de cette thèse s'y emploient. Le premier chapitre évalue dans quelle mesure les électeurs votent de manière expressive ou stratégique et les conséquences sur les résultats électoraux. Le deuxième chapitre étudie le type d'information susceptible d'affecter les comportements électoraux, et évalue l'impact du classement des candidats aux élections passées sur les résultats électoraux.

Comportement des Élus

Les politiques publiques ne dépendent pas seulement de la façon dont les préférences sont agrégées lors des élections, elles dépendent aussi de la manière dont les élus se comportent lorsqu'ils sont au pouvoir.

Tout d'abord, dans la pratique, de nombreuses politiques publiques sont décidées par des groupes d'individus, des assemblées législatives ou des conseils de politique monétaire par exemple. Dans ce cadre, les politiques choisies sont le résultat de négociations entre les membres du groupe. Baron and Ferejohn (1989) étudient le processus de décision d'un comité de législateurs parfaitement informés qui doivent choisir comment se répartir les ressources disponibles. Ils montrent que, à équilibre, une coalition se forme et les ressources sont réparties de manière inégale, la personne menant les négociations parvenant à s'octroyer une part plus élevée. Lorsqu'elles sont choisies à plusieurs, les politiques publiques mises en place dépendent de plusieurs facteurs: le degré d'impatience des législateurs, la règle de vote au sein du comité, ainsi que les caractéristiques du statu quo et en particulier s'il est supposé endogène ou exogène (Baron, 1996; Bowen et al., 2014).

De plus, les décisions des élus peuvent être influencées par les enjeux électoraux. Si les élections sont considérées comme un moyen de discipliner les élus (Barro, 1973; Ferejohn, 1986), elles peuvent également être à l'origine de comportements opportunistes. En faisant l'hypothèse que les élus sont égoïstes et les électeurs courttermistes et partiellement informés, Nordhaus (1975) montre que les élus sont incités à manipuler les trajectoires économiques à leur profit. En particulier, ils auront tendance à stimuler l'économie juste avant les élections afin de réduire le chômage, tout en négligeant les effets néfastes de l'inflation qui ne se matérialiseront qu'après l'élection. De tels cycles politico-économiques peuvent se produire même en supposant les électeurs rationnels, à condition que l'élection se joue entre deux partis avec des préférences politiques différentes (Alesina, 1987; Hibbs Jr, 1994). Dans ce cas, ce sont les forces partisans qui nourrissent le cycle et celui-ci est engendré par l'alternance des deux partis au pouvoir.

Dans la mesure où les politiques publiques peuvent être influencées à la fois par les négociations entre élus et les enjeux électoraux, l'échelle à laquelle la décision est prise joue un rôle décisif. Comme le montrent Battaglini and Coate (2008), les politiques publiques mises en place varient considérablement selon que la décision est prise par un.e décideur.e qui cherche à maximiser l'intérêt général ou par un comité dans lequel chaque législateur cherche à max-

imiser les ressources allouées à sa propre circonscription. Dans ce dernier cas, les politiques choisies sont le résultat de négociations entre les législateurs et elles sont impactées par leur volonté d'être réélus au niveau local. A l'équilibre, cela crée des taux d'imposition et des niveaux d'endettement trop élevés, et une provision insuffisante de biens publics. De même, Rodden (2002; 2006) souligne les risques du fédéralisme fiscal: si le pouvoir est donné aux États plutôt qu'au gouvernement central, les élus risquent de dépenser et d'emprunter de façon excessive, en essayant de faire peser le fardeau sur le gouvernement central et sur les autres États.

L'Économie Politique comme l'Économie Publique analyse depuis longtemps le niveau optimal auquel les décisions politiques devraient être prises. Cette question a été abordée très tôt dans les travaux de Tocqueville (1835), Tiebout (1956) et Oates (1972), et plus récemment formalisée par Alesina and Spolaore (1997; 2003). Leur modèle souligne que la taille des pays est le résultat d'un compromis coûts-bénéfices. D'un côté, les grands pays peuvent réaliser des économies d'échelle, réduisant ainsi le coût des biens publics par habitant; de l'autre, les grands pays font face à une plus grande hétérogénéité des préférences sur leur territoire et davantage d'individus ou de régions se trouvent en désaccord avec les politiques centrales. Alesina and Spolaore (1997; 2003) montrent également que lorsque les agents politiques décident de la taille des territoires en maximisant leur propre utilité, ils créent un nombre de territoires trop élevé en comparaison du nombre de territoires qu'un.e décideur.e créerait afin de maximiser l'utilité globale. Cela est dû au fait que les agents politiques n'internalisent pas les avantages, partagés par tous, que les grands territoires procurent en terme de réduction des coûts des biens publics. Par conséquent, les agents politiques qui préfèrent des politiques publiques différentes de celles mises en place par leur pays, décident de se séparer et de créer leur propre territoire.

L'urbanisme est un exemple de politique publique qui implique des tensions entre bénéfices globaux et préférences locales. Comme le montre l'économie urbaine, une hausse des constructions de logements pourrait être bénéfique, en particulier dans les zones à forte productivité: cela permettrait de faire baisser le coût des logements, de favoriser la mobilité des travailleurs et d'éviter l'étalement urbain (Glaeser and Gyourko, 2018; Hsieh and Moretti, 2019; Glaeser and Maré, 2001; Chetty et al., 2016; Glaeser and Kahn, 2010; Jones and Kammen, 2013). Cependant, il y a de fortes chances que les habitants de ces zones s'opposent à de nouvelles constructions, afin de préserver leur qualité de vie et éviter que des étrangers s'installent dans leur arrière-cour (voir Gyourko and Molloy, 2015 pour une revue de la littérature sur le NIM-

BYsm “*Not In My BackYard*”). Dans ce contexte, ignorant les bénéfices globaux, les élus locaux sont incités à garder le contrôle des politiques de logements et d’urbanisme et à s’opposer à leur centralisation.

Néanmoins, dans la pratique, nous avons encore une connaissance limitée des politiques publiques que les élus sont particulièrement réticents à partager. Les gouvernements centraux de nombreux pays développés cherchent à encourager la coopération intercommunale mais se voient confrontés à l’opposition de maires qui refusent de coopérer. Davantage d’analyses empiriques sont nécessaires pour en comprendre la raison: les maires refusent-ils de se regrouper parce qu’ils veulent protéger les intérêts locaux de leur commune? Ou sont-ils seulement motivés par leur intérêt personnel et veulent éviter de perdre leur pouvoir politique? Le troisième chapitre de cette thèse étudie cette question. Ce chapitre évalue l’impact de l’intégration intercommunale sur les communes qui ont été forcées de coopérer, afin de mieux comprendre les raisons pour lesquelles les maires étaient réticents à partager leurs politiques publiques. Une attention particulière est accordée aux politiques d’urbanisme afin de mettre en lumière le rôle éventuel du NIMBYsm dans la résistance des communes à l’intégration intercommunale.

Problématique et Plan de Thèse

Cette thèse s’appuie sur l’ensemble des apports théoriques discutés ci-dessus et présente trois études empiriques du comportement des agents politiques. A travers l’utilisation de méthodes quasi-expérimentales et de données administratives, l’objectif de cette thèse est d’établir des relations de causalité qui permettent d’éclairer les causes et les conséquences du comportement des électeurs et des élus. Dans un premier temps, et contrairement à l’hypothèse faite dans les modèles de vote, nous montrerons que les électeurs ne se soucient pas seulement de qui remportera l’élection lorsqu’ils choisissent pour quel.le candidat.e voter. Nous montrerons qu’ils attachent aussi de l’importance à voter de manière expressive pour leur.e candidat.e préféré.e, et que cela a un impact considérable sur la manière dont la règle majoritaire agrège les préférences. Dans un second temps, nous explorerons le type d’information susceptible d’affecter le comportement des électeurs et des candidats, et montrerons que les résultats électoraux passés affectent leurs choix présents, en facilitant la coordination des électeurs et des candidats mais aussi en générant des biais comportementaux. Enfin, nous étudierons le comportement des élus et montrerons que l’impact de l’intégration intercommunale sur les

communes permet d'expliquer la réticence des maires à partager leurs politiques publiques.

Chapitre 1 - Le Vote Expressif et ses Conséquences: Analyse des Seconds Tours de Scrutins Comptant Deux ou Trois Candidats (avec Vincent Pons)

Lorsque plus de deux candidats se présentent à une élection désignant un.e seul.e vainqueur.e, les citoyens qui soutiennent les candidats les moins bien placés font face à un choix difficile: voter pour leur.e candidat.e favori.e ou voter pour un.e autre candidat.e dont les chances de victoire sont plus grandes. En exprimant leur préférence réelle, les électeurs d'un même camp risquent de diviser leur soutien entre plusieurs candidats et de contribuer à la victoire d'un.e candidat.e plus éloigné.e de leurs préférences. Le résultat des élections dépend donc de la fraction des électeurs qui votent de façon expressive (uniquement en fonction de leurs préférences entre les différents candidats) ou de façon stratégique (en fonction du résultat attendu de l'élection).

Les études existantes sur le vote stratégique comparent les préférences des électeurs avec leurs choix de vote, et comptent le nombre d'électeurs qui votent pour l'un des principaux candidats plutôt que pour le.a candidat.e dont ils sont le plus proches. Les préférences réelles des électeurs sont cependant difficiles à observer : ainsi, ces études sont fondées sur des sondages dont la fiabilité peut être limitée (Blais et al., 2001; Hillygus, 2007), ou sur des hypothèses concernant la relation entre préférences et choix de vote (Kawai and Watanabe, 2013; Spenkuch, 2015). À l'inverse, ce chapitre repose exclusivement sur l'analyse de résultats électoraux officiels. Nous exploitons le seuil qui détermine la qualification des candidats pour le second tour des élections législatives et cantonales françaises, et comparons les résultats électoraux selon que deux ou trois candidats concourent au second tour.

Nous montrons que la présence d'un.e troisième candidat.e augmente la participation électorale et diminue le nombre de voix obtenues par les deux premiers candidats. Surtout, dans environ 1/5ème des élections, la présence d'un.e troisième candidat.e provoque la perte du ou de la candidat.e qui, parmi les deux premiers, était le plus proche idéologiquement. Nous montrons ainsi qu'une large fraction des électeurs préfère voter expressivement plutôt que stratégiquement et que, dans ce contexte, la règle majoritaire conduit souvent à des résultats sous-optimaux. Les candidats arrivés en troisième position pourraient éviter ces résultats en

se retirant de la course entre les premiers et seconds tours, mais très peu le font lorsque leur orientation politique est différente de celles des deux premiers candidats.

Contribution à la littérature. Une vaste littérature empirique cherche à mesurer l'étendue du vote stratégique. Afin de calculer la proportion d'électeurs votant pour un.e candidat.e autre que son.a favori.e, les études existantes comparent les préférences des électeurs avec leur choix de vote. Les études reposant sur les résultats de sondages estiment qu'une part faible des électeurs votent stratégiquement, généralement inférieure à 20% (e.g., Alvarez and Nagler, 2000; Blais et al., 2001; Hillygus, 2007; Kiewiet, 2013). Néanmoins, ces résultats peuvent être biaisés par des réponses erronées. Par exemple, afin de se situer du côté du ou de la gagnant.e, les électeurs peuvent déclarer avoir voté pour le.a vainqueur.e même lorsque ce n'est pas le cas (Wright, 1993; Atkeson, 1999), ce qui conduirait à une surestimation du comportement stratégique. Inversement, pour éviter d'être en situation de dissonance cognitive (Festinger, 1962), les électeurs peuvent faire en sorte que leur préférence déclarée corresponde à leur choix de vote, ce qui conduirait cette fois à sous-estimer le comportement stratégique.

Un second volet de la littérature exploite les résultats électoraux agrégés et mesure l'étendue du vote stratégique en imposant des hypothèses sur la relation entre les préférences des électeurs et les choix de vote. Kawai and Watanabe (2013) et Myatt and Fisher (2002) calibrent des modèles structurels afin, pour les premiers, d'estimer le nombre d'électeurs qui n'ont pas voté pour leur.e candidat.e préféré.e et, pour les seconds, d'estimer l'impact du vote stratégique sur le nombre de sièges remportés par un parti. Spenkuch (2018) s'intéresse aux élections législatives allemandes lors desquelles l'électeur.e dispose de deux voix: la première permet de voter pour un.e candidat.e de sa circonscription; la seconde permet de voter pour une liste de candidats au niveau de son État. Spenkuch (2018) compare les votes attribués aux listes sous la règle proportionnelle avec les votes attribués aux candidats sous la règle majoritaire, et infère qu'environ un tiers des électeurs se comporte de manière stratégique (voir également Spenkuch, 2015).

Plutôt que d'estimer les préférences des électeurs et de les comparer avec leurs choix de vote, ce chapitre repose exclusivement sur l'analyse de résultats électoraux officiels et compare les résultats électoraux selon que deux ou trois candidats se présentent. Notre méthode d'analyse nous permet d'apporter trois contributions majeures à la littérature. Premièrement, nous estimons l'impact de la présence d'un.e troisième candidat.e sur la participation et sur le nombre de voix reçues par les deux premiers candidats. Ce faisant, nous démontrons que

le.a troisième candidat.e obtient des votes à la fois d'électeurs qui auraient voté pour l'un des deux premiers s'il.elle n'avait pas été présent.e, ainsi que d'électeurs qui se seraient abstenus ou auraient voté blanc ou nul. Deuxièmement, nous pouvons estimer avec précision la fraction des élections dont le résultat final change en raison de ces comportements. Troisièmement, en plus d'étudier le comportement des électeurs, nous mettons en évidence les principaux facteurs qui influencent la décision des candidats de rester dans la course à l'issue du premier tour, ou au contraire de se désister.

Enfin, nous contribuons à la littérature théorique sur le vote stratégique en explorant dans quelle mesure nos résultats peuvent être expliqués par les modèles d'électeurs rationnels (par exemple, Myerson and Weber, 1993) et les modèles de groupe (Coate and Conlin, 2004; Feddersen and Sandroni, 2006; Bouton and Ogden, 2017), dans lesquels l'utilité d'un individu ou d'un groupe n'est affectée que par l'identité du ou de la vainqueur.e de l'élection.

Méthodologie. Nous exploitons la discontinuité générée par la règle de qualification des candidats pour le second tour des élections cantonales et législatives françaises. Lors de ces élections, sont qualifiés pour le second tour les deux premiers candidats ainsi que tout autre candidat obtenant un nombre de voix supérieur à 12,5% des inscrits. Nous utilisons la méthode de la régression sur discontinuités (*Regression Discontinuity Design*, RDD) et comparons les résultats dans les circonscriptions ou cantons dans lesquels le.a troisième candidat.e obtient un nombre de voix juste au-dessus du seuil de 12.5 et se qualifie donc pour le second tour, aux résultats de circonscriptions ou cantons dans lesquels le.a troisième candidat.e obtient un nombre de voix juste en-dessous du seuil et ne se qualifie donc pas pour le second tour. Cette méthode nous permet de comparer les résultats électoraux selon que deux ou trois candidats se présentent au second tour. Estimer l'impact de la présence d'un.e troisième candidat.e à travers cette méthode repose sur une hypothèse: que la distribution d'autres facteurs pouvant affecter les résultats soit continue autour du seuil de 12,5%, de sorte que le seul changement se produisant précisément à la discontinuité soit le fait que le.la troisième candidat.e se qualifie, ou non.

Chapitre 2 - Impact du Classement des Candidats sur la Coordination et Effets d'Entraînement: Analyse des Élections à Deux Tours (avec Vincent Pons)

Lors d'une élection, les électeurs et les candidats cherchent à prédire les préférences et les stratégies des autres afin de prendre leurs propres décisions. Ils peuvent utiliser les informations fournies par les sondages avant l'élection ou/et les informations tirées des élections précédentes. Si nous avons la preuve que les choix des agents politiques dépendent de leur niveau d'information (e.g., Hall and Snyder, 2015; Le Pennec and Pons, 2019), nous avons une compréhension encore limitée du type d'informations sur lesquelles ils s'appuient pour prendre leurs décisions.

Ce chapitre étudie une information spécifique, le classement des candidats: leur classement dans les sondages, leur classement aux élections précédentes ou au tour précédent lorsqu'il s'agit d'élections à deux tours. Tandis que le nombre de voix totalisées par chaque candidat dans les élections passées fournit des informations détaillées sur la distribution des préférences, les simples classements peuvent, en eux-mêmes, constituer des repères permettant aux agents de se coordonner. Lorsque plus de deux candidats sont en lice, les électeurs stratégiques peuvent utiliser le classement de ces candidats à des élections passées ou à des tours antérieurs comme un repère pour se coordonner sur le même sous-ensemble de candidats. Le classement antérieur peut aussi être utilisé par des partis politiques proches idéologiquement afin de déterminer quel.le candidat.e parmi les leurs devrait se désister. De plus, les effets du classement peuvent être renforcés par des biais comportementaux, comme par exemple celui de l'effet dit d'entraînement: les électeurs qui souhaitent voter pour le.a vainqueur.e pourraient décider de "prendre le train en marche" et se rallier aux candidats les mieux classés.

Dans ce chapitre, nous estimons l'impact du classement passé sur les décisions des électeurs et des candidats. Nous déterminons ensuite si cet impact s'explique par des comportements de coordination ou reflète plutôt des effets d'entraînement. Nous nous intéressons aux élections françaises à deux tours, dans lesquelles jusqu'à trois ou quatre candidats peuvent se qualifier au second tour, et nous étudions l'impact du classement des candidats au premier tour sur leurs performances au second tour.

Les résultats montrent qu'être classé.e marginalement premier.ère, deuxième ou troisième au premier tour augmente significativement la probabilité que le.la candidat.e se maintienne au

second tour et gagne l'élection. Nous constatons que les effets de classement sont beaucoup plus élevés lorsque les candidats qui sont juste au-dessus ou juste en-dessous l'un de l'autre ont la même orientation politique. Nous montrons également plusieurs statistiques comparatives qui suggèrent que les candidats, comme les électeurs, utilisent les classements pour se coordonner. Néanmoins, nous montrons que la coordination à elle seule ne suffit pas à expliquer les résultats obtenus: même lorsque deux candidats seulement se qualifient (et qu'il n'y a donc pas besoin de se coordonner), les classements du premier tour ont un impact fort sur le second tour, ce qui suggère que le désir de voter pour le vainqueur joue un rôle clé dans la décision des électeurs.

Contribution à la littérature. À la suite du papier de Lee (2008), une vaste littérature a examiné l'impact d'être classé premier (au lieu de deuxième) sur les élections futures et a montré que les candidats qui avaient gagné l'élection précédente de justesse bénéficiaient d'un avantage électoral lorsqu'ils se présentaient à nouveau (Eggers et al., 2015). Ces études ne distinguent pas l'effet d'être au pouvoir entre les deux élections de l'effet d'être étiqueté premier, effet que nous pouvons isoler en étudiant l'impact du classement entre les deux tours de la même élection. Complémentant les travaux sur l'effet d'être premier, Anagol and Fujiwara (2016) montrent que le fait d'arriver marginalement deuxième (au lieu de troisième) lors de l'élection précédente augmente également la probabilité de concourir à la prochaine élection et de la gagner. Ils interprètent ces effets comme résultant de la coordination entre électeurs. Ce chapitre apporte quatre contributions importantes à cette littérature.

Premièrement, nous pouvons mesurer l'effet d'arriver marginalement deuxième ou troisième, ainsi que l'effet d'arriver marginalement premier (indépendamment de l'avantage d'être au pouvoir). Deuxièmement, le court laps de temps entre le premier et le second tour (une semaine) nous permet d'isoler l'effet direct du classement sur les résultats, et de le séparer d'autres effets qui peuvent se matérialiser entre des élections distantes de plusieurs années, comme par exemple l'effet d'un accroissement de la notoriété des candidats les mieux classés. Troisièmement, nous observons l'ensemble des candidats qualifiés pour le second tour, qu'ils décident de se maintenir ou non. Cela nous permet d'étudier la coordination entre les partis politiques et d'évaluer son rôle dans l'effet observé des classements sur les résultats électoraux. Quatrièmement, nous pouvons isoler les effets d'entraînement créés par les classements en nous concentrant sur les seconds tours où deux candidats seulement sont qualifiés, et où la coordination n'a donc pas sa place.

Nos résultats sont en ligne avec les prédictions de modèles supposant que l'utilité des électeurs est plus grande lorsqu'ils votent pour le.a gagnant.e (Hinich, 1981; Callander, 2007). Il est aussi possible de faire le parallèle entre l'effet d'entraînement lié au classement des candidats et les effets de classement obtenus dans d'autres contextes: l'impact du classement des actifs sur les comportements boursiers (Hartzmark, 2015), l'impact du classement des hôpitaux sur leur nombre de patients (Pope, 2009), et celui du classement des employés sur leurs performances de ventes (Barankay, 2018).

Méthodologie. Afin d'isoler l'effet du classement de l'effet du nombre de voix, nous utilisons la méthode de la régression sur discontinuités (RDD). Nous comparons les résultats au second tour de candidats ayant reçu un nombre de voix quasi identique au premier tour, mais classés juste en-dessous ou juste au-dessus les uns des autres. Notre analyse est menée au niveau des candidats et utilise deux observations par élection, une pour le.la candidat.e classé.e juste en-dessous et l'autre pour le.la candidat.e classé.e juste au-dessus. La RDD s'appuie sur la différence entre le nombre de voix reçues par chaque candidat.e au premier tour, la discontinuité étant atteinte quand la différence est nulle. Nous faisons l'hypothèse que les caractéristiques des candidats sont distribuées de façon continue autour de la discontinuité, de sorte que le seul changement se produisant exactement à la discontinuité est le changement dans le classement des candidats au premier tour.

Chapitre 3 - Mieux vaut-il être Isolés ou Regroupés? Analyse des Conséquences de la Coopération Intercommunale

Au cours du siècle dernier, les gouvernements des pays développés ont cherché à consolider leurs territoires afin de réaliser des économies d'échelle et réduire ainsi le coût des biens publics. Ils ont, dans cette optique, mis en place de nombreuses réformes visant à encourager la coopération intercommunale. Néanmoins, les communes sont souvent réticentes à coopérer, ce qui peut ralentir ou bloquer le processus de consolidation. D'un côté, les communes qui coopèrent entre elles peuvent espérer réaliser des gains d'efficacité et bénéficier d'économies d'échelle; de l'autre, cette coopération entraîne une perte d'autonomie en matière de politique publique, perte qui peut s'avérer particulièrement coûteuse quand il s'agit de partager ses politiques publiques avec des communes qui ont des préférences différentes (Tiebout, 1956;

Alesina and Spolaore, 1997; Alesina et al., 2004).

Ce chapitre étudie les facteurs expliquant la résistance des communes. L'objectif est de déterminer si les communes sont réticentes à coopérer en raison des conséquences économiques locales provoquées par l'intégration, ou si leurs réticences sont uniquement liées à des raisons politiques ou idéologiques.

J'exploite la loi de 2010, qui en France, a forcé les communes non intégrées à rejoindre une intercommunalité. Cela me permet d'évaluer l'impact causal de la coopération intercommunale sur les communes réticentes à coopérer. Mes résultats montrent que les communes résistaient afin d'échapper aux conséquences économiques locales de l'intégration: les communes urbaines résistaient afin d'éviter la construction de nouveaux logements, tandis que les communes rurales résistaient afin d'éviter la perte de services publics locaux. Je constate premièrement que les communes contraintes d'entrer dans une intercommunalité ont connu une augmentation significative du nombre de permis de construire délivrés par an. Cet effet se manifeste principalement dans les communes densément construites et jouissant d'une forte demande immobilière: des communes où davantage de constructions accentuerait la congestion, et dans lesquelles les résidents ont donc de forte chance d'être sujets au syndrome NIMBY (*Not In My BackYard*). Deuxièmement, je montre, à travers l'exploitation de données sur les crèches et bibliothèques municipales, que les communes rurales ont subi une diminution du nombre de services publics de proximité. Enfin, mes résultats révèlent que les communes forcées de s'intégrer ont bénéficié de revenus fiscaux plus élevés et d'un meilleur accès aux transports publics, mais que ces bénéfices n'étaient pas suffisants pour compenser les effets négatifs de l'intégration.

Contribution à la littérature. La première littérature dans laquelle ce chapitre s'inscrit analyse les choix d'intégration des communes afin d'identifier les facteurs expliquant leur volonté de coopérer ou non. À l'aide de modèles économétriques structurels ou spatiaux, plusieurs études ont souligné que la décision de s'intégrer dépend des gains espérés en termes de revenus fiscaux, de la localisation future des services publics, ainsi que des caractéristiques socio-démographiques des communes et du degré d'hétérogénéité entre communes voisines (Gordon and Knight, 2009; Saarimaa and Tukiainen, 2014; Weese, 2015; Bel and Warner, 2016; Di Porto and Paty, 2018). Contrairement à ces études qui analysent le choix des communes en amont, ce chapitre évalue l'impact causal de la coopération intercommunale afin de mettre en lumière le rôle des conséquences directes de l'intégration dans la résistance des communes.

Plus proche de la méthode empirique utilisée dans ce chapitre, une deuxième littérature exploite des réformes gouvernementales afin d'identifier l'impact causal des fusions de commune sur le coût des services publics. Les résultats sont mitigés (Bel and Warner, 2015): alors que Reingewertz (2012) montre que les fusions en Israël ont entraîné des gains d'efficacité, Blom-Hansen et al. (2016) ne trouvent aucun effet des fusions récentes au Danemark. La grande majorité de ces études s'intéresse aux effets agrégés de l'intégration. Deux articles récents utilisent des données géocodées pour étudier les effets des fusions au sein de l'entité fusionnée afin d'identifier les gagnants et les perdants. Ils mesurent l'impact sur les emplois du secteur public local (Harjunen et al., 2019) et sur l'intensité de l'éclairage nocturne (Egger et al., 2017). À la différence de ces analyses, ce chapitre mesure l'impact de la coopération forcée (et non des fusions) sur les communes réticentes à s'intégrer et mesure les conséquences à la fois sur le marché du logement, les services publics locaux, les transports publics et les recettes fiscales.

Enfin, ce chapitre contribue à la littérature de l'économie urbaine. Plusieurs études empiriques récentes montrent l'impact négatif des limites imposées par les élus locaux sur les constructions de logement. Les restrictions de constructions dans les zones à forte demande réduisent la mobilité des travailleurs et donc leur productivité, empêchent les ménages de s'installer dans des endroits qui ouvrent davantage de perspectives, et favorisent l'étalement urbain (Glaeser and Gyourko, 2018; Hsieh and Moretti, 2019; Glaeser and Maré, 2001; Chetty et al., 2016; Glaeser and Kahn, 2010; Jones and Kammen, 2013). Malgré cela, les réglementations restrictives en matière de logement ont peu de chances d'évoluer si les décisions restent prises au niveau local, dans la mesure où la plupart des résidents considèrent toute nouvelle construction comme dérangeante (Glaeser, 2014) et refusent de voir des étrangers s'installer dans leur arrière-cour (voir Gyourko and Molloy, 2015 pour une revue de la littérature sur le NIMBYsm). Conformément à cet argument, les résultats de ce chapitre démontrent que les municipalités sont effectivement réticentes à l'idée de perdre le contrôle sur leurs politiques d'urbanisme. Les résultats suggèrent également que les restrictions locales peuvent être surmontées en transférant les décisions d'urbanisme et les politiques de logement au niveau intercommunal, dans la mesure où cela permet aux personnes extérieures à la commune et aux futurs résidents potentiels de participer à la prise de décision. Ces résultats soulignent la tension entre les préférences locales et les bénéfices globaux et mettent en évidence le risque dit de la "trappe locale", décrivant des situations où les décisions prises à l'échelle locale ignorent les conséquences globales (Purcell, 2006; Hankinson, 2018).

Méthodologie. Ce chapitre utilise la méthode des doubles différences afin d'évaluer l'impact de la coopération intercommunale sur les communes qui avaient refusé de coopérer et ont été forcées par la loi de 2010 d'entrer dans une intercommunalité. Plus précisément, je compare, avant et après la loi de 2010, les communes qui ne faisaient pas partie d'une intercommunalité en 2010 et qui ont donc été contraintes de s'intégrer (groupe de traitement) avec les communes qui faisaient déjà partie d'une intercommunalité en 2010 et n'ont donc pas été affectées par la loi (groupe de contrôle). Afin d'identifier l'impact de l'intégration sur les communes du groupe traitement, je fais l'hypothèse que, en l'absence de cette loi, les communes des groupes traitement et contrôles auraient évolué de la même manière après 2010.

Chapter 1

Expressive Voting and its Cost: Evidence from Runoffs with Two or Three Candidates¹ (with Vincent Pons)

Abstract

In French parliamentary and local elections, candidates ranked first and second in the first round automatically qualify for the second round, while a third candidate qualifies only when selected by more than 12.5 percent of registered citizens. Using a fuzzy RDD around this threshold, we find that the third candidate's presence substantially increases the share of registered citizens who vote for any candidate and reduces the vote share of the top two candidates. It disproportionately harms the candidate ideologically closest to the third and causes her defeat in one fifth of the races. Additional evidence suggests that these results are driven by voters who value voting expressively over voting strategically for the top-two candidate they dislike the least to ensure her victory; and by third candidates who, absent party-level agreements leading to their dropping out, value the benefits associated with competing in the second round more than influencing its outcome.

¹Published under the following reference: Pons, Vincent, and Clémence Tricaud. "Expressive voting and its cost: Evidence from runoffs with two or three candidates." *Econometrica* 86.5 (2018): 1621-1649. For suggestions that have improved this article, we are grateful to Daron Acemoglu, Abhijit Banerjee, Laurent Bouton, Pierre Boyer, Enrico Cantoni, Micael Castanheira, Bruno Crépon, Torun Dewan, Rafael Di Tella, Esther Duflo, Martial Foucault, Benjamin Marx, Nolan McCarty, Rebecca Morton, Benjamin Olken, Alessandro Riboni, Julio Rotemberg, Daniel Rubenson, Kenneth Shepsle, Daniel Smith, and James Snyder, as well as seminar participants at MIT, Harvard, the University of Chicago, Bocconi University, University of Maryland, Ecole Polytechnique, CREST, Sciences Po, and conference participants at NBER Political Economy Meeting, APSA, RIDGE/LACEA-PEG Workshop on Political Economy, Erasmus PE Workshop, D-TEA Workshop, Brussels PE Conference, Columbia PE Conference, PET Conference, and Toronto Political Behaviour Workshop. We thank Sebastian Calonico, Matias Cattaneo, Max Farrell, and Rocio Titiunik for guiding us through the use of their RDD Stata package "rdrobust" and for sharing their upgrades; Nicolas Sauger for his help with the collection of the 1981 and 1988 French parliamentary election results; Abel François for sharing his data on 1993, 1997, and 2002 candidates' campaign expenditures; Julia Cagé for sharing her panel dataset of local newspaper presence in France; Salomé Drouard and Eric Dubois for their excellent work of data entry and cleaning; Médiamétrie's senior officials – Philippe Tassi (deputy director general), David Bernier (radio project manager), and Jean-Pierre Panzani (TV project manager) – for sharing their data on radio and TV news audiences across France; and Anne-Marie Dussaix for introducing us to Médiamétrie.

1 Introduction

In an indirect democracy – the form of government prevalent across the West today – representatives rule on behalf of the people. In theory, their representativeness comes from being elected. In practice, it depends on the extent to which voters’ choices reflect their true preferences, and on the way in which the voting rule translates vote choices into election outcomes. These two conversions determine who is elected and which policies are enacted. Under plurality rule, when more than two candidates are running, citizens who support lower-ranked candidates face a difficult tradeoff: voting for their favorite, or for another candidate with higher chances of winning. By expressing their true preference, voters may split their support over multiple candidates and nominate less-preferred leaders. Hence, the result of the election depends on the extent to which voters are “expressive” – voting based on their preference among candidates only – or “strategic” (or “instrumentally rational”) – voting based on likely outcomes of the election.

In his groundbreaking work on strategic behavior, Duverger (1954) posits that voters do not want to waste their vote and will thus, in most cases, exclusively vote for two front-runners. Models by Palfrey (1989), Myerson and Weber (1993), Fey (1997), and Cox (1997) formalize this intuition: under plurality rule, when voters are instrumentally rational, an election with multiple candidates usually boils down to a two-candidate race, and of these two, the candidate who is preferred by the majority wins the election.¹

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

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

This paper uses a different approach: instead of estimating preferences, we focus on vote

¹This result applies to a number of settings. The model of Myerson and Weber (1993) applies to a wide range of single-winner electoral systems such as plurality rule, approval voting, and the Borda system. Cox (1994) extends the model to a multimember context. The case of dual ballot rule is studied in Bouton (2013) and Bouton and Gratton (2015).

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

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

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

The presence of the third candidate in the second round increases voter turnout by 4.0 percentage points and reduces the share of blank and null votes by 3.7 percentage points, resulting in an overall increase of the share of people casting a ballot for any of the candidates by 7.8 percentage points (14.2 percent). In addition, the presence of the third candidate decreases the vote share of the top two candidates (expressed as a fraction of registered citizens) by 6.9 percentage points (12.5 percent).

Our key results relate to the impact of the third candidate's presence on the outcome of the elections. We find that the vote shares of the top two candidates decrease in proportion to their proximity with the third on the left-right ideological axis. The impact remains equally strong and significant in subsets of elections where first-round results indicate that the third candidate is very unlikely to be a "front-runner" (to rank first or second) in the second round.

As a consequence, in 19.2 percent of the elections, the presence of the third candidate causes the loss of the candidate among the top two that is ideologically closest to her.² This candidate is the likely Condorcet winner of the second round: she would have won a two-candidate race against the other top-two candidate, absent the third, and she would most likely also win a two-candidate race against the third. Thus, the presence of the third candidate often results in an outcome that harms a majority of her supporters (those with a preference for the closest over the furthest top-two candidate) and a majority of voters.

Additional evidence suggests that our aggregate results are primarily driven by the behavior of two types of citizens: “loyal” supporters (whom we also designate as “loyals”), who abstain or vote blank or null when the third candidate is absent and vote for her when she is present, leading to increased participation; and “switchers”, who vote for one of the top two candidates when the third is absent but switch to the third when she is present, explaining the decreased vote share of the top two candidates and the impact on winner identity. We argue that these behaviors are difficult to rationalize within rational voter and group-rule utilitarian models, in particular when the third candidate appears to have slim chances of being a front-runner in the second round, and even when allowing for imperfect information, aggregate uncertainty, or dynamic strategic motives. Instead, our results suggest that voters’ choice of candidate, as well as their decision whether to vote or abstain, can only be satisfactorily explained by taking into account expressive benefits independent of the result of the election.

Because third candidates have the possibility to drop out of the race between the first and second rounds, our results also shed light on the motives underlying their behavior. We compare third candidates’ likelihood to drop out across different settings and gather additional descriptive evidence from press articles. When third candidates have the same political orientation as one of the top two candidates, party-level agreements internalize the cost that staying in the race would generate for the candidate of the sister party and result in an overwhelming majority of dropouts. Instead, when third candidates have a different orientation than both top two candidates, they tend to take the decision on their own, independently from any party’s instructions, and only drop out in rare circumstances. The fact that their decision to stay in the race decreases the chance of victory of the top-two candidate closest to them suggests that they often value the benefits associated with competing in the second round more than influencing its outcome.

The low likelihood of third-candidate dropouts when she has a different orientation than both top two results in the overrepresentation of this setting among complier districts, which likely contributes to voters’ difficulty to behave strategically in our sample. Overall, our results

²This estimate is obtained by restricting the analysis to elections where the three candidates are from different political orientations, where the third candidate is located either to the right or to the left of both top two candidates (since in these elections the candidate ideologically closest to the third is clearly identified), and where first-round results indicate that the third candidate’s chances of becoming a front-runner in the second round are low (so that the impact is not mechanically driven by the third candidate winning the election). More information in Sections 4.4 and 4.5.

suggest that plurality rule often produces suboptimal outcomes, due to the difficulty for parties of differing orientations to reach dropout agreements, and to a sufficiently large fraction of voters valuing expressive more than instrumental motives when they have to choose between more than two candidates of distinct orientations.

1.1 Contribution to the literature

A large literature studies how voting rules shape electoral outcomes and, in turn, how they affect voter behavior. Social choice theory has shown that no electoral system or voting rule is uniformly best under all criteria (Arrow, 1951),³ and that in any voting system, some voters have an incentive to misrepresent their true preferences in order to affect the outcome of the election (Gibbard, 1973; Satterthwaite, 1975). Building on this result, a normative literature has sought to identify voting rules that deliver outcomes best representing voters' preferences by most resisting strategic manipulation (e.g., Laslier, 2009; Balinski and Laraki, 2011). Under existing rules, a sufficiently large fraction of voters may need to engage in strategic manipulation in order to elect leaders who best correspond to citizens' preferences. If too many people vote according to their true preference instead of behaving strategically, then plurality rule may fail to choose the Condorcet winner, when one exists, thus decreasing the representativeness of the electoral outcome (Nurmi, 1983; Myerson and Weber, 1993).

A large empirical literature examines whether voters behave strategically or expressively. Small-scale laboratory experiments have provided direct evidence of the existence of strategic behaviors (e.g., Forsythe et al., 1993; Van der Straeten et al., 2010). Outside of the lab, Cox (1997) documents patterns consistent with strategic voting across electoral systems. Using RDD on population thresholds, Fujiwara (2011) and Eggers (2015) find that the top two candidates tend to get more votes under simple plurality than under runoff or proportional elections, in line with Duverger's prediction (but see Bordignon et al., 2016). Consistent with Myatt (2007)'s prediction that the amount of strategic voting increases with the level of information, Hall and Snyder (2015) find that higher levels of information in US primary elections decrease the number of votes and donations "wasted" on candidates unlikely to win. An important piece of information used by voters to coordinate comes from candidates' rankings in previous elections: second-place candidates are substantially more likely than close third-place candidates to run in, and win, a subsequent election (Anagol and Fujiwara, 2016).

To determine the actual proportion of voters voting for a candidate other than their preferred one, existing studies compare people's preferences and voting choices. Estimates based on surveys are typically low, below 20 percent (e.g., Alvarez and Nagler, 2000; Blais et al., 2001; Hillygus, 2007; Kiewiet, 2013), but potentially biased by misreporting. For instance, overreporting voting for the winner (Wright, 1993; Atkeson, 1999) may lead to overestimate

³Arrow's impossibility theorem states that when there are more than two alternatives, there is no social welfare function that satisfies the Pareto property and the Independence of Irrelevant Alternatives and which is not a dictatorship.

strategic behavior. Alternatively, to avoid cognitive dissonance (Festinger, 1962), people may adjust their stated preference to their voting choice, which would lead to underestimating it.

A second strand of the literature relies on aggregate electoral results and studies strategic voting by imposing assumptions on the mapping between voters' preferences and vote choices. Kawai and Watanabe (2013) and Myatt and Fisher (2002) calibrate structural models to estimate the number of voters who did not vote for their preferred candidate and the impact of strategic voting on the number of seats won by a party, respectively. In the context of the German split-ticket voting system, Spenkuch (2018) compares votes cast for party lists under a proportional rule with votes cast for individual candidates under plurality rule, and reports that about one third of voters behave strategically (also see Spenkuch, 2015).

Instead of estimating voter preferences and comparing them with their actual choices, this paper focuses on vote choices only. We compare electoral outcomes when two versus three candidates are competing.⁴ Methodologically, we draw on other studies that exploit vote-share thresholds to estimate the incumbency effect and other causal effects of interest (e.g., Lee, 2008; de la Cuesta and Imai, 2016). Our strategy allows us to make three important contributions to the literature. First, we estimate the impact of the presence of the third candidate on both participation and vote shares and demonstrate that the third candidate obtains votes both from voters who would have voted for the top two had she been absent and from supporters who would have abstained or voted blank or null. Second, we can precisely estimate the fraction of races whose final outcome changes as a result of these behaviors. Third, in addition to studying voters' behavior, we highlight the main factors affecting candidates' own decision to stay in the race or drop out.

We further contribute to the theoretical literature on strategic voting by exploring the extent to which our results can be explained by rational voter models (e.g., Myerson and Weber, 1993), and group-rule utilitarian models (Coate and Conlin, 2004; Feddersen and Sandroni, 2006; Bouton and Ogden, 2017), in which the individual or group's utility is only affected by who wins the election.

The remainder of the paper is organized as follows. We describe the data we use in Section 2 and our empirical framework in Section 3. Section 4 presents our empirical results. We interpret and discuss our results in Section 5. Section 6 concludes.

⁴Similarly to our setting, Blais et al. (2017) and Kiss (2015) study runoff elections in France and Hungary, respectively, in which more than two candidates can qualify. Differently from our strategy, they compare electoral outcomes in the first and second rounds, assuming that voters reveal their true preference in the first round. For a discussion of this assumption, see Piketty (2000) and Bouton and Gratton (2015).

2 Research setting

2.1 French parliamentary and local elections

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

Both parliamentary and local elections are held under a two-round plurality voting rule. In order to win directly in the first round, a candidate needs to obtain a number of votes greater than 50 percent of the candidate votes and 25 percent of the registered citizens. In the vast majority of districts, no candidate wins in the first round, and a second round takes place one week later. In the second round, the election is decided by simple plurality: the candidate who receives the largest vote share in the second round wins the election.

The two candidates who obtain the highest vote share in the first round automatically qualify for the second round. Other candidates qualify only if they obtain a first-round vote share higher than 12.5 percent of the registered citizens. This threshold does not have any other implication. All candidates qualified for the second round can decide to drop out of the race between rounds.

Our sample includes all parliamentary and local elections using the 12.5 percent qualification threshold: the eight parliamentary elections which took place since 1978 as well as the 2011 and 2015 local elections.⁵ The fact that this threshold is at a relatively high percentage means at most three candidates qualify for the second round in all but a handful of districts, which is ideal for our study design. In the elections we consider, the third candidate received more than 12.5 percent of votes in 1,215 districts (16.7 percent of our sample).⁶

⁵Each of the 10 elections we consider took place at a different date. Moreover, the local and parliamentary elections we study were never held at the same date as other types of elections such as presidential, mayoral, or regional ones.

⁶In local elections, the required vote share was 10 percent of the registered citizens until 2010, when the threshold was increased to 12.5 percent. The lower threshold resulted in more than three candidates qualifying in a large number of constituencies. One exception was made after the change: in the 2011 local elections, the threshold remained at 10 percent in the 9 cantons belonging to Mayotte département (0.6 percent of the 2011 observations). The threshold in parliamentary elections was increased from 5 to 10 percent in 1966 and from 10 to 12.5 percent in 1976.

2.2 Data

Our sample includes a total of 7,257 observations: 3,458 (47.7 percent) from local elections and 3,799 (52.3 percent) from parliamentary elections. Official results of local and parliamentary elections were digitized from printed booklets for the 1978, 1981, and 1988 parliamentary elections and obtained from the French Ministry of the Interior for all others. We exclude districts where only one round took place or those with fewer than three candidates in the first round.⁷ Table A1 in the Appendix gives the breakdown of the sample data by election type and year.

Table 1 presents some descriptive statistics on our sample.

Table 1: Summary statistics

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

In the average district, 7.78 candidates competed in the first round, and turnout was 58.2 percent. On average, 56.2 percent of the registered citizens cast a valid vote for one of the candidates. Valid voting for a candidate entails inserting a ballot pre-printed with the candidate's name in an envelope and putting this envelope in the ballot box. We term these "candidate votes". The difference between turnout and candidate votes arises from voters who cast a blank vote (putting an empty envelope in the ballot box) or a null vote (writing something on the ballot or inserting multiple ballots in the envelope). Turnout in the second round was slightly higher than in the first (58.8 percent on average), but the fraction of candidate votes was slightly lower (55.4 percent on average) due to an increased share of blank and null votes. The average number of candidates in the second round was 2.04, and there were three candidates in the second round in 453 districts (6.2 percent of the sample).

We further use the political label attributed to the candidates by the French Ministry of the Interior to allocate them to six political orientations: far-left, left, center, right, far-right, and

⁷We also exclude three elections where the second and third candidates in the first round obtained exactly the same number of votes. Here, the 12.5 percent threshold rule did not apply. Both candidates were allowed to move on to the second round, regardless of the number of votes they had obtained in the first.

other. In the elections we consider, the candidate who ranked third in the first round was on the left in 37.1 percent of the districts, on the right in 19.6 percent, on the far-right in 36.7 percent, and from another political orientation in the remaining 6.6 percent of the districts.⁸

2.3 Vote share of the third candidate

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

Voters who vote for the third candidate when she is present may either vote for one of the top two candidates or instead abstain or vote blank or null when she is absent. Thus, using a regression discontinuity design framework, we estimate the impact of the presence of the third candidate both on voter participation and on the top two candidates’ vote share. We then test whether her presence ultimately affects who wins the election.

⁸The Ministry of the Interior attributes political labels based on several indicators: candidates’ self-reported political affiliation, party endorsement, past candidacies, public declarations, local press, etc. We mapped political labels into the six political orientations, mainly based on the allocation chosen by Laurent de Boissieu in his blog “France Politique”: <http://www.france-politique.fr/>. We also used public declarations made by the candidates. Appendix I shows the mapping between labels and political orientations for each election.

3 Empirical strategy

3.1 Evaluation framework

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

Qualified third candidates can drop out of the race between the two rounds, making our regression discontinuity design fuzzy. Formally, we define the running variable X as the qualifying margin of the third candidate in the first round (the difference between her vote share, expressed as a fraction of the number of registered citizens, and the 12.5 percent threshold), the assignment variable D as a dummy equal to 1 if the third candidate qualifies for the second round ($X \geq 0$) and 0 otherwise ($X < 0$), and the treatment variable T as a dummy equal to 1 if the third candidate is present in the second round and 0 otherwise. We call compliers the districts in which the third candidate qualifies ($D = 1$) and runs in the second round ($T = 1$). We evaluate the impact of the presence of the third candidate in complier districts with the following specification:

$$Y_i = \alpha_1 + \tau T_i + \beta_1 X_i + \beta_2 X_i T_i + \mu_i \quad (1)$$

where Y_i is the outcome of interest in district i and T_i is instrumented with D_i as shown in the following first-stage equation:

$$T_i = \alpha_0 + \gamma D_i + \delta_1 X_i + \delta_2 X_i D_i + \varepsilon_i \quad (2)$$

Following Imbens and Lemieux (2008) and Calonico et al. (2014), our main specification uses a non-parametric approach, which amounts to fitting two linear regressions on districts respectively close to the left, and close to the right of the threshold. We test the robustness of our results to a quadratic specification, including X_i^2 and its interaction with T_i as regressors in equation [1] and X_i^2 and its interaction with D_i in equation [2]. Our estimation procedure follows Calonico et al. (2014), which provides robust confidence interval estimators. Our preferred specification uses the MSERD bandwidths developed by Calonico et al. (2018), which reduce potential bias the most. We also test the robustness of the main results to using the optimal bandwidths computed according to Imbens and Kalyanaraman (2012).⁹ The bandwidths used for the estimations are data-driven and therefore vary depending on the outcomes we consider. Instead, when we provide descriptive statistics on districts “close to the threshold”, we consider districts in which the vote share of the third candidate was within exactly 2 percentage points from the threshold. Thanks to our large sample size (7,257 districts), the number of districts close to the threshold is higher than 1,800.

⁹Figure A5 in the Appendix shows the robustness of our main results to bandwidth choice, both for linear and quadratic specifications.

On average, districts close to the threshold are of similar size than in the full sample, as shown in Table A2 in the Appendix. Close to the threshold, the average district is characterized by slightly higher first round turnout and third candidate share of candidate votes, and slightly lower number of candidates and total vote share of the top two candidates. But differences with the full sample remain relatively modest overall.

3.2 Identification

The 2SLS estimates obtained from equations [1] and [2] can be interpreted as a local average treatment effect conditional on the assumptions of the LATE theorem being satisfied (Imbens and Angrist, 1994).

First, *independence of the instrument* comes from the discrete rule of qualification. The identification assumption is that the distribution of potential confounders changes continuously around the 12.5 percent vote share threshold, so that the only discrete change occurring at this threshold is the shift in assignment status. Sorting of candidates across the threshold only threatens the validity of this assumption if it occurs at the cutoff, with potential losers pushed just above the threshold or potential winners pushed just below (de la Cuesta and Imai, 2016). Generally, this is unlikely, as it requires the ability to predict election outcomes and deploy campaign resources with extreme accuracy, and given that weather conditions on Election Day and other unpredictable events make the outcome of the election uncertain (Eggers et al., 2015).

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

To bring empirical support for the identification assumption, we check if there is a jump in the density of the running variable at the threshold (McCrary, 2008). As Figure A1 in the Appendix shows, we do not observe any. Figures B1 and B2 and Table B1 in Appendix B further show the lack of any significant jump at the cutoff for first round outcomes and for the assignment status predicted by these baseline variables, bringing additional support for the identification assumption.

Second, the *first stage* is strong. Figure 1 plots the treatment against the running variable, and Table 2 provides the formal estimates for the first stage. Columns (1) and (2) show the results obtained under the MSRED and IK optimal bandwidths, using a local linear regression. Columns (3) and (4) present the results using a quadratic specification. All four estimates are significant at the 1 percent level. In our preferred specification (column 1), we find that

the probability that the third candidate stays in the race jumps from 0 to approximately 55.2 percent at the threshold.

The third candidate tends to drop out of the race when she has the same political orientation as one of the top two candidates (in 91.1 percent of the cases, at the threshold); when her orientation differs from both the top two, she instead tends to run in the second round (in 85.2 percent of the cases). As a result, the latter case accounts for the vast majority of complier districts.¹⁰

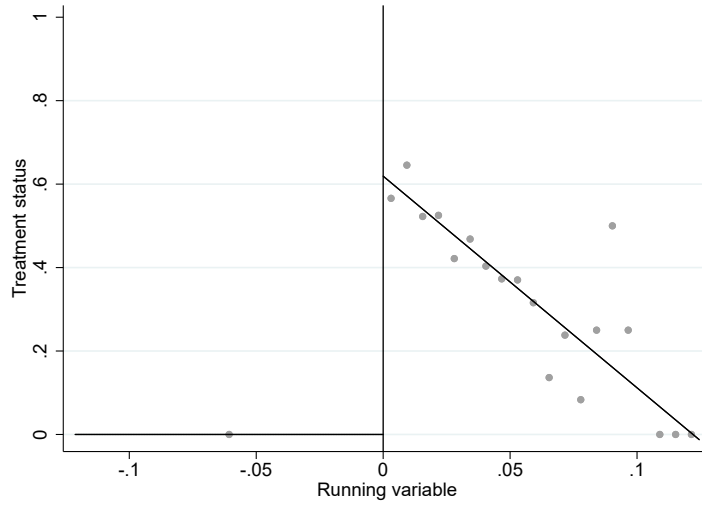
Third, *monotonicity* is fulfilled as the qualification rule did not generate any defier: no candidate who received a vote share lower than the qualification threshold was allowed to run in the second round.

Finally, the *exclusion restriction* requires that the qualification of the third candidate only affects second-round outcomes through the third candidate's presence in that round. We cannot entirely rule out the possibility that the third candidate's decision to drop out (which is only observed conditional on her qualification) disappoints some voters, or that it is interpreted as a signal on other candidates' characteristics, thus affecting voter behavior in the second round through other channels. Such violations of the exclusion restriction, however, are unlikely to drive the bulk of our results. First, the qualification of the third candidate has very small, and non-statistically significant, effects in districts where she has the same political orientation as one of the top two candidates and almost always drops out (Table A3 in the Appendix). Second, additional evidence on third candidates' dropouts and on their press coverage (which we discuss at greater length in Section 5.2) suggests that dropout decisions are unlikely to carry much private information and that they rarely generate adverse voter reaction.

Under these four conditions, the second-stage estimate can be interpreted as the causal impact of the treatment on complier districts, at the threshold. Estimating the impact of the presence of the third candidate on electoral outcomes amounts to comparing voters' behaviors in elections with two versus three candidates modulo two noticeable exceptions. First, in 5.5 percent of the elections near the discontinuity, the candidate ranking second in the first round dropped out of the race. Second, in 1.2 percent of the elections, the candidate ranking fourth in the first round also qualified for the second round, and she decided to run in three instances. Appendix C discusses these particular cases at greater length and shows that our results are not driven by them.

¹⁰While our analysis primarily seeks to estimate the impact of the third candidate's presence in the second round, the second part of our discussion (Section 5.2) examines at greater length the factors affecting the decision of the third candidate whether to drop out or stay in the race, thereby providing additional evidence on the characteristics of complier districts.

Figure 1: First stage



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

Table 2: First stage

Outcome	(1)	(2)	(3)	(4)
	Treatment status			
Assignment status	0.552*** (0.042)	0.611*** (0.030)	0.509*** (0.051)	0.566*** (0.043)
Robust p-value	0.000	0.000	0.000	0.000
Observations	1,541	3,579	2,142	3,579
Polynomial order	1	1	2	2
Bandwidth	0.017	0.038	0.023	0.038
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.00	0.00	0.00	0.00

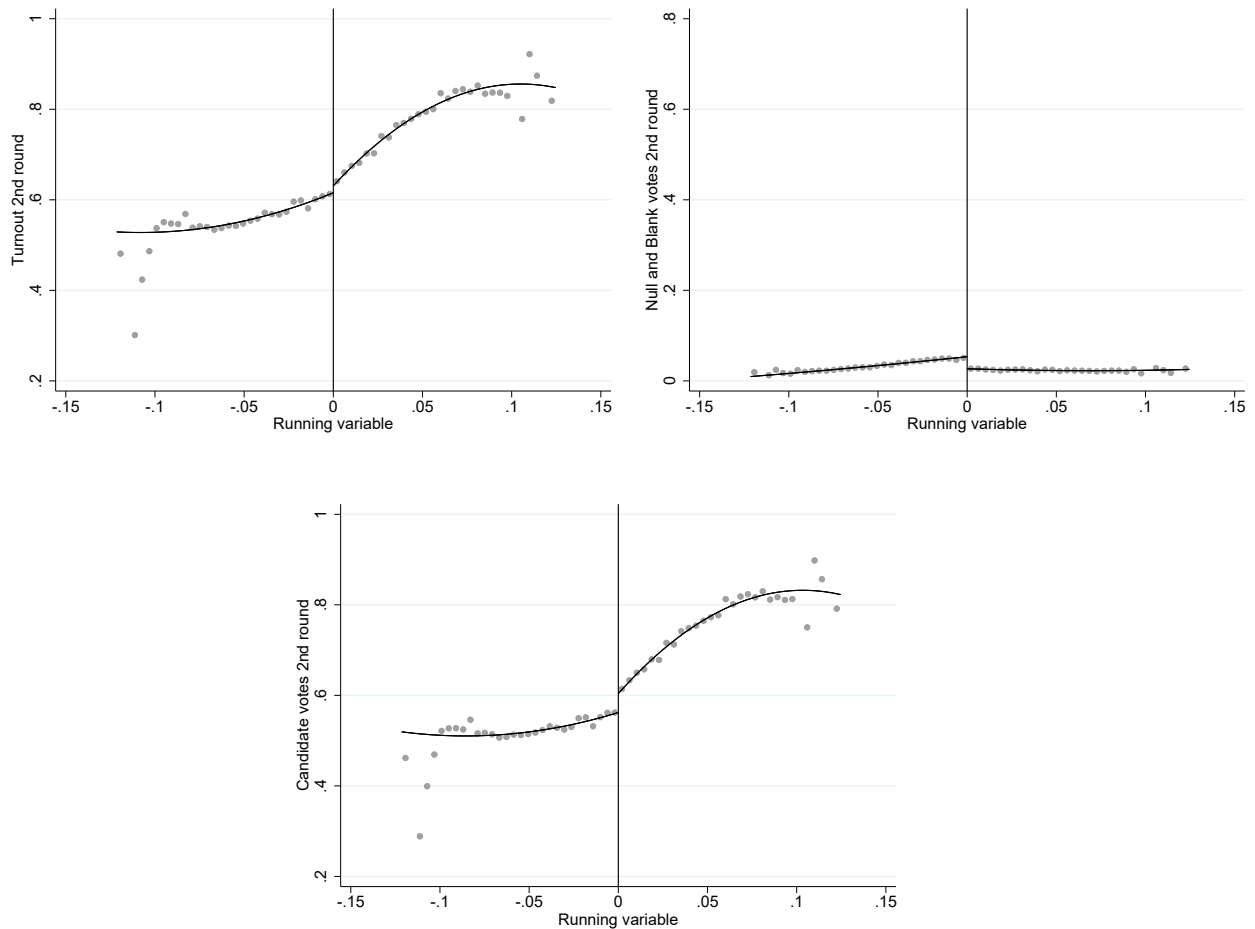
Notes: Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. The outcome is a dummy equal to 1 if the third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

4 Empirical results

4.1 Impact on participation and candidate votes

We consider three outcomes related to participation: turnout, the share of null and blank votes, and the share of candidate votes, all defined using the number of registered voters as the denominator. We begin with a graphical analysis before presenting formal estimates of treatment effects.

Figure 2: Impact on participation and candidate votes



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

In the graphs in Figure 2, each dot represents the average value of the outcome within a given bin of the running variable. To facilitate visualization, a quadratic polynomial is fitted on each side of the 12.5 percent threshold. We observe a clear discontinuity at the cutoff for each outcome: the presence of the third candidate has a large and positive impact on the share of registered citizens who vote and on the share of citizens who vote for one of the competing

candidates rather than casting a blank or a null vote.

Table 3 provides the formal estimates of the impacts using our preferred specification. On average, the presence of the third candidate in the second round increases turnout by 4.0 percentage points (6.7 percent), reduces the share of null and blank votes by 3.7 percentage points (78.7 percent),¹¹ and increases the share of candidate votes by 7.8 percentage points (14.2 percent). All effects are significant at the 5 percent level and the last two at the 1 percent level. As should be expected, the impact on the share of candidate votes corresponds to the sum of the absolute value of the impacts on the turnout rate and the share of blank and null votes.

Table 3: Impact on participation and candidate votes

Outcome	(1) Turnout	(2) Null and Blank votes 2nd round	(3) Candidate votes
3rd present	0.040** (0.017)	-0.037*** (0.004)	0.078*** (0.019)
Robust p-value	0.041	0.000	0.001
Observations	2,298	2,630	2,374
Polynomial order	1	1	1
Bandwidth	0.025	0.028	0.026
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.598	0.047	0.548

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

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

¹¹In the 2015 local elections, the only ones in which blank and null votes were counted separately, the impact on both outcomes was negative (Figure A2 and Table A4 in the Appendix).

Table 4: Impact on candidate votes

Outcome	(1)	(2)	(3)	(4)
	Candidate votes in the 2nd round			
3rd present	0.078*** (0.019)	0.075*** (0.016)	0.080*** (0.026)	0.080*** (0.024)
Robust p-value	0.001	0.000	0.005	0.007
Observations	2,374	3,454	3,039	3,454
Polynomial order	1	1	2	2
Bandwidth	0.026	0.037	0.033	0.037
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.548	0.541	0.544	0.541

Notes: The polynomial order is 1 in columns 1 and 2 and 2 in columns 3 and 4. The bandwidths are derived under the MSERD (columns 1 and 3) and IK (columns 2 and 4) procedures. Other notes as in Table 3.

4.2 Impact on votes going to the top two candidates

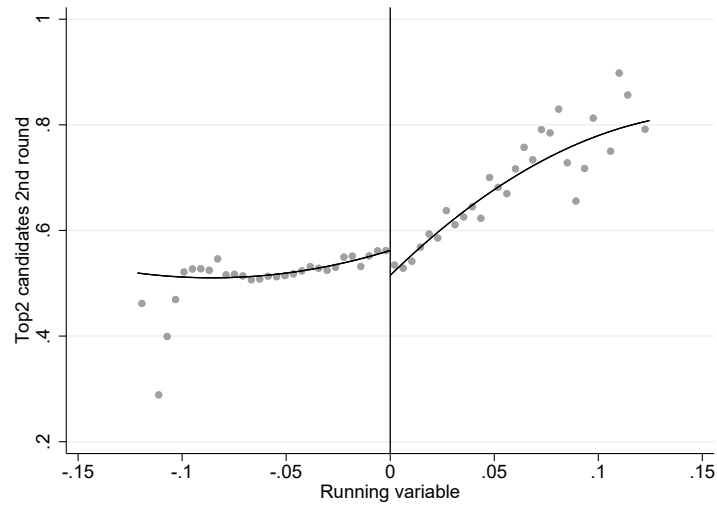
We now estimate the effect of the presence of the third candidate on the second round vote share of the candidates who placed first and second in the first round. If vote shares were defined using the number of candidate votes as the denominator, the total vote share of the top two candidates in the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

Figure 3 plots the vote share of the top two candidates against the running variable. The quadratic polynomial fit indicates a large downward jump at the cutoff.

Consistent with the graphical analysis, the estimates reported in Table 5 indicate a sizable and significant negative impact of the treatment on the vote share of the top two candidates in the second round. Our preferred specification (column 1) shows that the presence of the third candidate decreases the vote share of the top two candidates by 6.9 percentage points (12.5 percent) on average, an effect significant at the 1 percent level. The effect has a similar magnitude and remains statistically significant in other specifications.

We further estimate the impact on the vote share of the first and second candidates separately, and find that both decrease by a similar magnitude on average when the third candidate is present (Figure A3 and Table A5 in the Appendix).

Figure 3: Impact on votes going to the top two candidates



Notes as in Figure 2.

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

Outcome	(1)	(2)	(3)	(4)
	Vote share top 2 in the 2nd round			
3rd present	-0.069*** (0.020)	-0.072*** (0.015)	-0.061* (0.028)	-0.068* (0.023)
Robust p-value	0.003	0.002	0.074	0.085
Observations	2,250	3,704	2,726	3,704
Polynomial order	1	1	2	2
Bandwidth	0.024	0.040	0.030	0.040
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.551	0.541	0.546	0.541

Notes as in Table 4.

4.3 Impact on the top two candidates depending on political orientation

We now assess whether the presence of the third candidate systematically affects the outcome of the race. We first show that the vote shares of the top two candidates decrease in proportion to their ideological proximity with the third. We focus on elections in which the political orientations of the three candidates are well-identified and differ from one another, so that they can be ranked on the left-right axis. The resulting sample accounts for 75.7 percent of the districts where the third candidate qualifies and runs in the second round, near the discontinuity. We call A the candidate most to the left, C the candidate most to the right, and B the candidate located between A and C. We study three different settings, each characterized by the ideological position of the third candidate – on the left of the two others, right, or in the middle.¹²

As shown in Figure 4.1, when the third candidate is C, candidate B (who is closer to her) loses the most votes. The regression results are reported in Table A6 in the Appendix: while the vote share of candidate B decreases by 5.2 percentage points on average, the vote share of candidate A decreases by only 2.5 percentage points. The effect on an outcome defined as the difference between the vote shares of candidates A and B is equal to 2.4 percentage points but not significant.

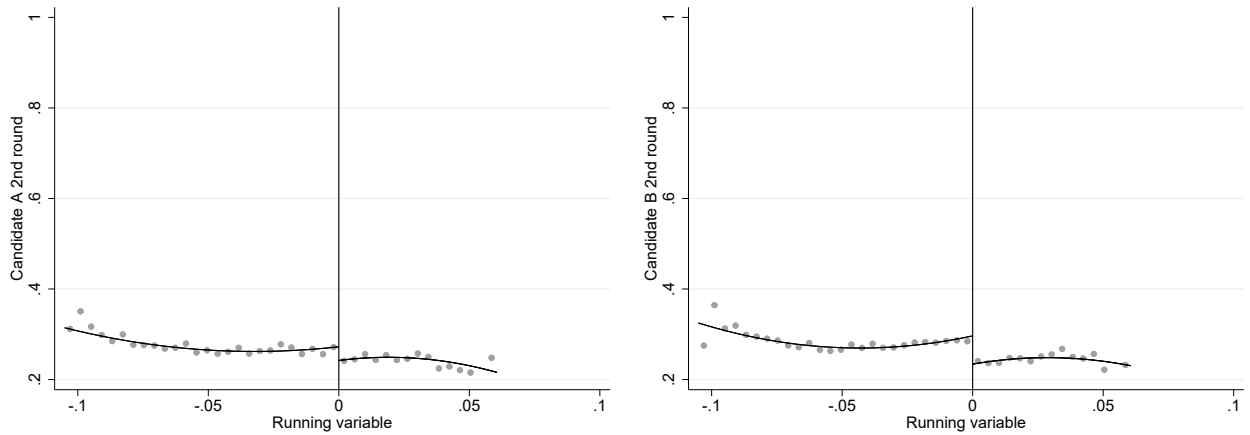
Symmetrically, when the third candidate is A, candidate B again loses the most from her presence (Figure 4.2). As shown in Table A7 in the Appendix, while the vote share of candidate B decreases by 8.6 percentage points on average, the vote share of candidate C is not significantly affected by the presence of candidate A. The effect on the difference between the vote shares of B and C is significant, and at the 1 percent level.

Finally, when the third candidate is located in the middle, both A and C lose a large number of votes (Figure 4.3). As shown in Table A8 in the Appendix, the presence of B decreases the vote share of candidate A by 7.1 percentage points on average and the vote share of candidate C by 5.6 percentage points. The effect on the difference between the vote shares of A and C is small and not statistically significant.

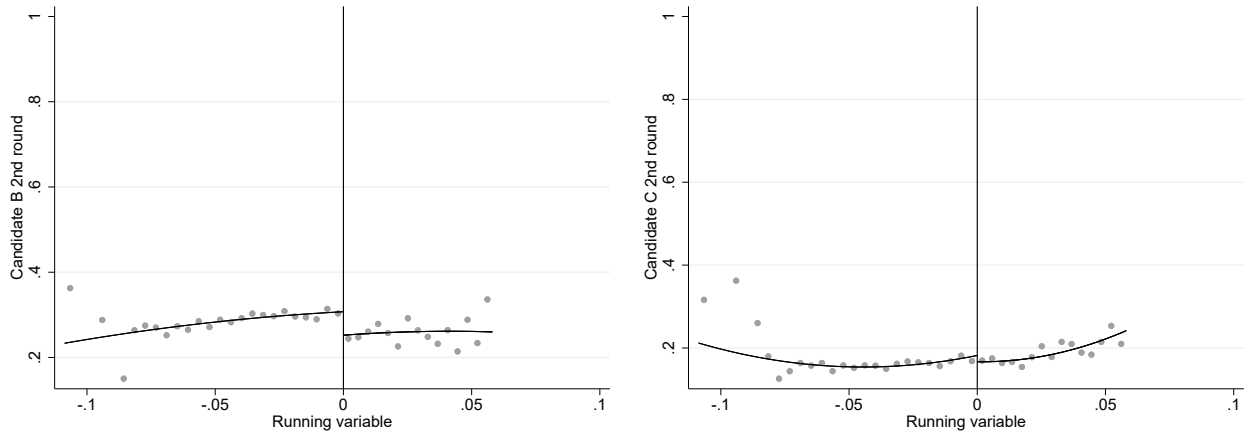
¹²In 97.3 percent of the elections corresponding to the first setting, the third candidate is on the far-right (C), one of the top two candidates is on the right (B), and the other is on the left (A). In 94.6 percent of the elections corresponding to the second setting, the third candidate is on the left (A), one of the top two candidates is on the right (B), and the other is on the far-right (C). In 62.7 percent of the elections corresponding to the third setting, the third candidate is on the right (B), one of the top two candidates is on the left (A), and the other is on the far-right (C). In 36.4 percent, the third candidate is on the center (B), one of the top two candidates is on the left (A), and the other is on the right (C).

Figure 4: Impact on the top two candidates depending on political orientation

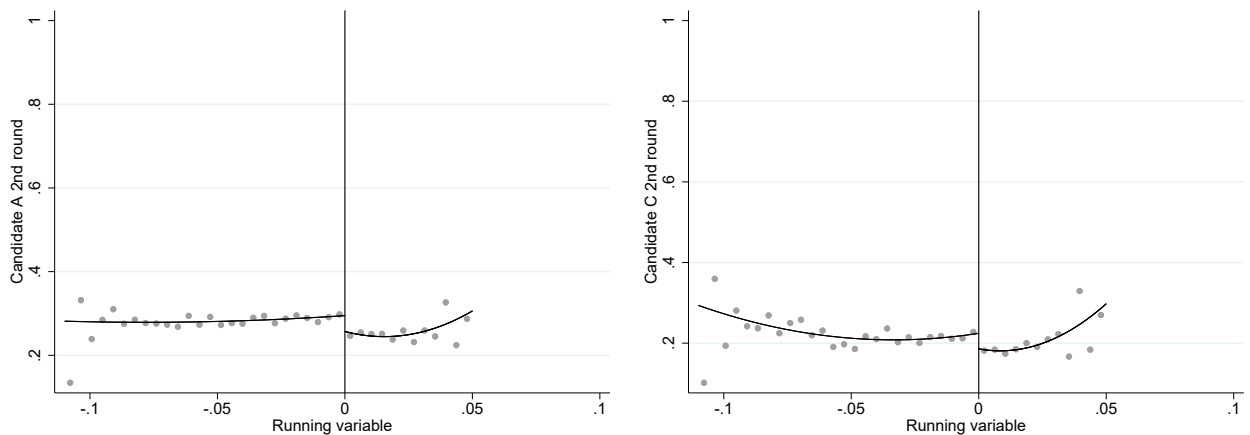
4.1. Impact on candidates A and B when the third candidate is C (1st setting)



4.2 Impact on candidates B and C when the third candidate is A (2nd setting)



4.3 Impact on candidates A and C when the third candidate is B (3rd setting)



Notes: Figure 4 includes the elections where the top three candidates have distinct political orientations. Figure 4.1 (resp. 4.2) includes only the elections where the third candidate is located to the right (resp. left) of both the first and the second candidates. Figure 4.3 includes only the elections where the third candidate is located between the first and the second candidates. Other notes as in Figure 2.

4.4 Impact on the top two candidates depending on the strength of the third

The rest of this section focuses on the first and second settings, where we can identify which candidate is ideologically closest to the third candidate and suffers the most from her presence. We now test whether the presence of the third candidate reduces the number of votes cast for the top two candidates even when she has low foreseeable chances of being a front-runner in the second round.

We estimate and compare the impact of the presence of the third candidate on the vote shares of the top two candidates in a series of subsamples. In each new subsample, we impose additional restrictions which, arguably, make it less plausible that the third candidate could pose a challenge on the top two.

The first subsample combines the first and second settings as defined in Section 4.3, without imposing any further restrictions: it includes all elections in which the three top candidates have distinct political orientations, and the third candidate is either on the left or the right of both top two candidates. In this sample, although the third candidate does rank behind the top two candidates in the second round in the vast majority of cases, she nonetheless wins in 11 elections near the threshold and ranks second in 28 (Table 6, column 3, sample 1).

To define the next subsamples, we consider the total voter support that each of the top three candidates may expect to receive in the second round, based on the votes obtained by other candidates of the same political orientations in the first round. A candidate from the left, for instance, may expect to receive votes not only from her supporters but also from supporters of other left-wing candidates who did not qualify for the second round. We thus define her “strength” as the sum of first-round vote shares of all candidates belonging to the left.

We restrict the second sample to observations from the first sample in which the third candidate’s strength is lower than that of each of the top two candidates. For example, if the third candidate is on the left, the second candidate is on the far-right and the first candidate is on the right, we consider only elections where the left candidates gathered fewer votes in total in the first round than those on the right or far-right. This restriction makes it arguably less likely that the third candidate could be a front-runner in the second round – and indeed, such a candidate never wins and ranks second in only 3 cases near the discontinuity in this sample (Table 6, column 3, sample 2).

Candidates’ strengths computed based on first-round results only provide imperfect information on the level of support that candidates can hope to receive in the second round, not least because not everyone votes sincerely in the first round. Thus, the third and fourth samples further impose a difference of at least five and ten percentage points, respectively, between the strength of the third candidate and the strength of each of the top two. In samples 3 and 4, the average gap between the strength of the third candidate and of each of the top two candidates is as large as 11.4 and 13.5 percentage points in the first round, close to the discontinuity. Hence,

these additional restrictions make it even less plausible that the third candidate had reasonable chances to be in contention for victory – and indeed, such a candidate never ranked first or second in those two subsamples (Table 6, column 3, samples 3 and 4).

As shown in Table 6, the impact of the presence of the third candidate on the vote share of the top two is robust across the four samples and close in magnitude: all estimates are significant at the 1 percent level and included between 8.0 and 13.2 percentage points (column 1).

Table 6: Impact on the vote share of the top two candidates depending on the strength of the third

	(1)	(2)	(3)	(4)	(5)
Impact 3rd present	Top 2 Candidates	<i>Bandwidth / Observations</i>	3rd becomes 1st/ 3rd becomes 2nd	Closest Candidate	Furthest Candidate
Sample 1	-0.082*** (0.017)	0.014 546	11 28	-0.061*** (0.008)	-0.027** (0.013)
Sample 2	-0.080*** (0.015)	0.019 609	0 3	-0.054*** (0.009)	-0.026** (0.010)
Sample 3	-0.088*** (0.020)	0.014 331	0 0	-0.050*** (0.010)	-0.037*** (0.013)
Sample 4	-0.132*** (0.029)	0.008 107	0 0	-0.060*** (0.016)	-0.074*** (0.018)

Notes: Column 2 gives the bandwidths used for the estimation of the impact on the vote share of the top two candidates as well as the number of observations lying in those bandwidths. Column 3 displays the number of cases where the third candidate ranks first or second in the second round in the elections included in the bandwidths defined in column 2. Other notes as in Table 3.

4.5 Impact on winner identity

We turn to the last and perhaps most important outcome, the winner of the election, and test whether the presence of the third candidate decreases the likelihood that the candidate ideologically closest to her wins the election. Formally, we use a dummy equal to 1 if the closest candidate wins (and 0 if she loses) as the outcome. We run the analysis for the elections in sample 2, in which the candidate ideologically closest to the third is clearly identified, and the third candidate never wins (as shown in Table 6), ensuring that the results are not artificially driven by elections in which the outcome takes the value 0 at the right of the threshold due to the victory of the third candidate.

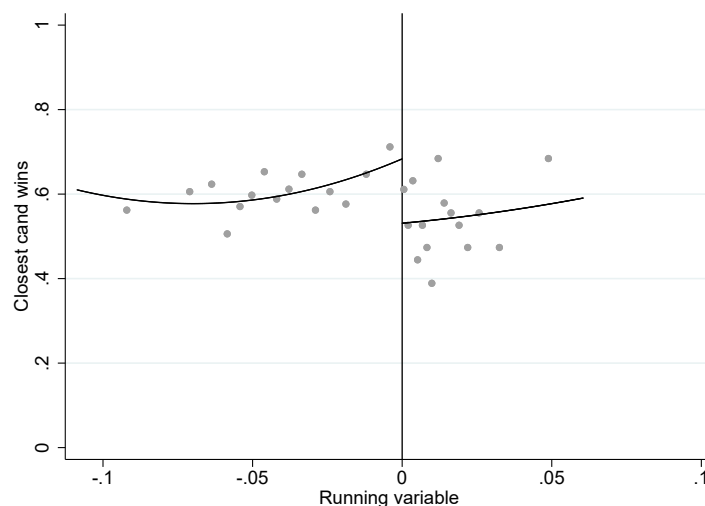
As shown in Figure 5, the presence of the third candidate has a negative impact on the probability that the closest candidate wins, thus harming the third candidate’s supporters who prefer the closest over the furthest candidate. The jump is large but noisier than in the other graphs, due to the fact that the outcome we consider is not a continuous variable but a dummy. Table 7 provides the formal estimates. In our preferred specification (column 1), we find a negative effect of 19.2 percentage points on average. This estimate is significant at the 10

percent level and robust to using the IK optimal bandwidth (column 2). This result means that, in around one-fifth of the elections we consider, the candidate among the top two that is ideologically closest to the third loses as a result of her presence whereas she would have won absent the third, in a two-candidate race against the other top-two candidate.

Given our results in Section 4.3, the candidate ideologically closest to the third candidate would most likely also win a two-candidate race against the third, making her the Condorcet winner of the second round. Indeed, in elections of the first setting, where the candidate who arrived third in the first round is C, B should be expected to win a race against C, as she obtained more votes in the first round and would attract relatively more voters of candidate A (based on estimates from Table A7 in the Appendix). Similarly, in elections of the second setting, in which the candidate who arrived third in the first round is A, B obtained more votes in the first round than A and she would attract relatively more voters of candidate C, in a race against A (based on estimates of Table A6 in the Appendix). In sum, in around one-fifth of the elections we consider, the presence of the third candidate causes the defeat of the likely Condorcet winner, harming the third candidate’s supporters and a majority of voters.

We next explore what these results tell us about the factors affecting voters’ participation and voting choice as well as candidates’ decision to compete or drop out.

Figure 5: Impact on the probability that the closest candidate wins



Note: We use elections of sample 2 as defined in Section 4.4. Dots represent the local averages of the probability that the candidate ideologically closest to the third wins in the second round. Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

Table 7: Impact on the probability that the closest candidate wins

Outcome	(1)	(2)	(3)	(4)
	Closest candidate wins			
3rd present	-0.192*	-0.178**	-0.101	-0.205
	(0.089)	(0.069)	(0.140)	(0.094)
Robust p-value	0.083	0.032	0.607	0.194
Observations	686	1,567	553	1,567
Polynomial order	1	1	2	2
Bandwidth	0.021	0.043	0.017	0.043
Band. method	MSERD	IK	MSERD	IK
Mean, left of the threshold	0.656	0.622	0.678	0.622

Notes: We use elections of sample 2 as defined in Section 4.4. The outcome is a dummy variable equal to 1 if the candidate ideologically closest to the third wins in the second round. Other notes as in Table 4.

5 Interpretation of the results and discussion

5.1 Voters' behavior

The effects of the third candidate's presence on electoral outcomes may in theory be driven by the response of both voters and other candidates. In our setting, however, there is only one week between the two rounds, leaving little time for the top two candidates to adjust their strategies to the presence of an additional competitor. While we lack data on candidates' precise political platforms, we collected data on their campaign expenditures for the 2011 and 2015 local elections and for the 1993, 1997, 2002, 2007, and 2012 parliamentary elections (collectively accounting for 77.8 percent of our sample).¹³ As shown in Appendix D, we find no impact of the presence of the third candidate on the top two candidates' campaign expenditures or contributions, suggesting that they do not intensify their political campaign when the third candidate is present and that our results are primarily driven by voters' response to changes in the electoral offer. Accordingly, we now discuss what our aggregate effects can teach us about individual voters' behavior and motivations.

5.1.1 Loyals and switchers

When the third candidate is absent, voters may choose to either vote for one of the top two or refrain through abstaining or voting blank or null. When the third candidate is present, they have a third possible option: voting for her. The corresponding two-by-three matrix (Figure 6) defines six different types of citizens. Our effects on participation and on the vote share

¹³All data come from the French National Commission on Campaign Accounts and Political Financing (CNC-CFP). Data on campaign expenditures for the 1993, 1997, and 2002 parliamentary elections were collected and digitized by Abel François and his co-authors (see Fauvelle-Aymar and François, 2005; Foucault and François, 2005).

of the top two candidates are driven by the four types of citizens (2, 3, 5, and 6) who behave differently depending on whether the third candidate is present or absent.

Figure 6: Matrix of citizens' behavior

3rd absent / 3rd present	Abstain or blank or null	Vote for one top-two candidate	Vote for the third candidate
Abstain or blank or null	Type 1	Type 3	Type 5
Vote for one top-two cand.	Type 2	Type 4	Type 6

The large increase in the number of candidate votes reported in Section 4.1 first suggests that the third candidate has a large number of “loyal” supporters who vote for her when she is present but abstain or vote blank or null otherwise (type 5 citizens). However, this effect might also result from the heightened mobilization of the top two candidates’ supporters, resulting in a larger number of type 3 citizens (who do not cast a candidate vote when the third is absent and vote for one of the top two when she is present) than type 2 citizens (who follow the opposite trajectory). Indeed, the presence of the third candidate in the second round reduces the winning margin by 5.8 percentage points on average (Figure A4 and Table A9 in the Appendix), and more contested elections may drive additional supporters of the top two candidates to the polls.

To assess the relative importance of both mechanisms, we compare the effects on participation in settings where one or the other channel is likely to dominate. In elections in which the top two candidates have distinct political orientations, the presence of the third candidate substantially increases closeness (Table A10 in the Appendix, column 3) by disproportionately decreasing the vote share of the front-runner ideologically closest to her. Instead, when the top two candidates have the same political orientation, they are equally distant from the third, so that the third candidate’s presence does not significantly affect closeness (Table A10 in the Appendix, column 2) but adds more diversity to the existing political offer. We find much larger effects on the number of candidate votes in races of the latter type (Table A11 in the Appendix), suggesting that the effects on participation are mostly driven by “loyals” voting for the third candidate.

We now interpret the large decrease in the total vote share of the top two candidates reported in Section 4.2. This effect is first driven by the behavior of another type of third candidate’s voters: “switchers”, who vote for one of the top two candidates when she is absent but switch to the third when she is present (type 6 citizens). The loss of these voters could be compensated by the increased mobilization of the top two candidates’ supporters, measured by the difference between the number of type 3 and type 2 citizens, but again, our test above suggests that this mechanism only has modest importance on average. The fact that the top two candidates’ vote shares decrease in proportion to their ideological proximity with the third brings additional support for the interpretation that switchers are responsible for most of the effect. It also rules out the possibility that switchers are simply noisy voters, randomly splitting their votes among competing candidates (such noisy voting should affect the top two candidates equally). Instead,

this result suggests that the candidate among the top two that most switchers prefer, and vote for when the third candidate is absent, is the one closest to the third.

The results in Sections 4.4 and 4.5 further imply that the number of switchers remains very high even when the third candidate has low foreseeable chances of being a front-runner in the second round, and that their choice to vote for the third candidate when she is present often causes the defeat of the top-two candidate they prefer.

5.1.2 Theoretical implications

When the third candidate has a reasonable chance of finishing first or second in the second round, switchers' behavior can easily be rationalized within pivotal models. Following Downs (1957a), these models posit that voters face a cost of voting and receive instrumental benefits depending on their likelihood to be pivotal and on the differential utility associated with one candidate defeating another (e.g., Palfrey and Rosenthal, 1983, 1985). Voters who face a sufficiently low cost of voting, whose instrumental benefits are maximized by voting for the third candidate, but who are not indifferent between the top two, will vote for one of the top two when the third is absent and for the third when she is present, as switchers do. The behavior of loyals is more difficult to rationalize, as it requires these voters to all receive sufficiently low instrumental benefits from choosing between the top two candidates. For loyals who abstain when the third candidate is absent, these benefits need to be lower than the cost of voting, itself bounded by the instrumental benefits of voting for the third candidate when she is present. For loyals who vote blank or null when the third candidate is absent, casting a vote for one of the top two candidates would not generate much additional cost, and the associated benefits thus need to be close to null (for instance, as a result of these voters being exactly indifferent between the top two candidates).

The behavior of loyals and switchers is even more difficult to rationalize when the third candidate is unlikely to be a front-runner in the second round. Rational voting models (e.g., Myerson and Weber, 1993) assume that voters care only about the winner of the election, predicting a positive vote share for the third candidate only when she and the second candidate have an equal probability to be in contention for victory. Similarly, in models positing that vote choices are driven by group rules rather than pivot probabilities (Coate and Conlin, 2004; Feddersen and Sandroni, 2006), groups' aggregate utility is assumed to depend only on who wins the election. As a result, in three candidates' races, the sincere equilibrium in which all voters choose their most-preferred candidate only exists when the third candidate is sufficiently strong. When her chances of victory are too low, we end up in a Duvergerian equilibrium where only the top two receive votes (Bouton and Ogden, 2017). Instead, we find that the number of switchers remains equally high in races where the third candidate's strength is lower than that of each of the top two candidates, the third never wins, and where switchers' behavior often leads to the defeat of their preferred top-two candidate. We now discuss four explanations that

could rationalize this behavior within voting models centered on instrumental motives.

First, in rational voter and group-rule utilitarian models, the emergence of a sincere equilibrium hinges both on the strength of the third candidate and on her supporters' differential utilities: how much more utility they derive from the victory of their favorite candidate than the victory of the closest candidate among the top two, and from the victory of that candidate than the victory of their least favorite candidate. A possible interpretation of our results is thus that the low chances of the third candidate are compensated by switchers being close to indifferent between the top two and deriving a much higher utility from the victory of the third candidate than the victory of the closest top-two. Bouton and Ogden (2017) show that, in runoff elections in which three candidates can compete in the second round, the conditions for sincere voting in the first round are relatively similar to the conditions for sincere voting in the second round, so that first round results provide information about the likelihood of a sincere equilibrium in the second round. In line with this model, the fact that the third candidate's first round vote share is sufficiently high to be at the threshold, in the elections we consider in our analysis, could signal preferences or other conditions conducive to sincere voting equilibrium in both the first and second rounds. On the other hand, based on first round results, districts close to the threshold are relatively similar on average to the full sample, as discussed in Section 3.1 and shown in Table A2 in the Appendix. In particular, the share of candidate votes obtained by the third candidate is only higher by 4.1 percentage points, in districts close to the threshold, and the top two candidates' vote share lower by 3.2 percentage points, suggesting that the likelihood of a sincere equilibrium is not much higher in these districts than in the full sample. In addition, while we do not observe the differential utility that switchers would obtain from the victory of the third candidate, the fact that they vote for one of the top two candidates when the third is absent suggests that they are not indifferent between the top two and that their preference for the closest candidate among the top two over their least favorite candidate is sufficiently strong to overcome their cost of voting.

Second, in rational voting models, aggregate uncertainty on the level of candidates' support can lead to an equilibrium in which more than two candidates receive a substantial fraction of votes (Myatt, 2007; Fisher and Myatt, 2017). But these models also predict that strategic coordination on the top two increases with the distance from contention, which is not what we observe: as shown in Section 4.4, the negative impact on the top two candidates' total vote share remains strong and stable in samples 3 and 4, which include elections with a gap of at least five and ten percentage points, respectively, between the third candidate's strength and that of each of the top two candidates. Building on previous models, Bouton et al. (2015) further demonstrate that a sincere equilibrium exists as long as voters conceive that the third candidate could win, with even a very low probability. While the third candidate never ranks first or second in samples 3 and 4, we cannot entirely exclude that some voters thought *ex ante* that she had a non-zero probability of winning. However, Bouton et al. (2015) show that the sincere equilibrium also requires that voters give a sufficiently large probability to the existence

of a state of nature where the third candidate has a stronger support than one of the top two, or that they give a positive probability to a full reversal of support between the third and the candidate ideologically closest. Both conditions seem unlikely to be satisfied when the third candidate lags far behind the top two in the first round. In addition, while existing models are mostly silent about equilibrium selection, one could expect such configuration to facilitate the coordination of instrumental voters on the Duvergerian equilibrium, which again we do not observe.

A third, complementary, way in which our results could be reconciled with instrumental motives is if voters had limited information about, or gave little attention to, the three candidates' first-round rankings and vote shares. Not collecting or not paying attention to freely available and directly relevant information may itself be considered at odds with strategic behavior. In addition, if switchers' voting behavior was driven by limited information, we would expect it to be less prevalent in districts where voters are more informed, which we do not observe. We proxy the level of information by the salience of the race (parliamentary elections being more salient than local elections) as well as three different measures of media exposure: local newspaper consumption, measured at the *département* level, and radio and TV news audiences, measured at the *département* and region level, respectively.¹⁴ We find that the impact on the top two candidates' vote share stays equally high in parliamentary elections and in districts with media exposure higher than the median or the second tercile (Tables F1, F4, and F5 in Appendix F), and that it remains unaffected by the gap between the strength of the third candidate and of each of the top two in these subsamples (Tables F2 and F6 to F8 in Appendix F).

A fourth possible interpretation for switchers' behavior that is consistent with instrumental motives is that these voters are not *short-term* but *long-term* instrumentally rational (Castanheira, 2003). Vote shares in the second round only determine who wins the election, and they are not taken into account for other purposes such as campaign expenditure reimbursement (which is based on first-round vote shares only). Hence, voters should not expect their candidate or party to benefit directly from a higher vote share in the second round. However, they may choose to vote for the third candidate in order to signal their preferences and affect the policies implemented by the winner, or to influence the opinions and future votes of other voters (Piketty, 2000). To the extent that dynamic strategic motives are driving voters' behavior, we should expect voters to trade off the impact of their vote on present elections and on future elections and policies. We should see fewer people vote for the third candidate when their vote is likely to matter more for the result of the current election: for instance, when the second round is expected to be close. Instead, we find that the impact on the vote share of the top

¹⁴Proxying the level of information by the salience of the race is a strategy also used by Hall and Snyder (2015). Data on local newspaper circulation were collected by Julia Cagé (see Cagé, 2017). We collected the data on radio and TV news audiences from Médiamétrie. Data on radio news audience are available at the *département* level on a yearly basis for the years 2003, 2007, 2011, 2012, and 2015. Data on TV news audience are available at the region level for the years 2010, 2011, 2012, and 2015. For each of the three measures of media exposure (local newspaper consumption, radio news audience, and TV news audience), we split our sample based on medians and terciles estimated separately for each election year, to control for time trends. See Appendix F for more details.

two candidates is equally strong in elections where the top two candidates were very close in the first round (Tables E1 and E2 in Appendix E), including in parliamentary elections and in districts with higher media exposure (Tables F3 and F9 to F11 in Appendix F). These results suggest that switchers are willing to decrease the vote share of the top-two candidate they prefer whether or not they expect the race to be close, which further suggests that dynamic strategic motives are unlikely to explain the bulk of our results.¹⁵

In summary, it is difficult to rationalize our results within rational or group rule-utilitarian voting models assuming that the individual or group's utility is only affected by who wins the election, even when allowing for imperfect information, aggregate uncertainty, or dynamic strategic motives. Instead, our results suggest that voters' choice of candidate, as well as their decision whether to vote or abstain might be better explained by taking into account expressive benefits independent of the election results. For many voters, the expressive utility of voting for their favorite candidate outweighs the cost of voting.¹⁶ For others, this expressive utility outweighs the instrumental cost of contributing to the victory of their *least* favorite candidate. A simple way to account for expressive benefits in a theoretical model would be to introduce a share of expressive voters who always vote for their preferred candidate, next to strategic voters behaving according to the predictions of existing canonical models. Such a hybrid model would be *prima facie* consistent with our comparative statics showing that the impact of the presence of the third candidate varies neither with her strength nor with the closeness of the race (Table 6 and Tables E1 and E2 in Appendix E). Alternatively, each voter's choice to behave strategically or expressively when choosing a candidate and deciding whether to vote or abstain could be endogenized in a full-fledged model of costly voting in elections with more than two candidates. Future work along these lines could build on Shayo and Harel (2012), whose theoretical framework includes a trade-off between instrumental and non-instrumental voting motives, and on Kawai and Watanabe (2013) and Fisher and Myatt (2017), who introduce a share of sincere voters in rational models of vote choice.

5.2 Candidates' decision to drop out

The negative impact of the third candidate's presence on the chances of victory of the top-two candidate ideologically closest to her sheds light on the motives underlying third candidates' behavior as much as voters'. Third candidates could prevent this effect by simply dropping out of the race between the first and second rounds. The fact that they do not drop out more systematically, especially when their own likelihood of being a front-runner is very low, implies that they often value the benefits associated with competing in the second round more than they

¹⁵Alternatively, switchers might continue voting for the third when this threatens the victory of the top-two candidate they prefer because they reason that the associated signal is all the more powerful as it is costly. This interpretation requires a large number of voters to be sufficiently sophisticated and to believe that candidates and other voters are sufficiently sophisticated as well to interpret their vote in this way, making it perhaps less likely.

¹⁶For other recent evidence on expressive motives of turnout, see Fiva and Smith (2017) and Ujhelyi et al. (2017).

care about influencing the race's outcome. In this last section, we provide additional evidence to better understand why, and under which circumstances, third candidates decide to stay in the race instead of dropping out. This evidence also allows us to better characterize complier districts, in which the effects of the third candidate's presence were estimated, and to bring additional support for the exclusion restriction underlying these estimates.

Formally, using our regression discontinuity design framework, we regress a dummy equal to 1 if the third candidate drops out on the assignment variable, and compare the effects in different configurations. The most important factor affecting candidates' decision to drop out is political orientation. As mentioned in Section 3.2, the third candidate drops out of most races (91.1 percent, at the threshold) in which she has the same orientation (far-left, left, center, right, or far-right) as one of the top two candidates. Conversely, she stays in the race in the large majority of elections (85.2 percent, at the threshold) in which she has a different orientation than both top two candidates (Figure G1 and Table G1 in Appendix G).

To better understand this difference, we gathered descriptive evidence from press articles covering instances of candidates dropping out. Using Factiva's research tool, we collected all articles released between the two rounds of all elections in our sample and containing the entity "désist".¹⁷ We obtained a total of 1,678 articles published in 86 different newspapers in election years 1997, 2002, 2007, 2011, 2012, and 2015 (more information in Appendix H). In each, we systematically coded the context in which the dropout took place (decision made by the party, existence of an agreement among parties, or decision made individually by the candidate); the reasons provided by the party or candidate (preventing the victory of another candidate or feeling ideologically close to a top-two); and whether the article mentions the reactions of the candidate's party, voters, or competing candidates. All statistics are reported in Tables H1 to H3 in Appendix H.

The most striking lessons of this investigation are as follows. First, articles covering dropouts are nearly ten times more likely to report that the third candidate's decision to drop out was the result of a party-level agreement when she has the same orientation as one of the top two candidates than when she has a different orientation than both. Conversely, articles are ten times more likely to mention that she took the decision to drop out on her own, independently from any party's instructions, when she has a different orientation than both top two candidates. Second, when the third candidate has a different orientation, 63.1 percent of articles report that she dropped out in order to prevent the victory of another candidate (against 29.2 percent when she has the same orientation as one of the top two), and none that she dropped out because she felt ideologically close to one of the top two candidates (against 15.9 percent).

These patterns suggest that third candidates who have a different orientation than both of the top two only drop out in rare circumstances, motivated by their aversion for one of the top two more than ideological proximity with the other. In general, they fail to internalize the cost that

¹⁷This entity is present in all forms of the verb "se désister" (to drop out) and in the noun "désistement" (dropout).

staying in the race generates for the ideologically closest top-two candidate. Instead, when the third candidate has the same orientation as one of the top two, ideological proximity (as well as, perhaps, the habit to govern together) helps their parties reach an agreement which internalizes this cost and, in an overwhelming majority of cases, results in the third candidate dropping out. Such agreements are often *département*-wide or even nationwide, so that dropouts of candidates of the two parties balance each other, and the cost for one candidate to drop out is mitigated by the increased likelihood of victory of another candidate of her party in another district.

Interestingly, when the third candidate has the same orientation as one of the top two, her likelihood to drop out is not only much higher, but also varies more with first-round results. Same-orientation third candidates are relatively more likely to drop out in competitive races, characterized by a small difference between the strengths of the top two candidates (Table G2 in Appendix G). This brings additional support for the interpretation that their decision to drop out is motivated by their party's desire to ensure the victory of a sister party's candidate. However, same-orientation third candidates are less likely to drop out when they are closer to the top two candidates in the first round (Table G4 in Appendix G), suggesting that candidates trade-off party instructions with their desire to stay in the race, and that their immediate individual interest is more likely to prevail when they have a chance of being a front-runner in the second round. Instead, the dropout decision of third candidates with a different orientation than both top two does not follow any such pattern (Tables G3 and G5 in Appendix G).¹⁸

Our analysis of the factors affecting third candidates' dropout decision has two implications for the interpretation of our effects on voters' behavior. First, voters' difficulty to behave strategically in our sample might be driven by the overrepresentation, among complier districts, of second-round races where the third candidate has a different orientation than both top two, resulting from the low likelihood of third-candidate dropouts in this setting. Second, this feature of our sample is not idiosyncratic. It results from the intrinsic difficulty for candidates of distinct orientations to find an agreement leading to the dropout of the lower ranked. The tradeoff between expressive and instrumental motives might well be particularly difficult for voters to solve when they have to choose between candidates of different orientations, but this tradeoff is also likely to occur systematically more often in this context.

¹⁸The third candidate's decision to drop out or stay in the race only occurs conditional on qualifying for the second round. Observing this decision could, in theory, affect voters' behavior directly, independently of the third candidate's presence, thus violating the exclusion restriction (previously discussed in Section 3.2). The evidence we collected on factors affecting dropouts suggests this is unlikely to be the case, in particular in elections where the third candidate has a different orientation than both top two, which account for the majority of our compliers, and which we focus on to measure the effects on winner identity and on the vote shares of the top two depending on their political orientation or on the strength of the third candidate (Sections 4.3 through 4.5). First, in these districts, the fact that the third candidate's decision does not systematically vary with first-round results alleviates the worry that it might increase the salience of some features of the race, including its level of competitiveness. Second, in addition to providing information, the decision of the candidate could provoke specific voters' reactions. However, less than one percent of dropout articles report an adverse reaction following the decision to drop out.

6 Conclusion

This paper highlights the motivations and consequences of citizens voting for lower-ranked candidates in elections held under plurality rule. Using a fuzzy regression discontinuity design around the qualification threshold for the second round of French local and parliamentary elections, we compare electoral outcomes when voters have to choose between two or three candidates.

The presence of the third candidate increases the share of registered citizens who vote for any candidate by 7.8 percentage points; it reduces the vote share of the top two candidates in proportion to their ideological proximity with the third by 6.9 percentage points. The latter impact remains equally strong when first-round results indicate that the third candidate is very unlikely to be a front-runner in the second round, and it causes the defeat of the Condorcet winner in one-fifth of the races.

The behavior of third candidates' voters, whether they are loyal supporters who abstain or vote blank or null when the third candidate is absent, or switchers who vote for one of the top two candidates in that case, is difficult to rationalize within canonical voting models. In both rational voter and group-rule-utilitarian models, the individual or group's utility is only affected by the election's winner. Instead, our results suggest that voters' choice of candidate, as well as their decision whether to vote or abstain, can only be satisfactorily explained by taking into account expressive benefits independent of the result. For many voters, the expressive utility of voting for their favorite candidate outweighs the cost of voting. For others, this expressive utility outweighs the instrumental cost of contributing to the victory of their *least* favorite candidate.

Anticipating voters' behavior, third candidates could drop out of the race between the first and the second rounds to prevent the defeat of the ideologically closest top-two candidate. However, our results suggest that third candidates often value the benefits associated with competing in the second round more than influencing its outcome. When the third candidate has a different political orientation than both top two candidates, dropouts remain an exception. They only become the rule when the third candidate has the same orientation as one of the top two candidates and she follows party-level agreements. In that case, the coordination taking place between parties of the same orientation reduces the likelihood that citizens split their support and waste their vote. This may contribute to explain political parties' very existence, a conjecture which should be tested more directly by future work.

Regardless, our overall results on voter and candidate behaviors suggest that plurality rule often leads to suboptimal outcomes, due to the combination of two phenomena. First, in a large number of elections voters have to choose between more than two candidates due to the difficulty for parties of differing orientations to reach an agreement. And second, a sufficiently large fraction of voters value expressive more than instrumental motives when confronted with such an electoral offer. These voters simply want to rally for their favorite, even if that means

their least favorite might win and bring in undesirable policies. Ultimately, our findings call into question the widespread use of the plurality rule to aggregate voter preferences.

Appendix

This Section presents the main appendix. Additional analysis can be found in the Supplemental Material available here.

Appendix A. Additional tables and figures

Table A1: Elections in the sample

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

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

Table A2: First round summary statistics: districts close to the threshold vs. full sample

	Summary statistics - 1st round					
	Full Sample			Close to the threshold		
	Mean	Sd	Obs	Mean	Sd	Obs
Registered voters	45,964	30,882	7,257	45,753	30,812	1,817
Turnout	0.582	0.124	7,257	0.619	0.100	1,817
Candidate votes	0.562	0.122	7,257	0.597	0.101	1,817
Blank and Null votes	0.019	0.011	7,257	0.022	0.011	1,817
Number of candidates	7.78	4.08	7,257	7.36	4.09	1,817
Vote share 3rd candidate (cand. votes)	0.167	0.053	7,257	0.208	0.036	1,817
Vote share top2 candidates (cand. votes)	0.657	0.098	7,257	0.625	0.086	1,817

Notes: Districts close to the threshold include all districts in which the first round vote share of the third candidate (as a fraction of registered citizens) was within exactly 2 percentage points from the threshold.

Table A3: Intend to treat estimates when the third candidate has the same orientation as one of the top two candidates

Outcome	(1) Turnout	(2) Null and Blank	(3) Candidate votes 2nd round	(4) Vote share top 2	(5) Closest cand wins
3rd qualifies	-0.015 (0.019)	-0.004 (0.003)	-0.010 (0.020)	-0.023 (0.020)	0.022 (0.068)
Robust p-value	0.398	0.314	0.526	0.261	0.611
Observations	687	824	690	718	820
Polynomial order	1	1	1	1	1
Bandwidth	0.024	0.029	0.024	0.025	0.030
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	0.676	0.033	0.643	0.641	0.718

Notes: Sample includes all elections where the third candidate has the same orientation as one of the top two candidates. In column (5), the sample is further restricted to elections where the candidate ideologically closest to the third is identified (we exclude the elections where the three candidates have the same orientation and elections where one of the top two candidates is from a non-classified orientation). Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. In Columns (1) to (4), each outcome uses the number of registered voters as the denominator. In Column (5), the outcome is a dummy equal to 1 if the candidate closest to the third wins the election. The dependent variable is a dummy equal to 1 if the third candidate is qualified. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table A4: Impact on blank and null votes separately for the 2015 local elections

Outcome	(1) Null and Blank votes	(2) Blank votes	(3) Null votes
	2nd round		
3rd present	-0.027*** (0.003)	-0.015*** (0.002)	-0.011*** (0.001)
Robust p-value	0.000	0.000	0.000
Observations	474	386	601
Polynomial order	1	1	1
Bandwidth	0.014	0.011	0.017
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.053	0.036	0.017

Notes: Sample includes only the 2015 local elections. Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. Each outcome uses the number of registered voters as the denominator. The variable of interest (the presence of a third candidate in the second round) is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table A5: Impact on the vote shares of the candidates who ranked first and second in the first round, taken separately

Outcome	(1) Vote share top 2	(2) Vote share 1st	(3) Vote share 2nd
	2nd round		
3rd present	-0.069*** (0.020)	-0.032** (0.011)	-0.035** (0.012)
Robust p-value	0.003	0.012	0.016
Observations	2,250	2,923	2,126
Polynomial order	1	1	1
Bandwidth	0.024	0.031	0.024
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.551	0.308	0.249

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

Table A6: Impact on candidates A and B when the third candidate is C (1st setting)

Outcome variable	(1)	(2)	(3)
	Top two cand.	Cand. A	Cand. B
3rd present	-0.077*** (0.016)	-0.025** (0.010)	-0.052*** (0.010)
Robust p-value	0.000	0.024	0.000
Observations	556	627	508
Polynomial order	1	1	1
Bandwidth	0.019	0.022	0.018
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.547	0.268	0.285

Notes: Sample includes the elections where the top three candidates have distinct political orientations and where the third candidate is located to the right of both the first and the second candidates. Other notes as in Table A5.

Table A7: Impact on candidates B and C when the third candidate is A (2nd setting)

Outcome variable	(1)	(2)	(3)
	Top two cand.	Cand. B	Cand. C
3rd present	-0.062** (0.020)	-0.086*** (0.019)	0.005 (0.014)
Robust p-value	0.021	0.000	0.639
Observations	136	160	187
Polynomial order	1	1	1
Bandwidth	0.012	0.013	0.015
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.478	0.300	0.169

Notes: Sample includes the elections where the top three candidates have distinct political orientations and where the third candidate is located to the left of both the first and the second candidates. Other notes as in Table A5.

Table A8: Impact on candidates A and C when the third candidate is B (3rd setting)

Outcome variable	(1)	(2)	(3)
	Top two cand. 2nd round	Cand. A	Cand. C
3rd present	-0.125*** (0.039)	-0.071*** (0.022)	-0.056 (0.036)
Robust p-value	0.004	0.004	0.202
Observations	145	152	133
Polynomial order	1	1	1
Bandwidth	0.015	0.016	0.013
Bandwidth method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.508	0.289	0.217

Notes: Sample includes the elections where the top three candidates have distinct political orientations and where the third candidate is located between the first and the second candidates. Other notes as in Table A5.

Table A9: Impact on the winning margin in the second round

Outcome	(1)
	Distance winner - 2nd candidate 2nd round
3rd present	-0.058*** (0.014)
Robust p-value	0.000
Observations	2,677
Polynomial order	1
Bandwidth	0.030
Band. method	MSERD
Mean, left of the threshold	0.155

Notes: The outcome variable is the vote share of the winner minus the vote share of the second candidate in the second round, as fractions of candidate votes. Other notes as in Table A5.

Table A10: Impact on the winning margin depending on the orientations of the top two candidates

Outcome	(1)	(2)	(3)
	Distance winner - 2nd candidate, 2nd round		
	Full Sample	Top 2 same orientation	Top 2 distinct orientations
3rd present	-0.058*** (0.014)	0.056 (0.045)	-0.080*** (0.016)
Robust p-value	0.000	0.407	0.000
Observations	2,677	164	2,303
Polynomial order	1	1	1
Bandwidth	0.030	0.024	0.028
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.155	0.136	0.156

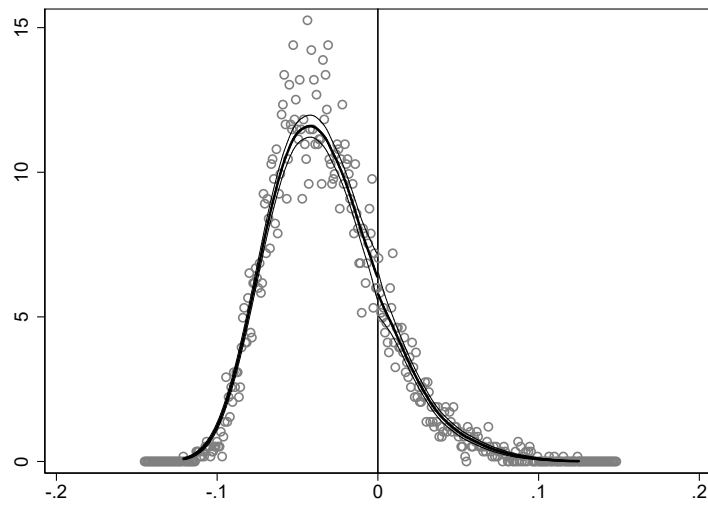
Notes: The outcome variable is the vote share of the winner minus the vote share of the second candidate in the second round, as fractions of candidate votes. Column 2 includes only the elections where the top two candidates have the same orientation. Column 3 includes only the elections where the top two candidates have distinct orientations. Other notes as in Table A5.

Table A11: Impact on participation depending on the orientations of the top two candidates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Turnout	Full Sample		Top 2 same orientation			Top 2 distinct orientations		
		BlankNull	Cand vot	Turnout	BlankNull	Cand vot	Turnout	BlankNull	Cand v
3rd present	0.040** (0.017)	-0.037*** (0.004)	0.078*** (0.019)	0.120*** (0.033)	-0.087*** (0.016)	0.207*** (0.038)	0.026 (0.020)	-0.032*** (0.003)	0.059* (0.021)
Robust p-value	0.041	0.000	0.001	0.002	0.000	0.000	0.215	0.000	0.014
Observations	2,298	2,630	2,374	301	286	274	1,981	2,076	1,998
Polyn. order	1	1	1	1	1	1	1	1	1
Bandwidth	0.025	0.028	0.026	0.033	0.031	0.030	0.024	0.025	0.024
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	0.598	0.047	0.548	0.533	0.097	0.439	0.605	0.042	0.562

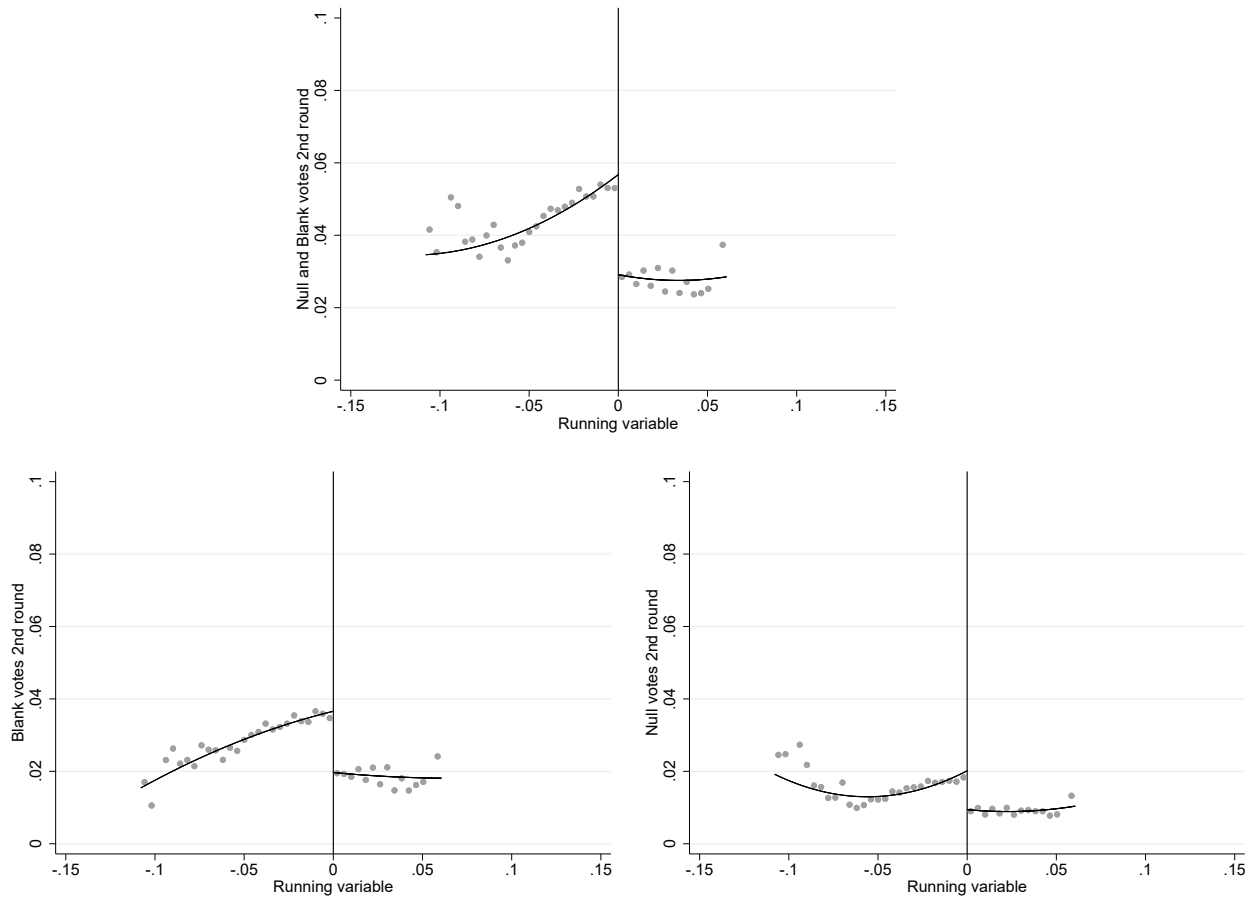
Notes: Columns 4, 5, and 6 include only elections where the top two candidates have the same orientation. Columns 7, 8, and 9 include only elections where the top two candidates have distinct orientations. Other notes as in Table A5.

Figure A1: McCrary test of the density of the running variable



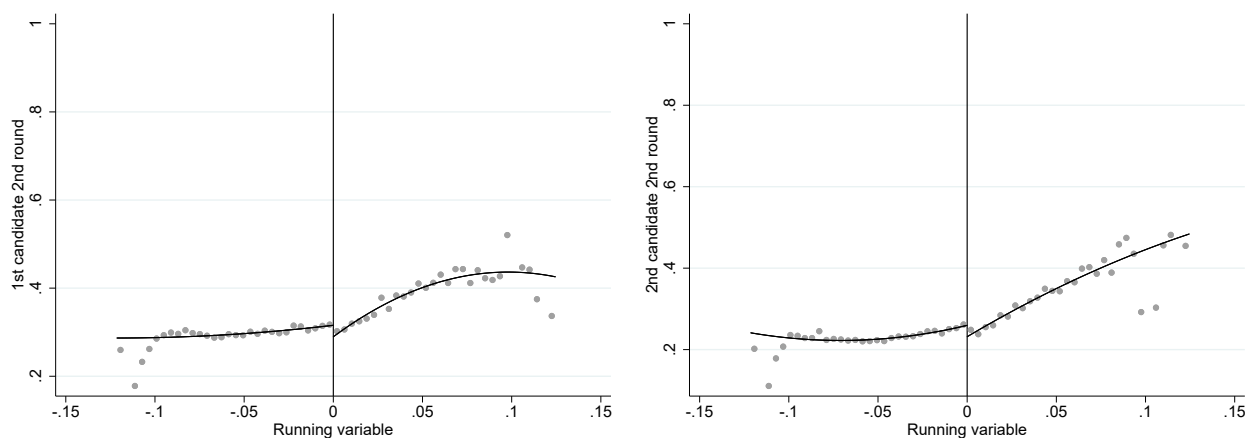
Notes: This Figure tests for a jump in the density of the running variable (the qualifying margin of the third-highest-ranking candidate in the first round) at the threshold. The solid line represents the density of the running variable. Thin lines represent the confidence intervals.

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



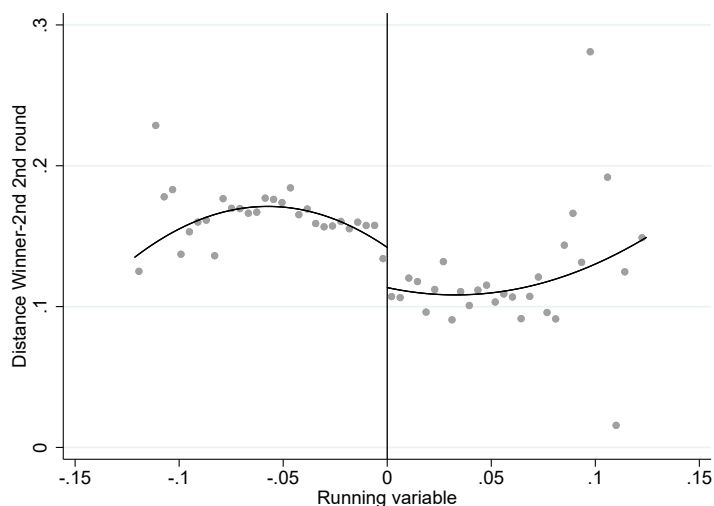
Notes: Sample includes only the 2015 local elections. Dots represent the local averages of the outcome variable (y-axis). Averages are calculated within 0.4 percentage-point-wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the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

Figure A3: Impact on the vote shares of the candidates who ranked first and second in the first round, taken separately



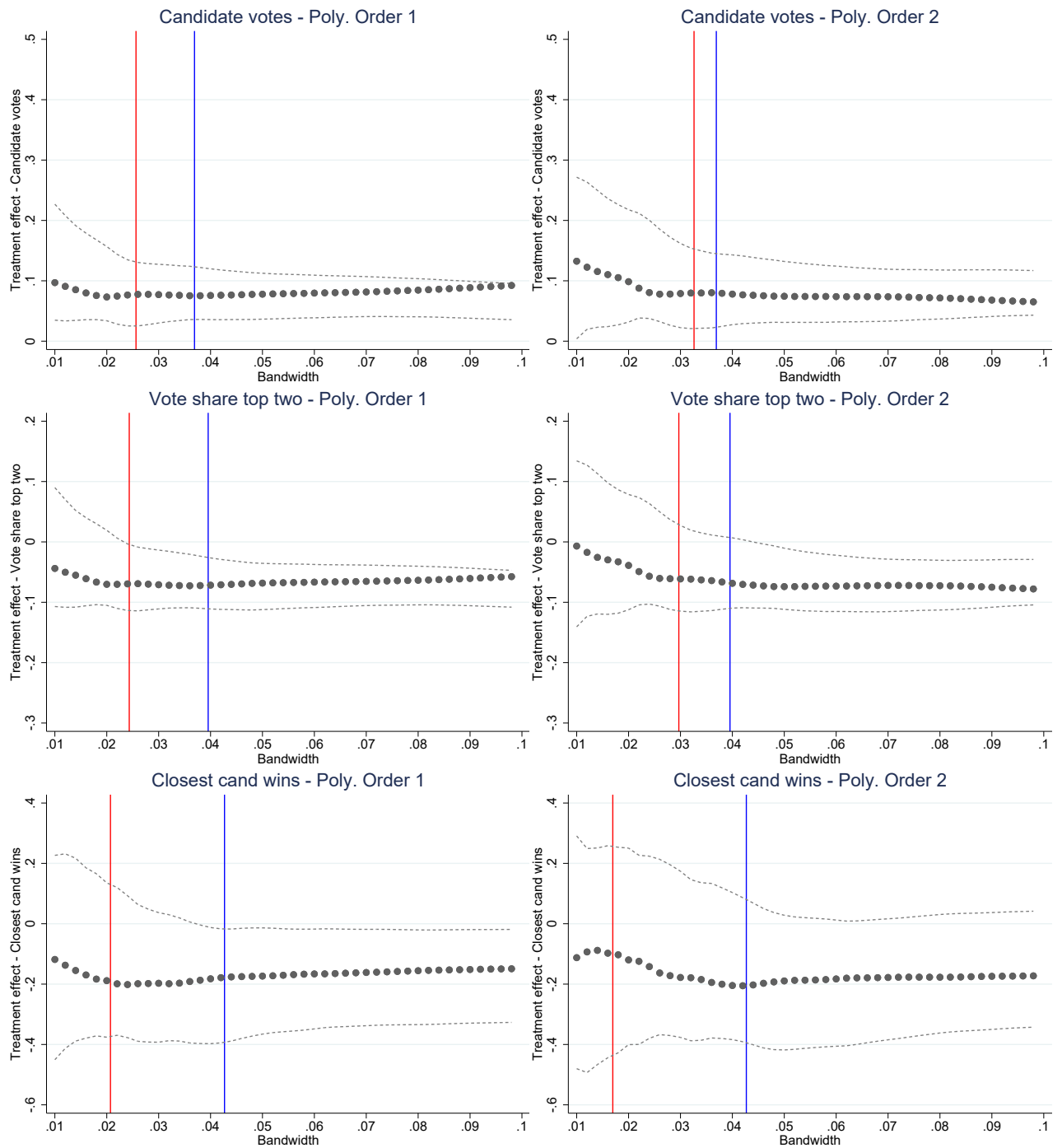
Notes: Dots represent the local averages of the vote share of the first (resp. second) candidate in the second round (y-axis). Vote shares are computed using the number of registered citizens as the denominator. Averages are calculated within 0.4 percentage-point-wide bins of the running variable (x-axis). The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured in percentage points. Continuous lines are a quadratic fit.

Figure A4: Impact on the winning margin in the second round



Notes: Dots represent the local averages of the difference between the share of candidate votes obtained by the winner and by the candidate who came in second in the second round. Other notes as in Figure A3.

Figure A5: Robustness of the main results to bandwidth choice



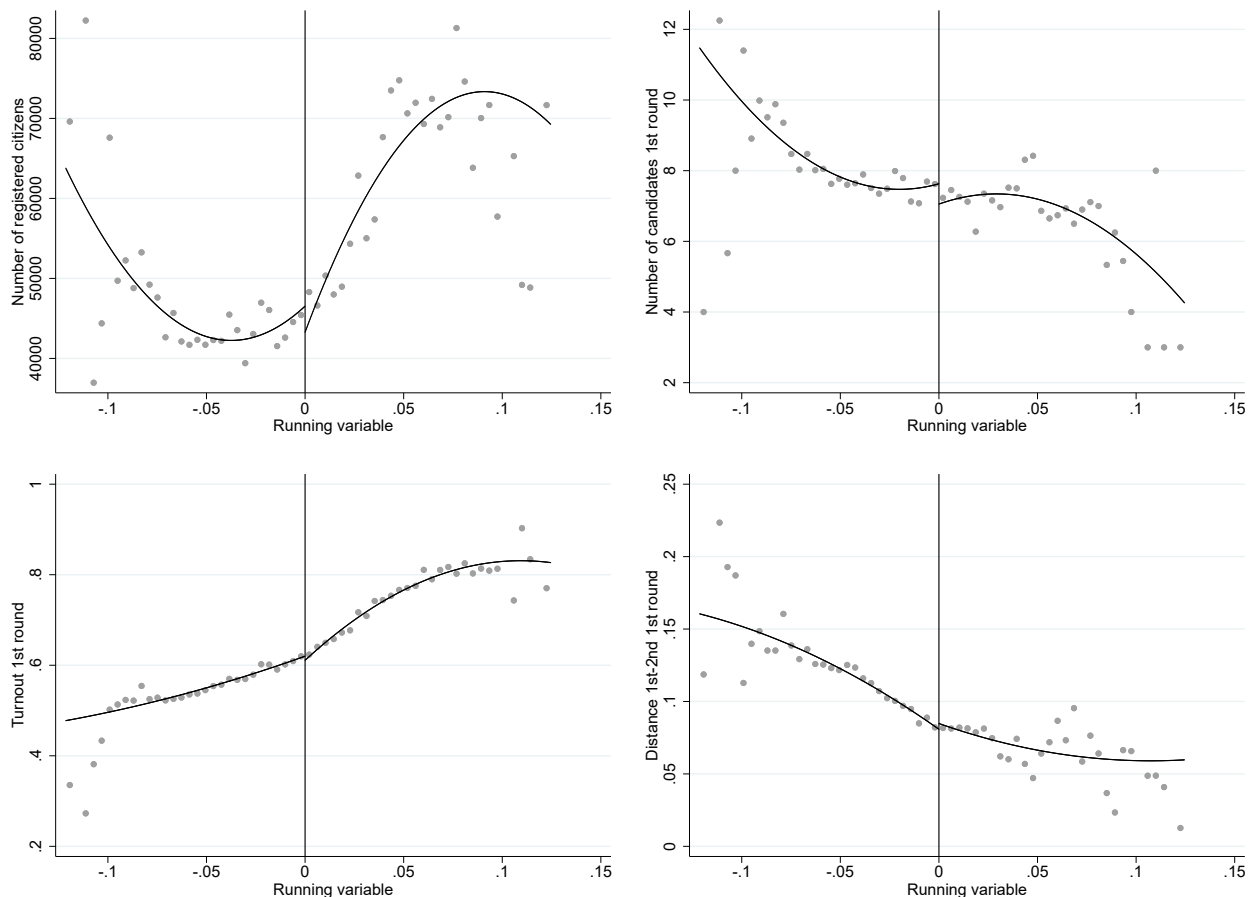
Notes: We show the sensitivity of our main results to bandwidth choice, using a linear or a quadratic specification. Dots represent the estimated treatment effect of the presence of the third candidate using different bandwidths (x-axis). Dotted lines represent the 95% robust confidence interval. We report all estimates for values of the bandwidth from 1 to 10 percentage points, in step of 0.2 percentage points. The vertical red (resp. blue) line gives the value of the MSERD (resp. IK) optimal bandwidth.

Appendix B: Placebo tests

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

As shown in Figure B1, there is no significant jump at the cutoff for any of these variables. The formal estimation confirms the absence of treatment effect. Columns 1 through 4 of Table B1 present the results obtained for these four outcomes under our preferred specification. None of the estimates is statistically significant at the standard levels. Hence, we cannot reject the null hypothesis that the treatment has no effect on these baseline variables.

Figure B1: Placebo tests on baseline variables



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

In addition, we conduct the following general test for imbalance. We regress the assignment variable D on a set of first-round variables including the four aforementioned variables as well

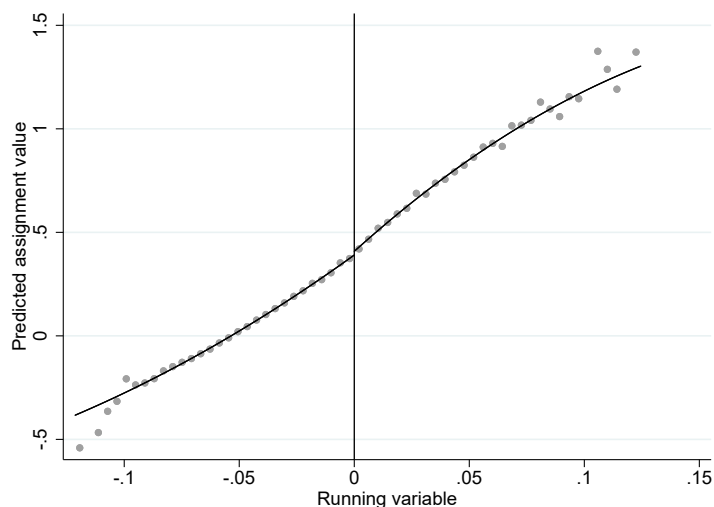
as share of candidate votes, vote share of each of the top three candidates, political label and orientation of the three candidates, number of candidates from the left, right, far-right, far-left, and center, number of candidates of a non classified orientation, dummies if there is at least one candidate from the far-left, left, center, right and far-right, and a dummy equal to one if the third candidate has the same orientation as one of the top two candidates. We then use the coefficients from this regression to predict assignment status, and test whether the predicted value jumps at the threshold. As shown in Figure B2, the assignment status predicted by base-line variables increases continuously as a function of the running variable and does not show any discontinuity at the threshold. This suggests that there is no systematic discontinuity in the preexisting observable districts' characteristics at the threshold. The formal estimate in Column 5 of Table B1 confirms this result: the coefficient is small (1.5 percentage points) and non-significant.

Table B1: Placebo tests

Outcome	(1) Nb reg. citizens	(2) Nb cand. 1st round	(3) Turnout 1st round	(4) Dist. 1-2 1st round	(5) Predicted assignment
3rd present	5,601 (5,639)	-0.14 (0.64)	0.003 (0.015)	0.003 (0.011)	0.015 (0.017)
Robust p-value	0.246	0.902	0.705	0.852	0.607
Observations	1,879	2,428	2,189	2,018	1,930
Polynomial order	1	1	1	1	1
Bandwidth	0.021	0.026	0.024	0.022	0.021
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD
Mean, left of the threshold	43,982	7.55	0.604	0.092	0.301

Notes: Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest (the presence of a third candidate in the second round), is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Figure B2: General balance test



Notes: Dots represent the local averages of the predicted assignment status (y-axis). Other notes as in Figure B1.

Appendix C. Robustness of the results to two special cases

See the additional Supplement, included in the replication material.

Appendix D. Campaign expenditures

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

Data on campaign expenditures are made publicly available by the CNCCFP. The CNCCFP was created in 1990. Hence, data for elections held before 1990 are not available. Official accounts for the most recent elections - 2011 and 2015 local elections, as well as 2007 and 2012 parliamentary elections - are available online on the CNCCFP website (<http://www.cnccfp.fr/index.php?art=58>). Official accounts for the 1993, 1997, and 2002 parliamentary elections were digitized from printed booklets by Abel François and his co-authors for their studies on the impact of electoral expenditures on turnout (Fauvelle-Aymar and François, 2005) and electoral results (Foucault and François, 2005). In total, we were able to gather data corresponding to 77.8 percent of our sample.¹⁹

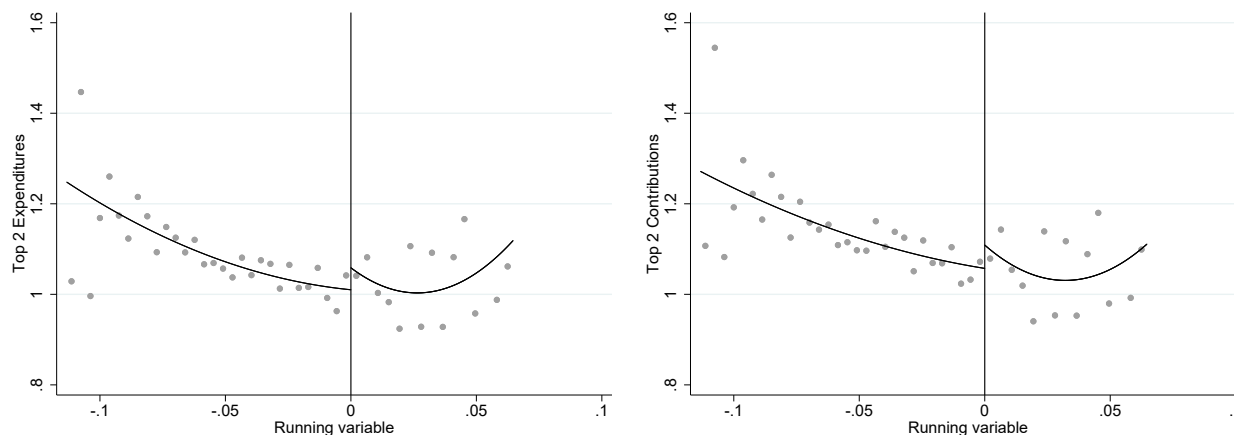
¹⁹Note that for the 2011 local elections, data are only available for districts exceeding 9,000 inhabitants. As a result, we observe the campaign expenditures for 74.4 percent of that election's races. In addition, for the

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

These data enable us to test whether the top two candidates increase their campaign expenditures in response to the presence of the third candidate. As we can see in Figure D1, the presence of the third candidate does not significantly affect the campaign expenditures of the top two candidates, or the contributions they receive to finance their campaign. Table D1 provides the formal estimates. Neither the effect on top two candidates' total expenditures nor the estimate on total contributions is statistically significant. The estimate on contributions received from candidates' political parties is significant at the 10 percent level, and positive. Nevertheless, the estimate on total contributions is small, not significant, and actually negative: the top two candidates do not receive significantly more money overall when the third candidate is present.

1993 parliamentary elections, data for two French territories overseas are missing. Finally, data are missing for candidates who received less than 1 percent of the candidate votes in the first round (in which case they do not need to release their accounts publicly), or because they did not release their campaign account on time (this happened in 1 case for the first candidate, 8 cases for the second candidate, and 34 cases for the third candidate).

Figure D1: Campaign expenditures of the top two candidates



Notes: One outlier has been removed to make the graph clearer (the district “Saint-Pierre-et-Miquelon” in the 1997 parliamentary elections). Dots represent the local averages of the outcome variable (y-axis). Averages are calculated within 0.4 percentage-point-wide bins of the running variable (x-axis). Each outcome uses the number of registered voters as the denominator. The running variable (qualifying margin of the third-highest-ranking candidate in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

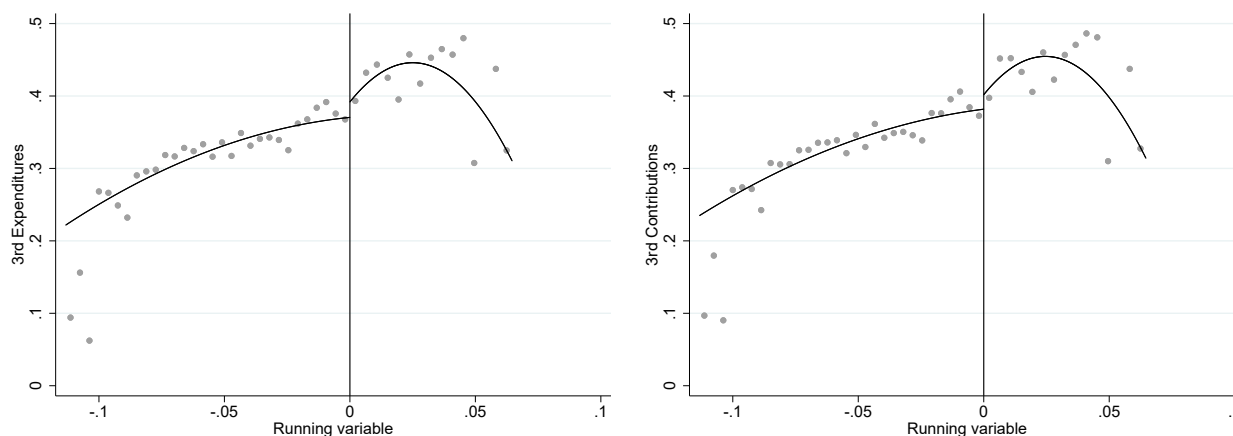
Table D1: Campaign expenditures of the top two candidates

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total expenditures	Total contrib.	Personal contrib.	Party’s contrib.	Natural advantages	Donations	Other contrib.	Balance
3rd present	0.013 (0.079)	-0.024 (0.092)	-0.078 (0.068)	0.080* (0.039)	0.003 (0.012)	0.003 (0.061)	-0.006 (0.020)	-0.010 (0.021)
Robust p-value	0.885	0.577	0.224	0.061	0.803	0.816	0.937	0.512
Observations	890	786	800	1,236	1,437	954	1,216	774
Polyn. order	1	1	1	1	1	1	1	1
Bandwidth	0.013	0.012	0.012	0.018	0.023	0.014	0.018	0.011
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	1.017	1.045	0.688	0.113	0.037	0.198	0.022	0.044

Notes: Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. Each outcome uses the number of registered voters as the denominator. The outcome “Other Contributions” (column 7) sums all the contributions received by the candidates that are not personal contributions, party contributions, natural advantages, or donations. In the 1993 parliamentary elections, natural advantages were not counted separately, and they are included in the other contributions. The variable of interest (the presence of a third candidate in the second round) is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

We now turn to the impact of the presence of the third candidate on her own overall campaign expenditures (which again sum up expenditures incurred before the first round and between the first and second rounds). We do not find any significant impact on the third candidate's total expenditures or on the total contributions she received (Figure D2 and Table D2). When we disentangle between the different sources of contributions, only two coefficients out of six are significant at the 10 percent level, and one is negative.

Figure D2: Campaign expenditures of the third candidate



Notes as in Figure D1.

Table D2: Campaign expenditures of the third candidate

Outcome	(1)	(2)	Third candidate			(6)	(7)	(8)
	Total expenditures	Total contrib.	Personal contrib.	Party's contrib.	Natural advantages	Donations	Other contrib.	Balance
3rd present	0.035 (0.037)	0.038 (0.040)	0.054* (0.035)	-0.031* (0.018)	-0.000 (0.007)	-0.004 (0.024)	0.000 (0.005)	-0.000 (0.011)
Robust p-value	0.441	0.429	0.090	0.074	0.939	0.679	0.961	0.958
Observations	929	842	831	679	1,034	679	901	701
Polyn. order	1	1	1	1	1	1	1	1
Bandwidth	0.014	0.013	0.012	0.010	0.017	0.010	0.013	0.010
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD	MSERD
Mean	0.382	0.397	0.307	0.023	0.015	0.043	0.005	0.010

Notes as in Table D1.

Appendix E: Impact on the top two candidates depending on the closeness of the race

We estimate the impact of the presence of the third candidate on the vote share of the top two candidates depending on the closeness of the race in the first round. In Table E1, closeness is defined as the difference in vote shares (as a fraction of candidate votes) between the first and second candidates. In Table E2, closeness is defined as the difference in strengths between the first and second candidates, where a candidate's strength is equal to the sum of first-round vote shares (as a fraction of candidate votes) of all candidates from the same orientation (see Section 4.4).

As defined in Section 4.4, sample 1 includes all elections in which the top three candidates have distinct political orientations and the third candidate is either on the left or the right of both top two candidates, so that the candidate ideologically closest to the third is clearly identified. We then consider two subsamples: one in which the distance between the top two candidates in the first round is smaller than 10 percentage points and one in which the distance is smaller than 5 percentage points. In those subsamples, the gap in vote shares (resp. strengths) between the first and second candidates is, on average, equal to 4.3 and 2.3 (resp. 4.7 and 2.5) percentage points, respectively, close to the discontinuity.

As shown in Tables E1 and E2, whatever the definition of closeness we use, the effect of switchers' behavior on the vote share of the top two candidates is robust across the three samples and strikingly close in magnitude. In Table E1, all estimates are significant at the 1 or 5 percent level and included between 7.7 and 8.2 percentage points (column 1). In Table E2, all estimates are significant at the 1 percent level and included between 8.2 and 10.4 percentage points (column 1).

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

Our results are robust to considering sample 2 as defined in Section 4.4 instead of sample 1.

Table E1: Impact on the top two candidates depending on the closeness of the race (defined in terms of vote shares)

	(1)	(2)	(3)	(4)
Impact 3rd present	Top 2 cand. 2nd round	<i>Bandwidth / Observations</i>	Closest candidate 2nd round	Furthest candidate 2nd round
Sample 1	-0.082*** (0.017)	0.014 546	-0.061*** (0.008)	-0.027** (0.013)
Sample 1 + distance $12 \leq 10$ pp	-0.080*** (0.021)	0.012 345	-0.056*** (0.009)	-0.031* (0.015)
Sample 1 + distance $12 \leq 5$ pp	-0.077** (0.025)	0.014 228	-0.057*** (0.010)	-0.024 (0.017)

Notes: The distance between the top two candidates is defined as the difference in vote shares (as a fraction of candidate votes) between the first and second candidates. Column 2 gives the bandwidths used for the estimation of the impact on the vote share of the top two candidates as well as the number of observations lying in those bandwidths. Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. Each outcome uses the number of registered voters as the denominator. The variable of interest (the presence of a third candidate in the second round) is instrumented by the assignment variable (whether the vote share of the third-highest-ranking candidate was higher than the cutoff). Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E2: Impact on the top two candidates depending on the closeness of the race (defined in terms of strengths)

	(1)	(2)	(3)	(4)
Impact 3rd present	Top 2 cand. 2nd round	<i>Bandwidth / Observations</i>	Closest candidate 2nd round	Furthest candidate 2nd round
Sample 1	-0.082*** (0.017)	0.014 546	-0.061*** (0.008)	-0.027** (0.013)
Sample 1 + distance $12 \leq 10$ pp	-0.094*** (0.020)	0.019 423	-0.059*** (0.010)	-0.034** (0.013)
Sample 1 + distance $12 \leq 5$ pp	-0.104*** (0.034)	0.015 189	-0.064*** (0.012)	-0.036 (0.024)

Notes: The distance between the top two candidates is defined as the difference in strengths between the first and second candidates. Other notes as in Table E1.

Appendix F. Impact on the top two candidates depending on voters' level of information

See the additional Supplement, included in the replication material.

Appendix G. Third candidate dropouts - RDD analysis

In this section, we use our regression discontinuity design framework to provide additional evidence on third candidates' dropout decision. In the graphs and tables below, the outcome is a dummy equal to 1 if the third candidate drops out of the race in the second round. By definition, it always takes value 0 at the left of the threshold.

The first graph in Figure G1 plots the probability to drop out against the running variable in the whole sample. Note that it is the exact mirror of the first-stage figure shown in Section 3.2. We then differentiate elections where the third candidate has the same political orientation as one of the top two candidates from elections where she does not. Table G1 provides the formal estimates: on average, the third candidate is much more likely to drop out when she has the same political orientation as one of the top two candidates (91.1 percent close to the threshold) than when she has a different orientation than both of them (14.8 percent close to the threshold).

We now assess whether the likelihood to drop out depends on first round results, focusing on elections where the candidate ideologically closest to the third is clearly identified: in the sample where the third candidate has the same orientation as one of the top two, we exclude the elections where the three candidates have the same orientation and elections where one of the top two candidates is from a non-classified orientation. In the sample where the third candidate has a different orientation than both top two, we only include the elections of sample 1 as defined in Section 4.4.

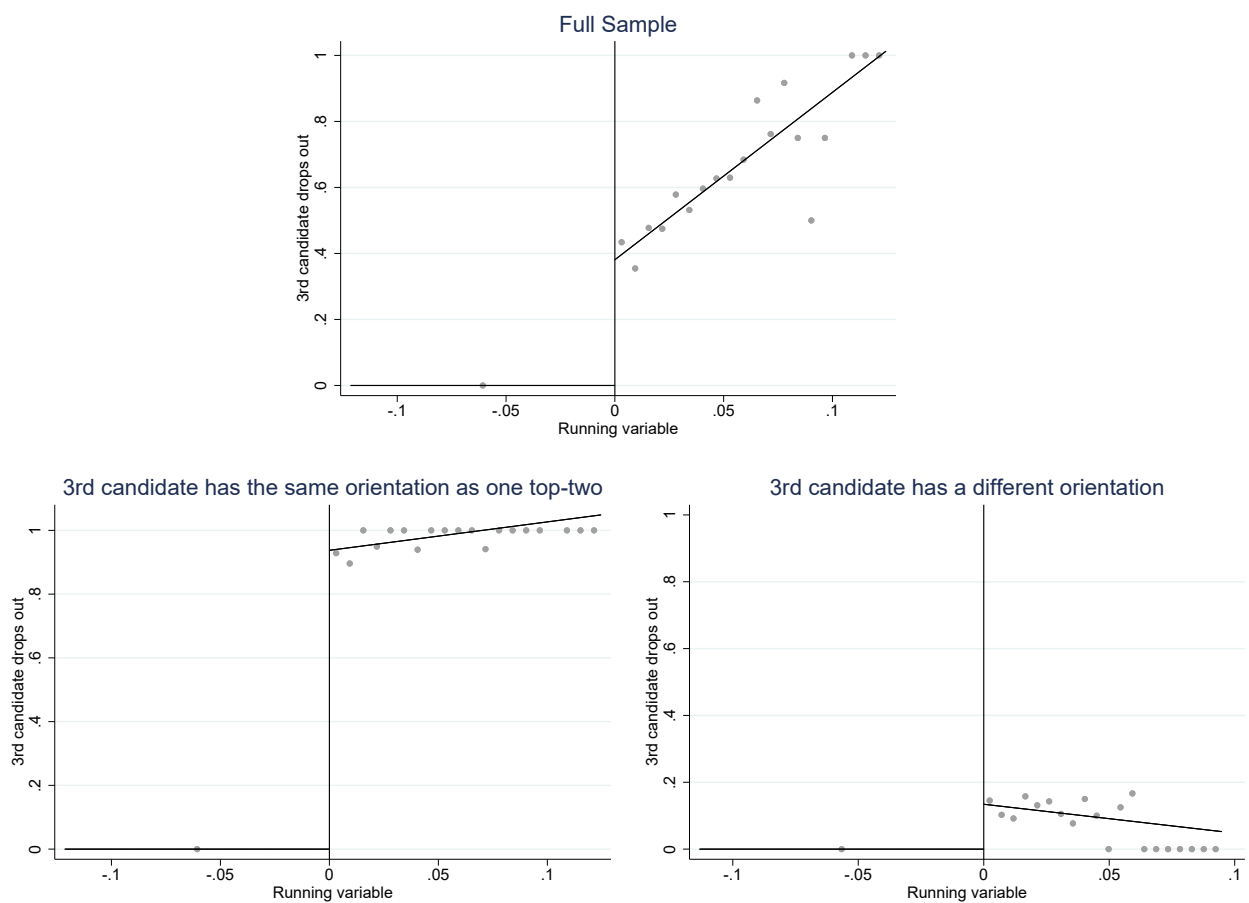
We first test whether the decision to drop out depends on the closeness of the race. As in Appendix E, we test for two definitions of closeness: difference between the vote shares of the first and second candidates in the first round and difference between their strengths. Both measures use the fraction of candidate votes as the denominator. When the third candidate has the same orientation as one of the top two, her probability to drop out does not vary when the vote shares of the top two candidates are close, but she does drop out more often when the difference between their strengths is small: the third candidate always drops out of the race when the gap is smaller than 10 or 5 percentage points (Table G2). When the third candidate has a different orientation, neither the gap in vote shares nor in strengths between the top two affects her decision to drop out: estimates across the four subsamples are comprised between 10.4 and 11.6 percentage points (Table G3, columns 2 to 5), compared with 13.2 for all elections of sample 1 (Table G3, column 1).

Finally, we test whether the probability that the third candidate drops out is affected by her distance with the top two candidates.

As shown in Table G4, when the third candidate has the same orientation as one of the top two, she is even more likely to drop out when she lags far behind the candidate ranked second in the first round: she drops out of the race in 95.8 percent of the cases when the gap between their vote shares is larger than 5 percentage points (column 2) and in all elections

when it is larger than 10 percentage points (column 3). Instead, when the third candidate has a different orientation, her probability to drop out is not higher when she has very low chances of becoming a front-runner in the second round: she drops out in 11.7 percent (resp. 10.6 percent) of the cases when the gap between her strength and the strength of each of the top two candidates is larger than 5 (resp. 10) percentage points, compared with 13.2 for all elections of sample 1 (Table G5).

Figure G1: Probability that the third candidate drops out depending on her political orientation



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

Table G1: Probability that the third candidate drops out depending on her political orientation

Outcome	(1)	(2)	(3)
	Full Sample	3rd candidate drops out	
		3rd same orientation	3rd different orientation
3rd qualifies	0.448*** (0.042)	0.911*** (0.026)	0.148*** (0.035)
Robust p-value	0.000	0.000	0.000
Observations	1,541	1,027	1,169
Polyn. order	1	1	1
Bandwidth	0.017	0.036	0.019
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.00	0.00	0.00

Notes: Column 1 includes all elections. Column 2 includes only the elections where the third candidate has the same political orientation as one of the top two candidates. Column 3 includes only the elections where the third candidate has a different orientation than both top two candidates. Standard errors are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. Each column reports the results from a separate local polynomial regression. The outcome is a dummy equal to 1 if the third candidate drops out of the race in the second round. The dependent variable is a dummy equal to 1 if the third candidate is qualified. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table G2: Probability that the third candidate drops out depending on the closeness of the race, when she has the same orientation as one top-two

Outcome	(1)	(2)	(3)	(4)	(5)
	3rd candidate drops out - same orientation				
	Closest Identified	Distance12 (vote share) $\leq 10\%$	Distance12 (vote share) $\leq 5\%$	Distance12 (strength) $\leq 10\%$	Distance12 (strength) $\leq 5\%$
3rd qualifies	0.914*** (0.025)	0.866*** (0.041)	0.880*** (0.051)	1.004*** (0.003)	1.012*** (0.011)
Robust p-value	0.000	0.000	0.000	0.000	0.000
Observations	966	671	402	384	185
Polyn. order	1	1	1	1	1
Bandwidth	0.034	0.043	0.045	0.040	0.040
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD
Mean, left of the threshold	0.00	0.00	0.00	0.00	0.00

Notes: Column 1 includes all elections where the third candidate has the same political orientation as one of the top two candidates and where the closest candidate is clearly identified (we exclude the elections where the three candidates have the same orientation and elections where one of the top two candidates has a non-classified orientation). Columns 2 and 3 include only elections where the difference in vote shares between the first and second candidates in the first round is lower than 10 and 5 percentage points, respectively. Columns 4 and 5 include only elections where the difference in strengths between the first and the second candidates is lower than 10 and 5 percentage points, respectively. In column 5, due to the small sample size, the command rdrobust could not compute the optimal bandwidth, and we used the same bandwidth as in column 4. Other notes as in Table G1.

Table G3: Probability that the third candidate drops out depending on the closeness of the race, when she has a different orientation than both top two

Outcome	(1)	(2)	(3)	(4)	(5)
	3rd candidate drops out - different orientation				
	Sample 1	Distance12 (vote share) $\leq 10\%$	Distance12 (vote share) $\leq 5\%$	Distance12 (strength) $\leq 10\%$	Distance12 (strength) $\leq 5\%$
3rd qualifies	0.132*** (0.040)	0.116** (0.051)	0.106** (0.058)	0.104** (0.044)	0.104 (0.082)
Robust p-value	0.002	0.023	0.047	0.041	0.220
Observations	518	345	215	388	196
Polyn. order	1	1	1	1	1
Bandwidth	0.013	0.012	0.013	0.018	0.016
Band. method	MSERD	MSERD	MSERD	MSERD	MSERD
Mean, left of the threshold	0.00	0.00	0.00	0.00	0.00

Notes: Column 1 includes only the elections of sample 1: all elections where the three candidates have distinct political orientations and the third candidate is either on the left or on the right of the two other candidates, making one of them the candidate ideologically closest to the third. Columns 2 and 3 include only elections where the difference in vote shares between the first and second candidates in the first round is lower than 10 and 5 percentage points, respectively. Columns 4 and 5 include only elections where the difference in strengths between the first and the second candidates is lower than 10 and 5 percentage points, respectively. Other notes as in Table G1.

Table G4: Probability that the third candidate drops out depending on her distance with the top two candidates, when she has the same orientation as one top-two

Outcome	(1)	(2)	(3)
	3rd candidate drops out - same orientation		
	Closest identified	Distance top two ≥ 5 pp	Distance top two ≥ 10 pp
3rd qualifies	0.914*** (0.025)	0.958*** (0.033)	1.000 (0.000)
Robust p-value	0.000	0.000	-
Observations	966	460	241
Polyn. order	1	1	1
Bandwidth	0.034	0.029	0.029
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.00	0.00	0.00

Notes: Column 2 (resp. 3) further restricts the sample to elections where the third candidate's vote share in the first round is lower than the vote share of the second candidate by at least 5 percentage points (resp. 10 percentage points). In column 3, due to the small sample size, the command rdrobust could not compute the optimal bandwidth, and we used the same bandwidth as in column 2. Other notes as in Table G2.

Table G5: Probability that the third candidate drops out depending on her distance with the top two candidates, when she has a different orientation than both top two

Outcome	(1)	(2)	(3)
	3rd candidate drops out - different orientation		
	Sample 1	Distance top two ≥ 5 pp	Distance top two ≥ 10 pp
3rd qualifies	0.132*** (0.040)	0.117** (0.053)	0.106 (0.089)
Robust p-value	0.002	0.045	0.503
Observations	518	433	102
Polyn. order	1	1	1
Bandwidth	0.013	0.018	0.007
Band. method	MSERD	MSERD	MSERD
Mean, left of the threshold	0.00	0.00	0.00

Notes: Column 2 (resp. 3) further restricts the sample to elections where the third candidate's strength is lower than that of each of the top two candidates by at least 5 percentage points (resp. 10 percentage points). Other notes as in Table G3.

Appendix H. Third candidate dropouts - Descriptive evidence from press articles

Using Factiva’s research tool (<https://www.dowjones.com/products/factiva>), we collected all press articles released between the two rounds of all elections in our sample and containing the entity “désist”. This entity is present in all forms of the verb “se désister” (to drop out) and in the noun “désistement” (dropout). We obtained a total of 1,678 articles published in 86 different newspapers in election years 1997, 2002, 2007, 2011, 2012, and 2015. Table H1 gives the breakdown of the articles collected by election type and year.

Table H1: Press articles collected by election type and year

	Year	Number of articles
Parliamentary elections	1997	11
	2002	26
	2007	240
	2012	631
	Total	908
Local elections	2011	263
	2015	507
	Total	770
Total		1,678

We read each article and kept only the articles providing information on third candidates’ decision to drop out. We discarded articles covering elections where the third candidate eventually stayed in the race, articles covering second candidates’ dropouts, articles reporting a dropout without giving any information on it, and articles commenting dropouts that occurred in past elections.

We are left with a total of 590 instances of third candidates dropping out. Note that an instance may be covered by several articles and that one article may cover several instances.

As shown in Table H2, for each instance, we first coded whether the article discusses dropouts at the national level (15.6 percent of the cases), or instead focuses on a particular département (8.6 percent of the cases) or a particular district (75.8 percent of the cases). Next, we classified each instance depending on whether the third candidate has the same orientation as one of the top two candidates (39.5 percent of the cases) or a different one (53.7 percent of the cases)²⁰. Dropouts from candidates who have a different orientation than both top two are

²⁰Forty instances are left unclassified (6.8 percent). In these cases, the dropout decision was made by the party at the national or département level without stating the exact configuration in which the third candidate was required to drop out, making it impossible to know whether it led to dropouts in elections where the third had the same orientation as one top-two or not. For instance, some articles report that left parties asked their candidates to

over-covered by the press, as they represent less than 15 percent of all dropouts in our sample but more than 50 percent of the instances covered by the press.

Table H2: Descriptive statistics

	Number of instances	Percentage
<i>Level</i>		
National	92	15.6
Département	51	8.6
District	447	75.8
Total	590	100
<i>Configuration</i>		
Same orientation	233	39.5
Different orientation	317	53.7
Unspecified	40	6.8
Total	590	100

Next, for each instance, we coded the context in which the dropout took place (decision made by the party, existence of an agreement among parties, or decision made individually by the candidate), the reasons provided by the party or candidate (preventing the victory of another candidate, or feeling ideologically close to a top-two), and whether the article mentions the reactions of the candidate’s party, voters, or competing candidates. Table H3 gives the statistics for the whole sample and separately for instances where the third candidate has the same orientation as one top-two or a different one. Note that the 40 instances for which the configuration is unspecified are included in the whole sample (first line in the tables below) but not in the breakdown by political orientations (second and third lines in the tables).

drop out if ranked third and if a far-right candidate was among the top two, but without giving any information on the orientation of the other top-two candidate.

Table H3: Statistics on the dropouts' context, reasons, and reactions

	<i>A. Context</i>			
	(1)	(2)	(3)	(4)
	Presence	Party's decision	Agreement	Candidate's decision
All	47.8	29.8	15.4	18.0
Same orientation	44.2	41.6	33.9	2.6
Different orientation	44.2	12.9	3.8	31.2

Notes: Column 1 gives the percentage of instances for which the article provides information on the context in which the dropout took place. Column 2 gives the percentage of dropouts decided by the party, and column 3 gives the percentage of dropouts that are part of an agreement among parties. All dropouts part of an agreement are also considered as dropouts decided by the party. Column 4 gives the percentage of dropouts decided individually by the candidate, independently from any party's instructions.

	<i>B. Reasons</i>			
	(1)	(2)	(3)	(4)
	Presence	Prevent victory	Ideological proximity	Other reasons
All	59.0	51.7	6.3	5.8
Same orientation	44.2	29.2	15.9	7.7
Different orientation	65.6	63.1	0.0	4.7

Notes: Column 1 gives the percentage of instances for which the article reports at least one reason provided by the candidate or party to justify the decision to drop out. Column 2 (resp. 3) gives the percentage of instances where the article reports that the candidate dropped out to prevent the victory of another candidate (resp. because the candidate or party felt ideologically close to a top-two). Column 4 gives the percentage of instances where the article reports another reason. It includes all reasons that concern less than 5 percent of the instances each.

	<i>C. Reactions</i>				
	(1)	(2)	(3)	(4)	(5)
	Presence	Exclusion	Voters critics	Furthest cand critics	Other reactions
All	10.7	5.1	0.5	4.1	5.4
Same orientation	2.2	0.0	0.4	0.9	1.7
Different orientation	18.3	9.5	0.6	6.9	8.8

Notes: Column 1 gives the percentage of instances for which the article reports at least one reaction following the dropout. Column 2 gives the percentage of cases where the party decided to exclude the candidate because she dropped out against the party's instructions. Column 3 (resp. 4) gives the percentage of cases where voters (resp. the furthest candidate among the top two) criticized the dropout decision. Column 5 gives the percentage of instances where the article reports another reaction. It includes all other reactions that concern less than 5 percent of the instances each, except voters' critics, as we discuss this statistics in the main text (see Section 5.2).

Appendix I. Political orientations

See the additional Supplement, included in the replication material.

Chapter 2

Coordination and Bandwagon Effects of Candidate Rankings: Evidence from Runoff Elections¹ (with Vincent Pons)

Abstract

To predict others' behavior and make their own choices, voters and candidates can rely on information provided by polls and past election results. We isolate the impact of candidates' rankings using an RDD in French local and parliamentary two-round elections, where up to 3 or 4 candidates can qualify for the second round. Candidates who barely ranked first in the first round are more likely to run in the second round (5.6pp), win (5.8pp), and win conditionally on running (2.9 to 5.9pp), than those who barely ranked second. The effects are even larger for ranking second instead of third (23.5, 9.9, and 6.9 to 12.2pp), and also present for ranking third instead of fourth (14.6, 2.2, and 3.0 to 5.0pp). The impact of rankings is largest when the candidates have the same political orientation (making coordination relatively more important and desirable), but it remains strong when two candidates only qualify for the second round (and there is no need for coordination). Overall, our evidence suggests that both coordination and bandwagon effects are important drivers of the behavior of candidates and voters and of election outcomes.

¹For suggestions that have improved this article, we are grateful to Alberto Alesina, Abhijit Banerjee, Laurent Bouton, Pierre Boyer, Enrico Cantoni, Gary Cox, Rafael Di Tella, Allan Drazen, Andrew Eggers, Jeffrey Frieden, Olivier Gossner, Benjamin Marx, Shom Mazumder, Benjamin Olken, Thomas Piketty, James Snyder, Matthew Weinzierl, as well as seminar participants at New York University, Stanford University, Paris School of Economics, University of Maryland, Texas A&M, Rice University, University of Bologna, Ecole Polytechnique, CREST, European University Institute, and conference participants at Economics and Politics Workshop, the International Conference on Mathematical Optimization for Fair Social Decisions, APSA, and EEA. We thank Sebastian Calonico, Matias Cattaneo, Max Farrell, and Rocio Titiunik for guiding us through the use of their RDD Stata package "rdrobust" and for sharing their upgrades; Nicolas Sauger for his help with the collection of the 1958, 1962, 1967, 1968, 1973, 1981 and 1988 French parliamentary election results; Abel François for sharing his data on 1993, 1997, and 2002 candidates' campaign expenditures; Paul-Adrien Hyppolite for the excellent research assistance he provided at the onset of the project; and Eric Dubois for his outstanding work of data entry and cleaning.

1 Introduction

In elections, some voters decide whether to participate or abstain and which candidate to vote for based only on their own preferences (e.g., Pons and Tricaud, 2018; Spenkuch, 2018). Others also take into account their expectations about the behavior of the rest of the electorate. For instance, voters might strategically decide to shift their support away from their preferred candidate towards a candidate that they like less but expect to have higher chances to be in contention for victory (e.g., Duverger, 1954; Myerson and Weber, 1993; Cox, 1997). Similarly, candidates can decide whether or not to enter the race based on the fraction of voters they expect to vote for them and for their competitors. Those who foresee that they will receive few votes or that their presence might divide the votes of their camp and lead to the victory of a candidate that is more distant ideologically might choose to stay out of the race. To form their expectations about others' preferences and strategies and make their own decisions, voters and candidates can rely on information provided by polls and past election results. Despite large evidence that political actors' overall level of information matters (e.g., Hall and Snyder, 2015; Le Pennec and Pons, 2019), little is known about which exact pieces of information they actually base their decisions on.

We focus on one specific piece of information, candidate rankings: the ordering of candidates in polls, in previous elections, or in the previous round (in elections with multiple rounds). While past and predicted vote shares provide detailed information on the distribution of preferences, rough-hewn rankings can serve as a coordination device in and of themselves. When more than two candidates are in the running, their past rankings can be used by strategic voters as a focal point to coordinate on the same subset of candidates and on the same equilibrium in a decentralized way, and by sister parties to determine which of their candidates should drop out. These coordination mechanisms can be reinforced by behavioral motives such as a bandwagon effect. Voters who desire to vote for the winner might decide to “jump on the bandwagon” and rally behind higher-ranked candidates if they derive intrinsic value from voting for candidates who won or had a higher rank in the past or if they rightly anticipate that these candidates are more likely to win in the future.

In this paper, we estimate the impact of rankings on voters and candidates' decisions and examine whether this impact reflects strategic coordination or bandwagon effects. To isolate the effect of rankings from the effect of past vote shares, we use a regression discontinuity design (RDD) and compare the likelihood of running, the likelihood of winning, and the vote share obtained by candidates who had previously received close-to-identical numbers of votes but ranked just below or just above one another.

We implement this strategy in French local and parliamentary elections from 1958 to 2017. In these elections, the winner is designated by a two-round plurality voting rule, and up to three or four first-round candidates can qualify for the second round. This enables us to measure the effect on second-round outcomes of ranking first in the first round (instead of second), second

(instead of third), and third (instead of fourth), and to test for impact heterogeneity by the number and type of other candidates qualified for the second round. In addition, all candidates qualified for the second round can decide to drop out of the race. We can thus estimate the impact of first-round rankings both on voters' choice of candidate in the second round and on candidates' decision to run.

We first run an RDD on the vote share difference between the top two candidates in the first round and find that ranking first instead of second increases candidates' likelihood to stay in the race by 5.6 percentage points and their likelihood to win by 5.8 percentage points. Arriving second instead of third has even larger effects, of respectively 23.5 and 9.9 percentage points. These effects are measured in races in which both the higher-ranked and lower-ranked candidates qualified for the second round, and they are all significant at the 1 percent level. Arriving third instead of fourth increases candidates' likelihood to run in the second round by 14.6 percentage points and their likelihood to win by 2.2 percentage points. These effects are significant at the 1 and 10 percent level respectively.

The overall effects on winning can be driven both by the effect on running (candidates' decision to stay in the race) and by an effect on vote shares and winning conditional on running (if voters rally behind higher-ranked candidates). Naturally, the anticipation of the latter effects may contribute to explain the effects on candidates' behavior. We cannot directly estimate the effects conditional on running by focusing on elections in which both the lower-ranked and higher-ranked candidates decide to enter the second round, as those who do are not randomly chosen. We derive bounds to deal with this selection issue. We find that arriving first instead of second increases candidates' vote share and likelihood of winning conditional on running by 1.3 to 4.0 percentage points and 2.9 to 5.9 percentage points respectively. The effects of ranking second instead of third (resp. third instead of fourth) are 4.0 to 14.7 and 6.9 to 12.2 percentage points (resp. 2.5 to 10.0 and 3.0 to 5.0 percentage points).

Second, we disentangle the mechanisms responsible for these effects. We show that they are unlikely to be driven by differences in campaign expenditures, press coverage, or other qualified candidates' decision to stay in the race or drop out. In addition, we find that the effects are much larger when the higher- and lower-ranked candidates have the same political orientation. For instance, ranking first instead of second increases the likelihood of running by 35.2 percentage points when the first and second candidates have the same orientation, due to alliances among their parties, and the likelihood of winning by 30.4 percentage points, compared to only 0.1 and 1.7 percentage points when they have distinct orientations. This can come from the fact that shared orientation makes voters' and candidates' coordination against ideologically distant candidates more desirable, but also from the fact that it makes rallying behind the higher-ranked candidate less costly, whatever the underlying motive.

To investigate the extent to which coordination explains our results, we focus on elections in which three candidates or more qualified for the second round (and rankings can be used to coordinate on a subset of them) and compare the effects of ranking first instead of second

depending on the challenge posed by the third candidate. We find that the effects on running and winning decrease with the gap between the second and third candidates' vote shares, suggesting that coordination (which is more critical when the gap is lower) explains part of the effects. In addition, the effects of ranking first are larger when the ideological distance between the top two candidates is lower than their distance with the third candidate (making coordination between the top two more desirable).

To test whether strategic coordination suffices to explain our results, we turn to elections in which the third candidate is *not* qualified. In these elections, there is no room or need for coordination: all voters should vote for their preferred candidate among the top two, and candidates do not risk contributing to the victory of a disliked competitor by running. Nonetheless, we find that the effects of ranking first instead of second remain large: it increases candidates' probability of running in the second round by 1.8 percentage points and their vote share and probability of winning conditional on running by 1.0 to 1.9 percentage points and 4.9 to 5.8 percentage points respectively. The overall effect of ranking first on winning when the third candidate is not qualified is 5.8 percentage points, which is very close to the effect in the full sample. These results indicate, first, that dropout agreements are not only driven by the desire to avoid the victory of a third candidate, second, that the desire to vote for the winner is an important driver of voter behavior, and third, that bandwagon effects sway many elections.

Our analysis builds on previous studies estimating the impact of candidate rankings across elections. Following Lee (2008), a large literature has examined the impact of ranking first (instead of second) on future elections and shown that close winners generally benefit from an incumbency advantage when they run again (e.g., Eggers et al., 2015). These studies have been unable to distinguish the effect of holding office from the pure effect of being labeled first, which we can instead isolate by studying two rounds of the same election. Complementing this work on the effect of ranking first, Anagol and Fujiwara (2016) show that arriving second (instead of third) in past elections also increases candidates' likelihood to run in the next elections and win them, and they attribute these effects to strategic coordination by voters. Our paper completes these studies in four important ways.

First, we can estimate the effect of arriving second or third, as well as the effect of arriving first (independently from incumbency advantage).

Second, the short time span (one week) between the first and second rounds helps us isolating the direct effect of rankings from reinforcing mechanisms which are more likely to matter when considering elections separated by several years, such as increased notoriety of the higher-ranked candidates and differential likelihood to be replaced by another candidate of their party.

Third, coordination between parties can lead them to ask candidates to stay out of the race. This important aspect of party coordination is difficult to study in general, when we only observe candidates who are actually competing, not those who considered entering but eventually stayed out. Instead, we observe the full set of candidates eligible to compete in the runoff,

whether or not they actually stay in the race. Indeed, qualification to the second round is entirely determined by first round results. This enables us to cleanly estimate and characterize the contribution of party coordination to the effects of rankings. We find evidence that coordination between parties of similar orientation is motivated by the desire to avoid the victory of a candidate of a different orientation as well as other motives, such as enforcing a deliberative ideal among allied parties and their supporters. These results echo prior work showing that parties favor candidates ranked higher by voters: Folke et al. (2016) and Meriläinen and Tukiainen (2016) both find that rankings in the number of preference votes obtained by politicians in open-list municipal elections affect their chances of promotion within their party.

Fourth, we can isolate the contribution of the bandwagon effect to the impact of rankings by focusing on second rounds in which only two candidates qualified and there is thus no need or even room for strategic coordination. Most existing measures of the bandwagon effect are based on survey data. Bartels (1985) finds that a large fraction of voters in U.S. primary elections intend to vote for the candidate they deem the most likely to win the nomination, but acknowledges that this correlation may also capture the effect of people's preferences on their assessment of candidates' chances. Wright (1993) and Atkeson (1999) find systematic overreporting voting for the winner in post-electoral surveys, suggesting that voters have a strong desire to be on the winning side. While the design of these studies addresses the concern of reverse causality, their result is consistent with alternative interpretations. For instance, it may capture an actual change in people's beliefs on the quality of the winning candidate which began before the election and continued afterwards, irrespective of any desire to be in the winning camp. Alternatively, when the winner is a non-mainstream candidate, the perceived acceptability of reporting voting for her may increase when it is known that she won even if voters' true likelihood to vote for her is unchanged. These alternative interpretations are less likely to account for the findings of Morton et al. (2015), who compare electoral results in French territories overseas between elections in which these territories voted before or after the overall election outcome had been made public. A limitation of this original natural experiment is that the voting reform took place in all overseas territories at once, making it difficult to disentangle its effect from concomitant factors.

We build on this preexisting work and bring causal evidence on the existence and size of the bandwagon effect using administrative electoral results of a large number of individual races.¹ The positive effects we find on vote shares and winning are in line with the predictions of models assuming that voters gain utility from voting for the winner (Hinich, 1981; Callander, 2007). The bandwagon effect of candidate rankings is akin to the effects measured in other contexts of asset rankings on trading behavior (Hartzmark, 2015), hospitals' rankings on their number

¹Similarly to our setting, Kiss and Simonovits (2014) study bandwagon effects in two-round elections in Hungary. Differently from our strategy, they compare the size of the difference between the first and second candidates' vote shares in the first and second rounds. They interpret the increase in the winning margin as evidence of a bandwagon effect of first round results on second round vote choices, but other differences between the first and second rounds than the availability of first round results could drive this pattern.

of patients and revenues (Pope, 2009), and employees' rankings on their sales performance (Barankay, 2018).

Methodologically, we draw on other studies that exploit vote-share thresholds to estimate causal effects of interest (for a review, see de la Cuesta and Imai, 2016). Our setting is closest to Pons and Tricaud (2018), which uses French runoffs to measure the impact of the qualification of the third candidate in the second round, with three important differences. First and foremost, that previous study assessed the extent to which voters behave expressively or strategically, as well as the consequences for electoral outcomes. Instead, the present study asks which information voters (and candidates) who want to behave strategically use to predict others' behavior, and it shows that rankings facilitate decentralized coordination by serving as focal points. Second, while Pons and Tricaud (2018) use district-level specifications, our analysis is conducted at the candidate level and uses two observations per district, corresponding to the higher-ranked and lower-ranked candidates. Third, the present study includes a larger number of elections as we do not need to restrict the analysis to parliamentary and local elections using the 12.5 percent qualification threshold.

The remainder of the paper is organized as follows. We provide more details on our setting and empirical strategy in Section 2. Section 3 presents our empirical results and Section 4 discusses the underlying mechanisms. Section 5 concludes.

2 Empirical strategy

2.1 Setting

Our sample includes parliamentary and local elections. Parliamentary elections elect the representatives of the French National Assembly, the lower house of the Parliament. France is divided into 577 constituencies, each of which elects a Member of Parliament every five years. Local elections determine the members of the departmental councils. France is divided into 101 départements, which have authority over education, social assistance, transportation, housing, culture, local development, and tourism. Each département is further divided into small constituencies, the cantons, which elect members of the departmental councils for a length of six years. All French territories participate in local elections, except for Paris and Lyon (where the departmental council is elected during municipal elections) and some French territories overseas. Until an electoral reform in 2013, each canton elected one departmental council member. Local elections took place every three years and, in each département, only half of the cantons were electing their council member in a given election. After the reform, all cantons participated in elections held every six years and each canton elected a ticket composed of a man and a woman. The reform further reduced the number of cantons from 4035 to 2054, to leave the total number of council members roughly unchanged. This new rule applied to the 2015 local elections, which are included in our sample. We consider a ticket as a single candidate in our

analysis, since the two candidates organize a common electoral campaign, run in the election under the same ticket, and get elected or defeated together.

Both parliamentary and local elections are held under a two-round plurality voting rule. In order to win directly in the first round, a candidate needs to obtain a number of votes greater than 50 percent of the candidate votes and 25 percent of the registered citizens. In the vast majority of races, no candidate wins in the first round, and a second round takes place one week later. In the second round, the election is decided by simple plurality: the candidate who receives the largest vote share in the second round wins the election.

The two candidates who obtain the highest vote share in the first round automatically qualify for the second round. Other candidates qualify if they obtain a first-round vote share higher than a predetermined threshold. This rule is essential for our study design, as it enables us to estimate the impact of ranking first instead of second (using all races in which no candidate wins in the first round) as well as the impact of ranking second instead of third, and third instead of fourth (using all races in which three or four candidates qualified for the second round).

Importantly, all candidates qualified for the second round can decide to drop out of the race between rounds. This allows us to estimate the impact of first-round rankings both on voters' choice of candidate in the second round and on candidates' decision to stay in the second round.

The qualification threshold changed over time: in local elections, the required vote share was 10 percent of the registered citizens until 2010, when the threshold was increased to 12.5 percent.² In parliamentary elections, the required vote share was 5 percent of the voters in 1958 and 1962, it was changed to 10 percent of the registered citizens in 1966, and 12.5 percent of the registered citizens in 1976.

Our sample includes 14 parliamentary elections and 12 local elections: all parliamentary elections of the Fifth Republic from 1958 to 2017 except for the 1986 election (which used proportionality rule), and all local elections from 1979 to 2015. We do not include local elections held before 1979 as the electoral rule allowed any candidate to run in the second round, irrespective of their vote share in the first round and even if they were absent from the first round.³

2.2 Data

Our sample includes a total of 22,556 races: 16,221 (71.9 percent) from local elections and 6,335 (28.1 percent) from parliamentary elections. Official results of local and parliamentary elections were digitized from printed booklets for the 1958 to 1988 parliamentary elections and for the 1979 to 1988 local elections, and obtained from the French Ministry of the Interior for the 1993 to 2017 parliamentary elections and for the 1992 to 2015 local elections. We exclude

²In the 2011 local elections, the threshold remained at 10 percent in the 9 cantons belonging to Mayotte département.

³Each of the 26 elections we consider took place at a different date. In 1988, both parliamentary and local elections were held, but in different months. The 2001 and 2008 elections took place at the same date as mayoral elections. Our results are robust to excluding these two elections.

racess where only one round took place and those with only one candidate in the first round. Table 1 gives the breakdown of the sample data by election type and year.

Table 1: Number of races by election type and year

Election type	Year	Nb of races	Election type	Year	Nb of races	
Parliamentary elections	1958	433	Local elections	1979	1,086	
	1962	374		1982	1,061	
	1967	405		1985	1,230	
	1968	319		1988	1,177	
	1973	430		1992	1,425	
	1978	423		1994	1,369	
	1981	334		1998	1,513	
	1988	455		2001	1,301	
	1993	497		2004	1,516	
	1997	565		2008	1,074	
	2002	519		2011	1,564	
	2007	467		2015	1,905	
	2012	541				
	2017	573				
	Total	6,335		Total	16,221	
	Total	22,556				

To measure the impact of ranking first instead of second (henceforth “1vs2”), we further exclude races in which two of the top three candidates obtained an identical number of votes in the first round (sample 1).⁴ Indeed, we do not have any way to choose which candidate to treat as first, when the top two candidates obtained the same number of votes, and which candidate to compare to the first, when the two candidates ranked below her obtained the same number of votes. To measure the impact of ranking second instead of third (henceforth “2vs3”), we restrict our sample to races where three candidates at least competed in the first round and the third candidate qualified for the second round, and we exclude races in which two of the top four candidates obtained an identical number of votes in the first round (sample 2). To measure the impact of ranking third instead of fourth (henceforth “3vs4”), we restrict our sample to races where four candidates at least competed in the first round and the third and fourth candidates qualified for the second round, and we exclude races in which two candidates among the second, third, fourth and fifth obtained an identical number of votes in the first round (sample 3).

Thanks to the large set of elections we consider and the large number of races in each election, we have a large number of close races: the vote share difference between the candidates ranked first and second (resp. second and third, third and fourth) is lower than 2 percentage

⁴By “two of the top three candidates”, we mean two of the top two candidates (i.e. the top two) if two candidates only competed in the first round, and two of the top three candidates if three candidates or more competed in the first round. The same applies to the next restrictions.

points in 2,581 races in sample 1, 1,874 races in sample 2, and 758 races in sample 3.

Table 2 presents some descriptive statistics on our full sample.

Table 2: Summary statistics

	Mean	Sd	Min	Max	Obs.
<i>Panel A. 1st round</i>					
Registered voters	28,295	28,157	258	200,205	22,556
Turnout	0.636	0.125	0.094	0.921	22,556
Candidate votes	0.613	0.122	0.093	0.914	22,556
Number of candidates	6.5	3.1	2	48	22,556
<i>Panel B. 2nd round</i>					
Turnout	0.628	0.134	0.117	0.968	22,556
Candidate votes	0.595	0.138	0.103	0.963	22,556
Number of candidates	2.1	0.4	1	6	22,556

In the average race, 63.6 percent of registered citizens turned out in the first round and 61.3 percent cast a valid vote for one of the candidates (henceforth “candidate votes”), as opposed to casting a blank or null vote.⁵ On average, 6.5 candidates competed in the first round. Turnout in the second round was slightly higher than in the first round (62.8 percent on average) but the fraction of candidate votes was slightly lower (59.5 percent on average). The number of candidates competing in the second round ranged from 1 to 6 with an average of 2.1.

Tables A1, A2, and A3 in the Appendix present descriptive statistics on all races in samples 1, 2, and 3 as well as close races between the first and second candidates in sample 1, the second and third in sample 2, and the third and fourth in sample 3. Overall, samples 1, 2, and 3, and close races in these samples are very similar to the full sample. The most noticeable differences are as follows: in sample 1, turnout and the share of candidate votes in the second round were slightly higher in close races. In sample 2, the average number of candidates in the first round was slightly higher in close races but the number of candidates in the second round was almost identical, and turnout and the share of candidate votes were slightly lower in close races in both rounds but the increase between rounds was similar as in the full sample.

All the statistics shown in Tables 2, A1, A2, and A3 are at the race level. Importantly, the analysis below is conducted at the candidate level and uses exactly two observations per race, for the higher-ranked and lower-ranked candidates. We use the political label attributed to the candidates by the French Ministry of the Interior to allocate them to six political orientations: far-left, left, center, right, far-right, and other.⁶

⁵Valid voting for a candidate entails inserting a ballot pre-printed with the candidate’s name in an envelope and putting this envelope in the ballot box. Casting a blank vote means putting an empty envelope in the ballot box, and a null vote writing something on the ballot or inserting multiple ballots in the envelope.

⁶The Ministry of the Interior attributes political labels based on several indicators: candidates’ self-reported political affiliation, party endorsement, past candidacies, public declarations, local press, etc. We mapped political labels into the six political orientations, mainly based on the allocation chosen by Laurent de Boissieu in his blog

2.3 Evaluation framework

We exploit close races to estimate the impact of candidates' first round rankings on their second round outcomes. To measure the impact of ranking 1vs2, we use two observations per race, corresponding to the candidates arrived first and second, and define the running variable X_1 as the difference between each candidate's vote share and the vote share of the other top-two candidate. For the candidate ranked first, the running variable is equal to her vote share minus the vote share of the candidate ranked second, and for the candidate ranked second it is equal to her vote share minus the vote share of the candidate ranked first:

$$X_1 = \begin{cases} \text{voteshare}_1 - \text{voteshare}_2 & \text{if ranked 1st} \\ \text{voteshare}_2 - \text{voteshare}_1 & \text{if ranked 2nd} \end{cases}$$

Similarly, for 2vs3 and 3vs4, we define the running variables X_2 and X_3 as:

$$X_2 = \begin{cases} \text{voteshare}_2 - \text{voteshare}_3 & \text{if ranked 2nd} \\ \text{voteshare}_3 - \text{voteshare}_2 & \text{if ranked 3rd} \end{cases}$$

$$X_3 = \begin{cases} \text{voteshare}_3 - \text{voteshare}_4 & \text{if ranked 3rd} \\ \text{voteshare}_4 - \text{voteshare}_3 & \text{if ranked 4th} \end{cases}$$

We define the treatment variable T as a dummy equal to 1 if the candidate was better ranked in the first round ($X > 0$) and 0 otherwise, and evaluate the impact of higher rank with the following specification:

$$Y_i = \alpha_1 + \tau T_i + \beta_1 X_i + \beta_2 X_i T_i + \mu_i \quad (1)$$

where Y_i is the outcome of interest for candidate i . We run this specification separately for 1vs2, 2vs3, and 3vs4. It estimates the impact of rankings at the limit, when both candidates have an identical vote share. Therefore, it enables us to isolate the impact of ranking from the impact of differential vote share.

Following Imbens and Lemieux (2008) and Calonico et al. (2014), our main specification uses a non-parametric approach, which amounts to fitting two linear regressions on candidates respectively close to the left, and close to the right of the threshold. We test the robustness of our results to a quadratic specification, including X_i^2 and its interaction with T_i as regressors in equation [1].

Since we use two observations per race (one on the left, and one on the right of the threshold), we cluster our standard errors at the race level (district x year level). Our results are left unchanged if we instead cluster at the district level (results available upon request).

“France Politique”: <http://www.france-politique.fr/>. We also used public declarations made by the candidates. Appendix D shows the mapping between labels and political orientations for each election.

Our estimation procedure follows Calonico et al. (2014), which provides robust confidence interval estimators. Our preferred specification uses the MSERD bandwidths developed by Calonico et al. (2018), which reduce potential bias the most. We also test the robustness of the main results to using the optimal bandwidths computed according to Imbens and Kalyanaraman (2012) and to using tighter bandwidths by dividing the MSERD bandwidths by 2.

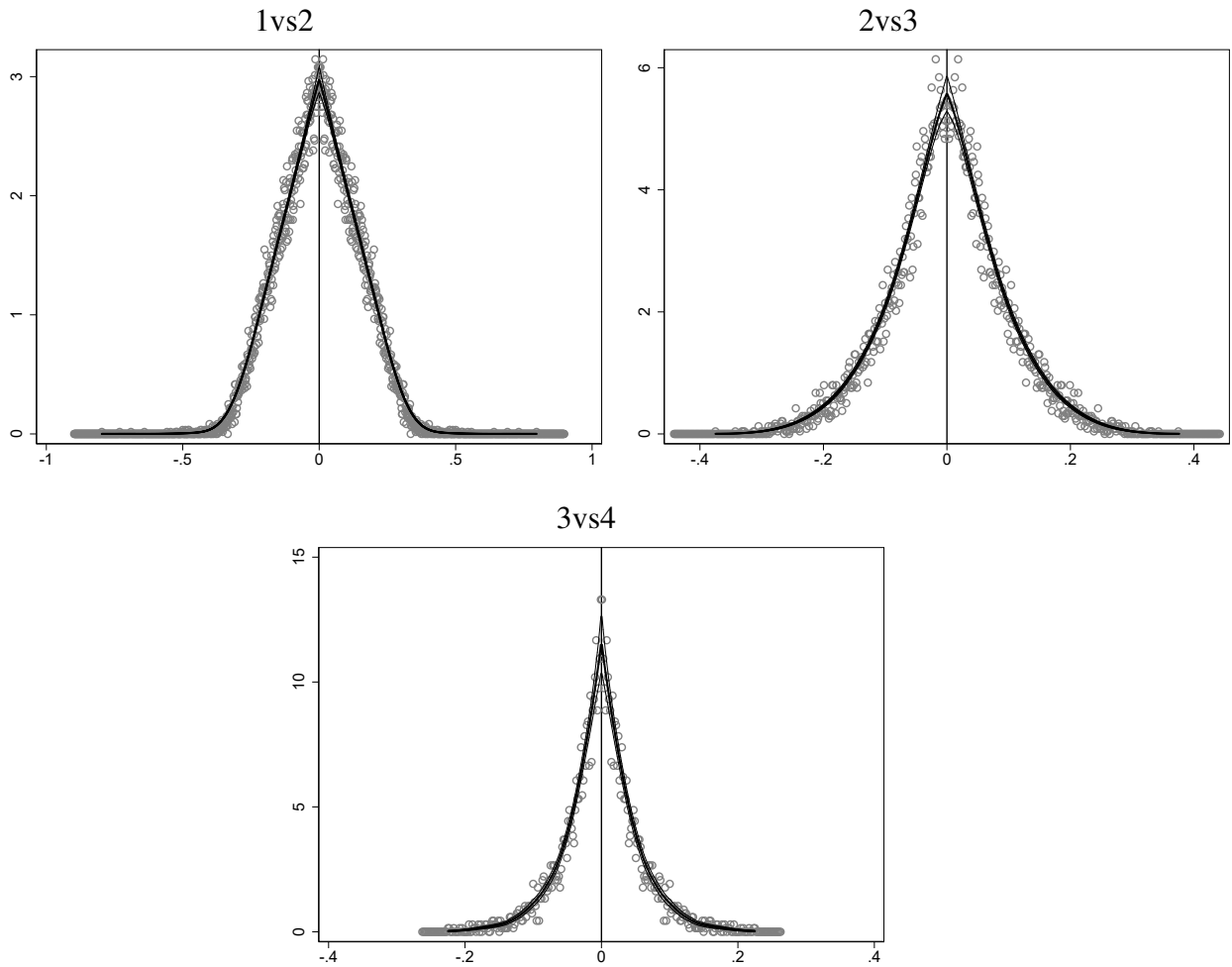
The bandwidths used for the estimations are data-driven and therefore vary depending on the outcomes and samples we consider.

2.4 Identification assumption

The identification assumption is that the distribution of candidate characteristics changes continuously around the threshold, so that the only discrete change occurring at this threshold is the shift in candidates' ranking. Sorting of candidates across the threshold only threatens the validity of this assumption if it occurs at the cutoff, with candidates of a particular type pushed just above or just below the threshold (de la Cuesta and Imai, 2016). Generally, this is unlikely, as it requires the ability to predict election outcomes and deploy campaign resources with extreme accuracy, and given that weather conditions on Election Day and other unpredictable events make the outcome of the election uncertain (Eggers et al., 2015). In our setting, manipulation of the threshold is perhaps even more unlikely than in other RDDs using vote share thresholds as very limited information is available about voters' intentions in the first round of French parliamentary or local races. In particular, district-level polls are very rare during parliamentary elections, and nonexistent during local elections, due to small district size and limited campaign funding.

To bring empirical support for the identification assumption, it is customary for RDDs to check if there is a jump in the density of the running variable at the threshold using a test designed by McCrary (2008). In our setting, this test is satisfied by construction since we consider the same set of races on both sides of the threshold and in each race the higher- and lower-ranked candidate are equally distant to the cutoff (see Figure 1).

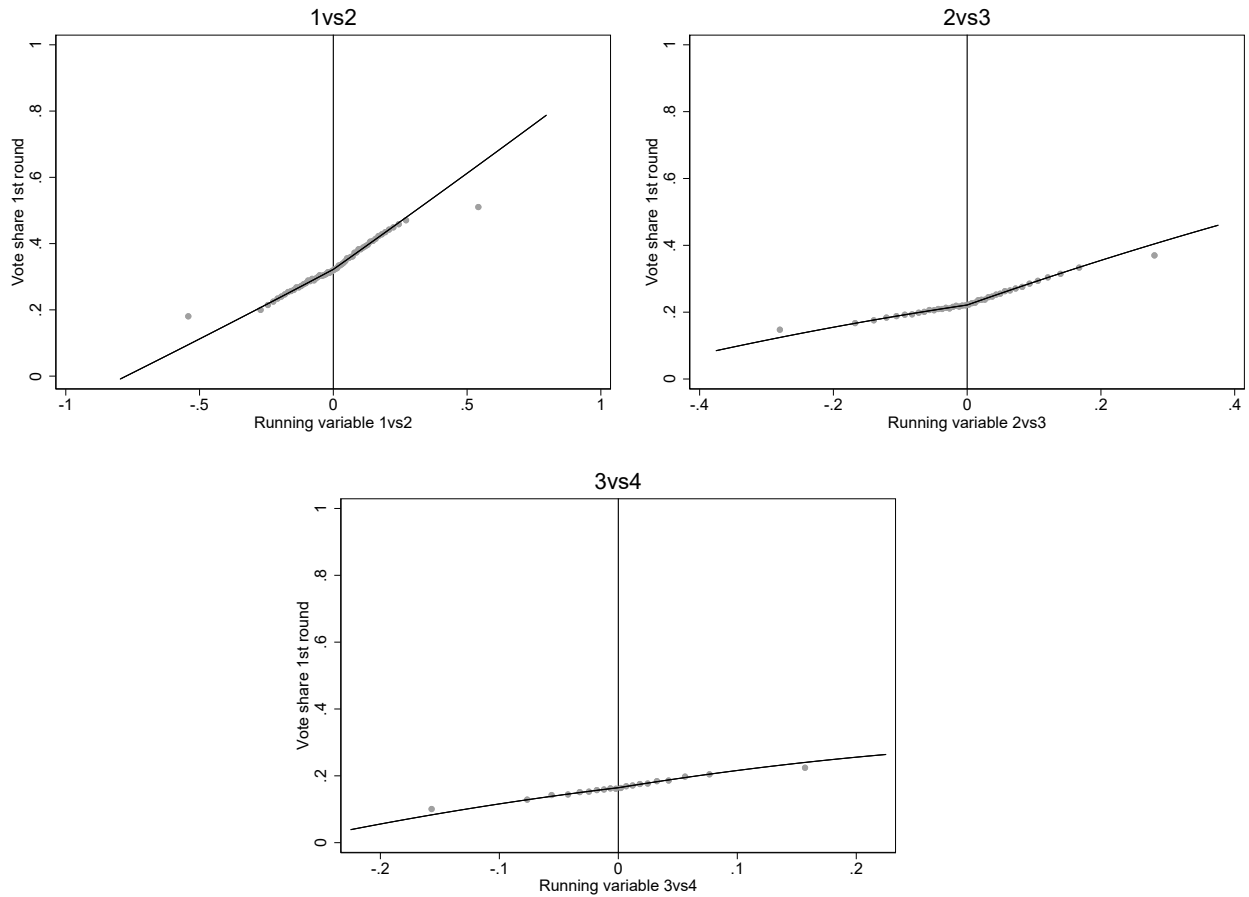
Figure 1: Density of the running variable



Similarly, first round outcomes such as district size, the total number of candidates, voter turnout, or the candidate's vote share are smooth by construction at the threshold. Figure 2 plots the candidate vote share in the first round against the running variable. For this graph as well as for graphs showing the effects of rankings, each dot represents the average value of the outcome within a given bin of the running variable. Observations corresponding to higher-ranked candidates are on the right of the threshold, and those corresponding to lower-ranked candidates on the left. To facilitate visualization, a quadratic polynomial is fitted on each side of the threshold.

We observe that in sample 1, on average, the candidates ranked first and second in the first round received around 30 percent of candidate votes at the threshold. In sample 2 (resp. 3), the first round vote share of candidates ranked second and third (resp. third and fourth) was 20 percent (resp. 18 percent) at the threshold.

Figure 2: Vote share in the first round



Notes: Dots represent the local averages of the candidate’s vote share in the first round (y-axis). Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the vote share difference between the two candidates in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

To bring additional support for the identification assumption, we consider variables whose distribution at the threshold is not mechanically symmetric: the candidate’s gender, whether she ran in the previous election, in the same département and then in the same district,⁷ whether she won a race in the previous election, in the same département and then in the same district, whether she belongs to a political party (as opposed to running as an independent), a set of six dummies indicating her political orientation, whether this orientation is the same as the incumbent’s, the number of other candidates of her orientation who were present in the first round, her strength in the first round, defined as the sum of first-round vote shares of all candidates of the same orientation, and the average strength of her orientation at the national level in the first round.⁸ We conduct the following general test for imbalance. We regress the assignment

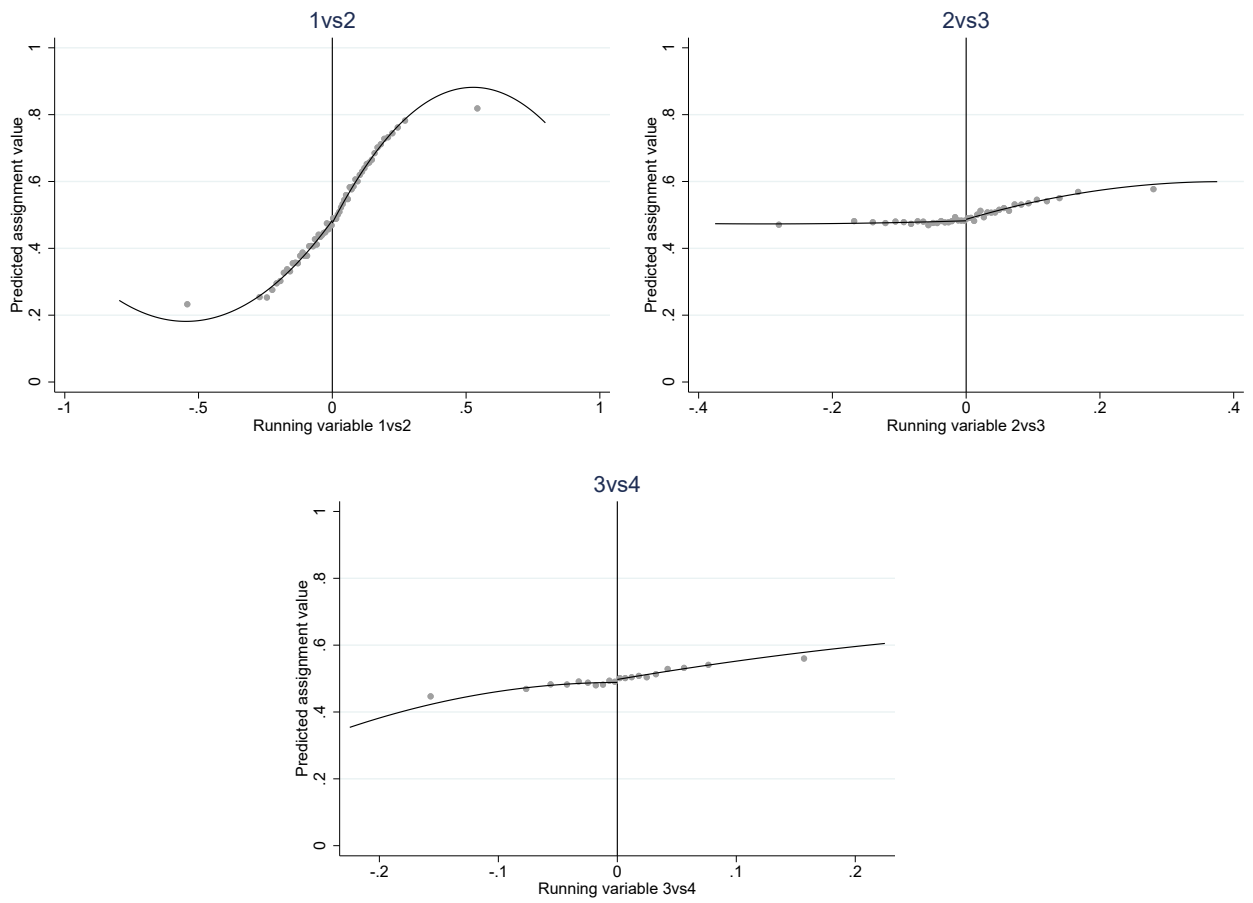
⁷These are two distinct variables because some candidates ran in the previous election in the same département but in a different district.

⁸The average strength of a candidate’s orientation is computed over all districts in which at least one first round candidate has this orientation. It is missing for non-classified candidates.

variable D on these variables, use the coefficients from this regression to predict assignment status for each candidate, and test whether the predicted value jumps at the threshold. To avoid dropping observations, for each regressor, we include a dummy equal to one when the variable is missing and replace missings by 0s. Figure 3 shows the lack of any jump at the cutoff for predicted assignment to first rank (instead of second), second rank (instead of third), and third rank (instead of fourth). As shown in Table 3 the coefficients are small and non-significant.

We also examine whether there is a discontinuity in any of the variables used to predict assignment, taken individually (the corresponding graphs and tables are included in Appendix B). For 1vs2, one coefficient out of 16 is significant: the probability to be on the center. For 2vs3, the coefficient for the probability to be on the left is significant. Finally, for 3vs4, the probability to be on the left and the probability that the candidate won a race in the previous election in the same département are both significant. Overall, one coefficient out of 48 is significant at the 1 percent level, 3 are significant at the 5 percent level, and 4 at the 10 percent level. Since the general balance test shows no discontinuity, we are confident that there is no systematic sorting of candidates at the threshold. In addition, the results shown in the rest of the paper are robust in sign, magnitude, and statistical significance to controlling for baseline variables that are statistically significant.

Figure 3: General balance test



Notes: Dots represent the local averages of the predicted assignment status (y-axis). Other notes as in Figure 2.

Table 3: General balance test

Outcome	(1)	(2)	(3)
	Predicted treatment		
	1vs2	2vs3	3vs4
	(sample 1)	(sample 2)	(sample 3)
Treatment	0.000 (0.006)	0.001 (0.004)	0.008 (0.007)
Robust p-value	0.814	0.978	0.317
Observations left	11,840	5,145	1,199
Observations right	11,840	5,145	1,199
Polyn. order	1	1	1
Bandwidth	0.104	0.065	0.037
Mean, left of threshold	0.477	0.486	0.491

Notes: Standard errors clustered at the race level are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. The unit of observation is the candidate. Each column reports the results from a separate local polynomial regression. The outcome is the value of the treatment predicted by the following baseline variables: the candidate's gender, whether she ran in the previous election, in the same département and then in the same district, whether she won a race in the previous election, in the same département and then in the same district, whether she belongs to a political party (as opposed to running as an independent), a set of six dummies indicating her political orientation, whether this orientation is the same as the incumbent's, the number of candidates of her orientation who were present in the first round, her strength in the first round, and the average strength of her orientation at the national level in the first round. The independent variable is a dummy equal to 1 if the candidate is better ranked in the first round. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1 and the bandwidths are derived under the MSERD procedure. The mean, left of the threshold gives the value of the outcome for the lower ranked candidate at the threshold. It is computed by taking as outcome the probability of winning with all values on the right of the threshold replaced by 0.

3 Empirical results

3.1 Impact on winning

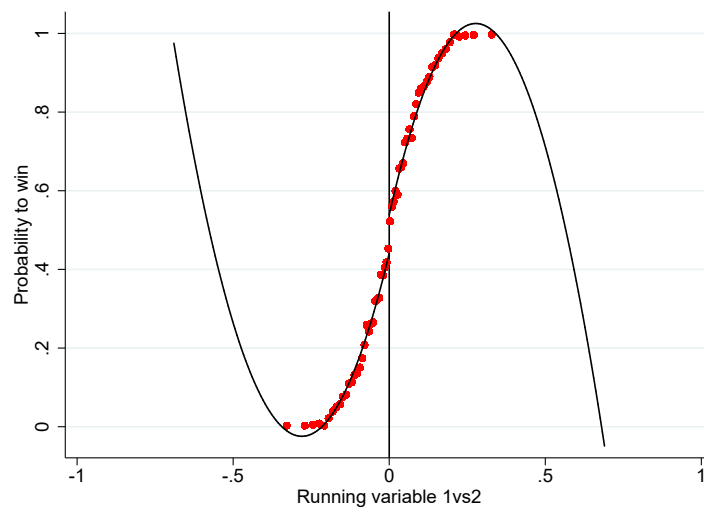
We first measure the impact of candidates' first-round rankings on their unconditional likelihood to win the race: an outcome defined whether the candidate participates in the second round or not, and equal to 1 if the candidate wins, and 0 if she enters the second round and loses or if she drops out between rounds.

Figure 4 plots the likelihood that the first and second candidates win the election against the running variable. We observe a clear discontinuity at the cutoff: ranking 1vs2 in the first round has a large and positive impact on winning. Figure 5 shows an even larger jump for the impact of ranking 2vs3. A jump remains visible on Figure 6, which shows the impact of ranking 3vs4, but it is smaller: very few candidates ranked third and fourth in the first round are in a position to win the second round, limiting the scope for impact.

Table 4 provides the formal estimates of the effects using our preferred specification. On

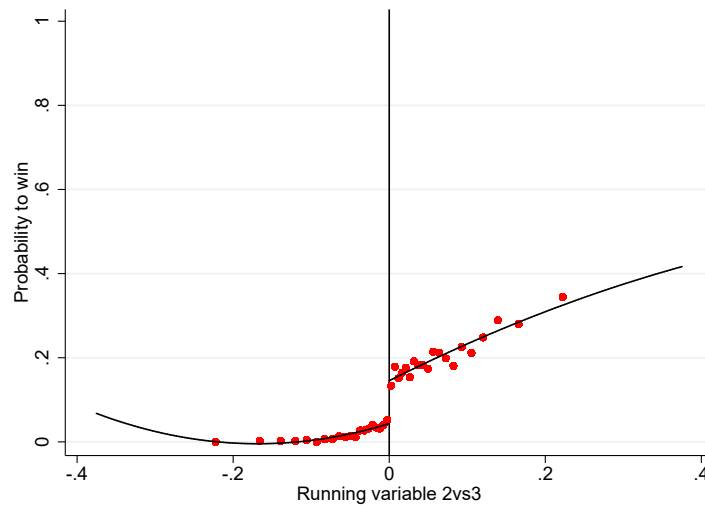
average, ranking 1vs2 in the first round increases the likelihood to win the election by 5.8 percentage points (column 1), which represents a 12.7 percent increase compared to the mean chance of victory of close second candidates at the threshold (45.6 percent). Ranking 2vs3 has an even larger effect, of 9.9 percentage points (column 2): it more than doubles the likelihood of victory of close third candidates (4.8 percent). The effect of ranking 3vs4 is smaller in magnitude (2.2 percentage points, column 3), but it amounts to a three-fold increase compared to the very small fraction of races won by close fourth candidates (0.6 percent). The effects of ranking 1vs2 and 2vs3 are significant at the 1 percent level and the effect of ranking 3vs4 at the 10 percent level.

Figure 4: Impact on winning 1vs2



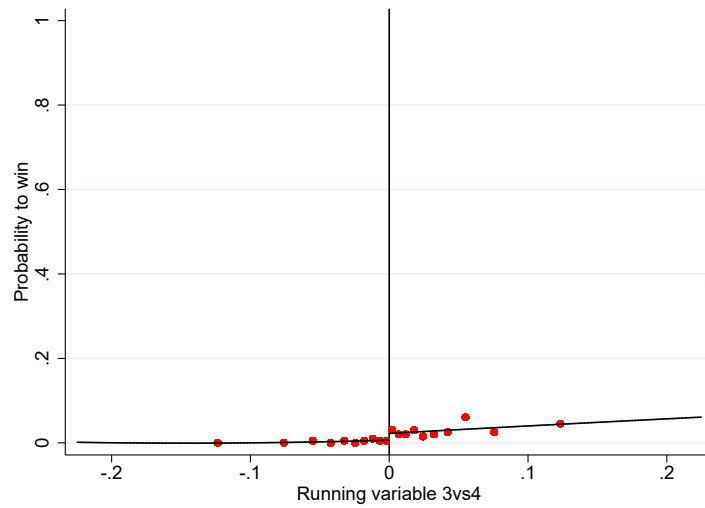
Notes: Dots represent the local averages of the probability that the candidate wins in the second round (y-axis). Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the vote share difference between the two candidates in the first round) is measured as percentage points. The graph is truncated at 70 percent on the x-axis to accommodate for a single outlier. Continuous lines are a quadratic fit.

Figure 5: Impact on winning 2vs3



Dots represent the local averages of the probability that the candidate wins in the second round (y-axis). Averages are calculated within quantile-spaced bins of the running variable (x-axis). The running variable (the vote share difference between the two candidates in the first round) is measured as percentage points. Continuous lines are a quadratic fit.

Figure 6: Impact on winning 3vs4



Notes as in Figure 5.

Table 4: Impact on winning

Outcome	(1)	(2)	(3)
	Probability to win in the 2 nd round		
	1vs2	2vs3	3vs4
	(sample 1)	(sample 2)	(sample 3)
Treatment	0.058*** (0.017)	0.099*** (0.013)	0.022* (0.011)
Robust p-value	0.004	0.000	0.051
Observations left	8,018	4,375	1,119
Observations right	8,018	4,375	1,119
Polyn. order	1	1	1
Bandwidth	0.066	0.052	0.033
Mean, left of threshold	0.456	0.048	0.006

Notes: Standard errors clustered at the race level are in parentheses. Statistical significance is computed based on the robust p-value and ***, **, and * indicate significance at 1, 5, and 10%, respectively. The unit of observation is the candidate. Each column reports the results from a separate local polynomial regression. The outcome is a dummy equal to 1 if the candidate wins the election. The independent variable is a dummy equal to 1 if the candidate is better ranked in the first round. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1 and the bandwidths are derived under the MSERD procedure. The mean, left of the threshold gives the value of the outcome for the lower ranked candidate at the threshold. It is computed by taking as outcome the probability of winning with all values on the right of the threshold replaced by 0.

To probe the robustness of the results to alternative specification and bandwidth choices, we estimate the treatment impacts using a quadratic specification (Table C1 in Appendix C, columns 2, 4 and 6), the optimal bandwidths computed according to Imbens and Kalyanaraman (2012) (Table C2, columns 2, 4 and 6), and tighter bandwidths obtained by dividing the MSERD bandwidths by 2 (Table C3, columns 2, 4 and 6). All regressions use Calonico et al. (2014)'s estimation procedure. The estimates obtained using these different specifications are very close in magnitude and they remain statistically significant at the same level. In addition, the effects of ranking 2vs3 are robust to excluding races in which the second candidate is close to the first candidate in the first round and the effects of ranking 3vs4 to excluding races in which the third candidate is close to the second, indicating that our estimates are not driven by cases in which several vote share discontinuities overlap.

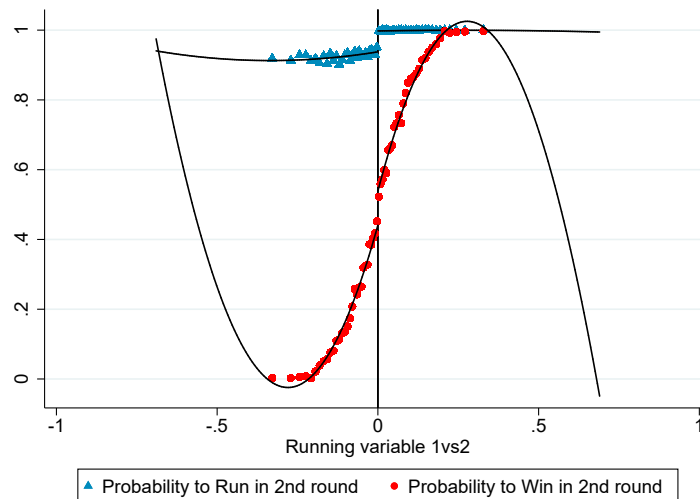
The effects of rankings on winning the race can result both from an increased likelihood to run in the second round, as any qualified candidate can decide to drop out and winning requires staying in the race, and from an increased likelihood to win the election, conditional on running, if voters rally behind higher-ranked candidates. We now use our RDD framework to estimate the effects of rankings on both outcomes and disentangle these two channels. We also estimate the impact on vote shares conditional on running, to determine which fraction of voters drives the impact on winning conditional on running.

3.2 Impact on running

We begin again with a graphical analysis. Figure 7 plots both the likelihood of running (in blue) and the likelihood of winning (in red, replicating Figure 4) of the first and second candidates against the running variable. The quadratic polynomial fit for running indicates a large upward jump at the cutoff. The jump is even more spectacular for ranking 2vs3 (Figure 8) and 3vs4 (Figure 9), and in both cases it is larger than the discontinuity observed for winning.

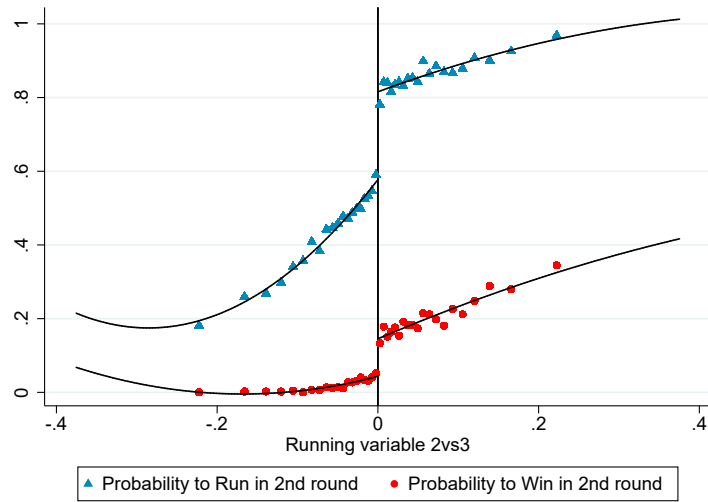
Consistent with the graphical analysis, the estimates reported in Table 5 indicate that ranking 1vs2 increases candidates' likelihood to run in the second round by 5.6 percentage points (6.0 percent of the mean at the threshold on the left): while 5.9 percent of close second candidates decide not to enter the second round, all first place candidates do (column 1). Ranking 2vs3 and 3vs4 has larger effects still: it increases running by 23.5 percentage points (40.9 percent) and 14.6 percentage points (48.0 percent) respectively (columns 3 and 5). All three effects are significant at the 1 percent level.

Figure 7: Impact on winning and running 1vs2



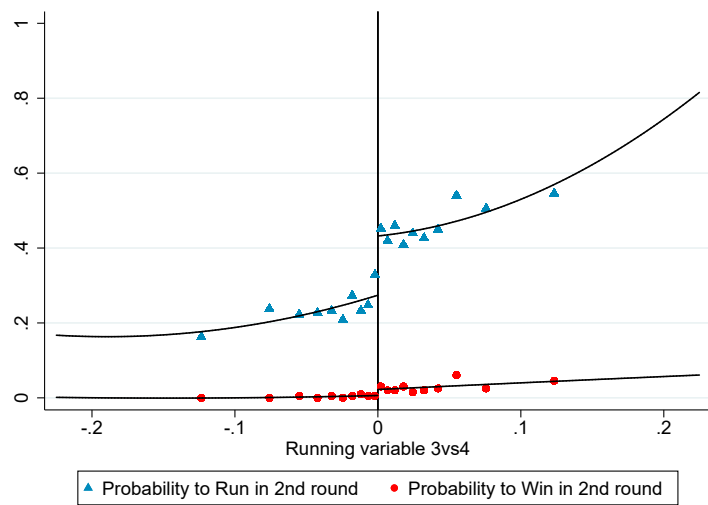
Notes: Triangles (resp. circles) represent the local averages of the probability that the candidate runs (resp. wins) in the second round (y-axis). Other notes as in Figure 4.

Figure 8: Impact on winning and running 2vs3



Notes: Triangles (resp. circles) represent the local averages of the probability that the candidate runs (resp. wins) in the second round (y-axis). Other notes as in Figure 5.

Figure 9: Impact on winning and running 3vs4



Notes as in Figure 8.

Table 5: Impact on winning and running

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	1vs2		2vs3		3vs4	
	(sample 1)		(sample 2)		(sample 3)	
	Run	Win	Run	Win	Run	Win
Treatment	0.056***	0.058***	0.235***	0.099***	0.146***	0.022*
	(0.005)	(0.017)	(0.018)	(0.013)	(0.040)	(0.011)
Robust p-value	0.000	0.004	0.000	0.000	0.003	0.051
Observations left	12,208	8,018	5,350	4,375	1,157	1,119
Observations right	12,208	8,018	5,350	4,375	1,157	1,119
Polyn. order	1	1	1	1	1	1
Bandwidth	0.109	0.066	0.068	0.052	0.036	0.033
Mean, left of threshold	0.941	0.456	0.574	0.048	0.304	0.006

Notes: In columns 1, 3 and 5 (resp. 2, 4 and 6), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. Other notes as in Table 4.

Once again, these effects have a similar magnitude and remain statistically significant when using a quadratic specification or the two alternative bandwidths (see Appendix C), and the effect of ranking 2vs3 (resp. 3vs4) is robust to excluding races in which the second candidate was close to the first candidate in the first round (resp. the third candidate was close to the second).

3.3 Impact on winning and vote shares conditional on running

We now turn to the second channel which might underlie the impacts of rankings on winning: an increased vote share and likelihood of winning *conditional on running* in the second round, either because active voters rally behind higher-ranked candidates or because these candidates manage to mobilize a larger fraction of their supporters.

To estimate these effects, we cannot simply run an RDD on elections in which both the lower-ranked and higher-ranked candidates decide to enter the second round. Indeed, the fact that close candidates qualified for the second round are similar at the threshold does not imply that close candidates who decide to run in the second round are similar as well.

To address this selection issue, we follow Anagol and Fujiwara (2016), who adapt Lee (2009)'s bounds method to RDDs. To estimate the impact of ranking 1vs2 on the likelihood of winning conditional on running, we first decompose it mathematically into observed and unobserved components.

Using the potential outcomes framework, we define R_0 and R_1 as binary variables indicating if the candidate runs in the second round when $T = 0$ (the candidate ranked second in the first round) and $T = 1$ (the candidate ranked first in the first round), respectively. In the data, we only observe $R = TR_1 + (1 - T)R_0$: we know whether the candidate ranked first decides to run in the second round but not whether she would have run if ranked second, and conversely.

Similarly, we define W_0 and W_1 as binary variables indicating if the candidate wins in the second round conditional on running when $T = 0$ and $T = 1$, respectively. We only observe $W = R[TW_1 + (1 - T)W_0]$: when the candidate does not run in the second round ($R = 0$), she does not win ($W = 0$) and we do not observe whether she would have won if she had stayed in the race. When she runs in the second round ($R = 1$), we observe whether the candidate ranked first wins the election but not whether she would have won if ranked second, and conversely.

We further define four types of candidates: “always takers,” who always run in the second round, whether they ranked first or second in the first round; “never takers,” who never run in the second round; “compliers,” who run in the second round if ranked first but not if ranked second; and “defiers,” who run in the second round if ranked second but not if ranked first. The key assumption we use to derive bounds is that there are no defiers: all candidates who ranked second and enter the second round would also have run if ranked first. Under this assumption, we have that $R_1 \geq R_0$ and we can write the impact on the unconditional likelihood of winning (estimated in Section 3.1) as the sum of the impact on running in the second round (estimated in Section 3.2), multiplied by the likelihood that close second-place compliers would win if they entered the race; and the impact on the likelihood of winning conditional on running (for compliers and always takers), multiplied by the probability of running of first-place candidates at the thresh-

$$\underbrace{E(W_1R_1 - W_0R_0|x = 0)}_{RD \text{ effect on } W} = \underbrace{Prob(R_1 > R_0|x = 0)}_{RD \text{ effect on } R} \cdot \underbrace{E(W_0|x = 0, R_1 > R_0)}_{Unobservable}$$

old:

$$+ \underbrace{E[W_1 - W_0|x = 0, R_1 = 1]}_{\text{Effect on win cond on being always-taker or complier}} \cdot \underbrace{E(R_1|x = 0)}_{\lim_{x \downarrow 0} E[R|x]}$$

From this expression, we get:

Effect on win cond on being always-taker or complier

$$\underbrace{E[W_1 - W_0|x = 0, R_1 = 1]}_{\text{Effect on win cond on being always-taker or complier}} = \underbrace{\frac{1}{\lim_{x \downarrow 0} E[R|x]}}_{\lim_{x \downarrow 0} E[R|x]} \left[\underbrace{E(W_1R_1 - W_0R_0|x = 0)}_{RD \text{ effect on } W} - \underbrace{Prob(R_1 > R_0|x = 0) \cdot E(W_0|x = 0, R_1 > R_0)}_{\substack{RD \text{ effect on } R \\ Unobservable}} \right] \quad (2)$$

$E(W_0|x = 0, R_1 > R_0)$ is the likelihood that compliers would win if they entered the race, absent treatment (i.e., when they rank second). By definition, compliers do not run when they rank second (but only when they rank first). This term is thus unobservable. Since all the other terms on the right-hand side of equation [2] are observed, we can derive bounds on the effect on winning conditional on running by making assumptions about this term.

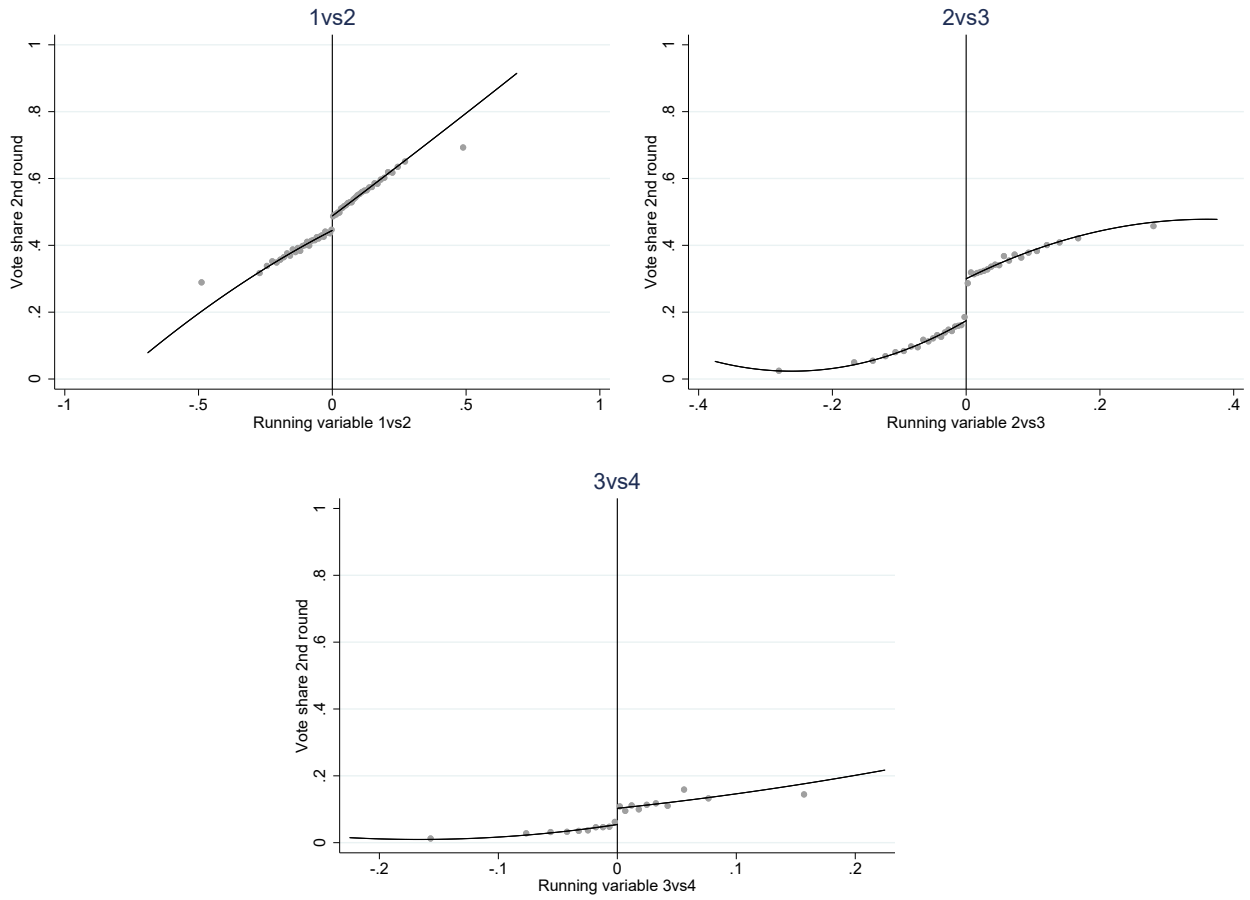
To obtain an upper bound, we set $E(W_0|x = 0, R_1 > R_0) = 0$, as the largest possible effect occurs if we assume that close second-ranked compliers would never win in the second round

if they decided to run. To obtain a lower bound, we replace the unobservable term by the probability that close first-ranked candidates who do choose to stay in the race win the election: 51.8 percent. The choice of this high probability (which is higher than the probability of victory of close second-rank candidates who actually run in the second round: 48.4 percent) makes our lower bound conservative.

We use the same method to derive bounds on the impact of ranking 2vs3 and 3vs4 on the likelihood of winning conditional on running. The probabilities that close second-ranked and third-ranked compliers win the election, which we use to replace the unobservable term when computing the lower bounds of both impacts are 18.3 and 6.2, respectively, which is much higher than the probability of victory of close third-ranked (resp. fourth-ranked) candidates who do run in the second round: 8.5 (resp. 1.6).

To derive bounds on the effects on second round vote shares conditional on running, we replace the RD effect on the unconditional likelihood of winning by the RD effect on unconditional vote shares (an outcome equal to 0 if the candidate drops out between rounds), in Equation [2]. This effect corresponds to the jumps observed on Figure 10, which plots unconditional vote shares of the lower-ranked and higher-ranked candidates against the running variable. In addition, to derive the lower bound 1vs2, we replace the unobservable term by the vote share obtained in the second round by close first-ranked compliers: 48.6 percent. Again, we use the same method for 2vs3 and 3vs4. The second round vote share of close second-ranked and third-ranked compliers, which we use to compute their lower bounds, are 37.0 and 23.1 respectively.

Figure 10: Vote share in the second round



Notes: Dots represent the local averages of the predicted assignment status (y-axis). Other notes as in Figure 2.

Finally, we use a bootstrapping procedure to estimate the standard errors of the bounds: we draw a sample from our data with replacement, compute the lower and upper bounds as indicated above, repeat these two steps a very large number of times, and estimate the empirical standard deviation of both bounds.

Table 6 provides the resulting bounds and bootstrapped standard errors of the effects of ranking 1vs2, 2vs3 and 3vs4 on conditional vote shares and likelihood of winning.

Table 6: Bounds on winning and vote share, conditional on running

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	1vs2		2vs3		3vs4	
	(sample 1)		(sample 2)		(sample 3)	
	Win	Vote share	Win	Vote share	Win	Vote share
Upper bound	0.059	0.040	0.122	0.147	0.050	0.100
Boot. std error	(0.025)**	(0.004)***	(0.017)***	(0.012)***	(0.023)**	(0.023)***
Lower bound	0.029	0.013	0.069	0.040	0.030	0.025
Boot. std error	(0.024)	(0.003)***	(0.014)***	(0.005)***	(0.018)*	(0.012)**
Mean $T = 0$	0.484	0.472	0.085	0.311	0.016	0.197

Notes: ***, **, and * indicate significance at 1, 5, and 10%, respectively. The mean, left of the threshold gives the value of the outcome for the lower ranked candidate at the threshold, conditional on running in the second round. It is computed by taking as outcome the probability of winning or the vote share, with all values on the right of the threshold replaced by 0.

As shown in column 1, conditional on running in the second round, ranking 1vs2 in the first round increases the likelihood of winning by 2.9 to 5.9 percentage points (6.0 to 12.2 percent of the mean for candidates ranked second who run in the second round at the threshold). The upper bound is significant at the 5 percent level, but the lower bound is not. The effect on vote shares conditional on running is 1.3 to 4.0 percentage points (2.8 to 8.5 percent), where both the upper and lower bounds are significant at the 1 percent level (column 2).

Ranking 2vs3 has larger effects, conditional on running. First, it increases the likelihood of winning by 6.9 to 12.2 percentage points (81.2 to 143.5 percent, column 3). In other words, ranking 2vs3 roughly doubles candidates' likelihood of winning, conditional on running. Second, it increases the conditional second round vote share by 4.0 to 14.7 percentage points (12.9 to 47.3 percent, column 4). Both the upper and lower bounds of both effects are significant at the 1 percent level.

Finally, ranking 3vs4 increases the conditional likelihood of winning by 3.0 to 5.0 percentage points, which corresponds to a two-fold or three-fold increase, compared to the mean at the threshold on the left (column 5). It increases the second round vote share by 2.5 to 10.0 percentage points (12.7 to 50.8 percent), conditional on running (column 6). All these bounds are significant at the 10 percent level or at a higher level.

Overall, our results so far indicate that the effects of first round rankings on candidates' likelihood to win the second round are driven both by higher-ranked candidates' higher likelihood to stay in the race and by voters rallying behind them. Importantly, these two effects may be linked. Indeed, candidates may decide whether or not to stay in the race based on the comparison between their expected likelihood to win, which depends in part on voters' response to rankings, and the costs of running, such as time and money spent campaigning. As a result, lower-ranked candidates' higher likelihood to drop out may in part reflect the (accurate) anticipation of facing an electoral disadvantage in the second round.

To the extent that candidates adjust their decisions to their expectations about voter behavior, any mechanism affecting voters' response to rankings may help explain candidates' own response. Therefore, the next section discusses the mechanisms underlying the behavior of candidates and voters jointly.

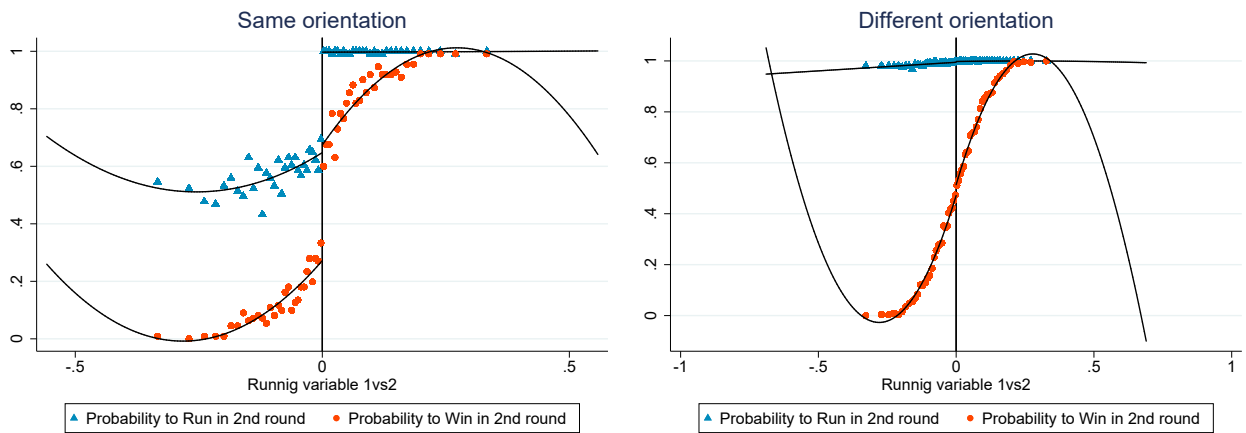
4 Mechanisms

4.1 Impact depending on political orientation

To investigate the mechanisms underlying the large effects of ranking 1vs2, 2vs3, and 3vs4 on running in the second round, winning, and winning and vote shares conditional on running, we first compare effect size when the higher- and lower-ranked candidates have the same political orientation or, instead, distinct orientations.

As shown on Figures 11, 12 and 13, the effects of rankings on running and winning are much larger in races in which candidates have the same orientation. When the first and second candidates have the same orientation, ranking 1vs2 increases the likelihood of running and winning by 35.2 and 30.4 percentage points (Table 7, columns 2 and 5). Both estimates are significant at the 1 percent level. Instead, the effects are close to zero and not significant when they have distinct orientations (columns 3 and 6). We find a similar difference, although not as important, for ranking 2vs3: its effects on running and winning are 62.7 and 22.3 percentage points, significant at the 1 percent level, when the second and third candidates have the same orientation. When they have distinct orientations, the effects remain significant at the 5 percent level but decrease to 5.2 and 4.1 percentage points (Table 8). Finally, when the third and fourth candidates have the same orientation, the effect of ranking 3vs4 on running is 40.1 percentage points and significant at the 1 percent level, and the effect on winning 4.0 percentage points and not significant. Both point estimates are lower and not significant when they have distinct orientations (Table 9).

Figure 11: Impact depending on political orientation 1vs2



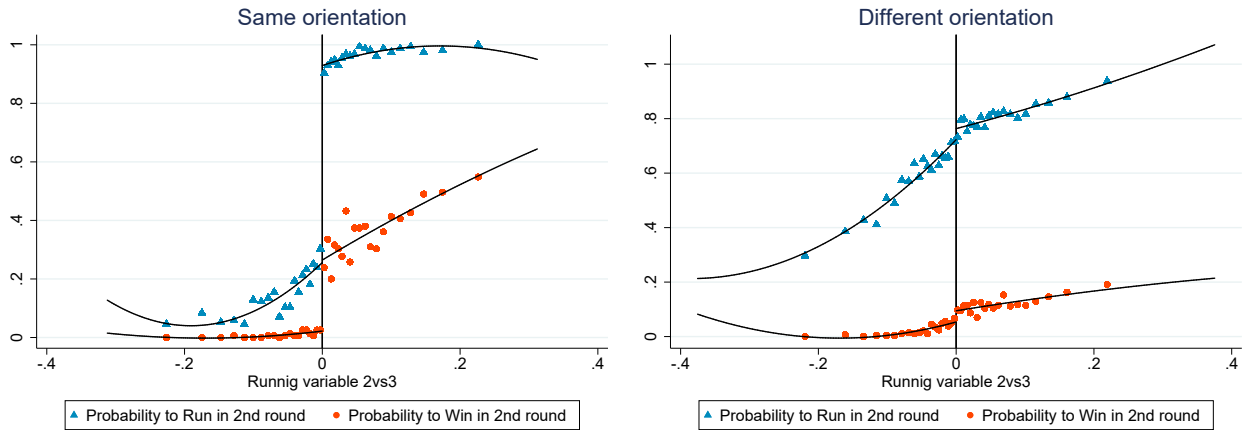
Notes as in Figure 7.

Table 7: Impact depending on political orientation 1vs2

Outcome	Probability to run 1vs2			Probability to win 1vs2		
	Full	Same	Diff	Full	Same	Diff
Treatment	0.056*** (0.005)	0.352*** (0.022)	0.001 (0.002)	0.058*** (0.017)	0.304*** (0.038)	0.017 (0.018)
Robust p-value	0.000	0.000	0.693	0.004	0.000	0.623
Observations left	12,208	2,046	7,219	8,018	1,394	7,243
Observations right	12,208	2,046	7,219	8,018	1,394	7,243
Polyn. order	1	1	1	1	1	1
Bandwidth	0.109	0.121	0.071	0.066	0.076	0.072
Mean, left of threshold	0.941	0.647	0.996	0.456	0.315	0.480

Notes: In columns 1, 2 and 3 (resp. 3, 4 and 5), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. In columns 2 and 5 (resp. 3 and 6), the two candidates have the same orientation (resp. distinct orientations). Other notes as in Table 4.

Figure 12: Impact depending on political orientation 2vs3



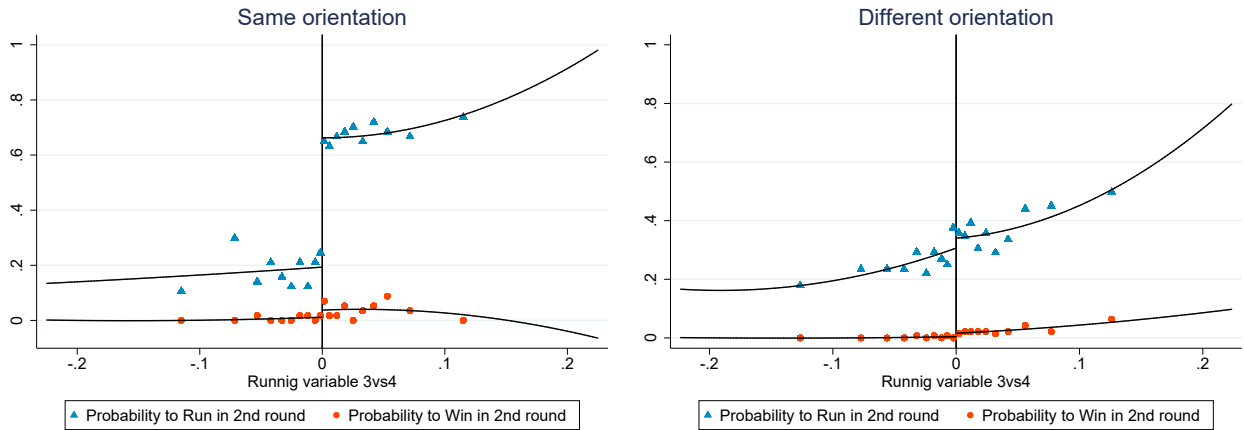
Notes as in Figure 7.

Table 8: Impact depending on political orientation 2vs3

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	Probability to run 2vs3			Probability to win 2vs3		
	Full	Same	Diff	Full	Same	Diff
Treatment	0.235*** (0.018)	0.627*** (0.029)	0.052** (0.021)	0.099*** (0.013)	0.223*** (0.027)	0.041** (0.013)
Robust p-value	0.000	0.000	0.045	0.000	0.000	0.011
Observations left	5,350	1,489	3,727	4,375	1,343	3,490
Observations right	5,350	1,489	3,727	4,375	1,343	3,490
Polyn. order	1	1	1	1	1	1
Bandwidth	0.068	0.055	0.073	0.052	0.048	0.066
Mean, left of threshold	0.574	0.286	0.706	0.048	0.023	0.059

Notes: In columns 1, 2 and 3 (resp. 3, 4 and 5), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. In columns 2 and 5 (resp. 3 and 6), the two candidates have the same orientation (resp. distinct orientations). Other notes as in Table 4.

Figure 13: Impact depending on political orientation 3vs4



Notes as in Figure 7.

Table 9: Impact depending on political orientation 3vs4

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	Probability to run 3vs4			Probability to win 3vs4		
	Full	Same	Diff	Full	Same	Diff
Treatment	0.146*** (0.040)	0.401*** (0.065)	0.026 (0.049)	0.022* (0.011)	0.040 (0.027)	0.014 (0.009)
Robust p-value	0.003	0.000	0.758	0.051	0.127	0.156
Observations left	1,157	349	812	1,119	325	847
Observations right	1,157	349	812	1,119	325	847
Polyn. order	1	1	1	1	1	1
Bandwidth	0.036	0.038	0.035	0.033	0.034	0.037
Mean, left of threshold	0.304	0.233	0.330	0.006	0.013	0.002

Notes as in Table 7.

A possible interpretation is that the effects of rankings are driven by coordination: strategic voters use them to all coordinate on the same subset of candidates in a decentralized way, and parties to decide which candidates should drop out of the race. Shared political orientation makes coordination more *desirable*: it increases the value that the lower-ranked candidate and her supporters associate with the victory of the higher-ranked candidate against ideologically distant candidates and makes them more willing to contribute to it (by dropping out and voting for her, respectively), resulting in larger effects of rankings. But other interpretations are possible, since shared orientation also makes it *less costly* for voters to rally behind the higher-ranked candidate and easier for sister parties to reach dropout agreements (Pons and Tricaud, 2018), whatever the underlying motive. Dropout agreements between sister parties often involve multiple constituencies and are part of a long-term collaborative equilibrium which may be strengthened by ideological proximity and the habit to govern together.

In the next two sections, we focus on the impact of ranking 1vs2 and consider separately races in which a third candidate qualified or failed to qualify, to disentangle the different possible mechanisms at play.

4.2 The role of coordination

To investigate the extent to which coordination explains the effects of ranking 1vs2, we focus on elections in which three candidates or more qualified for the second round (sample 2). In these elections, the top two candidates and their supporters might want to coordinate against lower-ranked candidates and use rankings to do so. We conduct two distinct tests.

First, the first and second candidates and their supporters should be more willing to coordinate when the candidate ranked third is stronger and more likely to challenge the victory of one of them. If coordination against the third candidate drives our results, we should thus expect the second candidate to be more likely to drop out of the race and voters to be more likely to rally behind the first when the third candidate's vote share is closer to the second candidate's. Consistent with this prediction, Table 10 shows that the effects of ranking 1vs2 on entering the second round and winning are larger when the gap in first round vote shares between the second and third candidates is lower than 5 percentage points than in the full sample (columns 1 through 4). Effect size further increases when the gap is lower than 2.5 percentage points (columns 5 and 6).⁹

Table 10: Impact on 1vs2 depending on the strength of the 3rd

Outcome	(1)	(2)	(3) 1vs2		(5)	(6)
	Full		Gap<5%		Gap<2.5%	
	Run	Win	Run	Win	Run	Win
Treatment	0.096*** (0.010)	0.053 (0.025)	0.130*** (0.017)	0.099* (0.040)	0.185*** (0.030)	0.150** (0.046)
Robust p-value	0.000	0.119	0.000	0.065	0.000	0.012
Observations left	4,505	3,544	1,950	1,496	810	1,073
Observations right	4,505	3,544	1,950	1,496	810	1,073
Polyn. order	1	1	1	1	1	1
Bandwidth	0.087	0.065	0.089	0.066	0.065	0.089
Mean, left of threshold	0.899	0.450	0.864	0.392	0.807	0.350

Notes: Sample only includes the races where the third candidate is qualified for the second round. In columns 2 and 5 (resp. 3 and 6) the sample is further restricted to elections where the gap in the vote share between the candidates ranked second and third in the first round is lower than 5 (resp. 2.5) percentage points. In columns 1, 2 and 3 (resp. 4,5 and 6), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. Other notes as in Table 4.

⁹We observe the same patterns when we restrict the sample to races in which the top two candidates have the same orientation (see Table A4 in the Appendix).

Second, the top two candidates and their supporters should be more likely to coordinate together (instead of coordinating with other candidates and groups of voters) when their ideological distance is relatively lower than their distance with the third candidate. To the extent that our results are driven by coordination, we should first expect the effects to be larger when the third candidate has a different orientation than both top two than when she has the same orientation, in races in which the top two candidates have the same orientation. The results shown in Table 11 are aligned with this prediction: ranking 1vs2 increases the likelihood of running by 12.8 percentage points when the third candidate has the same orientation and 48.0 percentage points when she has a different orientation (columns 3 and 5); its effects on the likelihood of winning are -3.1 percentage points (which is not statistically significant) and 45.2 percentage points respectively (columns 4 and 6). When the top two candidates have distinct orientations, we should expect larger effects when the third candidate is on the right or on the left of both of them, on the left-right axis, than when she has the same orientation as one of them or is located in between. Support for this prediction is weaker as none of the effects found on running and winning in these two cases is significant (Table 12).

Overall, effect size heterogeneity in races in which three candidates or more qualified for the second round supports the interpretation that coordination by candidates and voters explains at least part of the effects of ranking 1vs2. To test whether strategic coordination can explain them entirely, we now turn to races in which the third candidate is *not* qualified for the second round (races of sample 1 where the third candidate received a vote share below the qualification threshold in the first round).

Table 11: Impact on 1vs2 depending on the political orientation of the 3rd - same orientation

Outcome	(1)	(2)	(3) (4) (5) (6)			
	Full		1vs2 - same orientation		3rd diff	
	Run	Win	Run	Win	Run	Win
Treatment	0.420*** (0.036)	0.371*** (0.045)	0.128** (0.049)	-0.031 (0.122)	0.480*** (0.041)	0.452*** (0.044)
Robust p-value	0.000	0.000	0.023	0.523	0.000	0.000
Observations left	869	854	177	138	707	802
Observations right	869	854	177	138	707	802
Polyn. order	1	1	1	1	1	1
Bandwidth	0.071	0.069	0.088	0.063	0.070	0.082
Mean, left of threshold	0.580	0.274	0.874	0.567	0.521	0.220

Notes: Sample includes only the races where the third candidate is qualified for the second round and the top two candidates have the same political orientation. In columns 3 and 4 (resp. 5 and 6) the sample is further restricted to elections where the third candidate has the same political orientation as the top two (resp. has a different political orientation). In columns 1, 3 and 5 (resp. 2,4 and 6), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. Other notes as in Table 4.

Table 12: Impact on 1vs2 depending on the political orientation of the 3rd - distinct orientations

Outcome	(1)	(2)	(3) (4) (5) (6)			
	1vs2 - distinct orientations					
	Full		3rd same or middle		3rd on the left or right	
	Run	Win	Run	Win	Run	Win
Treatment	0.003 (0.005)	-0.021 (0.026)	-0.006 (0.004)	-0.003 (0.027)	0.028 (0.015)	-0.020 (0.056)
Robust p-value	0.745	0.277	0.140	0.782	0.103	0.470
Observations left	2,843	3,158	1,658	2,952	801	658
Observations right	2,843	3,158	1,658	2,952	801	658
Polyn. order	1	1	1	1	1	1
Bandwidth	0.069	0.078	0.050	0.101	0.098	0.076
Mean, left of threshold	0.991	0.491	1.002	0.491	0.964	0.456

Notes: Sample includes only the races where the third candidate is qualified for the second round and the top two candidates have distinct political orientations. In columns 3 and 4 (resp. 5 and 6) the sample is further restricted to elections where the third candidate has the same political orientation as one of the top two or has a different orientation and is located in the middle of the top two on the left-right axis (resp. has a different political orientation and is located either on the right or on the left of the top two). In columns 1, 3 and 5 (resp. 2, 4 and 6), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. Other notes as in Table 4.

4.3 Party norms and bandwagon effect

When the third candidate is not qualified for the second round, there is no need or even room for the top two candidates and their voters to coordinate against a lower-ranked candidate. Nonetheless, the effects of ranking 1vs2 remain substantial. As shown in Table 13, it increases candidates' likelihood of running and winning by 1.8 and 5.8 percentage points overall (columns 1 and 4). These estimates are significant at the 1 percent and 5 percent level respectively.

When the first and second candidates have distinct orientations, none of them drops out between rounds, at the threshold (column 3). Ranking 1vs2 increases the likelihood of winning by 4.9 percentage points (column 6), which is at the margin of statistical significance (p-value 0.108).

When the top two candidates have the same orientation, the first candidate always enters the second round but the second drops out in 18.7 percent of the races, at the threshold. This difference is significant at the 1 percent level (column 2). Candidate dropouts often result from agreements between left-wing parties. These parties argue that they want to follow the first round choice of their supporters instead of allowing voters supporting candidates eliminated after the first round to decide of the outcome of the race between the remaining candidates. In sum, dropouts are motivated by the enforcement of a deliberative ideal in the subgroup of parties and voters on the left. Dropout agreements when the top two candidates have the same orientation and the third is absent may further be explained by the desire to save on campaign

efforts, avoid a campaign where negative arguments could hurt the long-term reputation of both competitors, and enforce national agreements allocating a certain number of seats to each of the allied parties. Indeed, in areas where they are enforced, the dropout agreements ensure that roughly half of the races are won by the candidates of either of the competing parties, at the threshold. Whatever their exact motivation is, these dropouts result in a large effect of rankings on the outcome of the race: The effect on winning is almost as large as the effect on running (16.8 percentage points), and significant at the 5 percent level.

Table 13: Impact on 1vs2 in races where the 3rd is not qualified

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	1vs2 absent 3rd					
	Probability to run 1vs2			Probability to win 1vs2		
	Full	Same	Diff	Full	Same	Diff
Treatment	0.018*** (0.004)	0.187*** (0.031)	-0.000 (0.000)	0.058** (0.021)	0.168** (0.054)	0.049 (0.022)
Robust p-value	0.000	0.000	0.269	0.034	0.015	0.108
Observations left	7,440	759	3,129	4,979	689	4,661
Observations right	7,440	759	3,129	4,979	689	4,661
Polyn. order	1	1	1	1	1	1
Bandwidth	0.120	0.125	0.051	0.075	0.111	0.078
Mean, left of threshold	0.982	0.813	1.000	0.471	0.418	0.476

Notes: Sample includes only the races where the third candidate is not qualified for the second round. In columns 2 and 5 (resp. 3 and 6) the sample is further restricted to elections where the two candidates have the same orientation (resp. distinct orientations). In columns 1, 2 and 3 (resp. 4, 5 and 6), the outcome is a dummy equal to 1 if the candidate runs (resp. wins) in the second round. Other notes as in Table 4.

To test whether voters respond to the top two candidates' first round rankings as well, in races in which the third candidate is not qualified for the second round, Table 14 derives bounds for the effects on winning and vote share conditional on running. We find that ranking 1vs2 increases candidates' likelihood of winning by 4.9 to 5.8 percentages points overall (column 1). The lower and upper bounds are both significant at the 10 percent level. These effects can be driven by active voters rallying behind the first candidate as well as a larger mobilization of this candidate's supporters. The behavior of these voters cannot be explained by the desire to coordinate against lower-ranked candidates (who, again, are not present, as they are not qualified). Instead, the most likely interpretation is that these voters derive intrinsic value from siding with the winner of the first round, or that they desire to vote for the winner of the race and rightly anticipate that the candidate ranked first in the first round has increased chances of also winning the second round.

Interestingly, the fraction of voters whose choice of candidate is based on these behavioral motives is relatively small on average: the effect on vote shares is between 1.0 and 1.9 percentages points (column 4), where both the lower and upper bounds are significant at 1 per-

cent. This fraction is sufficient to sway a larger fraction of close elections, demonstrating the importance of bandwagon effect for election outcomes.

The effect on winning conditional on running is observed not only when the top two candidates have the same orientation (column 2) but also when they have distinct orientations (column 3). This is a remarkable result as it indicates that the desire to vote for the winner affects electoral outcomes in a substantial number of races, even when the ideological distance between candidates is important. However, the lower and upper bounds on the effects on vote shares are small and not significant in that case (column 6). Instead, the conditional effect on vote shares is very large when both candidates have the same orientation, with lower and upper bounds of 7.6 and 18.7 percentage points, both significant at the 1 percent level (column 5).

An alternative interpretation for the effects of rankings on voter behavior is that preferences include a common value component and voters update their beliefs on candidates' quality based on the first round choices of others (e.g., Feddersen and Pesendorfer, 1997; Deltas et al., 2016), which benefits the candidate arrived first on average. However, our effects are estimated at the threshold, when the first and second candidates received exactly the same vote share and their rankings do not contain any additional information on voters' private signals. For this mechanism to operate, one would thus also need to assume that voters lack information on candidates' exact vote shares and that they wrongly believe that the first candidate received substantially more votes. Instead, one could expect voters who update their beliefs on candidates' quality in such a sophisticated way to also gather precise information on election results and pay attention to it. This makes this interpretation less plausible than the bandwagon mechanism even if we cannot entirely rule out that it contributes to the results.

Table 14: Bounds on winning and vote share 1vs2 in races where the 3rd is not qualified

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	Full	Win Same	1vs2 Diff	absent Full	3rd Vote share Same	Diff
Upper bound	0.058	0.168	0.049	0.019	0.187	0.002
Boot. std error	(0.030)*	(0.082)**	(0.029)*	(0.004)***	(0.032)***	(0.003)
Lower bound	0.049	0.059	0.049	0.010	0.076	0.002
Boot. std error	(0.029)*	(0.068)	(0.029)*	(0.003)***	(0.016)***	(0.003)
Mean $T = 0$	0.480	0.526	0.477	0.499	0.500	0.499

4.4 Discussion of alternative mechanisms

We have so far attributed the effects of rankings on candidates' likelihood of winning and on their vote shares conditional on running to choices made by voters. We now discuss three alternative mechanisms which could also explain these effects. First, we examine whether these effects might be driven by campaign choices made by the higher- and lower-ranked candidates

between the two rounds.

While we lack data on candidates' precise political platforms, we collected data on their campaign expenditures for the 1992 to 2015 local elections and for the 1993 to 2017 parliamentary elections (collectively accounting for 65.8 percent of our sample).¹⁰ We do not measure candidates' expenditures between rounds separately, but only know the total amounts of money they received and spent over the entire course of the campaign. We measure the impact of rankings on these two outcomes divided by the number of registered citizens in the district. The effects, shown in Table 15, are small overall and not significant, even though better-ranked candidates are more likely to run in the second round. The lack of systematic impact of rankings on total campaign expenditures and contributions is perhaps not too surprising, since the first and second rounds are separated by one week only.

Table 15: Impact on campaign expenditures and contributions

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	1vs2		2vs3		3vs4	
	Expend.	Contrib.	Expend.	Contrib.	Expend.	Contrib.
Treatment	-0.009	-0.015	0.034	0.033	0.017	0.007
	(0.012)	(0.014)	(0.022)	(0.023)	(0.079)	(0.081)
R. p-value	0.365	0.210	0.123	0.145	0.781	0.936
Obs l	5,109	4,889	1,544	1,570	92	92
Obs r	5,109	4,889	1,544	1,570	92	92
Polyn.	1	1	1	1	1	1
Bdw	0.084	0.080	0.053	0.055	0.018	0.018
Mean	0.583	0.609	0.417	0.427	0.350	0.362

Notes: Sample includes only the elections for which campaign expenditure data are available. In column 1 and 2 (resp. 3 and 4, 5 and 6) we further restrict the analysis to races where campaign expenditures and contributions are available both for the candidate ranked first and the candidate ranked second (resp. second and third, third and forth). In columns 1, 3 and 5 (resp. 2, 4 and 6) the outcome is the candidate's total expenditures (resp. contributions) spent (resp. received) during the electoral campaign. Other notes as in Table 4.

Second, we check whether the effects might be driven by choices made by a third political actor, different from voters and the higher- and lower-ranked candidates on whom we have focused so far: other candidates qualified for the second round. These candidates' decision to stay in the race or drop out between rounds might depend on the rankings of top candidates and it might in turn affect the higher- and lower-ranked candidates' vote shares and likelihood of winning. For instance, if third candidates are more likely to drop out of the race when

¹⁰All data come from the French National Commission on Campaign Accounts and Political Financing (CNCCFP). Data on campaign expenditures for recent years are available on the Commission's website (<http://www.cnccfp.fr/index.php?art=584>). We collected and digitized the data for the 1992, 1994, 1998, 2001 and 2004 local elections. Data on campaign expenditures for the 1993, 1997, and 2002 parliamentary elections were collected and digitized by Abel François and his co-authors (see Fauvelle-Aymar and François, 2005; Foucault and François, 2005). Data are only available for candidates who received more than 1 percent of the candidate votes in the first round and, in local elections, for cantons above the 9,000 inhabitants threshold.

the candidate ideologically closest to them among the top two is ranked first than when she is ranked second, then that candidate should receive more votes by the third candidate’s supporters when ranked first.

To examine this mechanism in a systematic way, we define two outcomes at the candidate level: a dummy equal to 1 if a lower-ranked candidate with the same orientation is present in the second round, and the number of such candidates. Both outcomes directly reflect dropout decisions of lower-ranked candidates.¹¹ For ranking 1vs2 (resp. 2vs3 and 3vs4), we consider candidates ranked third and below (resp. fourth and below and fifth and below).

The effects are shown in Tables 16, 17, and 18: ranking 1vs2, 2vs3, or 3vs4 does not have any significant effect on the presence of lower ranked candidates of the same orientation in the second round (columns 1 and 3). We test the robustness of this result in the subsample of races in which such effects are most likely to occur: races in which the two candidates of interest have distinct political orientations and at least one lower-ranked candidate qualified (columns 2 and 4 of each table). Again, we do not find any significant impact, except for a positive effect of ranking 3vs4 on the likelihood that a lower-ranked candidate of the same orientation is present, which is significant at the 10 percent level (Table 18, column 2).

Table 16: Impact on the presence of same-orientation lower ranked candidates - 1 vs2

Outcome	(1)	(2)	(3)	(4)
	Dummy lower ranked Full	Subsample	Number of lower ranked Full	Subsample
Treatment	-0.002 (0.005)	-0.011 (0.013)	-0.003 (0.005)	-0.017 (0.014)
Robust p-value	0.510	0.401	0.393	0.223
Observations left	11,441	2,804	11,181	2,676
Observations right	11,442	2,804	11,181	2,676
Polyn. order	1	1	1	1
Bandwidth	0.100	0.068	0.098	0.064
Mean, left of threshold	0.034	0.069	0.037	0.073

Notes: In columns 2 and 4 we only include races where the third candidate is qualified and the top two candidates have distinct political orientations. In columns 1 and 2 the outcome is a dummy equal to 1 if a lower ranked candidate who has the same orientation as the candidate is running in the second round. In columns 3 and 4 the outcome is the number of lower ranked candidates who have the same orientation as the candidate and are running in the second round. Other notes as in Table 4.

¹¹Effects on these outcomes should not be driven by the likelihood that a lower-ranked candidate with the same orientation qualified for the second round or the number of such candidates, as we do not observe any discontinuity at the cutoff for these outcomes, as should be expected (see the placebo checks in Section 2.4).

Table 17: Impact on the presence of same-orientation lower ranked candidates - 2vs3

Outcome	(1)	(2)	(3)	(4)
	Dummy lower ranked Full	Subsample	Number of lower ranked Full	Subsample
Treatment	-0.004 (0.005)	-0.022 (0.028)	-0.005 (0.006)	-0.024 (0.030)
Robust p-value	0.471	0.434	0.449	0.421
Observations left	5,082	693	4,851	687
Observations right	5,082	693	4,851	687
Polyn. order	1	1	1	1
Bandwidth	0.064	0.047	0.060	0.047
Mean, left of threshold	0.022	0.074	0.023	0.077

Notes: In columns 2 and 4 we only include races where the fourth candidate is qualified and the candidates ranked second and third have distinct political orientations. In columns 1 and 3 the outcome is a dummy equal to 1 if a lower ranked candidate who has the same orientation as the candidate is running in the second round. In columns 2 and 4 the outcome is the number of lower ranked candidates who have the same orientation as the candidate and are running in the second round. Other notes as in Table 4.

Table 18: Impact on the presence of same-orientation lower ranked candidates - 3vs4

Outcome	(1)	(2)	(3)	(4)
	Dummy lower ranked Full	Subsample	Number of lower ranked Full	Subsample
Treatment	0.013 (0.009)	0.080* (0.050)	0.011 (0.008)	0.065 (0.044)
Robust p-value	0.112	0.082	0.167	0.155
Observations left	1,188	195	1,317	246
Observations right	1,188	195	1,317	246
Polyn. order	1	1	1	1
Bandwidth	0.037	0.041	0.044	0.057
Mean, left of threshold	0.006	0.023	0.006	0.023

Notes: In columns 2 and 4 we only include races where the fifth candidate is qualified and the candidates ranked third and fourth have distinct political orientations. In columns 1 and 3 the outcome is a dummy equal to 1 if a lower ranked candidate who has the same orientation as the candidate is running in the second round. In columns 2 and 4 the outcome is the number of lower ranked candidates who have the same orientation as the candidate and are running in the second round. Other notes as in Table 4.

Third, voters may rally behind higher-ranked candidates as a result of larger coverage of these candidates by the media. To test for differential media coverage, we used Factiva’s research tool and collected all press articles released between the two rounds of parliamentary elections since 1997 and containing the entities “élection,” “électoral,” “législative,” “candidat,” or “circonscription” as well as all press articles released between the two rounds of local elections since 1998 and containing the entities “élection,” “électoral,” “cantonale,” “candidat,”

“canton,” or “cantons.” In total, these elections account for 51.1 percent of our sample.¹² We obtained a total of 76,673 articles. We measure the impact of ranking 1vs2, 2vs3, or 3vs4 on three different outcomes: the total number of articles mentioning the candidate’s first and last names at least once; the total number of mentions (counting twice the articles in which the candidate is mentioned twice, thrice the articles in which they are mentioned thrice, etc.); and the total number of articles mentioning the candidate in the title. As shown in Table 19, ranking 1vs2, 2vs3, or 3vs4 does not have any significant effect on any of these outcomes.

We conclude that rankings’ effects on electoral outcomes are driven neither by differential campaign expenditures, nor by differential press coverage, nor by dropout decisions of other candidates.

Table 19: Impact on press coverage

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Articles	1vs2 Quotes	Titles	Articles	2vs3 Quotes	Titles	Articles	3vs4 Quotes	Titles
Treatment	-0.151 (0.516)	-0.020 (0.928)	0.035 (0.046)	0.065 (0.593)	0.053 (1.018)	0.048 (0.035)	0.056 (0.140)	0.110 (0.247)	0.006 (0.006)
R. p-value	0.777	0.974	0.458	0.926	0.842	0.284	0.743	0.656	0.314
Obs left	5,193	5,062	6,272	1,405	1,494	1,464	134	125	125
Obs right	5,193	5,062	6,272	1,405	1,494	1,464	134	125	125
Polyn.	1	1	1	1	1	1	1	1	1
Bdw	0.086	0.083	0.109	0.044	0.048	0.047	0.020	0.018	0.018
Mean	4.369	7.080	0.232	2.043	3.110	0.041	0.160	0.287	0.006

Notes: Sample includes only the elections for which newspaper articles are available. In columns 1, 4 and 7, the outcome is the total number of articles mentioning the candidate at least once. In columns 2, 5, and 8, the outcome is the total number of mentions. In columns 3, 6, and 9, the outcome is the total number of articles mentioning the candidate in the title. Other notes as in Table 4.

5 Conclusion

This paper shows that past candidate rankings have large effects on future electoral outcomes. Using a regression discontinuity design in French two-round parliamentary and local elections since 1958, we find that arriving first in the first round increases candidates’ likelihood to run in the second round by 5.6 percentage points, compared to arriving second, and that arriving second and third increases running by 23.5 and 14.6 percentage points respectively, compared to arriving third and fourth. In addition to being more likely to stay in the race, higher-ranked candidates obtain larger vote shares and they are more likely to win, conditional on running. Voters rallying behind the candidate ranked first increase her conditional vote share by 1.3 to

¹²Press articles are only available on Factiva from the end of the 1990s onwards. The number of newspapers covered and the total number of articles are much lower in the earlier years. Since a disproportional fraction of races of sample 3, used to measure the impact of ranking 3vs4, come from these earlier elections, the average number of citations for these candidates is very low.

4.0 percentage points and her likelihood of winning by 2.9 to 5.9 percentage points. The effects of arriving second instead of third are even larger – 4.0 to 14.7 and 6.9 to 12.2 percentage points –, and arriving third instead of fourth also has significant effects on vote shares and winning, conditional on running.

Overall, candidates and voters' combined response to rankings generates large effects on candidates' likelihood to win: arriving first instead of second, second instead of third, and third instead of fourth increases winning by 5.8, 9.9, and 2.2 percentage points, respectively.

The effect of ranking first is larger when the third candidate is more likely to challenge the top two candidates and when the top two candidates are of the same political orientation, suggesting that coordination by parties and voters against other candidates qualified for the second round drives part of the effects. Our evidence first indicates that rankings help strategic voters to focus on the same subset of candidates and on the same equilibrium in a decentralized way. This is an important result, given that multiple strategic equilibria usually exist, when there are three candidates or more (Myerson and Weber, 1993). Second, rankings also facilitate parties' coordination, leading to a decrease in the number of candidates. This is at least as important given many voters' propensity to vote expressively, when they have to choose between more than two candidates, and the resulting risk of suboptimal electoral outcomes such as a defeat of the Condorcet winner (Pons and Tricaud, 2018). Dropout agreements based on rankings can help addressing this issue and increasing the representativeness of elected leaders by reducing the number of alternatives and often bringing it down to two.

But the effects of ranking first instead of second remain large in elections in which the third candidate is *not* qualified, showing that strategic coordination cannot explain everything. In this case, party-level agreements lead the second candidate to drop out in 18.7 percent of the races, when she has the same orientation as the first, and voters rallying behind the first increase her vote share by 1.0 to 1.9 percentage points and her likelihood of winning by 4.9 to 5.8 percentage points on average, conditional on running. We infer that dropout agreements between parties can be motivated by other reasons than strategic coordination, such as enforcing national agreements which allocate a certain number of seats to each party, and that behavioral motives such as bandwagon effect can greatly affect voter behavior and electoral outcomes.

This last result is perhaps more unsettling. Mainstream political economy models predict that election outcomes and policies implemented by elected leaders correspond to voters' preferences. In citizen-candidate models, the candidate proposing the platform preferred by the largest group of voters gets elected (Osborne and Slivinski, 1996; Besley and Coate, 1997), and in the voter median theorem, competing parties align their platforms with the policy preference of the voter who is the most representative by virtue of being located in the median (Downs, 1957b). Instead, we find that a large number of elections are swayed by a relatively small fraction of voters driven by their desire to vote for the winner instead of substantial differences between candidates such as valence and policy platforms.

This result also has implications for the choice of an optimal voting rule. A large literature

compares voters' incentives to misrepresent their true preferences and strategically adjust their choices to the expected behavior of others under different voting rules (e.g., Laslier, 2009; Balinski and Laraki, 2011; Dasgupta and Maskin, 2019). Our findings indicate that voters' actual preferences may themselves depend on others' behavior. This phenomenon, and the fact that it affects the outcome of many races, adds a new layer of complexity to the problem of preferences' aggregation.

Chapter 3

Better Alone? Evidence on the Costs of Intermunicipal Cooperation¹

Abstract

What is the optimal jurisdiction size? While central governments tend to encourage intermunicipal cooperation in order to achieve economies of scale, municipalities are often reluctant to integrate. Exploiting a 2010 reform in France that forced non-integrated municipalities to enter an intermunicipal community, this paper provides causal evidence that resistance is driven by local costs of integration. Using a difference-in-differences strategy, I first find that municipalities forced to integrate experienced a 12.4 percent increase in the number of building permits delivered, raising congestion in urban municipalities. Second, I find that rural municipalities experienced a decrease in the number of public service facilities located within their territory, increasing the distance to public services for their residents. Additional results suggest that these costs are sufficiently high to offset the benefits of integration in terms of better access to public transport and higher fiscal revenues. These findings shed new light on the factors explaining municipalities' resistance while stressing the consequences of changing the scale of decision making.

¹I am grateful to Philippe Aghion, Alberto Alesina, Pierre Boyer, Julia Cagé, Guillaume Chapelle, Pierre-Philippe Combes, Bruno Crépon, Allan Drazen, Xavier D'Haultfoeuille, Antoine Ferey, Jeffrey Frieden, Edward Glaeser, Francis Kramarz, Isabelle Méjean, Vincent Pons, Pia Raffler, Alessandro Riboni, Benoit Schmutz, James Snyder, Stefanie Stantcheva, Mathias Thoenig, Clémentine Van Effenterre, and Romain Wacziarg for helpful comments, as well as seminar participants at Harvard, Ecole Polytechnique, CREST, LIEPP-Sciences Po, Paris School of Economics, Collège de France, and conference participants at CREST Workshop on Political Economy, ECHOPPE Workshop, Transatlantic Theory Workshop, Munich Political Economy and Public Economics Workshop, PEARL, and UEA. I thank Brice Fabre for sharing his data on municipalities' accounts prior to 2010, the team in charge of the building permits database at the Ministry of Sustainable Development for their help with the data, Bernard Pelamourgues for the CAF data on daycare, Cécile Queffelec for the Ministry of Culture data on public libraries, and Laurent Chevereau for the CEREMA data on public transport. I gratefully acknowledge Harvard University for its hospitality and the Labex Ecodec for financial support.

1 Introduction

Over the last century, central governments in developed countries have sought to consolidate their local jurisdictions in order to achieve economies of scale in the provision of public goods. In particular, they tend to set reforms to encourage intermunicipal cooperation. However, municipalities are often reluctant to cooperate, slowing down or even blocking the consolidation process.¹ On the one hand, cooperating municipalities benefit from efficiency gains and economies of scale. On the other, the loss of autonomy over public policies can be costly, most notably when integration implies cooperating with other municipalities that have different preferences (Tiebout, 1956; Alesina and Spolaore, 1997; Alesina et al., 2004).

This paper studies the factors explaining municipalities' resistance. In particular, it seeks to assess whether resistance can be explained by actual local costs of integration, beyond ideological or political considerations. For instance, a municipality cooperating with neighbors that have different preferences regarding urban planning or the location of public goods may end up with more construction than desired, or farther away from public services. For these municipalities, the costs of integration can outweigh their share of the benefits, leading them to oppose integration. Moreover, as they do not internalize the aggregate benefits of cooperation, they can refuse to integrate even when it would improve welfare overall. Absent appropriate mechanisms to compensate them, this may stall cooperation as a whole.

Identifying the costs, if any, that municipalities face when integrating is empirically challenging. First, despite the fact that intermunicipal cooperation can take various forms, most of the empirical literature has focused on mergers. While these studies can perhaps identify the overall impact of consolidation, it is difficult to use merger laws to assess the impact on individual municipalities. Indeed, after the merge, administrative data are collected at the post-merger level, preventing researchers from using such data to compare the situation of pre-existing municipalities before and after integration. Second, when municipalities do not fully merge but instead share some public policies, cooperation is usually voluntary. Hence, we observe the consequences of integration only for municipalities that have chosen to cooperate, and for which the costs are likely to be the lowest.

This paper is the first to causally identify the effects of integration for resisting municipalities. To do so, I exploit a 2010 reform in France that made intermunicipal cooperation mandatory, forcing around 1,800 municipalities to enter an intermunicipal community. Importantly, when entering an intermunicipal community, municipalities do not disappear (as they do in mergers), but coexist with this new higher level of governance. Using a difference-in-differences strategy, I compare, before and after 2010, municipalities that were forced to inte-

¹Despite large financial incentives, voluntary mergers in Japan resulted in a higher number of municipalities than what the government would have chosen to impose (Weese, 2015); other countries such as Denmark or Sweden instead decided to force mergers (Mouritzen, 2010; Lidström, 2010); in the US, after a wave of consolidation in the 19th century, municipalities became more and more reluctant to integrate (Jackson, 1987); in France, the government failed to impose mergers in the 1970s and then strove to promote the formation of intermunicipal communities.

grate with municipalities that have been part of an intermunicipal community for a long time. This strategy enables me to measure the causal impact of integration on the municipalities that resisted integration the most, and for which the costs are likely to be the highest. I then compare these results with the consequences of integration faced by municipalities that instead voluntarily integrated before the law, using a staggered adoption design where the date of the treatment corresponds to the year when a community was integrated.² Through this comparison, I can infer the reasons why municipalities forced to integrate opposed integration in the first place.

Municipalities that are part of the same community share two main policies, over which losing autonomy can be costly. First, French intermunicipal communities are in charge of urban planning and thus decide where and how much to build in each member municipality. As discussed by the literature on housing restrictions and NIMBYism,³ new constructions impose several costs on residents that can explain why a municipality might prefer to keep control over its own housing supply: More housing in high-demand and densely-built municipalities is likely to create congestion in the use of local amenities; it may also increase population's heterogeneity in municipalities whose neighbors are different, by bringing in minorities and/or poorer households; and it might decrease housing prices, threatening the value of residents' assets in municipalities with a large share of homeowners. These costs explain why local housing restrictions are difficult to overcome when decisions are made by the residents (Glaeser and Gyourko, 2018; Glaeser, 2014). They may also explain why municipalities are reluctant to share such policies with outsiders. Second, municipalities that are part of the same community jointly finance and provide some public services: waste management, road maintenance, public transport, and social, cultural, and sport facilities. While pooling resources for large-scale services such as public transport seems generally beneficial, sharing decisions over the location of local public service facilities can be costly for some municipalities. Indeed, with the aim of achieving economies of scale, intermunicipal communities seek to rationalize the offer of local public services and thus tend to concentrate resources on facilities located in high-density places. As a result, low-density municipalities might lose some public service facilities, increasing the distance to public services for their residents.

I show that both dimensions are important in explaining municipalities' resistance, but that the costs of integration are not the same for urban and rural municipalities.

Exploring first the consequences of the loss of autonomy over urban planning, I find that municipalities forced to enter an intermunicipal community experienced an increase of 12.4 percent in the number of building permits delivered per year, on average. To investigate further whether such effect can explain their resistance, I compare this result with the impact of inte-

²In light of the recent literature focusing on the issues associated with the staggered adoption design (Borusyak and Jaravel, 2017; Goodman-Bacon, 2018), I also use an alternative estimation procedure developed by de Chaisemartin and D'Haultfoeuille (2019). See Section 4.1.2 for more details.

³NIMBYism ("Not In My BackYard") dates back to Frieden (1979) and captures the idea that residents tend to oppose new constructions in their locality, as they do not want a stranger to settle in their backyard. See Gyourko and Molloy (2015) for a review.

gration for municipalities that instead voluntarily joined an intermunicipal community before the law. I find that municipalities that voluntarily integrated did not experience any change in their housing supply following integration. Hence, only municipalities that did not want to enter an intermunicipal community faced an increase in construction. This finding supports the view that resisting municipalities refused to integrate to avoid an increase in housing supply.⁴ It also provides evidence that their decision was driven by actual costs of integration, beyond ideological or political considerations.

I then carry additional analyses to assess which specific costs implied by a rise in housing supply explain their opposition. Results show that the increase in building permits is driven by municipalities where a rise in construction is the most likely to create congestion: those where the demand for housing is high and that are already densely built. In contrast, the impact on housing is not stronger for municipalities whose neighbors are more different, in terms of residents' income, share of immigrants, or political preferences, and the results are similarly not driven by municipalities where the share of homeowners is particularly large. This suggests that municipalities that opposed integration to keep control over their housing supply are mainly urban municipalities trying to avoid congestion costs, rather than municipalities fearing an increase in population heterogeneity or a housing price decline. In line with this interpretation, I find that despite its effect on construction, integration did not lead to a significant decrease in housing prices.

Second, I assess the consequences of forced integration on local public services. To do so, I gathered local-level data on two different local public services: the number of daycare facilities and public libraries located in each municipality. For both services, evidence suggest that resisting municipalities experienced a decrease in the number of facilities located within their territory, implying that integration increased the distance to public services for their residents. This effect is entirely driven by rural municipalities, that ended up with 20 percent fewer facilities on average relative to control municipalities. These results suggest that, while urban municipalities resisted integration to prevent further constructions from raising congestion, rural municipalities resisted to avoid losing local public services.

In the last part of the paper, I turn to the analysis of the benefits of integration. I explore whether resisting municipalities also benefited less from the gains of integration, which would contribute to explaining why they resisted longer. I first look at the impact of integration on public transport. By enhancing cooperation and enabling municipalities to pool resources, integration is likely to help neighboring municipalities build larger and more efficient public

⁴Section 6 discusses further the differential impact between municipalities that voluntarily integrated and those that resisted. First, I provide additional evidence that this is not driven by the fact that resisting municipalities were forced to enter already-formed communities where the other members could have punished them for having resisted. In particular, the impact is similar for resisting municipalities that entered existing communities as for those that created new ones, supporting the fact that they would have experienced the same effects had they integrated earlier. Second, I show that municipalities that resisted integration typically ended up with less bargaining power inside their IC compared to the average municipality, making them less able to fight IC's decisions. This helps explain why they experienced costs other municipalities were able to avoid.

transport networks. As a result, joining an intermunicipal community might increase a municipality's probability of being connected. In line with this argument, I find that municipalities forced to enter an intermunicipal community became twice as likely to have access to public transport.⁵ This is similar to the effect experienced by municipalities that voluntarily joined a community before the law, showing that resisting municipalities enjoyed the same benefits in terms of public transport as other municipalities.

Finally, I measure the impact of integration on municipalities' fiscal revenues and find that resisting municipalities experienced an increase of 13.0 percent per year in their total revenues per capita, on average. As with public transport, this increase is comparable to the increase experienced by municipalities that integrated voluntarily before the law. Assuming that municipalities anticipated such gain,⁶ this suggests that they did not oppose integration because they anticipated getting a lower share of the benefits, but rather that the benefits were not high enough to compensate for the congestion costs faced by urban municipalities and the loss of local public services faced by rural municipalities.

This paper provides new evidence on the factors explaining municipalities' opposition to consolidation and shows that resistance is driven by local costs of integration. These findings help understand consolidation failures and could help policymakers design better compensation schemes to implement consolidation policies more effectively. In particular, it is important to take into account that rural and urban municipalities do not face the same costs of integration.

This paper relates to several strands of literature. The first analyzes patterns of integration to study the factors explaining jurisdictions' choice to cooperate. Using structural or spatial econometric models, papers have emphasized the role of expected change in fiscal revenues, expected distance to public services, and municipalities' socio-demographic characteristics, as well as the role of neighbors' characteristics and thus heterogeneity (Gordon and Knight, 2009; Saarimaa and Tukiainen, 2014; Weese, 2015; Bel and Warner, 2016; Di Porto and Paty, 2018).⁷ In contrast, this paper provides new evidence on the factors explaining resistance by directly assessing the causal impact of integration on resisting municipalities.

Closer to my empirical strategy, a second strand of literature seeks to identify the causal impact of merger laws on cost reduction in the newly formed jurisdictions. The results are mixed (Bel and Warner, 2015): Although Reingewertz (2012) finds evidence of overall efficiency

⁵This result is obtained using data on regular public transport networks that are managed at the municipal or intermunicipal level, excluding the ones managed at the departmental or regional level. Municipal networks account for half of French public transport networks and 90 percent of all public transit trips. See Section 2.3 for more information on public transport data.

⁶It is likely that municipalities could anticipate the change in revenues they would face after integration. The first parameter to take into account is the state transfer awarded to the community they would join, which was public information, as the goal of the transfers were to incentivize cooperation. The second parameter is the tax rate and tax base of other municipalities, which are quite easy to learn given the geographical proximity between municipalities inside the same IC.

⁷Papers using different settings and identification strategies include Sorensen (2006) and Lapointe (2018). The first one uses surveys among local politicians in Norway and shows that their decisions to merge are mainly driven by the expected changes in revenues. The second paper studies voters' preferences over secession in Canada and find that income and language differences affect the likelihood of secession.

gains in Israel, Blom-Hansen et al. (2016) find no effect of recent merger laws in Denmark. The vast majority of these papers focus on the aggregate effects of consolidation. Two recent papers use geocoded data to study the distributional impact of mergers on local public sector jobs (Harjunen et al., 2019) and night-light intensity (Egger et al., 2017). In contrast, this paper measures the impact of forced collaboration (not mergers) on resisting municipalities and provides evidence on both the local costs and benefits of integration by looking at the effects on the housing market, local public services, public transport, and fiscal revenues.

Third, I contribute to the literature on housing restrictions. Recent empirical papers show that regulations in high-demand places are responsible for lowering productivity, preventing households from moving to opportunity and fostering urban sprawl (Glaeser and Gyourko, 2018; Hsieh and Moretti, 2019; Glaeser and Maré, 2001; Chetty et al., 2016; Glaeser and Kahn, 2010; Jones and Kammen, 2013). Despite these consequences, housing regulations are unlikely to lessen if decisions are made by residents at the local level, as most residents consider a new construction project nothing but a bother (Glaeser, 2014). In line with this argument, results show that municipalities are reluctant to lose control over urban planning. They also suggest that transferring urban planning to a higher level, and thus allowing outsiders and potential future residents to participate in the decision making, helps overcome local housing restrictions.⁸ These results stress the tension between local preferences and aggregate effects and highlight the risk of the local trap, where decisions made at the local scale ignore macro consequences (Purcell, 2006; Hankinson, 2018).

The remainder of the paper is organized as follows. Section 2 presents the institutional framework and the data. Section 3 describes the empirical strategy. I present the main results on housing supply and local public services in Section 4 and the results on the benefits of integration in Section 5. Section 6 discusses alternative interpretations and external validity. Section 7 concludes.

⁸Consistent with this finding, cross-sectional evidence in the US shows that municipalities with ward-based representatives do more exclusionary zoning than those with at-large representatives (Clingermayer, 1994).

2 Institutional background and data

2.1 Intermunicipal cooperation in France

France is divided into about 36,000 municipalities, accounting for 40 percent of all municipalities in Europe.⁹ Municipalities are the lowest and third tier of local government and represent 11 percent of total public spending. They are responsible for local urban planning; social housing; the provision of nurseries and primary schools; social, sport, and cultural facilities; municipal roads; and public transport. Municipalities' revenues come mainly from local taxation (54%) and state transfers (23%). They raise four local taxes: the local business tax paid by firms, the housing tax paid by all residents on the cadastral value of their accommodations (whether they are renters or owners), the property tax paid by owners, and the land tax. Each municipality is governed by a municipal council chaired by the mayor. Elections for municipal councilors take place every six years.

In the 1970s, the French government intended to pass a law that would have reduced the number of municipalities by 20 percent through mergers. But mayors massively blocked the reform and only few mergers took place; the number of municipalities was reduced by only 3 percent. Following this failure, the government decided to pivot toward promoting the creation of a new administrative structure: intermunicipal communities (**IC**).¹⁰ When entering an IC, the municipality does not disappear as it does in mergers, but continues to exist under a new level of local governance and share some public services with the other municipalities part of the same community.

First, by law, intermunicipal communities are in charge of “territory and economic development,” or the promotion of local businesses and urban planning. Since the 1980s decentralization laws, French municipalities are in charge of the delivery of building permits, and they need to produce a planning and development plan (*plan local d'urbanisme*). Once part of an IC, the local planning made by the municipality becomes subject to guidelines set by the community through the housing planning plan (*plan local d'habitat*) and the territorial coherence plan (*schéma de cohérence territoriale*). Hence, even if the municipality remains administratively in charge of delivering the building permits, the overall planning is done by the intermunicipal community, which decides where and how much to build in each municipality within the community.

On top of urban planning, municipalities decide which additional public services to transfer

⁹The average municipality size is 1,800 inhabitants, compared to 5,500 and 8,990 for the average municipality in the European Union and in the US, respectively.

¹⁰Intermunicipal communities are widespread in Europe (Hulst et al., 2009). In France, they are called EPCI, for *Etablissement public de coopération intercommunale*. France counts four types of ICs: *Communautes de Communes* (CC), *Communautes d'Agglomérations* (CA), *Communautes Urbaines* (CU) and *Métropoles*. The definition of these different ICs depends mainly on the number of mandatory competences to be transferred to the community and the size of the municipalities. The vast majority of municipalities are part of CC (more than 80 percent), the form of cooperation I describe below and the one that gives the most freedom to municipalities in deciding which public services to transfer.

to the IC. The most commonly delegated ones are services for which cooperation is likely to result in economies of scale: waste management, road maintenance, public transport, and social, cultural and sport facilities. Finally, ICs' revenues consist of state transfers and local taxes. The intermunicipal community can either apply an additional tax rate to the four municipal tax rates, or set a single business tax rate while still applying an additional tax rate on the three other local tax rates. In the latter case, the municipalities lose the right to set their own business tax. Municipalities choose the tax system when creating the community.

The IC is run by a board made of members of the municipal councils of all participating municipalities. The number of seats held by a municipality inside the intermunicipal council is proportional to the municipal population. By law, each municipality has at least one seat, and no municipality has more than half of the seats. Once the allocation of the seats is decided, each municipal council elects the municipal councilors who will be part of the intermunicipal council.¹¹ Then, the intermunicipal council elects its president. Decisions over which public services to transfer to the community or over which tax system to adopt require the approval of either (a) two thirds of the municipal councils representing more than half of the IC's total population or (b) at least half of the municipal councils representing more than two thirds of the population, as well as the approval of all municipal councils that represent more than one fourth of the IC's population. Then, day-to-day decisions about urban planning or the delivery of public services are made by the intermunicipal council, by majority. Hence, when joining an intermunicipal community, a municipality loses power over policies, the more so the smaller its population compared with the population of the other municipalities from the same IC.

Until 2010, municipalities were free to decide whether to create or join an IC or remain outside, but the financial incentives to integrate into an IC were high: Since the "Chevènement law" in 1999, ICs receive a state transfer on top of the individual transfers received by each municipality. The latter remains unchanged whether the municipality integrates or not. This law marked a turning point: whereas half of the municipalities were part of an IC in 1999, 95 percent were integrated by 2010. In 2010, France counted 2,611 ICs, and a community gathered on average 13 municipalities and 22,192 inhabitants. Still, 5 percent of the municipalities, or about 1,800, remained isolated in 2010.

¹¹In municipalities of more than 1,000 inhabitants and only since the 2014 election, the members of the intermunicipal council are directly elected by voters during municipal elections. Each list presents a ballot containing the list of candidates for both the municipal council and the intermunicipal council. As members of the intermunicipal council have to be municipal councilors, the two lists are typically very similar.

2.2 The 2010 law

In December 2010, a new law passed requiring that **(1) all municipalities must be part of an intermunicipal community** (2) all intermunicipal communities must contain at least 5,000 inhabitants. This law followed a report issued by the French Court of Auditors (*Cour des comptes*) indicating that, in general, ICs were too small to achieve economies of scale. The goal of the reform was thus to draw a new map of intermunicipal communities and organize the territory more effectively.¹²

This law forced the 5 percent of municipalities that were isolated to enter a community. The only exceptions were Paris and municipalities in three départements around it,¹³ as well as a few islands that make up one municipality. It also forced existing small intermunicipal communities to merge with other communities in order to comply with the 5,000 threshold. This threshold did not apply for ICs located in mountain zones, where municipalities are typically far from each other. In this paper, I focus on the first aspect of the law and look at what happened to the municipalities that were not part of an IC before 2010 and were thus forced to integrate.

Figure 1 shows the geographic distribution of French municipalities depending on their integration status in 2010. Red indicates that the municipality was still not integrated in 2010; blue indicates that the municipality was already part of an IC. The gray areas are municipalities excluded from the sample, as further explained in Section 3.1. Even if many isolated municipalities are located in the center-north region of the country, red municipalities are present across the whole country and in all French regions.

The implementation of the reform took place between 2011 and 2014 and was conducted by the departmental prefect.¹⁴ In 2010, municipalities forced to enter an IC shared a common border with 2.4 intermunicipal communities on average. They could choose which IC to join, but the decision had to be approved by the prefect. When possible, and if allowed by the prefect, they could also create a new IC with neighboring isolated municipalities. Finally, if a municipality was still not integrated in 2013, the prefect could force its integration in the IC she chose. At the end, 77 percent of the isolated municipalities entered an IC between 2011 and 2013, and 23 percent of them entered an IC chosen by the prefect in 2014. The majority (73 percent) joined an existing community, whereas the remaining ones created new ICs.¹⁵

¹²The 2010 law also changed the rule for the allocation of seats inside the intermunicipal council, as well as the way municipal councilors are elected. These changes became effective only after the 2014 municipal elections. At this time, all municipalities were already part of an IC, and it thus affected all municipalities the same way.

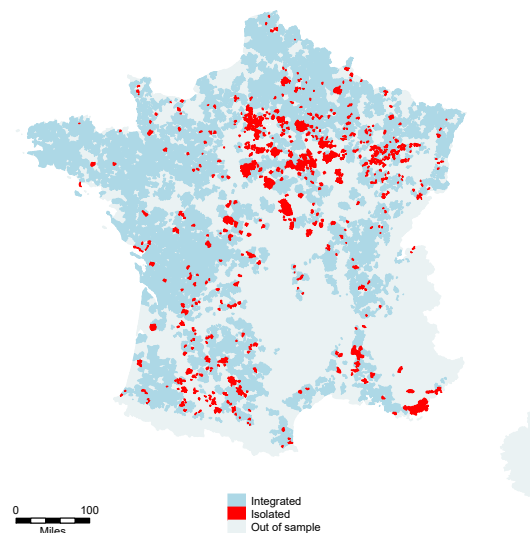
¹³Paris and municipalities in three départements around Paris (Hauts-de-Seine, Seine-Saint-Denis, and Val-de-Marne) were exempt from this law as they are part of the “Great Paris Project,” which has a goal of consolidating the Paris metropolitan area.

¹⁴The prefect is appointed by the French president and represents the state to local governments in the départements. Her main responsibilities include coordinating the police forces, applying immigration rules, and ensuring that local policies are in line with national ones.

¹⁵As shown in Section 6, the effects are similar whether municipalities entered an existing IC or created a new one. This may sound surprising, as we could have expected resisting municipalities forming new ICs to be able to coordinate to avoid integration costs. However, once in a community, municipalities have no choice but to share some policies, and coordinating on the status quo might be difficult when different types of municipalities face

On average, municipalities forced to integrate represented about 5 percent of the population of the community they joined. In 2014, the average intermunicipal community included 17 municipalities and 29,197 inhabitants.

Figure 1: French municipalities depending on their integration status in 2010



Notes: Municipalities in red were not part of an intermunicipal community in 2010. Municipalities in blue were already integrated. Grey areas represent municipalities excluded from the sample of analysis, as explained in Section 3.1.

2.3 Data

This study combines several databases at the municipality level.

Municipalities' characteristics. The broad range of municipalities' characteristics I use for the descriptive statistics as well as the heterogeneity analysis come from various sources. Municipalities' socio-demographic characteristics, such as population size or the share of immigrants, are available from censuses conducted by the National Institute of Statistics and Economic Studies (INSEE). They also provide the number of housing units and the share of homeowners in each municipality. On top of the censuses, INSEE provides a classification of the municipalities as urban or rural, based on the share of built land and population density, as well as the list and composition of French urban areas. Household taxable income data aggregated at the municipality level are provided by the Ministry of Finance and extracted from income tax declarations. Municipal electoral results and presidential electoral results at the municipality level are obtained from the Ministry of the Interior.

Housing building permits. Housing building permits data come from the Ministry of Sustainable Development (sit@del2 database). The dataset contains the number of housing

different costs. As further discussed in Section 6, this result provides support for the fact that municipalities would have experienced the same effects had they integrated earlier, and thus that the costs I measure explain why they resisted.

building permits delivered every year in each municipality.¹⁶ More precisely, it provides the “number of authorized constructions,” or the number of housing units allowed for construction. Hence, if a building of 10 apartments was approved, the dataset registers 10 authorized housing units, even if only 1 permit was delivered in practice. The database also indicates whether the unit is a house or an apartment and whether the construction takes place on empty land or as an extension to an existing building.

The ministry collects the data from the local institutions in charge of delivering the building permits. Integration into an IC could affect the data collection if, for instance, the information is transmitted more consistently or thoroughly once the IC is in charge. This, however, is unlikely to explain the results I find. First, this would be particularly relevant for small and remote municipalities where the information is difficult to collect. Instead, as shown in Section 4.1.3, the increase in building permits is particularly strong for urban and central municipalities. Moreover, I find that contrary to municipalities forced to integrate, municipalities that voluntarily entered an IC did not experience a similar increase, showing that there is no mechanical effect of entering an IC on the number of housing building permits reported in the database (see Section 4.1.2).¹⁷

Housing prices. Building on Combes et al. (2018) and following INSEE’s guidelines (see Gouriéroux and Laferrère, 2009; Musiedlak and Vignolles, 2016), I measure housing prices using indices estimated at the municipality level based on official transactions records. I perform the analysis separately for the Parisian region of Ile-de-France and for the rest of France, as the two databases come from two distinct notary associations and do not define dwellings’ characteristics in the same way. The data are made available by the Ministry of Sustainable Development for every even year since 2000. They are available until 2014 for Ile-de-France and until 2016 for the rest of the country. To construct the indices, I regress separately for each year the log of the price per square meter on the characteristics of the house or apartment. I then compute the indices as the average of the residuals for each municipality and year after adding the regression constant. Since I center the explanatory variables, the resulting indices can be interpreted as the price per square meter of a reference dwelling. Appendix D provides further details on the construction of the indices.

Local public services. I gathered data on two local public services: daycare and municipal libraries. Daycare data come from the Family Allowance Agency (CAF), which gives the number of daycare facilities and the total number of daycare spots available in each municipality for every year over the 2007-2017 period. Data on municipal libraries come from a yearly survey run by the Ministry of Culture from 2009 to 2017. The dataset lists all municipal libraries and their location, enabling me to compute the number of libraries in each municipality for

¹⁶I focus on ordinary housing, excluding residences providing particular services, such as medical or retirement residences.

¹⁷Note also that data at the year level are particularly reliable: The ministry collects the information every month and corrects past information if they realize any permits were omitted. Moreover, those in charge of collection told me that they did not notice any differences when collecting the data from ICs or municipalities.

each year. Unfortunately, when the Ministry of Culture began to survey local public libraries in 2009, only 7 départements took part in the survey. The analysis is thus restricted to those départements for this particular outcome.

Public transport. I obtained information on municipalities' access to public transport thanks to the CEREMA¹⁸ that surveys public transport operators every year in all French regions except Ile-de-France. During the survey, operators report the list of municipalities served by their network. Hence, I know for each year which municipalities have access to a public transport network. It includes all regular public transport networks that are managed at the municipal or intermunicipal level. They represent about half of French public transport networks (the other half being managed at the departmental or regional level) and account for 90 percent of all public transit trips. Specific and temporary transports such as school bus services or transport services during market days are not included in the database. Data are available for the period 1995-2017.

Fiscal revenues. To measure municipalities' fiscal revenues, I use municipalities' and ICs' annual accounts that are made publicly available by the French Ministry of Economy and Finance (DGFIP) from 2010 to 2017. For years prior to 2010, the data come from Fabre (2017).

3 Empirical Strategy

3.1 Treatment and control groups

This paper uses a difference-in-differences strategy in order to assess the impact of integration on municipalities that resisted integration and were forced to enter an intermunicipal community. More precisely, I compare, before and after the 2010 law, municipalities that were not part of an intermunicipal community in 2010 and thus forced to integrate (treatment group) with municipalities that were already part of an IC in 2010 and thus not affected by the law (control group).

I consider a balanced panel of municipalities focusing on metropolitan France, excluding overseas territories and Corsica. I also exclude municipalities exempt from the law (Paris, municipalities in the 3 départements around Paris, and a few islands) as well as municipalities located in mountain zones where the law applied differently.¹⁹

To make sure that control municipalities are not affected by their shift in integration status

¹⁸CEREMA stands for *Centre d'Etudes et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement*, which means in English "Center for Studies and Expertise on Risk, the Environment, Mobility and Development". It operates under the authority of the Ministry of Sustainable Development and of the Ministry of Territorial Cohesion.

¹⁹As noted, because those municipalities are typically small and far from each others, the 2010 law applied differently. First, the prefect had less leverage to force municipalities to enter an IC, as her decision had to be approved by the local "mountain committee," which slowed down the process: By 2013, only 60 percent of them were integrated, compared to 77 percent outside of mountain zones. Second, communities in mountain zones were not subject to the 5,000 threshold, meaning that isolated municipalities were free to create or join small ICs.

during the period of analysis, I restrict the control group to municipalities that were already integrated since 1999 and focus on the time period around the 2010 law, from 2004 to 2018. Hence, I exclude from the main sample of analysis the 45 percent of French municipalities that entered an IC between 1999 and 2010.²⁰ Note that the results are robust both in terms of significance and magnitude to varying the latest date of integration of the control municipalities, ranging from 1999 to including all municipalities already integrated in 2010 (see Appendix B1).

The final sample is composed of a balanced panel of 16,362 municipalities: 15,065 in the control group and 1,297 in the treatment group (8 percent). To illustrate the sample, Figure 2 plots an indicator variable equal to one if the municipality is part of an IC, separately for the control group (blue line) and the treatment group (red line). By construction, all municipalities in the control group belong to an IC since 1999 and over the whole period. In contrast, no municipality in the treatment group belongs to an IC before 2010. After the law, they gradually enter a community: 34 percent of the treated municipalities enter an IC during the first two years following the law, 43 percent in 2013, and the remaining 23 percent in 2014. By 2014, all municipalities from the sample are part of an IC.²¹

Table 1 provides some descriptive statistics on the two groups in 2010.²² Although it first seems that treated municipalities are smaller on average (first line), once I remove the 31 municipalities with a population higher than the one of the largest treated municipality from the control group, the average population size is similar in the two groups: 1,558 in the control group and 1,626 in the treatment group (second line).

Similarly, the population growth between 1999 and 2010, the population density, and the share of urban municipalities are quite comparable across the two groups. Turning to the population composition, the main difference between the two groups comes from the average share of executives in the municipality: 7.0 percent on average in the treatment group against 5.2 percent in the control group. Accordingly, residents are on average richer: The average annual taxable income per capita is 14,209 euros in 2010 in the treated municipalities versus 12,633 euros in the control municipalities. Note, however, that the standard deviation in the treatment group is large (4,432), indicating that it encompasses both rich and poor municipalities. Finally, treated municipalities are more likely to have a right-wing mayor (62.1 vs. 57.1), but the turnout rate in the 2008 municipal election or the probability of having a mayor not affiliated with any party (“NC,” standing for “non-classified”) is comparable.²³

Overall, the two groups are quite similar based on observables. This is reassuring for the

²⁰This definition of the sample also excludes less than 1 percent of municipalities whose integration status changed over the 1999-2010 period: Isolated municipalities in 2010 that were part of an IC at some point between 1999 and 2010, as well as municipalities that were part of an IC in 2010 but briefly isolated at some point between 1999 and 2010. The results remain unchanged if I include them.

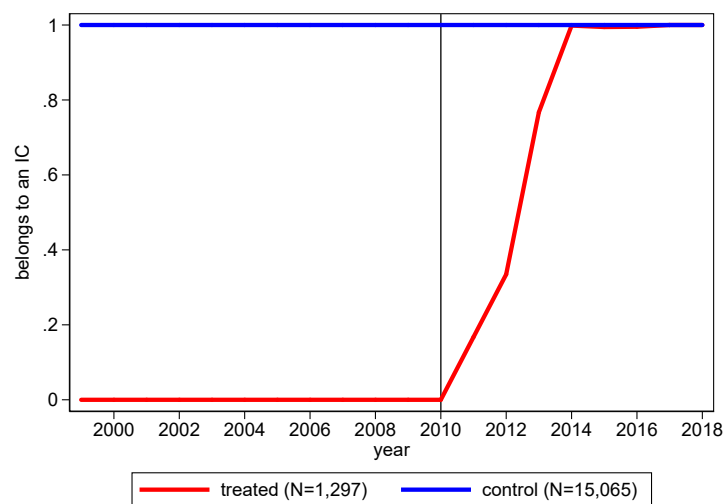
²¹There are only two exceptions: one municipality that integrated in 2015 and another in 2017. Both went to court to challenge the decision made by the prefect, but finally had to comply with it.

²²The t-tests of the differences between the control and treatment groups are displayed in Appendix B3, along with the matching exercise (see Section 3.2).

²³Descriptive statics using municipalities’ characteristics in 1999 instead of 2010 display comparable patterns (see Table A1 in the Appendix).

identification strategy, as we would not expect completely different groups to display parallel trends in the outcomes of interest. It also illustrates that there is not one type of resisting municipality, and that the explanation for why treated municipalities resisted integration cannot be based solely on their socio-demographic characteristics. Instead, it suggests that it is key to assess the impact of entering an IC on these municipalities to understand why they resisted integration.

Figure 2: Integration status



Notes: This graph plots the probability of being part of an intermunicipal community for the treatment group (red line) and the control group (blue line) separately.

Table 1: Descriptive statistics – 2010

	Control (N=15,065)				Treatment (N=1,297)			
	mean	sd	min	max	mean	sd	min	max
population	1,943	11,692	7	852,395	1,626	4,661	15	70,829
<i>w/out largest</i>	1,558	<i>4,114</i>	7	<i>66,002</i>	1,626	<i>4,661</i>	<i>15</i>	<i>70,829</i>
Δ population	0.09	0.14	-0.47	3.50	0.09	0.15	-0.41	1.18
density	155	449	1	9,876	161	540	2	6,931
urban	0.22	0.42	0.00	1.00	0.20	0.40	0.00	1.00
urban area	0.61	0.49	0.00	1.00	0.64	0.48	0.00	1.00
core urban area	0.12	0.32	0.00	1.00	0.09	0.29	0.00	1.00
immigrants	0.04	0.04	0.00	0.67	0.05	0.04	0.00	0.34
unemployed	0.09	0.04	0.00	0.42	0.08	0.04	0.00	0.57
child	0.08	0.02	0.00	0.23	0.07	0.02	0.00	0.21
farmers	0.04	0.05	0.00	0.60	0.04	0.05	0.00	0.44
executives	0.05	0.04	0.00	0.56	0.07	0.07	0.00	0.43
workers	0.17	0.07	0.00	0.67	0.15	0.08	0.00	1.00
retired	0.29	0.10	0.00	1.00	0.28	0.10	0.00	0.79
no diploma	0.19	0.07	0.00	0.61	0.17	0.07	0.00	0.46
baccalaureate	0.15	0.04	0.00	0.56	0.16	0.04	0.00	0.32
high education	0.07	0.05	0.00	0.54	0.09	0.07	0.00	0.51
residents' income	12,633	2,926	3,237	66,868	14,209	4,432	5,495	61,360
turnout mun	0.78	0.09	0.37	1.00	0.76	0.10	0.43	1.00
right-wing mayor	0.57	0.50	0.00	1.00	0.62	0.49	0.00	1.00
left-wing mayor	0.34	0.473	0.00	1.00	0.27	0.44	0.00	1.00
NC mayor	0.09	0.29	0.00	1.00	0.11	0.31	0.00	1.00
% vote right pdt	0.56	0.11	0.14	1.00	0.60	0.11	0.15	1.00
% vote far-right	0.13	0.05	0.00	0.52	0.14	0.06	0.00	0.47

Notes: Data on the municipal population, age, education, and occupational composition comes from the 2008 census, which applies to the 2010 year. The variation of the population (line 3) is computed as the variation in the number of inhabitants between the 1999 and 2008 census. Indicator variables for whether the municipality is urban, part of an urban area, or located in the urban core are based on the INSEE 2010 classification. The municipal turnout rate and political orientation of the mayors are based on the results of the 2008 municipal elections. The last two variables give the vote share at the municipal level of the right and the far-right during the 2007 presidential elections, in the second and first round respectively.

3.2 Specification and identification

I estimate the following specification for all municipalities in the sample over the 2004-2018 period:

$$Y_{mt} = \alpha + \beta 1_{\{t>2010\}} 1_{\{treated_m=1\}} + \delta_t + \theta_m + \varepsilon_{mt} \quad (1)$$

where m stands for the municipality and t for the year. Y_{mt} represents the outcome considered. Outcomes are standardized: I divide each outcome by the 2010 municipal population and then multiply by 10,000. Hence, Y_{mt} is, for instance, the number of housing building permits delivered in municipality m during year t per 10,000 inhabitants using the 2010 population. $1_{\{t>2010\}}$ is an indicator variable equal to one for years after the reform, starting in 2011. $1_{\{treated_m=1\}}$ is an indicator variable equal to one for municipalities that were not yet integrated in 2010 and thus forced to join an IC. δ_t and θ_m are time and municipality fixed effects, respectively. The inclusion of municipality fixed effects controls for any time-invariant unobserved factors, while the inclusion of year fixed effects captures changes over time that affect all municipalities the same way. Standard errors are clustered at the municipality level.²⁴

The identification assumption is that absent the law, municipalities would have evolved the same way in the control and in the treatment groups. Under this assumption, the main coefficient of interest β captures any deviation from a parallel evolution in the outcome of interest between the treatment and the control groups due to the 2010 law. I outline below the additional analyses and robustness checks I performed to alleviate the identification concerns and provide support for the common-trend assumption.

Selection into treatment. The first concern is the selection into treatment: Municipalities in the control group chose to integrate early, whereas municipalities in the treatment group chose to resist. The question is whether we can expect those two groups to display parallel trends in the outcomes of interest. First, as seen in Section 3.1, control and treated municipalities are quite similar based on socio-demographic characteristics. Moreover, to make sure that the results are not driven by differences in observables, I test the robustness of the main estimates to using propensity score matching. As shown in Appendix B3, matching leaves the significance and magnitude of the results almost unchanged.

Still, one could question whether control municipalities experience different trends because they are part of an intermunicipal community. Removing municipalities that integrated after 1999 alleviates the concern that control municipalities' shifts in integration status affect their trends during the period of interest. Nevertheless, being part of an IC might make control municipalities evolve differently across time and react differently to shocks. This is ultimately an empirical question, and while the parallel trend assumption cannot be directly tested, I run several analyses to provide support for it.

First, when presenting the results, I systematically test for the presence of pre-trends by plotting for each outcome the coefficients of the following leads-and-lags regression:

²⁴I also run specifications with standard errors clustered at the IC level, considering municipalities' IC either in 2014 or 2018. While, as expected, the standard errors are generally slightly larger, the significance of the main results is left unchanged (see Appendix B2).

$$Y_{mt} = \alpha + \sum_{k=2004}^{2018} \beta_k 1_{\{t=k\}} 1_{\{treated_m=1\}} + \delta_t + \theta_m + \varepsilon_{mt} \quad (2)$$

where $1_{\{t=k\}}$ is an indicator variable equal to 1 for year k . All coefficients are normalized relative to 2010. On top of visually inspecting the pre-trends, I also test for the significance of the pre-treatment estimates as well as their joint significance. Second, I estimate the impact of a series of placebo reforms on the outcomes of interest by pretending that treated municipalities were forced to enter an IC before 2010 and excluding the actual treatment period from the estimation (see Figure B4). The results of the placebo exercise are discussed in the main text for each outcome.²⁵

Control municipalities affected. We can interpret the estimated β of equation (1) as the causal impact of integration on the treated municipalities, provided that only municipalities in the treatment group are affected by the law. However, some municipalities in the control group might also be impacted—namely, those that are part of an intermunicipal community where a treated municipality entered, as well as those that are part of ICs that fall below the 5,000-inhabitant threshold (and that had thus to enter larger ICs or whose IC had to merge with another one). Appendix B5 reports the main results using an alternative control group that excludes all municipalities whose IC changed as a result of the 2010 law. I also test the robustness of the results to excluding all control municipalities sharing a common border with a treated municipalities, to address potential spillovers. The effects are very similar, suggesting that the estimated effects capture the impact on the treated municipalities only.

4 Main results: the costs of integration

4.1 Impact on housing supply

4.1.1 Impact on municipalities forced to integrate

In this section, I estimate the impact of integration on the number of housing building permits delivered in municipalities forced to enter an intermunicipal community. The outcome is the number of housing units allowed for construction in the municipality in a given year per 10,000 inhabitants (using the 2010 population).

Figure 3 plots the coefficients of the leads-and-lags regression, each coefficient corresponding to the interaction between an indicator variable for the year and an indicator variable equal to one for municipalities in the treatment group. Vertical lines represent the 95-percent-confidence intervals. First, no coefficient before 2010 is significant, and all coefficients prior to the law are close to 0 in magnitude. The F-stat for the joint significance of the pre-reform

²⁵Unfortunately, I cannot run the placebo tests on daycare and public libraries, given the few pre-treatment periods I have in the data. This analysis is thus restricted to the number of building permits, fiscal revenues, and public transport.

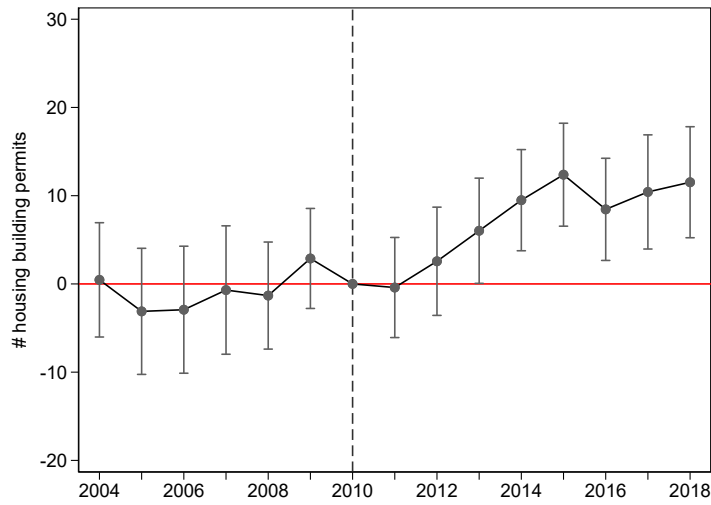
estimates is equal to 0.670 (p-value 0.673), confirming the absence of pre-trend. Instead, after the law, we observe a large increase in the number of building permits delivered in treated municipalities compared to control municipalities. The increase starts in 2012, consistent with the fact that most municipalities forced to integrate joined an IC after 2011.²⁶

Table 2 provides the formal estimate. On average, integration into an intermunicipal community increased by 8.2 the number of housing building permits delivered per 10,000 inhabitants per year in treated municipalities (column 1). The estimate is significant at the 1-percent level. As, on average, municipalities that resisted integration were delivering 66.2 building permits per year per 10,000 inhabitants before 2010, this represents an increase of 12.4 percent. The size of the impact is very similar when using propensity score matching (+7.9), and the estimate remains significant at the 1-percent level (see Table B3.3 in Appendix B3). Moreover, in line with the absence of pre-trends, Figure B4 in Appendix B shows that none of the placebo reforms have a significant effect on housing supply, supporting the fact that the estimate is capturing the impact of the 2010 law rather the impact of factors that systematically affect treated and control municipalities differently.

As shown in columns 2 and 3, the increase is mainly driven by new constructions: On average, municipalities forced to enter an IC experienced an increase of 6.8 in the number of building permits delivered to build new housing units on empty land, while they experienced an increase of only 1.4 in the number of building permits delivered to expand existing units. This result shows that resisting municipalities had to expand their share of built land once integrated, which is consistent with municipalities losing control over both the number of building permits delivered and the land-use policy. As nothing prevented municipalities from building more before the law, this rise in construction is unlikely to reflect their preferences; it is best interpreted as a cost of their loss of autonomy.

²⁶Moreover, when splitting the treatment group according to the exact year treated municipalities entered an IC, we see that the increase starts after their integration, confirming that the effect captures the impact of entering an IC rather than the influence of other factors impacting the treated and control groups differently after 2010 (see Figure A1 in the Appendix).

Figure 3: Impact on housing building permits



Notes: This graph plots the estimates and 95-percent confidence intervals from the leads-and-lags regression (see equation (2)).

Table 2: Impact on housing building permits

	(1)	(2)	(3)
Outcome	Number of building permits per 10,000 inhabitants		
	All	New	Extensions
Treatment	8.228*** (1.609)	6.862*** (1.477)	1.366*** (0.520)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	245,430	245,430	245,430
Mean DepVar	66.203	55.154	11.049
Sd DepVar	94.440	86.470	35.995

Notes: Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. The last two columns distinguish between two types of building permits: whether the permit allows the construction of a new unit on empty land (column 2) or the extension of an existing housing unit (column 3).

4.1.2 Impact on municipalities that voluntarily integrated

I now compare this effect with the impact of integration for municipalities that joined an IC voluntarily before the law. If the increase in housing supply explains why treated municipalities resisted, we should see a differential impact for municipalities that instead decided to integrate.

I consider municipalities that were previously excluded from the main sample of analysis: municipalities that joined an IC between 2000 and 2010.²⁷ I focus on metropolitan France and consider all municipalities that voluntarily integrated, including municipalities located in the Paris area, islands, and mountain zones.²⁸ This results in a balanced panel of 14,702 municipalities. To estimate the impact of their voluntary integration, I use a staggered adoption design, where the date of the treatment is the year when the municipality first joined an IC. I run the analysis over the period from 1999 to 2018: In 1999, no municipality in this sample is part of an IC, whereas all are part of a community starting in 2010. Hence, I observe municipalities up to 11 years prior to the integration (for municipalities integrating in 2010) and up to 18 years after the integration (for municipalities integrating in 2000). I estimate the following dynamic specification:

$$Y_{mt} = \sum_{d=-5}^{d=5} \beta_d 1_{\{t=t_{m0}+d\}} + \gamma_1 1_{\{t < t_{m0}-5\}} + \gamma_2 1_{\{t > t_{m0}+5\}} + \delta_t + \theta_m + \varepsilon_{mt} \quad (3)$$

where the year of integration of municipality m is denoted t_{m0} , and d indexes time-to-integration (negative before integration and positive after). The β_d s measure the difference between municipalities that are part of an IC and municipalities not already part of a community, for each of the 5 years preceding and following the integration. γ_1 (resp. γ_2) estimates the effect for being more than 5 years before (resp. after) the integration. All coefficients are normalized relative to the year preceding the integration ($d = -1$). As in the last estimations, regressions include time and municipality fixed effects, and the standard errors are clustered at the municipality level.

In light of the recent literature focusing on the issues associated with event study design (Borusyak and Jaravel, 2017; Goodman-Bacon, 2018; de Chaisemartin and D'Haultfoeuille, 2019), I also use an alternative estimation procedure developed by de Chaisemartin and D'Haultfoeuille (2019). As shown by the authors, two-way fixed effect regressions identify weighted sums of the average treatment effects in each group and period, with weights that may be negative, making the results often not robust to heterogeneous treatment effects across groups and over

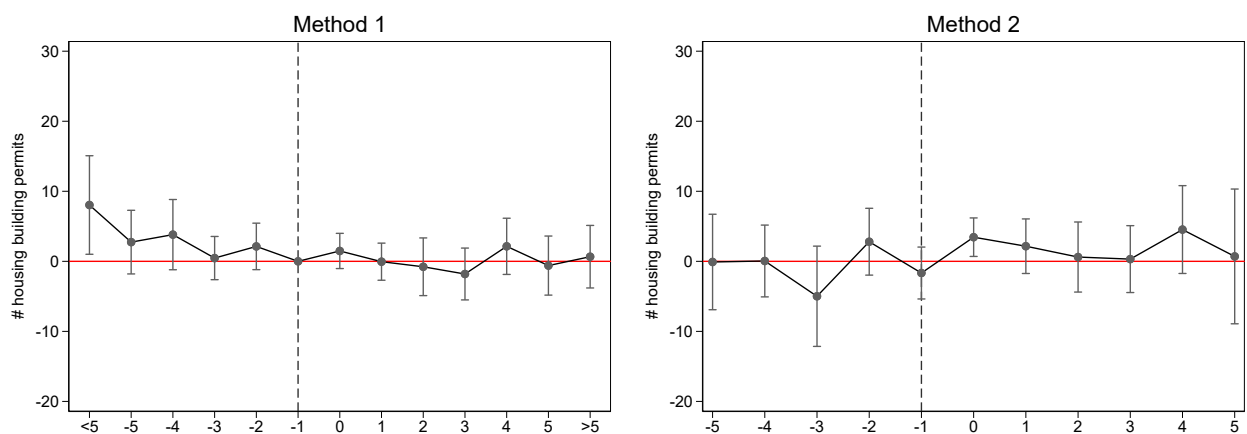
²⁷Because the municipal composition of French ICs is only available from 1999, we cannot know which year municipalities that integrated earlier entered an IC. I thus focus only on municipalities joining an IC between 2000 and 2010. Also, as the structure of ICs changed following the Chevènement law in 1999, municipalities that integrated voluntarily after 1999 entered ICs more comparable to the ICs municipalities forced to enter joined, making the comparison more relevant.

²⁸In the main analysis, I exploit the 2010 law that forced isolated municipalities to enter an IC in order to assess the impact of integration on resisting municipalities. I thus had to exclude law exceptions: municipalities located in the Paris area, islands, and mountain zones. In this section I assess the impact of voluntary integration, and I can thus consider all municipalities that decided to join an IC.

time. Their new estimand deals with this issue. In the context of staggered adoption designs, the coefficient at the date of treatment ($t = 0$) is estimated by comparing trends between $t - 1$ and t for units that switched to treatment in t compared with units that are not yet treated in t . Similarly, the coefficient at $t = 1$ is obtained by comparing trends between $t - 1$ and $t + 1$ for units that switched to treatment in t compared with units not already treated in $t + 1$, and so on. To implement their estimation, one group needs to remain untreated during the whole period. I thus run the analysis from 1999 to 2009 only, so that municipalities that integrate in 2010 are never treated during the period of analysis.²⁹

Figure 4 plots the coefficients, taking as outcome the number of housing building permits delivered per year per 10,000 inhabitants³⁰ and applying the same scale as in Figure 3. The left-hand side graph shows what I obtain using the regular staggered adoption design (referred to as method 1) and the right-hand side graph what I obtain using de Chaisemartin and D’Haultfoeuille (2019)’s method (referred to as method 2).

Figure 4: Impact on housing building permits for municipalities voluntarily integrating



Notes: This graph plots the estimates and 95-percent confidence intervals from the linear regression based on (3) (left-hand side graph) and using de Chaisemartin and D’Haultfoeuille (2019)’s method, implemented using the Stata command `did_multipleGT`, available on SSC repository (right-hand side graph).

²⁹For more details on de Chaisemartin and D’Haultfoeuille (2019)’s method for staggered adoption design, see Section 5.2 of their paper. Note that this method allows me to compute the estimates of being a given number of years before or after the treatment, but not to compute the estimates for being more than 5 years before or after the date of the treatment, which are thus absent in the right-hand side graph below.

³⁰As I consider municipalities integrating from 2000, I divide the outcome by the 1999 population instead of the 2010 population.

Using the first method, no coefficient, except the one associated with being more than 5 years before the integration, is significant at the 5- or even 10-percent level. In particular, no coefficient after the integration is significant, and all are close to 0 in magnitude. The test for the joint significance of the pre-integration estimates confirms the absence of pre-trends (F-stat of 1.78 with a pvalue of 0.11), and the test for the joint significance of post-integration estimates confirms the absence of impact (F-stat of 1.39 with a pvalue of 0.20). On the second graph, only the coefficient at $t = 0$ is significant at the 5-percent level, but the estimated effect is small in magnitude (3.4), and no other post-treatment estimate is significant. We can thus conclude that joining an IC had no significant impact on housing supply for municipalities that integrated voluntarily.

Hence, only municipalities that did not want to enter an IC experienced a rise in construction. This differential impact supports the view that treated municipalities opposed integration to avoid a rise in housing supply. It also suggests that mayors' decision to resist was driven by actual costs of integration, beyond purely ideological or political considerations. Section 6 discusses further the differential impact between resisting municipalities and those that voluntarily integrated and provides additional evidence supporting this interpretation.

4.1.3 Further evidence on the costs linked to housing supply

As noted in the literature on housing restrictions, several costs can explain why municipalities would oppose new constructions, depending on their characteristics. First, construction may create congestion in high-demand and densely-built places by limiting the space available and increasing the number of people using local amenities (Saiz, 2010; Hilber and Robert-Nicoud, 2013). Second, it might increase the population's heterogeneity in municipalities surrounded by different neighbors by bringing in poorer households and/or minorities (Rolleston, 1987; Bates and Santerre, 1994). Third, through its effect on housing prices, new constructions might decrease the value of residents' assets in municipalities with a large share of homeowners (Fischel, 2001; Ortalo-Magné and Prat, 2014).

To determine which of these costs best explain municipalities' opposition to construction, I explore which types of municipalities are driving the impact on housing supply.³¹ To do so, I perform heterogeneity analyses along several municipalities' characteristics measured in 2010, prior to the reform. For clarity, I present the graphs and tables obtained by splitting the sample at the median value of the heterogeneity variable. In Appendix C, I also report the results from the estimation of equation (1) where I add an interaction term between the treatment variable and the continuous heterogeneity variable. The corresponding coefficient measures the change

³¹The rationale for this test is the following: Places driving the effect are those that are the most likely to have resisted cooperation to avoid an increase in housing supply. Depending on the characteristics of those places, I can infer which of the costs associated with construction they were trying to avoid by resisting cooperation. For instance, if municipalities driving the results are mostly high-demand and densely-built places, rather than municipalities surrounded by poor neighbors, it is more likely that municipalities resisted an increase in construction to avoid congestion rather than to keep poorer people out.

in the treatment effect coming from an increase of 1 standard deviation in the heterogeneity variable.

Congestion and NIMBY costs. If congestion costs explain why municipalities were opposed to a rise in housing supply, we should see that the effect is stronger in high-demand and densely-built municipalities, where new construction is the most likely to increase congestion. Indeed, as pointed out by Gyourko and Molloy (2015), more desirable locations will have a larger share of developed land and, consequently, will be more likely to oppose new constructions in order to prevent further development from raising congestion disamenities.

Consistent with the monocentric model,³² the demand for housing in French municipalities is the highest for those municipalities closer to the core of the urban area (Combes et al., 2018), reflecting how connected the municipality is to the local center of employment.³³ Table 3 and Figure A2 in the Appendix display the impact on housing supply depending on whether the municipality is part of an urban area and, moreover, whether it is located in the core of the urban area.

As shown in columns 2 and 3, the entire effect is driven by treated municipalities that are part of an urban area: After integration, they experienced an average increase of 10.6 in the number of building permits delivered per year per 10,000 inhabitants, an effect significant at the 1-percent level, while municipalities outside any urban area experienced an average increase of 3.1 only, which is not significant. The effect is even stronger for municipalities located in the core, as the number of buildings permits per 10,000 inhabitants increased by 24.9 on average per year (37.7 percent), an impact significant at the 1-percent level. In line with this result, Table C1 in Appendix C shows that a shorter Euclidean distance to the core is indeed associated with a stronger effect of integration on housing supply. Finally, using residents' income as an alternative measures of desirability, I also find that the effect is stronger in municipalities where residents are richer on average (see Figure A3 and Table A2 in the Appendix and Tables C2 in Appendix C).

Focusing on municipalities inside an urban area, I next look at whether the effect is stronger the more densely built the municipality is. Table 4 and Figure A4 in the Appendix show the impact on housing for rural and urban municipalities separately, the definition of an urban municipality being based on the share of built land and population density.³⁴ Although the coef-

³²The monocentric model predicts that the housing demand is the highest in the center (Alonso, 1964). This is consistent with most housing markets in developed countries, with the notable exception of the US, where the more desirable places are usually located in the suburbs (see, for instance, Jackson, 1987).

³³Close to the definition of American metropolitan areas, a French urban area is defined as a group of neighboring municipalities encompassing an urban core (urban unit) providing at least 5,000 jobs, and by rural districts or urban units (urban periphery) among which at least 40 percent of the employed resident population works in the core or in the municipalities attracted by this core. I use the 2010 urban area breakdown and consider both large (providing at least 10,000 jobs) and medium (providing between 5,000 and 10,000 jobs) urban areas. I test the robustness to using the 1999 breakdown, and the results are left unchanged. In 2010, France counted 372 urban areas, representing two thirds of the municipalities and 85 percent of the population.

³⁴Based on international standards, an urban municipality is a municipality belonging to an urban unit, defined as a continuously built up zone (no cut of more than 200 meters between two constructions) counting at least 2,000 inhabitants.

ficients for urban municipalities are noisier—urban municipalities represent around 27 percent of the sample—the increase after 2010 is larger in magnitude: The number of building permits per 10,000 inhabitants increased by 15.9 per year in urban municipalities forced to integrate (23.6 percent, column 2) against 9.9 in rural municipalities (15.3 percent, column 3). I further split municipalities that are part of an urban area according to a continuous measure of housing density, equal to the number of housing units per square km in 2010. As shown in columns 3 and 4 and on Figure A4 in the Appendix, the impact is higher for municipalities above the housing density median: They experienced an increase of 14.1 building permits per 10,000 inhabitants per year (column 5), against 9.4 for treated municipalities below the median (column 4). Table C3 in Appendix C indicates that a one-standard-deviation increase in housing density increases the treatment effect by 8.6. This corresponds to increasing the effect by more than half compared with the impact on housing for an average treated municipality in the urban area.

Table 3: Impact on housing supply – urban area

Outcome	(1)	(2)	(3)	(4)
	Number of building permits per 10,000 inhabitants			
	All	Outside	Inside	Core
Treatment	8.228*** (1.609)	3.134 (2.733)	10.634*** (1.929)	24.888*** (5.622)
Municipality FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	245,430	94,305	151,125	28,665
Mean DepVar	66.203	67.802	65.288	66.107
Sd DepVar	94.440	102.100	89.761	108.505

Notes: Column 2 (resp. 3) includes only those municipalities that are not part (resp. part) of an urban area. Column 4 considers only those located in the core of the urban area. The municipal composition of urban areas is based on INSEE 2010 breakdown and I consider both large (providing at least 10,000 jobs) and medium (providing between 5,000 and 10,000 jobs) urban areas. Other notes as in Table 2.

Table 4: Impact on housing supply inside urban areas – housing density

	(1)	(2)	(3)	(4)	(5)
Outcome	Number of housing building permits per 10,000 inhabitants				
	All	Rural	Urban	Housing density median	
				Below	Above
Treatment	10.634*** (1.929)	9.939*** (2.102)	15.902*** (3.971)	9.415*** (2.395)	14.139*** (2.919)
Mun. FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Obs.	151,125	105,180	45,945	75,540	75,585
Mean	65.288	64.490	67.469	66.330	64.062
Sd	89.761	85.479	100.535	89.241	90.370

Notes: The sample includes only municipalities inside an urban area. Columns 2 to 5 add further restrictions. Among them, column 2 (resp. 3) includes only rural (resp. urban) municipalities, and column 4 (resp. 5) includes only municipalities with a housing density in 2010 below (resp. above) the median. Other notes as in Table 2.

Neighbors' characteristics. If municipalities were resisting new construction to avoid different people coming in, we should see that the effect is stronger for municipalities surrounded by neighboring municipalities that are more different from them.

I consider several heterogeneity dimensions: residents' average annual taxable income, the share of unemployed workers, and the share of immigrants. For each, I construct a ratio equal to the value for the municipality divided by the average value in neighboring municipalities, weighted by their population. For instance, the immigration ratio indicates whether the proportion of immigrants in a given municipality in 2010 was greater than, less than, or equal to the average proportion in surrounding municipalities. I define surrounding municipalities as encompassing all other municipalities from the same département. The results are robust if I instead define surrounding municipalities as direct neighbors—i.e., municipalities sharing a common border.

Table 5 reports the impact on housing supply depending on whether the municipality is above or below the median value of the ratio. If municipalities were resisting new construction to prevent poorer people from coming in, we should see a stronger effect the larger the average resident's income in the municipality compared with its neighbors, and thus the higher the income ratio. Instead, the two coefficients are close and, if anything, the impact is even larger for municipalities below the median (8.9 vs. 7.3, columns 2 vs. 3). Conversely, if municipalities resisted because they expected more immigrants or unemployed workers to come, we should see that the effect is stronger the smaller the share of immigrants or unemployed in the municipality compared with its neighbors, and thus the smaller the ratios. Again, the effect is actually goes in the opposite direction, as the impact is larger above the median (columns 4 to 7).

Table 5: Impact on housing supply – neighbors’ characteristics

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Number of housing building permits per 10,000 inhabitants						
	All	Median income ratio		Median immigrants ratio		Median unemployed ratio	
		Below	Above	Below	Above	Below	Above
Treatment	8.228*** (1.609)	8.934*** (2.139)	7.349*** (2.285)	6.483*** (2.008)	9.962*** (2.542)	4.406** (2.216)	12.042*** (2.333)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	245,430	123,195	121,995	122,445	122,985	122,535	122,895
Mean DepVar	66.203	59.188	70.940	61.057	71.719	68.351	64.032
Sd DepVar	94.440	80.847	101.636	87.041	101.496	97.203	91.523

Notes: For each municipality, the ratio divides the value of the heterogeneity variable in the municipality by the average value in the other municipalities from the same département, weighted by their population. Column 2 (resp. 3) includes only municipalities below (resp. above) the median value of the ratio using the residents’ average annual taxable income. Data on taxable income in 2010 are missing for the 16 municipalities with the smallest populations, in order to protect residents’ privacy. Column 4 (resp. 5) includes only municipalities below (resp. above) the median value of the ratio using the share of immigrants in 2010. Columns 6 and 7 repeat the same exercise using the share of unemployed in 2010. Other notes as in Table 2.

The effect is even surprisingly large for municipalities where the share of unemployed workers is higher compared to neighboring municipalities (column 6). However, the estimate obtained by interacting the treatment variable with the unemployment ratio is small and not significant, indicating that this result should not be overinterpreted. Similarly, the interactions with the income and immigrant ratios provide small and non-significant estimates (see Table C4 in Appendix C). Overall, these results suggest that municipalities whose neighbors are more different in terms of income level, immigration, and unemployment status are not driving the effect.³⁵

Next, I consider two measures related to political heterogeneity. I first look at whether, in 2010, the mayor had the same political orientation as the member of parliament elected by the legislative district the municipality belongs to. If not, this indicates that the municipality was not politically aligned with the median voter of its district.³⁶ Second, I consider another measure based on the share of voters in the municipality that voted for the right-wing candidate in the second round of the 2007 presidential election, where a candidate from the right faced a

³⁵Here, I consider all immigrants no matter their country of origin. The effects are similar if I consider non-European immigration only.

³⁶The last parliamentary election before 2010 took place in 2007. At this time, metropolitan France counted 551 legislative districts (*circonscriptions*), each encompassing 128 municipalities on average. In 2010, 58 percent of the mayors had the same political orientation as the member of parliament representing them (63 percent in the treatment group).

candidate from the left. I compute the absolute value of the difference between the right vote share in the municipality and the average right vote share in neighboring municipalities. The higher this difference, the less politically aligned the municipality is with its neighbors. As shown in Table A3 in the Appendix, the impact on housing is similar whether the mayor shares the same orientation as the member of parliament or not, and even a bit larger when she does (9.1 vs. 7.4, columns 2 and 3). Moreover, while the impact is slightly larger for municipalities above the median value of the absolute difference in terms of right vote share (9.6 vs. 6.4, columns 4 and 5), the coefficient corresponding to the interaction of the heterogeneity variable with the treatment variable is small and not significant (Table C4 in Appendix C). This suggests that municipalities less politically aligned with their neighbors are not driving the effect either.

Homeowners and housing prices. Finally, if homeowners' fear of a housing price decline explains why municipalities resisted new construction, we should see that the effect on housing is higher the larger the share of homeowners in the municipality. Instead, the impact of integration on housing supply is not stronger for treated municipalities above the median (Table 6), and the interaction between the treatment variable and the share of homeowners is negative and not significant (Table C5 in Appendix C).

Table 6: Impact on housing supply – share of homeowners

Outcome	(1)	(2)	(3)
	Number of housing building permits All	Median % homeowners	
		Below	Above
Treatment	8.046*** (1.607)	9.079*** (2.848)	7.891*** (1.892)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	245,430	117,405	128,025
Mean DepVar	66.203	72.555	62.046
Sd DepVar	94.440	102.776	88.325

Notes: Column 4 (resp. 5) includes only municipalities below (resp. above) the median value of the share of homeowners in 2010. Other notes as in Table 2.

These results are consistent with the lack of empirical evidence supporting the hypothesis that housing regulations are stronger in places with a higher rate of homeownership (e.g., Brueckner, 1998; Glaeser and Ward, 2009; Hilber and Robert-Nicoud, 2013), as well as with recent evidence suggesting that renters can express as much NIMBYism as homeowners, particularly in high-rent places (Hankinson, 2018).

I further investigate the actual impact of integration on municipalities' housing prices. To do so, I use indices based on official housing transactions, which can be interpreted as the price per square meter of a reference house or apartment (see Section 2.3 and Appendix D for more details). As data on housing transactions for the Parisian region of Ile-de-France and for the rest of France come from two distinct databases using different measures for housing characteristics, I present the results for the two parts of France separately. I have transaction data for every even year until 2014 for the Parisian region and until 2016 for the rest of France. Since housing prices are forward looking (Chapelle and Eyméoud, 2019), we can expect them to reflect any changes in expectations about housing value shortly after the 2010 reform. I consider a balanced sample of municipalities, restricting the analysis to municipalities where at least one transaction took place during every even year. This gives a sample of 455 municipalities in the Parisian region and 6,732 in the rest of France (43.9 percent of the initial sample in total). Considering instead all municipalities with at least one transaction during the period of interest almost doubles the sample size and leaves the results unchanged (Figure D1 in Appendix D).

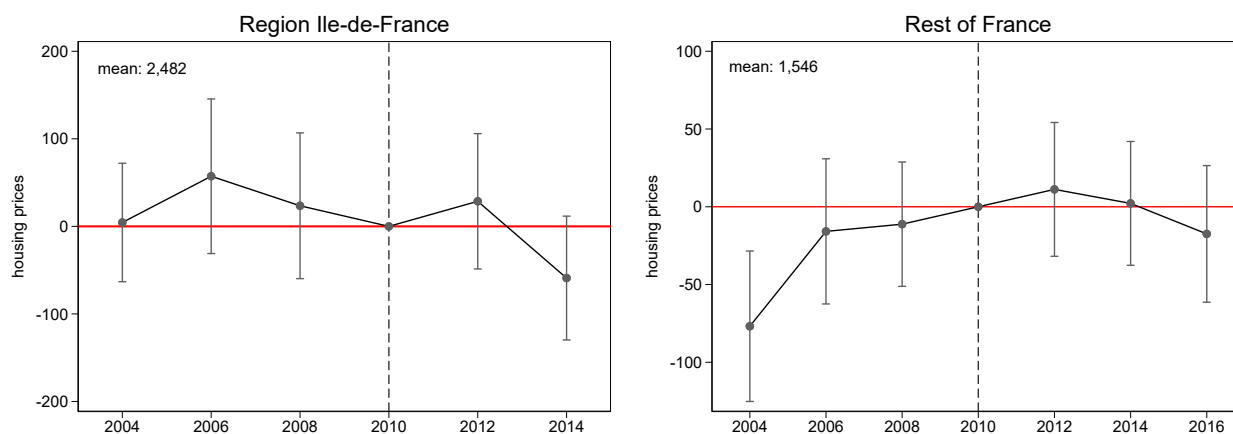
As shown in Figure 5, we see a small decrease in housing prices in 2014 for treated municipalities located in the Parisian region, but the estimated impact is not significant and small in magnitude (-36 euros per square meter). Compared to the average price in the Parisian region in the treatment group before the law (2,482), this would represent a decrease of only 1.5 percent. The decrease is even smaller for the rest of France, suggesting that overall, integration had no significant impact on housing prices for municipalities forced to join an IC. These results are robust to using an alternative version of the indices where I include a larger set of houses' and apartments' characteristics in the hedonic regressions (see Figure D2 in Appendix D). The results also remain similar when restricting the analysis to municipalities inside the urban area, which are driving the impact on building permits (see Figure D3 in Appendix D).

The absence of effect on prices do not rule out the homeowner hypothesis: They may have wrongly believed that prices would go down and, as argued by Fischel (2001), homeowners are above all risk averse and thus tend to oppose any housing projects, even those that may end up having positive effects on housing value growth. Nevertheless, the facts that the effect does not vary with the homeownership rate and that we do not see a significant change in prices suggest that municipalities were not resisting mainly to avoid a price decline. Moreover, the absence of effect on prices is consistent with high-demand places driving the impact on construction: Increasing the housing supply in places where the demand is the highest is less likely to lead to a decrease in prices.

To conclude, the heterogeneity analysis along with the estimated effect on housing prices

suggest that municipalities that opposed integration to keep control over their housing supply are mainly urban municipalities trying to avoid congestion costs, rather than municipalities fearing an increase in population heterogeneity or a housing price decline.

Figure 5: Impact on housing prices



Notes: The sample includes only municipalities where at least one housing transaction took place each even year over the period studied. The graph on the left-hand side includes only municipalities located in the Parisian region Ile-de-France, while the graph on the right-hand side includes all the other municipalities. On each graph, the average value of the price index in the treatment group before 2010 is displayed on the top left corner. Other notes as in Figure 3.

4.2 Impact on local public services

In order to achieve economies of scale, ICs seek to rationalize the offer of local public services and will thus tend to concentrate resources on a subset of public service facilities. In particular, they are likely to concentrate resources on facilities located in central and densely-populated areas, closing facilities or preventing new ones from opening in other areas. As a result, low-density municipalities may lose local public service facilities when entering a community, which is costly, as it increases the distance to public services for their residents.

To explore the consequences of integration on the presence of local public services, I gathered data at the municipal level on two different local public services: daycare facilities and public libraries. For both, I look at whether municipalities forced to join an IC experienced a loss in the number of facilities located within their territory, taking as outcome the number of facilities per 10,000 inhabitants (using the 2010 population). Note that the data available are less comprehensive than the ones on building permits for two reasons. First, data on daycare and libraries start in 2007 and 2009 respectively, which does not leave enough time before the law to compare the main results with the effects of integration for municipalities that voluntarily integrated between 2000 and 2010. Moreover, data on libraries are available for 7 départements only.³⁷ Nevertheless, the results go in the exact same direction for both daycare and libraries,

³⁷The 7 départements are located in different parts of the countries: Aisne (in the north of France), Finistère

strengthening the conclusions we can draw about the effects of integration on the presence of local public services in municipalities forced to integrate.

In Figure 6 as in Figure 7, the graph on the full sample suggests that, on average, municipalities forced to integrate experienced a decrease in the number of public service facilities compared to control municipalities. Moreover, in both cases, the entire effect is driven by rural municipalities: While the number of daycare facilities and public libraries decreases in rural municipalities after 2010, we do not see any decline for urban municipalities and, if anything, the trend is even slightly increasing after the law.

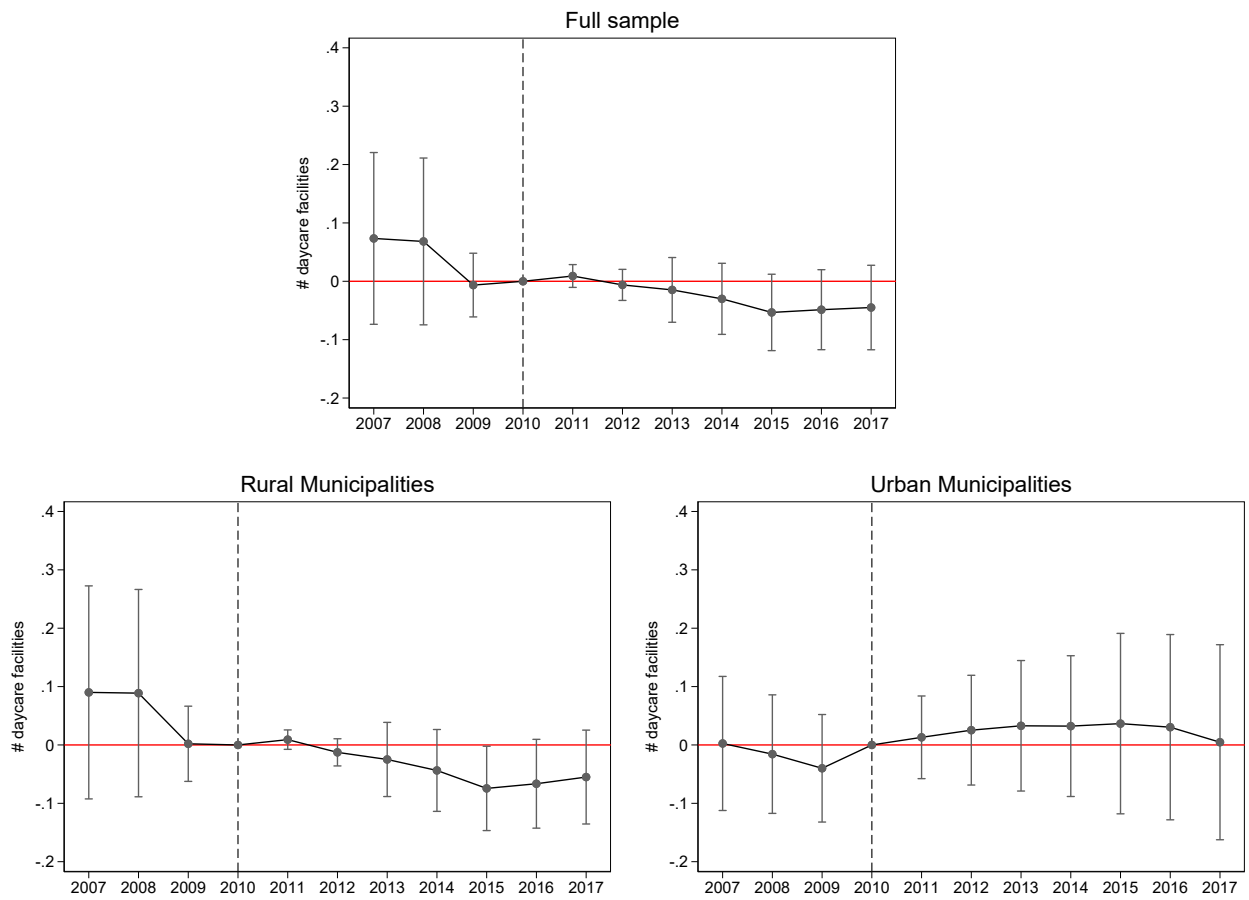
Table 7 provides the estimates. Even if no coefficient is significant at the standard levels for daycare facilities (columns 1 to 3), in line with the graphical evidence, the coefficient for rural municipalities is negative, whereas the coefficient for urban municipalities is positive. Taken at face value, the estimate suggests that in the average year post 2010, rural municipalities forced to integrate had 38.6 percent less daycare facilities located within their territory compared to rural control municipalities. The patterns are similar if we take as outcome the total number of daycare spots available at the municipal level rather the number of facilities (see Figure A5 and Table A4 in the Appendix).

Turning to public libraries (columns 4 to 6 of Table 7), the coefficients corresponding to the full sample and to rural municipalities are both negative and significant at the 10-percent level, while the coefficient for urban municipalities is close to 0 and not significant. The magnitude of the impact indicates that, in the average year after 2010, rural municipalities forced to integrate had 1.1 public libraries less per 10,000 inhabitants, corresponding to a decrease of 21.3 percent. Focusing on the last year of analysis, the estimate suggests that, in 2017, the number of public libraries per 10,000 inhabitants was lowered by 2.1 (39.6 percent) in rural municipalities forced to integrate compared to rural control municipalities.

These findings are consistent with ICs concentrating resources in denser municipalities and help explain rural municipalities' resistance to integration.

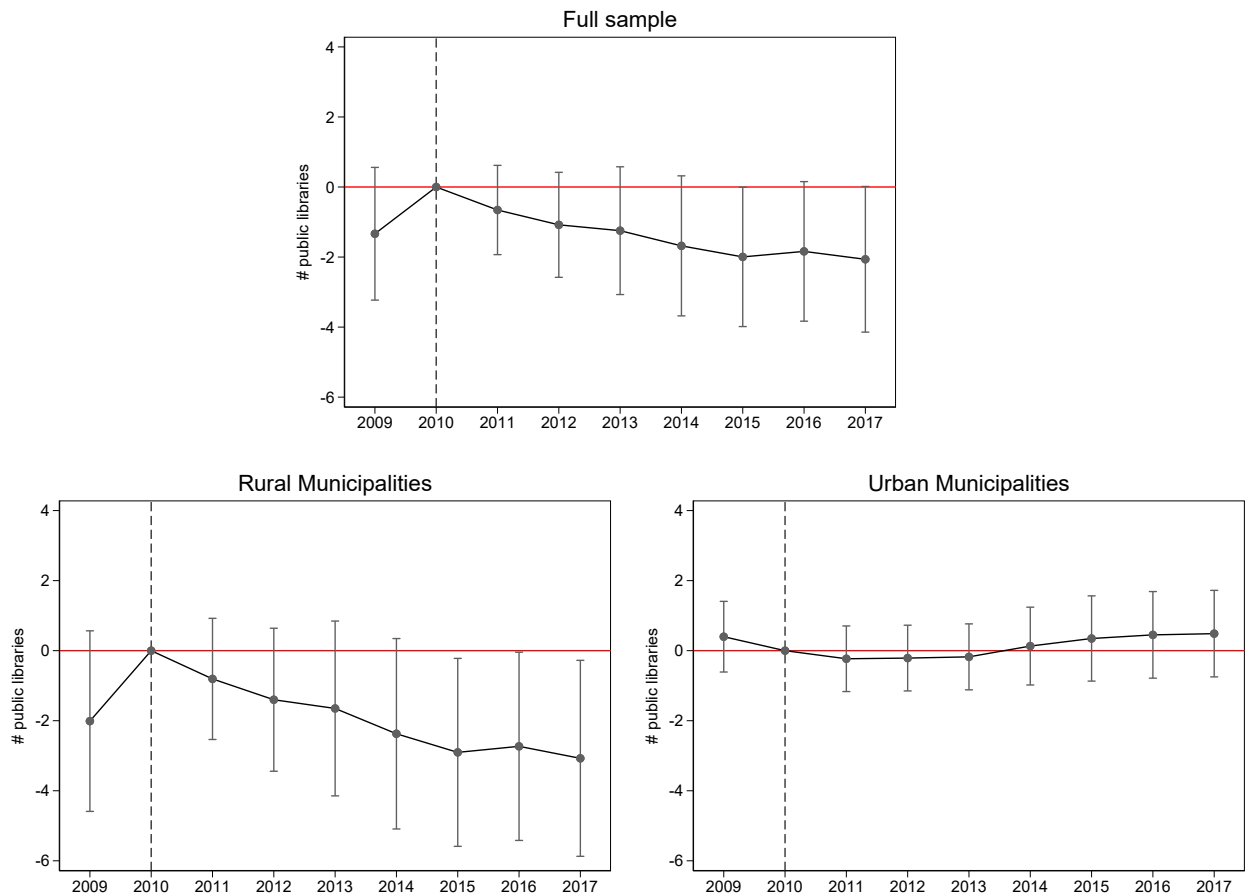
(west), Drôme (southeast), Gironde and Dordogne (southwest), and Essonne and Val d'Oise (center north). They represent 9 percent of the municipalities in the main sample of analysis.

Figure 6: Impact on daycare facilities



Notes: The graph on the bottom left-hand side includes only rural municipalities, while the graph on the bottom right-hand side includes only urban municipalities. Other notes as in Figure 3.

Figure 7: Impact on public libraries



Notes as in Figure 6.

Table 7: Impact on local public services

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	Number of facilities per 10,000 inhabitants					
	Daycare facilities			Public libraries		
	All	Rural	Urban	All	Rural	Urban
Treatment	-0.061 (0.046)	-0.083 (0.055)	0.038 (0.063)	-0.841* (0.465)	-1.129* (0.627)	-0.086 (0.312)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	179,982	140,272	39,710	12,825	9,900	2,925
Mean	0.438	0.210	1.326	4.395	5.302	2.045
Sd	2.174	2.211	1.762	9.611	11.026	3.095

Notes: In columns 1 to 3, the outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. In columns 4 to 6, the outcome is the number of public libraries located in the municipality per 10,000 inhabitants (using the 2010 population), the period of analysis is 2009-2017, and the sample is restricted to 7 départements for which data are available starting in 2009. Columns 2 and 5 include only rural municipalities, while column 3 and 6 includes only urban municipalities. Other notes as in Table 2.

5 Evidence on the benefits of integration

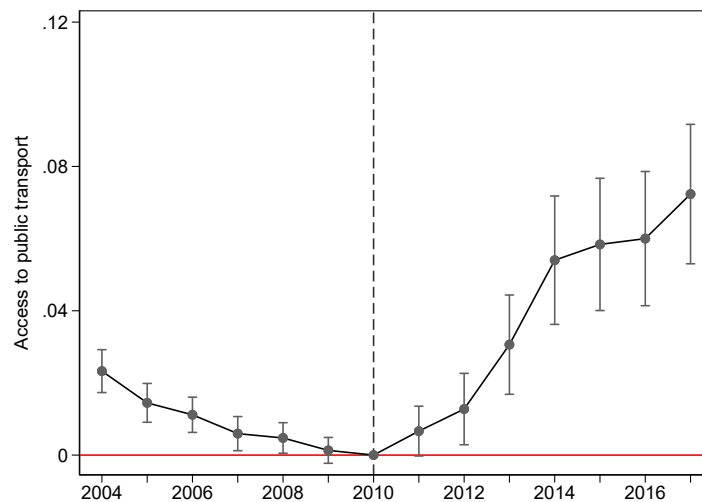
Results so far suggest that urban municipalities resisted cooperation to prevent more construction from raising congestion, while rural municipalities resisted to avoid losing local public services. This section explores whether, on top of experiencing costs of integration, resisting municipalities also benefited less from eventual gains of cooperation, which would contribute to explaining why they resisted longer. I first investigate the impact on public transport, a large-scale public service that integration is likely to improve, and then turn to the impact on fiscal revenues.

5.1 Impact on public transport

By enhancing cooperation and enabling municipalities to pool resources, integration is likely to help neighboring municipalities build larger and more efficient public transport networks. As a result, joining an intermunicipal community might increase a municipality's probability of being connected. To investigate whether municipalities forced to join an IC enjoyed such a benefit, I estimate the impact of forced integration on an indicator variable equal to one if the municipality has access to public transport. This analysis excludes municipalities located in the Parisian region of Ile-de-France, for which data on public transport are not available (representing 20.2 percent of treated municipalities).

As shown in Figure 8, pre-reform coefficients are significant, and the decreasing pre-trend suggests that before 2010, access to public transport increased more rapidly for control municipalities than for treated municipalities. One plausible explanation is that coordination inside ICs helped control municipalities develop transport networks more rapidly. In contrast, the large increasing trend after 2010 suggests that, after the law, access to public transport increased more rapidly for municipalities forced to join an IC.

Figure 8: Impact on public transport



Notes as in Figure 3.

Table 8 indicates that resisting municipalities' probability of having access to public transport was 3.3 percentage points higher in the average year post 2010, an effect significant at the 1-percent level. Given that only 2.4 percent of treated municipalities were connected to a transport network before 2010, it means that their entry into an intermunicipal community more than doubled their probability of having access to public transport. In 2017, the point estimate reaches 7.2, corresponding to more than tripling the probability of being connected.³⁸ The point estimate is larger for urban municipalities (10.9, column 3 of Table 8) than for rural municipalities (2.1, column 3 of Table 8), but the magnitude relative to the pre-treatment mean is similar, and both estimates are significant at the 1-percent level. In line with the presence of pre-trends, the coefficients associated to the placebo reforms before 2010 are significant but small and negative, contrasting with the large and positive effect of the true reform (Figure B4 in Appendix B).

³⁸The fraction of treated municipalities with access to a public transport network prior to 2010 is particularly low for two reasons. First, treated municipalities were not part of an IC before 2010, making them less likely to be part of a transport network. As a comparison, the share of control municipalities with access to a public transport network prior to 2010 is equal to 13.0 percent. Second, the data I use cover the regular public transport networks that are managed at the municipal or intermunicipal level, excluding the ones managed at the departmental or regional level (see Section 2.3).

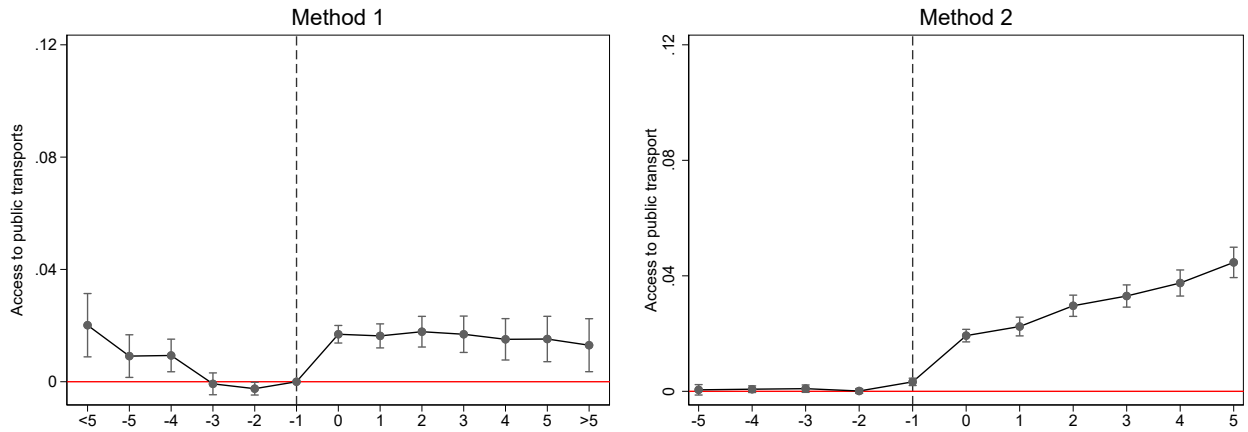
Table 8: Impact on public transport

	(1)	(2)	(3)
Outcome	Access to public transport		
	All	Rural	Urban
Treatment	0.033*** (0.007)	0.021*** (0.007)	0.109*** (0.025)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	220,948	173,684	47,264
Mean DepVar	0.024	0.014	0.078
Sd DepVar	0.153	0.116	0.269

Notes: The outcome is an indicator variable equal to one if the municipality has access to public transport. The sample excludes municipalities located in the Parisian region, Ile-de-France, for which the data are not available. Column 2 includes only rural municipalities, while column 3 includes only urban municipalities. Other notes as in Table 2.

I then use the same estimation methods as in Section 4.1.2 to estimate the impact of integration on public transport for municipalities that instead voluntarily joined an IC before 2010. The first graph in Figure 9 uses the regular staggered adoption design (method 1) and shows that on average, the probability of having access to public transport was 2 percentage points higher in the years following integration for municipalities that voluntarily integrated. The estimated effect is larger using de Chaisemartin and D’Haultfoeuille (2019)’s method: As shown on the second graph (method 2), the effect is increasing over time and reaches up to 4 percentage points 4 years after integration. This is a sizable effect, given that only 2.8 percent of the municipalities that integrated between 2000 and 2010 had access to public transport in 1999. Taken together, these results suggest that integration had a large positive effect on public transport for both municipalities that voluntarily joined an IC and those that resisted integration.

Figure 9: Impact on public transport for municipalities voluntary integrating



Notes as in Figure 4.

5.2 Impact on fiscal revenues

Finally, I study the impact of integration on municipalities' resources. When entering an IC, municipalities' fiscal revenues can be impacted in two ways. First, while municipalities' own state transfers do not change after integration, once part of an IC, municipalities also benefit from the additional state transfers allotted to the community they are now part of. Second, municipalities' tax revenues can change. In particular, if the business tax is decided at the IC level, as is the case in about half of the ICs, the tax rate and tax base might differ from the municipality's own rate and base prior to integration, which can lead to either a decrease or increase in tax revenues, depending on the gap in tax rate and tax base compared with the other members of the community.

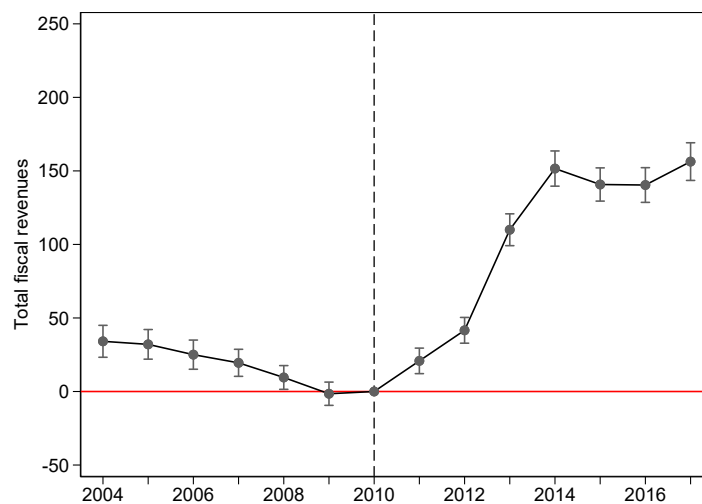
To investigate whether municipalities forced to integrate benefited from larger resources after their integration, I estimate the impact of forced integration on municipalities' total fiscal revenues, which I compute as follows. If a municipality is isolated, its fiscal revenues are made up of the municipality's own tax revenues and state transfers. If a municipality is part of an intermunicipal community, I compute its fiscal revenues as the sum of the municipality's own fiscal revenues and the total IC's fiscal revenues scaled by the share of the IC population the municipality represents.³⁹ Finally, I divide the total fiscal revenues by the municipal population in 2010.

As shown in Figure 10, contrasting with the small decreasing trend before 2010, we see a large increasing trend after the law, suggesting that treated municipalities experienced an increase in their fiscal revenues compared to control municipalities following their integration.

³⁹If the IC gives direct transfers to one of the member municipalities—for instance, as part of the “solidarity transfers” given to poorer municipalities in the community—I consider these transfers as part of the revenues of that municipality only, but not as part of the total IC revenues shared among all members.

Table 9 indicates that on average, integration increased fiscal revenues by 91.8 euros per capita per year (13.0 percent) in municipalities forced to join an IC, an effect significant at the 1-percent level. As shown in columns 2 and 3, the increase compared to the average fiscal revenues prior to 2010 is similar for rural and urban municipalities: 13.3 and 12.5 percent respectively. As for public transport, the estimates associated to placebo reforms are small and negative, working against the large and positive effect of the true reform (Figure B4 in Appendix B).

Figure 10: Impact on fiscal revenues



Notes as in Figure 3.

Table 9: Impact on fiscal revenues

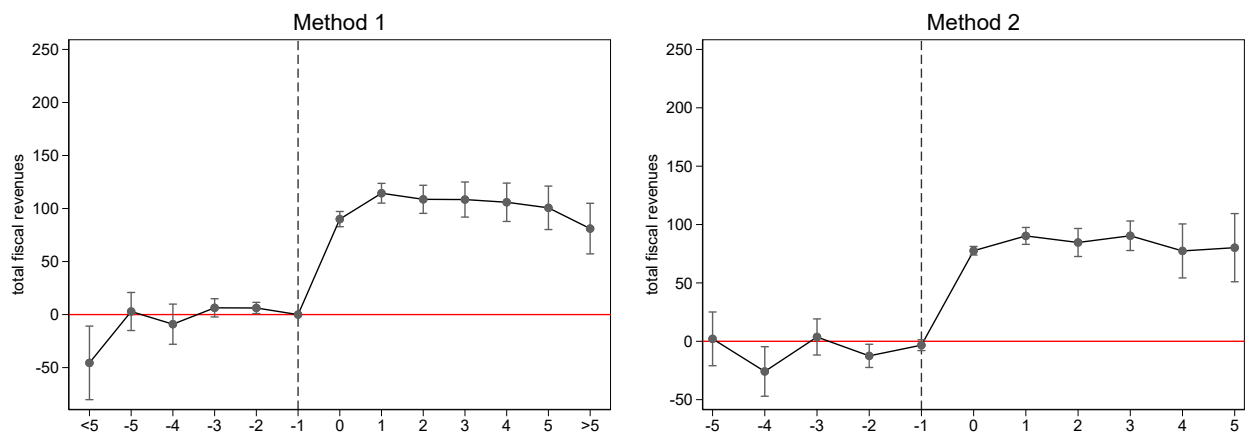
Outcome	(1)	(2) Fiscal revenues per capita	
	All	Rural	Urban
Treatment	91.8*** (5.1)	82.9*** (5.4)	128.5*** (12.8)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	227,766	177,618	50,148
Mean DepVar	705.7	624.0	1,025.3
Sd DepVar	431.9	329.7	602.4

Notes: The outcome is the municipality's total fiscal revenues per capita, as defined in the text. I exclude from the analysis few municipalities for which the data are missing for at least one year over the period 2004-2017 (0.6 percent of the sample). Column 2 includes only rural municipalities, while column 3 includes only urban municipalities. Other notes as in Table 2.

I then compare this effect to the one experienced by municipalities that instead decided to join an IC voluntarily before the law, again using the same estimation methods as in Section

4.2.1. As shown in Figure 11, both estimation methods suggest that they experienced an increase in fiscal revenues of about 100 euros per capita after their integration (or about 15.2 percent).⁴⁰ Hence, municipalities forced to integrate gained about the same amount of revenue as municipalities that chose to integrate. Assuming that municipalities could anticipate such effects,⁴¹ these results suggest that municipalities forced to integrate did not oppose integration because they believed they would get a lower share of the benefits, but rather that these benefits were not high enough to compensate for the costs of integration.

Figure 11: Impact on fiscal revenues for municipalities voluntary integrating



Notes: Data on fiscal revenues are only available starting in 2002. The analysis is thus restricted to municipalities integrating between 2003 and 2010. I further exclude from the analysis a few municipalities for which the data are missing for at least one year over the period 2002-2017 (1.4 percent of the sample). In method 1, the period of analysis goes from 2002 to 2017, whereas in method 2 it goes from 2002 to 2009. Other notes as in Figure 4.

⁴⁰Data on municipalities' and IC's fiscal revenues are available starting in 2002 only. I thus focus the analysis on municipalities that voluntarily integrated between 2003 and 2010. For the regular staggered adoption design, the period of analysis goes from 2002 to 2017, while in using de Chaisemartin and D'Haultfoeuille (2019)'s method I restrict to the period from 2002 to 2009 (see Section 4.2.1). To compute the magnitude of the effect, I compare the increase of 100 euros to the average fiscal revenues in 2002.

⁴¹It is likely that municipalities could anticipate the change in revenues they would face after integration: The record of state transfers awarded to ICs was public information, as the goal was to incentivize cooperation, and the tax rate and tax base of other municipalities are quite easy to learn given the geographical proximity between municipalities inside the same IC.

6 Discussion

6.1 Interpretation

While municipalities forced to integrate experienced the same benefits of integration as municipalities that voluntarily integrated, they faced some costs in terms of housing supply and loss of local public services that the other municipalities did not, at least in terms of construction. I have interpreted those results as evidence that municipalities resisted integration to avoid such costs. This interpretation assumes that they correctly anticipated the consequences of integration and that they would have experienced the same had they integrated earlier.

Alternatively, one could argue that resisting municipalities faced such effects because they entered an intermunicipal community later on and were forced to do so. As a result, they might have entered more organized ICs that were more able to impose costs on them, and municipalities already integrated in those communities might have wanted to punish them for having resisted so long. This would imply that forced municipalities would not have experienced the same had they voluntarily integrated earlier, casting doubt on whether the results presented in this paper can explain their resistance. I test this alternative interpretation by looking at the impact of integration depending on the types of ICs the resisting municipalities joined after the 2010 law.

First, if this is the correct interpretation, we should see that the effect of forced integration on building permits and local public services are driven by municipalities that entered existing intermunicipal communities after the law, as opposed to municipalities that created new ICs after 2010. Instead, as shown in Table 10, the 27 percent of treated municipalities that created new ICs also experienced a large increase in the number of building permits delivered (+10.0 vs. +8.2 for the full sample, columns 2 and 1). The same conclusion holds when focusing on municipalities inside the urban area, which drive the impact on housing (columns 6 vs. 5). When entering a new IC, the integration process of resisting municipalities resembles the integration process of municipalities that integrated early on, supporting the view that municipalities that opposed integration would have experienced the same had they integrated earlier. Tables A5 and A6 in the Appendix provides the results for daycare facilities and public libraries, for the full sample and focusing on rural municipalities only. While the impact on both services is on average smaller in magnitude for municipalities entering new ICs, it remains negative. Moreover, the effect on daycare facilities for rural municipalities entering new ICs is high and even becomes significant.⁴²

Second, focusing on the larger sample of municipalities that entered existing ICs, I look at the impact depending on whether resisting municipalities had the choice of joining several communities when they were forced to integrate in 2010. If punishment for having refused

⁴²The results on public services are, however, difficult to interpret, given the general lack of precision for the effects on daycare facilities and the extremely small sample available for public libraries (there are only 15 municipalities that entered a new IC and for which data on libraries are available).

to enter earlier was driving the results, we should see that municipalities that had common borders with several communities experienced lower costs. While the effect on housing is a bit smaller on average for municipalities that had the choice between at least two ICs (5.2 vs 8.2, columns 3 vs. 1), the impact is very similar when focusing on municipalities inside an urban area (9.2 vs. 10.6, columns 7 vs. 5), and all estimates remain significant at the 1-percent level. Moreover, as shown in Tables A5 and A6 in the Appendix, the impact on daycare facilities and public libraries remains the same depending on the number of options the municipalities had, both overall and focusing on rural municipalities only.

Finally, I study the impact of integration for resisting municipalities that joined existing communities made of recently integrated neighbors. In particular, I consider communities whose members integrated later than 2000 and were thus quite recently created in 2010. If entering communities that had been in existence for a long time was driving the results, we should see a smaller impact for those municipalities. Instead, the impact on building permits and public services remain as high in magnitude as for the full sample (see columns 4 vs. 1 and 8 vs. 5 in Table 10, and Table A5 and A6 in the Appendix).

Table 10: Impact on housing depending on the integration process

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of building permits per 10,000 inhabitants							
	Full sample				Municipalities inside an urban area			
	All	New IC	Existing IC		All	New IC	Existing IC	
		≥ 2 choices	Recent			≥ 2 choices	Recent	
Treatment	8.228*** (1.609)	10.024*** (3.061)	5.233*** (1.853)	8.852*** (3.305)	10.634*** (1.929)	8.803*** (3.409)	9.182*** (2.302)	15.730*** (3.690)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	245,430	231,225	239,055	231,015	151,125	142,665	146,385	141,750
Mean	66.203	65.660	67.273	62.718	65.288	64.439	67.071	59.873
Sd	94.440	92.082	95.565	98.893	89.761	86.693	91.364	88.177

Notes: Columns 5 to 8 focus on municipalities inside an urban area. Columns 2 and 6 includes only treated municipalities that joined an existing IC after the law. Columns 2 and 6 include only treated municipalities that created a new IC in 2010. Columns 3 and 7 include only treated municipalities that entered an existing IC and that had the choice between at least two of them in 2010. Columns 4 and 8 include only treated municipalities that entered an existing IC in which all other members integrated after 2000. Other notes as in Table 2.

6.2 External validity

The results above further support the view that municipalities forced to integrate opposed co-operation knowing that they would face some costs in terms of housing supply and local public

services. Conversely, it also suggests that municipalities that voluntarily integrated entered an IC knowing they would be able to avoid such costs. This begs the question: Why or how were those municipalities able to avoid them?

A first possible explanation is that those costs of integration apply only to a particular type of municipality, which would limit the external validity of the results. However, based on observable characteristics and as shown in Section 3.1, resisting municipalities are on average quite similar to municipalities that integrated the earliest. Moreover, as shown in Section 4, different types of resisting municipalities faced different costs, urban municipalities suffering from more construction and rural ones from a decrease in local public services. This shows that there is not one type of cost impacting a specific type of municipality, but rather that integration costs can apply to a wide range of municipalities.

Hence, a most likely explanation is that municipalities that voluntarily integrated knew that they would get a higher bargaining power among their neighbors and would thus be able to oppose costly IC decisions. This is consistent with forced municipalities having less bargaining power in their intermunicipal communities today compared with the average municipality. As the number of seats a municipality gets in the intermunicipal council is proportional to its population, I measure a municipality's bargaining power as the ratio between the share of the IC's total population living in that municipality divided by the population share an average municipality from the same IC represents. Hence, the greater this ratio, the larger the share of seats the municipality has compared with the average number of seats others get. I look at the composition of ICs at the end of the period of analysis, in 2018. While the average value of the ratio is 1 by construction, it is on average equal to only 0.80 for municipalities that were forced to integrate. Those municipalities are also more likely to end up in an IC that encompasses a big city (of more than 10,000 inhabitants): 36.6 percent, versus 32.2 on average. This supports the view that municipalities resisted knowing that once they became part of a community, they would not be able to prevent their neighbors from imposing new construction or decreasing the availability of public services within their territory.

All in all, this suggests that the costs identified in this paper can explain resistance beyond the specific case of municipalities forced to enter an IC in 2010. In particular, they might explain why the vast majority of French municipalities recently opposed new laws aimed at increasing the size and competences of intermunicipal communities. An increase in the size of ICs would lead many municipalities to lose bargaining power and thus to suffer from costs they have been able to avoid up to now.⁴³ Moreover, as most forms of cooperation among local jurisdictions imply sharing urban planning policies and public services, these findings can help

⁴³In 2015, a new law passed requiring that ICs should be at least 15,000 inhabitants by 2017 and requiring municipalities to share more public services. There was widespread complaint from mayors, leading to multiple revisions of the law (e.g., <https://www.maire-info.com/organisation-gestion-communale/elus-locaux/un-fort-sentiment-de-perte-d-autonomie-chez-les-maires-selon-le-cevipof-article-22423>). In the face of such resistance, the French president announced he might reconsider the law (see for instance: <https://www.amf.asso.fr/documents-vers-une-revision-la-loi-notre-/39240>).

explain resistance against integration beyond the case of French municipalities.

7 Conclusion

This paper is the first to causally identify the effects of integration on resisting municipalities. Exploiting a 2010 reform in France that forced non-integrated municipalities to enter an inter-municipal community, I measure the impact of integration on resisting municipalities. Comparing those effects with those experienced by municipalities that instead voluntarily integrated before the law, I infer the factors explaining why forced municipalities opposed integration in the first place. I provide evidence that resistance is driven by actual local costs of integration: Urban municipalities resisted cooperation to prevent further constructions from raising congestion, while rural municipalities resisted to avoid losing local public services.

First, I find that municipalities forced to integrate experienced an increase of 12.4 percent in the number of building permits delivered per year. On the contrary, municipalities that integrated voluntarily did not experience any change in their housing supply following their integration. This finding supports the view that resisting municipalities refused to integrate to avoid an increase in housing supply and provides evidence that their decision was driven by actual costs of integration, beyond ideological or political considerations. Further analyses show that the impact is mainly driven by municipalities where the demand for housing is high and that are already densely built, for which more construction is likely to raise congestion.

I then assess the impact on local public services. To do so, I gathered data at the local level on two different local public services: daycare facilities and public libraries. For both, I find that rural municipalities forced to enter an intermunicipal community experienced a decrease in the number of facilities located on their territory, implying that integration increased the distance to public services for their residents.

Finally, I explore the benefits of integration. Results show that once in a community, municipalities forced to integrate became twice as likely to have access to public transport and that their fiscal revenues increased by 13.0 percent per year on average. I find that these gains are similar to the ones municipalities that voluntarily integrated experienced. These results suggest that resisting municipalities did not oppose integration anticipating they would get a lower share of the benefits, but rather that these benefits were not high enough to compensate for the local costs of integration.

This paper provides new evidence on the factors explaining municipalities' reluctance to integrate. These findings could help policymakers design better compensation schemes to implement consolidation policies at the municipal level more effectively. In particular, the results suggest that rural and urban municipalities should be compensated differently, as they do not face the same costs of integration. Moreover, by identifying the local costs of integration, this paper provides the first step toward a welfare analysis of consolidation reforms and thus opens new avenues for research.

These results might also help understand consolidation failures in general, from the reluctance of countries to expand the European Union to the aspiration of Catalonia to secede from Spain. While the specific costs identified in this paper might not apply in these settings, my findings suggest that we should go beyond ideological and political considerations and identify the actual costs of integration in order to better understand what is at stake.

Appendix

Contents

A. Additional tables and figures

B. Robustness checks

C. Housing supply – heterogeneity analysis

D. Housing price indices

A. Additional tables and figures

Table A1: Descriptive statics using municipalities' characteristics in 1999

	Control (N=15,065)				Treatment (N=1,297)			
	mean	sd	min	max	mean	sd	min	max
population	1,851	11,095	2	798,430	1,516	4,434	18	67,304
<i>w/out largest</i>	1,481	4,038	2	65,042	1,516	4,434	18	67,304
urban	0.19	0.39	0.00	1.00	0.18	0.38	0.00	1.00
urban area	0.54	0.50	0.00	1.00	0.57	0.50	0.00	1.00
core urban area	0.11	0.31	0.00	1.00	0.08	0.27	0.00	1.00
unemployed	0.10	0.05	0.00	0.50	0.10	0.05	0.00	0.35
child	0.07	0.02	0.00	0.26	0.07	0.02	0.00	0.18
farmers	0.05	0.06	0.00	1.00	0.05	0.06	0.00	0.50
executives	0.04	0.03	0.00	0.50	0.05	0.06	0.00	0.43
workers	0.17	0.07	0.00	0.67	0.16	0.07	0.00	0.60
retired	0.26	0.08	0.00	1.00	0.25	0.09	0.00	0.75
no diploma	0.22	0.08	0.00	0.63	0.20	0.08	0.03	0.68
baccalaureate	0.11	0.03	0.00	0.43	0.12	0.04	0.00	0.29
high education	0.05	0.03	0.00	0.46	0.06	0.06	0.00	0.41
residents' income	7,173	1,854	1,937	38,895	8,428	3,198	2,739	30,590

Notes: Data on the municipal population, age, education, and occupational composition comes from the 1999 census. Indicator variables for whether the municipality is urban, part of an urban area, or urban core are based on the INSEE 1999 classification. Residents' income corresponds to the 2000 taxable income data.

Table A2: Impact on housing – residents’ income

	(1)	(2)	(3)
Outcome	Number of housing building permits		
	All	Residents’ income median	
		Below	Above
Treatment	8.228*** (1.609)	4.431* (2.265)	9.273*** (2.174)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	245,430	122580	122610
Mean DepVar	66.203	59.172	70.028
Sd DepVar	94.440	85.053	98.247

Notes: Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. Column 2 (resp. 3) includes only municipalities below (resp. above) the median of the average residents’ taxable income in 2010. Data on taxable income in 2010 are missing for the 16 municipalities with the smallest populations, to protect residents’ privacy.

Table A3: Impact on housing – political alignment

	(1)	(2)	(3)	(4)	(5)
Outcome	Number of housing building permits per 10,000 inhabitants				
	All	Same orientation		Vote share distance median	
		Yes	No	Below	Above
Treatment	8.228*** (1.609)	9.093*** (2.075)	7.408*** (2.588)	6.372*** (2.429)	9.636*** (2.148)
Municipality FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	245,430	139,545	100,125	122385	123045
Mean DepVar	66.203	66.172	66.253	71.593	62.184
Sd DepVar	94.440	95.290	92.935	99.830	90.012

Notes: Column 2 (resp. 3) includes only municipalities whose mayor had the same orientation (resp. a different orientation) as the member of parliament of their district in 2010. Column 4 (resp. 5) includes only municipalities below (resp. above) the median value of the absolute difference in the right vote share in the 2007 presidential election between the municipality and the other municipalities from the same département. Other notes as in Table A2.

Table A4: Impact on the number of daycare spots available

Outcome	(1)	(2)	(3)
	Number of daycare spots		
	All	Rural	Urban
Treatment	-1.273 (0.877)	-1.863* (1.036)	1.487 (1.388)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	179,982	140,272	39,710
Mean DepVar	11.414	4.519	38.391
Sd DepVar	52.408	50.936	49.294

Notes: The outcome is the total number of daycare spots available in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. Column 2 (resp.3) include only rural (resp. urban) municipalities. Other notes as in Table A2.

Table A5: Impact on daycare depending on the integration process

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of daycare facilities per 10,000 inhabitants							
	Full sample				Rural municipalities			
	All	New IC	Existing IC		All	New IC	Existing IC	
		≥ 2 choices	Recent			≥ 2 choices	Recent	
Treatment	-0.061 (0.046)	-0.025 (0.035)	-0.075 (0.065)	-0.189 (0.131)	-0.083 (0.055)	-0.065*** (0.022)	-0.096 (0.078)	-0.205 (0.164)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	179,982	169,565	175,307	169,411	140,272	131,890	136,708	131,835
Mean	0.438	0.495	0.371	0.583	0.210	0.125	0.228	0.414
Sd	2.174	1.330	2.395	3.568	2.211	0.838	2.540	3.910

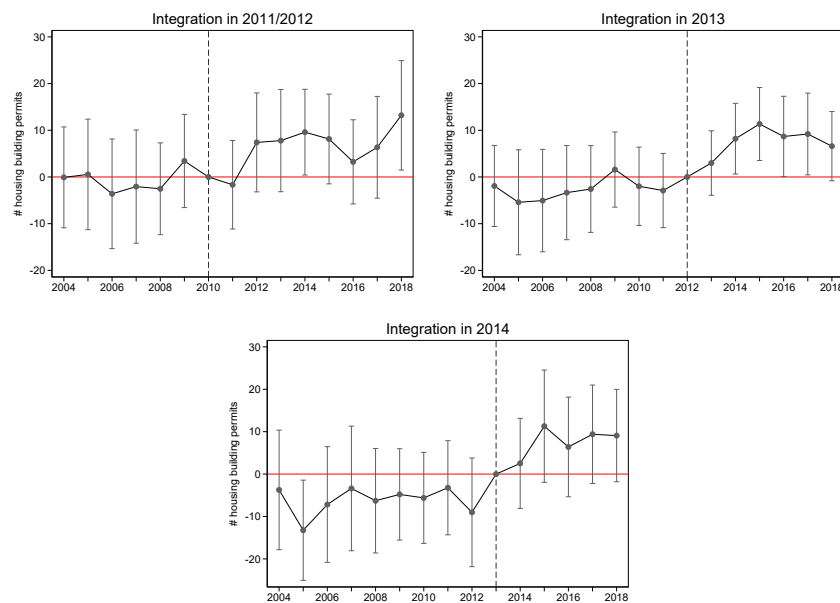
Notes: The outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. Columns 5 to 8 include only rural municipalities. Columns 2 and 6 include only treated municipalities that created a new IC in 2010. Columns 3 and 7 include only treated municipalities that entered an existing IC and that had the choice between at least two of them in 2010. Columns 4 and 8 include only treated municipalities that entered an existing IC in which all other members integrated after 2000. Other notes as in Table A2.

Table A6: Impact on public libraries depending on the integration process

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Number of public libraries per 10,000 inhabitants							
	Full sample				Rural municipalities			
	All	New IC	Existing IC		All	New IC	Existing IC	
		≥ 2 choices	Recent			≥ 2 choices	Recent	
Treatment	-0.841*	-0.402	-1.011*	-1.020	-1.129*	-0.541	-1.375*	-1.467
	(0.465)	(0.246)	(0.596)	(0.727)	(0.627)	(0.346)	(0.811)	(0.993)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	12,825	11,925	12,582	12,339	9,900	9,243	9,720	9,549
Mean	4.395	5.072	4.740	4.662	5.302	6.590	5.823	5.697
Sd	9.611	5.163	10.674	11.040	11.026	5.564	12.293	12.706

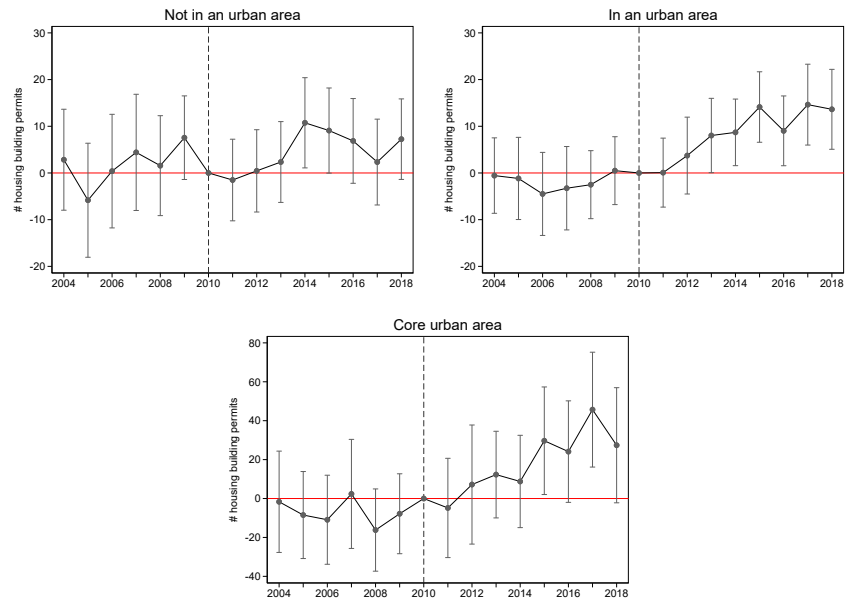
Notes: The outcome is the number of public libraries located in the municipality, per 10,000 inhabitants (using the 2010 population). The period of analysis is 2009-2017 and the sample is restricted to 7 départements for which data are available starting in 2009. Columns 5 to 8 include only rural municipalities. Columns 2 and 6 include only treated municipalities that created a new IC in 2010. Results should be interpreted with caution for this restriction, as it includes only 15 treated municipalities for the full sample and 10 treated municipalities for the rural sample, the only municipalities for which data on libraries are available. Columns 3 and 7 include only treated municipalities that entered an existing IC and that had the choice between at least two of them in 2010. Columns 4 and 8 include only treated municipalities that entered an existing IC in which all other members integrated after 2000. Other notes as in Table A2.

Figure A1: Impact on housing depending on the exact year of integration



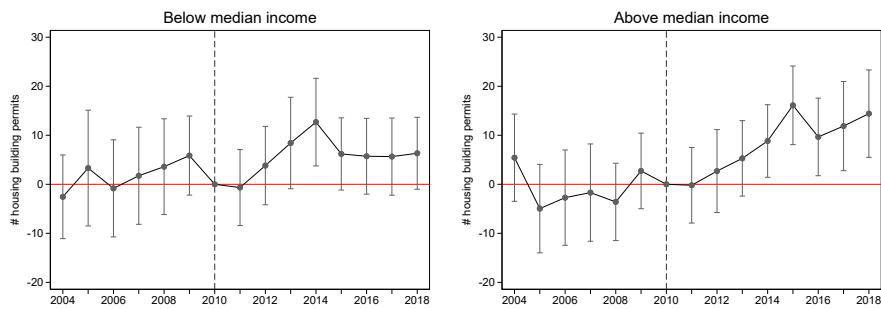
Notes: These graphs plot the estimates and 95-percent confidence intervals from the leads-and-lags regression, using as outcome the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). The first graph includes only treated municipalities that entered an IC in 2011 or 2012. The second (resp. third) graph includes only treated municipalities that entered an IC in 2013 (resp. 2014).

Figure A2: Impact on housing – urban area



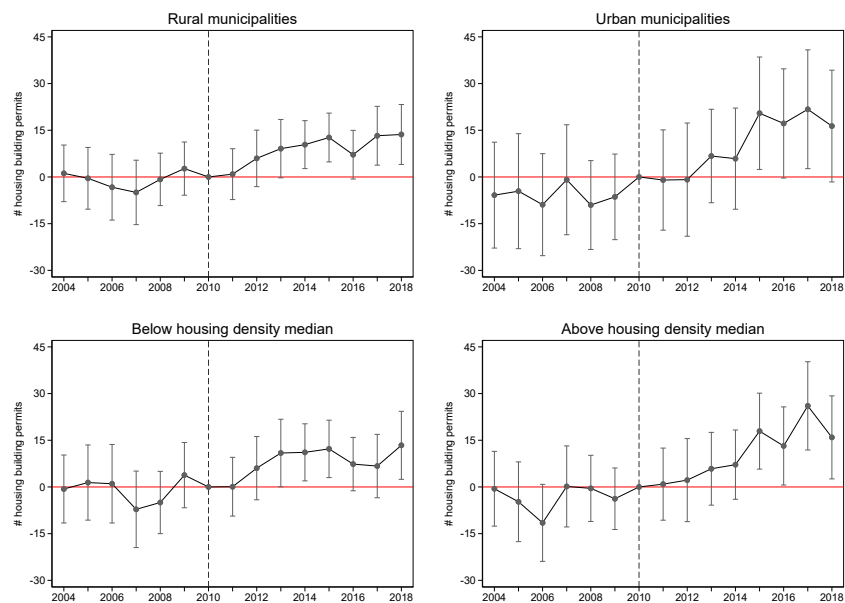
Notes: The first graph includes only municipalities that are not part of an urban area. The second graph includes only municipalities that are part of an urban area. The third graph includes only municipalities located in the core of the urban area (using a different scale given the magnitude of the results). Other notes as in Figure A1.

Figure A3: Impact on housing – residents' income



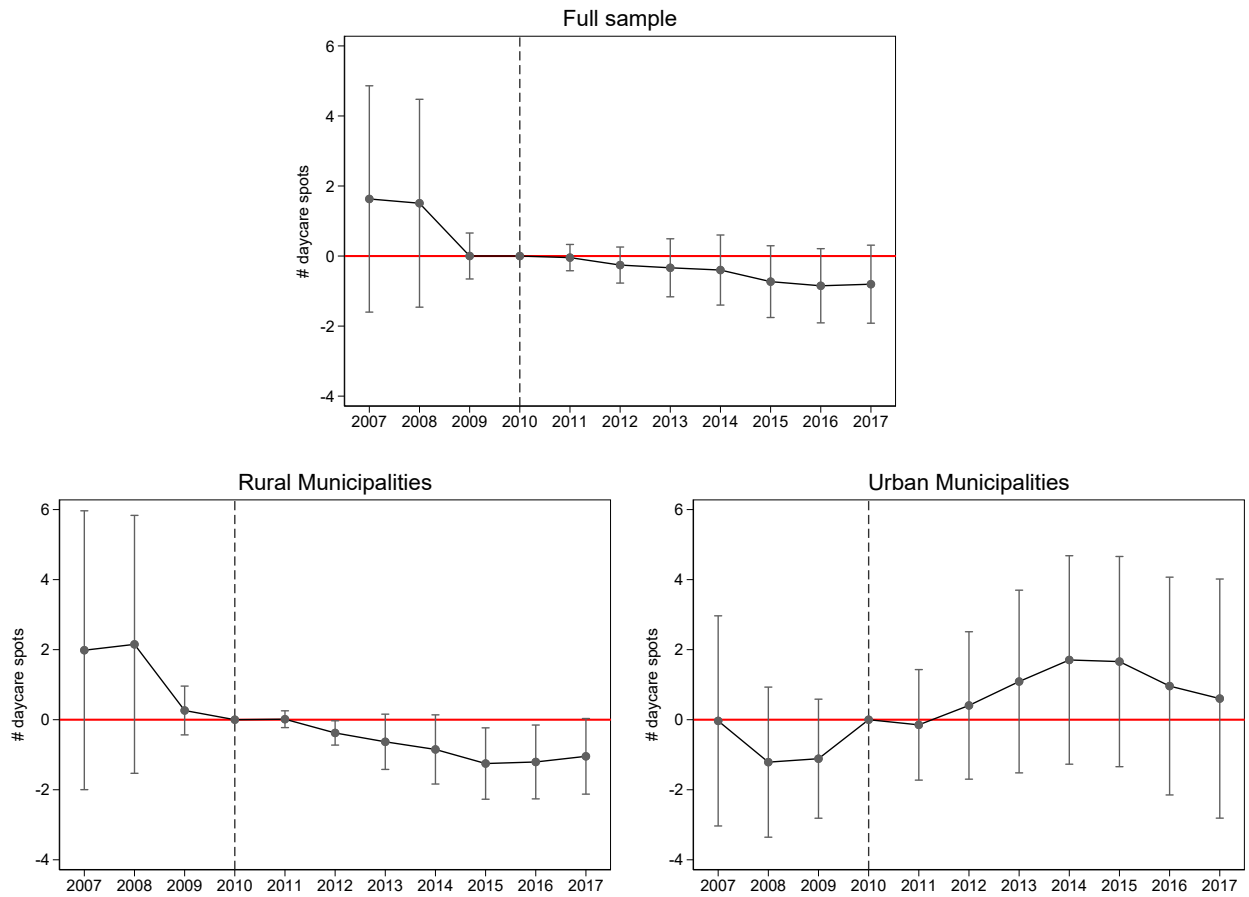
Notes: The left-hand side (resp. right-hand side) includes only municipalities below (resp. above) the median of the average residents' taxable income in 2010. Other notes as in Figure A1.

Figure A4: Impact on housing, inside the urban area – housing density



Notes: All graphs focus on municipalities inside an urban area. The first (resp. second) graph includes only rural (resp. urban) municipalities. The third (resp. fourth) graph includes only municipalities with a housing density in 2010 below (resp. above) the median. Other notes as in Figure A1.

Figure A5: Impact on the number of daycare spots available



Notes: The outcome is the total number of daycare spots available in the municipality per 10,000 inhabitants (using the 2010 population). The graph on the bottom left-hand side includes only rural municipalities, while the graph on the bottom right-hand side includes only urban municipalities.

B. Robustness checks

B1. Impact depending on the latest date of integration of the control municipalities

For each outcome, I test the robustness of the results to varying the control group depending on the latest date of integration of the control municipalities. In each table below, the first column gives the baseline estimate restricting the control group to municipalities integrated since 1999. The next columns provide the estimates obtained by considering instead all municipalities integrated since 2002, 2004, 2006, 2008, and 2010, respectively. Hence, in the last column, I include all municipalities already integrated in 2010. Note that each time I vary the date, the treatment group also slightly changes, as I remove municipalities that were isolated in 2010 but have been integrated at some point between the latest date of integration of the control group and 2010 (a group involving less than 1 percent of the isolated municipalities).

Table B1.1: Housing supply

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Number of building permits per 10,000 inhabitants					
Latest integration	1999	2002	2004	2006	2008	2010
Treatment	8.228*** (1.609)	8.255*** (1.578)	7.660*** (1.574)	7.455*** (1.567)	7.453*** (1.567)	7.639*** (1.565)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	245,430	335,295	388,950	406,695	415,560	426,270
Mean DepVar	66.203	65.902	66.074	66.003	66.213	66.194
Sd DepVar	94.440	94.059	94.232	94.135	94.654	94.619

Notes: Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. Column 1 reproduces the baseline estimate as reported in the paper, using as control group municipalities integrated since 1999. The next columns include in the control group all municipalities integrated since 2002, 2004, 2006, 2008, and 2010, respectively.

Table B1.2: Daycare

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Number of daycare facilities per 10,000 inhabitants					
Latest integration	1999	2002	2004	2006	2008	2010
Treatment	-0.061 (0.046)	-0.062 (0.045)	-0.069 (0.045)	-0.070 (0.044)	-0.068 (0.044)	-0.067 (0.044)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	179,982	245,883	285,230	298,243	304,744	312,598
Mean DepVar	0.438	0.434	0.439	0.439	0.443	0.446
Sd DepVar	2.174	2.162	2.160	2.156	2.155	2.155

Notes: The outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. Other notes as in Table B1.1.

Table B1.3: Public libraries

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Number of public libraries per 10,000 inhabitants					
Latest integration	1999	2002	2004	2006	2008	2010
Treatment	-0.841* (0.465)	-0.907** (0.434)	-1.035** (0.426)	-1.048** (0.426)	-1.044** (0.426)	-1.038** (0.422)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,825	17,397	22,032	23,013	23,184	23,841
Mean DepVar	4.395	4.224	4.285	4.285	4.285	4.253
Sd DepVar	9.611	9.373	9.392	9.392	9.392	9.361

Notes: The outcome is the number of public libraries located in the municipality per 10,000 inhabitants (using the 2010 population). The period of analysis is 2009-2017, and the sample is restricted to 7 départements for which data are available starting in 2009. Other notes as in Table B1.1.

Table B1.4: Public transport

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Access to public transport					
Latest integration	1999	2002	2004	2006	2008	2010
Treatment	0.033*** (0.007)	0.033*** (0.007)	0.035*** (0.007)	0.035*** (0.007)	0.034*** (0.007)	0.033*** (0.007)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	220,948	303,758	350,532	364,644	372,624	381,808
Mean DepVar	0.024	0.024	0.024	0.023	0.026	0.026
Sd DepVar	0.153	0.152	0.152	0.151	0.160	0.160

Notes: The outcome is an indicator variable equal to one if the municipality has access to public transport. The sample excludes municipalities located in the Parisian region of Ile-de-France, for which the data are not available. Other notes as in Table B1.1.

Table B1.5: Fiscal revenues

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Fiscal revenues per capita					
Latest integration	1999	2002	2004	2006	2008	2010
Treatment	91.8*** (5.1)	92.5*** (5.0)	91.1*** (4.9)	91.0*** (4.9)	90.6*** (4.9)	88.6*** (4.9)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	227,766	311,262	360,990	377,426	385,644	395,640
Mean DepVar	705.7	709.8	710.9	710.7	711.0	711.3
Sd DepVar	431.9	438.7	438.9	438.3	437.9	437.7

Notes: The outcome is the municipality's total fiscal revenues per capita, as defined in Section 5.2. I exclude from the analysis few municipalities for which the data are missing for at least one year over the period 2004-2017 (0.6 percent of the sample). Other notes as in Table B1.1.

B2. Clusters at the IC level

Table B2.1: Housing supply

	(1)	(2)	(3)
Outcome	Number of building permits		
Cluster	Municipality	IC 2014	IC 2018
Treatment	8.228*** (1.609)	8.228*** (2.354)	8.228*** (2.457)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	245,430	245,430	245,430
Mean DepVar	66.203	66.203	66.203
Sd DepVar	94.440	94.440	94.440

Notes: Standard errors are in parentheses. ***, **, and * indicate significance at 1, 5, and 10% respectively. In column 1, standard errors are clustered at the municipality level. In column 2 (resp. 3), standard errors are clustered at the IC level, considering the IC the municipality belongs to in 2014 (resp. 2018). The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction.

Table B2.2: Local public services

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Number of facilities per 10,000 inhabitants					
	Daycare facilities			Public libraries		
Cluster	Municipality	IC 2014	IC 2018	Municipality	IC 2014	IC 2018
Treatment	-0.061 (0.046)	-0.061 (0.046)	-0.061 (0.047)	-0.841* (0.465)	-0.841** (0.369)	-0.841** (0.372)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	179,982	179,982	179,982	12,825	12,825	12,825
Mean	0.438	0.438	0.438	4.395	4.395	4.395
Sd	2.174	2.174	2.174	9.611	9.611	9.611

Notes: In columns 1 to 3, the outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. In columns 4 to 6, the outcome is the number of public libraries located in the municipality per 10,000 inhabitants (using the 2010 population), the period of analysis is 2009-2017, and the sample is restricted to 7 départements for which data are available starting in 2009. Other notes as in Table B2.1.

Table B2.3: Public transport and fiscal revenues

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Access to public transport			Fiscal revenues per capita		
Cluster	Municipality	IC 2014	IC 2018	Municipality	IC 2014	IC 2018
Treatment	0.033*** (0.007)	0.033** (0.013)	0.033*** (0.012)	91.8*** (5.1)	91.8*** (9.5)	91.8*** (9.4)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	220,948	220,948	220,948	227,766	227,766	227,766
Mean	0.024	0.024	0.024	705.7	705.7	705.7
Sd	0.153	0.153	0.153	431.9	431.9	431.9

Notes: In columns 1 to 3, the outcome is an indicator variable equal to one if the municipality has access to public transport, and the sample excludes municipalities located in the Parisian region of Ile-de-France, for which the data are not available. In columns 4 to 6, the outcome is the municipality's total fiscal revenues per capita, as defined in Section 5.2, and I exclude from the analysis few municipalities for which the data are missing for at least one year over the period 2004-2017 (0.6 percent of the sample). Other notes as in Table B2.1.

B3. Matching

I test the robustness of the main results to using propensity score matching. I match control and treated municipalities on the following socio-demographic characteristics, considered in 2010: The number of inhabitants, population growth since 1999, population density, whether the municipality is urban, whether it belongs to an urban area, whether it belongs to the core of the urban area, the share of immigrants, the share of unemployed workers, the share of children, the share of the active population being farmers, executives, workers, or retired, the share of the population with no diploma, with the baccalaureate, or with higher education, and the average residents' annual taxable income. Descriptive statics on these variables are presented in Table 1.

In Tables B3.1 and B3.2, I first report the differences between the control and treatment groups along with the t-tests for each variable, with and without using matching, respectively. The next tables provide the estimates. For each outcome, the first column gives the baseline estimate, the second column the estimate obtained using propensity score matching, and the third column the estimate using matching and requiring common support on top.

Table B3.1: T-tests without matching

	Mean control	Mean treatment	Difference	P-value
population	1,943	1,626	-317	0.049**
Δ population	0.087	0.089	0.002	0.5628
density	154.9	160.6	5.7	0.7143
urban mun	0.222	0.204	-0.019	0.1122
urban area	0.614	0.636	0.022	0.1132
core urban area	0.119	0.091	-0.028	0.001***
immigrants	0.036	0.045	0.008	0.000***
unemployed	0.088	0.082	-0.005	0.000***
child	0.076	0.073	-0.002	0.001***
farmers	0.038	0.036	-0.002	0.3234
executives	0.052	0.070	0.019	0.000***
workers	0.169	0.152	-0.017	0.000***
retired	0.286	0.280	-0.006	0.035**
no diploma	0.185	0.173	-0.013	0.000***
baccalaureate	0.153	0.156	0.004	0.002***
high education	0.072	0.088	0.017	0.000***
residents' income	12,633	14,209	1,576	0.000***

Notes: Data on the municipal population, age, education, and occupational composition comes from the 2008 census, which applies to the 2010 year. The variation of the population (line 2) is computed as the variation in the number of inhabitants between the 1999 and 2008 census. Indicator variables for whether the municipality is urban, part of an urban area, or located in the urban core are based on the INSEE 2010 classification. ***, **, and * indicate significance at 1, 5, and 10% respectively.

Table B3.2: T-tests with matching

	Mean control	Mean treatment	Difference	P-value
population	1,581	1,626	45	0.745
Δ population	0.091	0.089	-0.002	0.662
density	146.4	160.6	14.2	0.359
urban mun	0.201	0.204	0.002	0.843
urban area	0.622	0.636	0.014	0.328
core urban area	0.093	0.091	-0.002	0.813
immigrants	0.044	0.045	0.001	0.664
unemployed	0.084	0.082	-0.002	0.147
child	0.074	0.073	-0.001	0.311
farmers	0.036	0.037	-0.001	0.5728
executives	0.063	0.070	0.007	0.000***
workers	0.158	0.152	-0.006	0.008***
retired	0.282	0.280	-0.002	0.466
no diploma	0.177	0.173	-0.005	0.020**
baccalaureate	0.155	0.156	0.001	0.293
high education	0.081	0.088	0.007	0.001***
residents' income	14,000	14,518	518	0.000***

Notes: Same notes as in Table B3.1.

Table B3.3: Impact using matching – housing supply

	(1)	(2)	(3)
Outcome	Number of building permits		
Matching	No	Yes	+ common support
Treatment	8.228*** (1.609)	7.859*** (1.745)	8.012*** (1.737)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	245,430	245,235	244,740
Mean DepVar	66.203	66.203	66.203
Sd DepVar	94.440	94.440	94.440

Notes: Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. In column 2, the analysis is performed using propensity score matching, and it excludes 16 observations for which the data on taxable income is missing. Column 3 repeat the same exercise including only observations in the control group that share common support with observations in the treatment group, with regards to the characteristics displayed in Table B3.1 and B3.2.

Table B3.4: Impact using matching – local public services

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome	Number of facilities per 10,000 inhabitants					
Matching	Daycare facilities			Public libraries		
	No	Yes	+ common support	No	Yes	+ common support
Treatment	-0.061 (0.046)	-0.054 (0.052)	-0.056 (0.051)	-0.841* (0.465)	-0.508 (0.725)	-0.485 (0.801)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	179,982	134,780	133,941	12,825	6,256	5,949
Mean	0.438	0.438	0.438	4.395	4.395	4.395
Sd	2.174	2.174	2.174	9.611	9.611	9.611

Notes: In columns 1 to 3, the outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. In columns 4 to 6, the outcome is the number of public libraries located in the municipality per 10,000 inhabitants (using the 2010 population), the period of analysis is 2009-2017, and the sample is restricted to 7 départements for which data are available starting in 2009. Other notes as in Table B3.3.

Table B3.5: Impact using matching – public transport and fiscal revenues

Outcome	(1)	(2)	(3)	(4)	(5)	(6)
	Access to public transport			Fiscal revenues per capita		
Matching	No	Yes	+ common support	No	Yes	+ common support
Treatment	0.033*** (0.007)	0.033*** (0.007)	0.033*** (0.007)	91.8*** (5.1)	85.4*** (5.1)	85.2*** (5.1)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	220,948	220,682	220,430	227,766	227,584	227,108
Mean	0.024	0.024	0.024	705.7	705.7	705.7
Sd	0.153	0.153	0.153	431.9	431.9	431.9

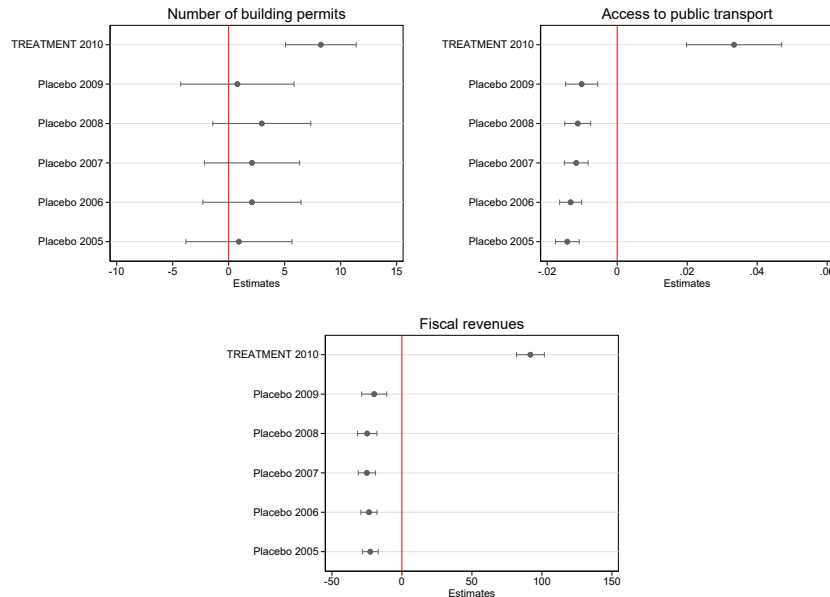
Notes: In columns 1 to 3, the outcome is an indicator variable equal to one if the municipality has access to public transport and the sample excludes municipalities located in the Parisian region of Ile-de-France, for which the data are not available. In columns 4 to 6, the outcome is the municipality's total fiscal revenues per capita, as defined in Section 5.2, and I exclude from the analysis few municipalities for which the data are missing for at least one year over the period 2004-2017 (0.6 percent of the sample). Other notes as in Table B2.3.

B4. Placebo tests

In this Section, I test the impact of a series a placebo reform. More precisely, I consider only the pre-treatment period from 2004 to 2010, and I run the same specification as the one described in Section 3.2, pretending that the law passed in 2005, 2006, 2007, 2008, or 2009. The graph below provides the estimates obtained for each placebo reform, along with the impact of the true reform (first coefficient on each graph). Unfortunately, I cannot run these placebo tests on daycare and public libraries, given the few pre-treatment periods I have in the data. This analysis is thus restricted to the number of building permits, access to public transport, and fiscal revenues.

As seen Figure B4, no coefficient associated with placebo reforms is significant at the standard level for the number of building permits. For public transport and fiscal revenues, consistent with the presence of decreasing pre-trends (see Section 5), the placebo estimates are significant but negative, which contrasts even more with the positive effect of the true reform. All in all, these results support the fact that the main results are capturing the impact of the 2010 law rather than the impact of factors that systematically affect treated and control municipalities differently.

Table B4: Placebo tests



Notes: The figure shows the impact of a series of placebo reforms on the number of building permits, the probability to have access to public transport and municipalities' fiscal revenues. In each graph, the first coefficient refers to the impact of the 2010 law, while the other estimates give the impact of a placebo reform implemented in 2005, 2006, 2007, 2008, and 2009, respectively. When estimating the impact of the placebo reforms, I include only the pre-treatment period. Horizontal lines are 95-percent confidence intervals.

B5. Alternative control groups

Table B5.1: Impact using alternative control groups – housing supply

	(1)	(2)	(3)	(4)	(5)
Outcome	Number of building permits per 10,000 inhabitants				
Control group	Group 1	Group 2	Group 3	Group 4	Group 5
Treatment	8.228*** (1.609)	8.696*** (1.631)	6.873*** (1.673)	8.622*** (1.616)	7.325*** (1.674)
Municipality FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	245,430	189,990	125,805	223,845	122,025
Mean DepVar	66.203	66.203	66.203	66.203	66.203
Sd DepVar	94.440	94.440	94.440	94.440	94.440

Notes: Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. In column 1, the control group is one used in the baseline estimation, including all municipalities integrated since 1999 (designated as “Group 1”). In column 2, I exclude from the control group all municipalities that were in an IC that received a treated municipality as a result of the 2010 law (“Group 2”). In column 3, I exclude more generally all municipalities that were in ICs that changed as a result of the 2010 law (“Group 3”). In column 4, I exclude control municipalities that share a common border with a treated municipality (“Group 4”). In column 5, I exclude both control municipalities whose ICs changed as well as neighbors (“Group 5”).

Table B5.2: Impact using alternative control groups – local public services

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Outcome	Number of facilities per 10,000 inhabitants									
	Daycare facilities					Public libraries				
Control	Group 1	Group 2	Group 3	Group 4	Group 5	Group 1	Group 2	Group 3	Group 4	Group 5
Treat.	-0.061	-0.047	-0.036	-0.057	-0.034	-0.841*	-0.992**	-1.128**	-0.866*	-1.147**
	(0.046)	(0.046)	(0.048)	(0.046)	(0.048)	(0.465)	(0.479)	(0.476)	(0.467)	(0.476)
Mun FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	179,982	139,326	92,257	164,153	89,485	12,825	9,270	7,146	11,736	6,930
Mean	0.438	0.438	0.438	0.438	0.438	4.395	4.395	4.395	4.395	4.395
Sd	2.174	2.174	2.174	2.174	2.174	9.611	9.611	9.611	9.611	9.611

Notes: In columns 1 to 3, the outcome is the number of daycare facilities located in the municipality per 10,000 inhabitants (using the 2010 population), and the period of analysis is 2007-2017. In columns 4 to 6, the outcome is the number of public libraries located in the municipality per 10,000 inhabitants (using the 2010 population), the period of analysis is 2009-2017, and the sample is restricted to 7 départements for which data are available starting in 2009. Other notes as in Table B5.1.

Table B5.3: Impact using alternative control groups – public transport

	(1)	(2)	(3)	(4)	(5)
Outcome	Access to public transport				
Control group	Group 1	Group 2	Group 3	Group 4	Group 5
Treatment	0.033***	0.038***	0.051***	0.034***	0.051***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Municipality FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	220,948	170,912	111,944	202,174	108,878
Mean DepVar	0.024	0.024	0.024	0.024	0.024
Sd DepVar	0.153	0.153	0.153	0.153	0.153

Notes: The outcome is an indicator variable equal to one if the municipality has access to public transport. The sample excludes municipalities located in the Parisian region of Ile-de-France, for which the data are not available. Other notes as in Table B5.1.

Table B5.4: Impact using alternative control groups – fiscal revenues

	(1)	(2)	(3)	(4)	(5)
Outcome	Fiscal revenues per capita				
Control group	Group 1	Group 2	Group 3	Group 4	Group 5
Treatment	91.8*** (5.1)	89.2*** (5.1)	94.5*** (5.2)	91.6*** (5.1)	95.0*** (5.2)
Municipality FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	227,766	176,638	117,054	207,886	113,540
Mean DepVar	705.7	705.7	705.7	705.7	705.7
Sd DepVar	431.9	431.9	431.9	431.9	431.9

Notes: The outcome is the municipality's total fiscal revenues per capita, as defined in Section 5.2. I exclude from the analysis few municipalities for which the data are missing for at least one year over the period 2004-2017 (0.6 percent of the sample). Other notes as in Table B5.1.

C. Housing supply – heterogeneity analysis

In the main text, I report the impact on housing supply obtained by splitting the sample according to the median value of the heterogeneity variable I consider. In this section, I explore the heterogeneity of the treatment impact by adding an interaction between the treatment variable and the heterogeneity variable in equation (1). I standardize each heterogeneity variable, subtracting its mean and dividing it by its standard error. As a result, in each table below, the impact of the treatment can be interpreted as the impact for a municipality with an average value of the heterogeneity variable, and the estimate associated to the interaction can be interpreted as the change in the treatment effect due to a one standard deviation increase in the heterogeneity variable.

Table C1: Distance to the core of the urban area

	(1)	(2)
Outcome	Number of housing building permits	
Heterogeneity	Distance to urban core	
	Ratio	Log ratio
Treatment	9.213*** (2.032)	10.204*** (2.036)
Interaction	-3.301 (2.028)	-6.361*** (2.110)
Municipality FE	Yes	Yes
Time FE	Yes	Yes
Observations	121,230	121,020
Mean DepVar	64.842	64.842
Sd DepVar	90.756	90.756

Notes: The sample includes only municipalities part of an urban area. Standard errors are in parentheses and are clustered at the municipality level. ***, **, and * indicate significance at 1, 5, and 10% respectively. The mean of the dependent variable gives the average of the outcome variable in the treatment group during the pre-reform period (from 2004 to 2010). The outcome is the number of housing building permits delivered in the municipality per year per 10,000 inhabitants (using the 2010 population). It counts the total number of housing units allowed for construction. The heterogeneity variable is the ratio between the municipality's euclidean distance to the core divided by the average distance to the core of the other municipalities from the same urban area. In column 2, I consider the log of the ratio, thus excluding 15 municipalities in the control group that constitute the core of their urban area and for which the distance is thus equal to 0.

Table C2: Residents' income

	(1)
Outcome	Number of housing building permits
Heterogeneity	Residents' income
Treatment	6.562*** (1.668)
Interaction	5.416*** (1.950)
Municipality FE	Yes
Time FE	Yes
Observations	245,190
Mean DepVar	66.136
Sd DepVar	93.870

Notes: The heterogeneity variable is the average residents' taxable income in 2010. Data on taxable income in 2010 are missing for 16 municipalities with the smallest populations, to protect residents' privacy. Other notes as in Table C1.

Table C3: Housing density

For this heterogeneity test, in the second column, I also include the interaction between the treatment variable and the square value of the housing density. Indeed, even if we expect the impact on housing to be larger in denser municipalities, we might not expect the densest places to experience the largest increase, as they may be too dense to be able to increase the housing supply as much as elsewhere. As a result, the effect is likely to rise non-monotonically with the housing density. The results confirm this hypothesis: While in the first column the interaction is close to zero and not significant, in the second column the interaction is large and significant, and the estimate associated to the interaction with the square value is negative and significant.

	(1)	(2)
Outcome	Number of housing building permits	Housing density
Heterogeneity		
Treatment	10.524*** (1.903)	12.282*** (1.952)
Interaction	-0.288 (2.002)	10.028*** (3.824)
Interaction ²		-1.456*** (0.487)
Municipality FE	Yes	Yes
Time FE	Yes	Yes
Observations	151,125	151,125
Mean DepVar	65.288	65.288
Sd DepVar	89.761	89.761

Notes: The heterogeneity variable is the number of housing units per square kilometer in 2010. In column 2, I also include the interaction between the treatment variable and the square of the housing density. Other notes as in Table C1.

Table C4: Neighbors' characteristics

	(1)	(2)	(3)	(4)
Outcome	Number of housing building permits			
Heterogeneity	Ratio revenues	Ratio immigrants	Ratio unemployed	Vote share distance
Treatment	7.941*** (1.616)	7.916*** (1.604)	8.102*** (1.610)	8.239*** (1.609)
Interaction	2.033 (1.730)	-1.904 (2.365)	1.191 (1.775)	1.358 (2.002)
Municipality FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	245,190	245,430	245,430	245,430
Mean DepVar	66.136	66.203	66.203	66.203
Sd DepVar	93.870	94.440	94.440	94.440

Notes: Each ratio (columns 1 to 3) divides the value of the heterogeneity variable in the municipality by the average value in the other municipalities from the same département, weighted by their population. In column 1, the heterogeneity variable used to compute the ratio is the residents' average annual taxable income. Data on taxable income in 2010 are missing for 16 municipalities with the smallest populations, in order to protect residents' privacy. In column 2 (resp. 3), the heterogeneity variable is the share of immigrants (resp. unemployed) in 2010. In column 4, the heterogeneity variable is the absolute difference in the right vote share in the 2007 presidential election, between the municipality and the other municipalities from the same département. Other notes as in Table C1.

Table C5: Share of homeowners

	(1)
Outcome	Number of housing building permits
Heterogeneity	Share of homeowners
Treatment	8.895*** (1.601)
Interaction	-1.405 (1.764)
Municipality FE	Yes
Time FE	Yes
Observations	245,430
Mean DepVar	66.203
Sd DepVar	94.440

Notes: The heterogeneity variable is the share of homeowners in the municipality in 2010. Other notes as in Table C1.

D. Housing price indices

Description of the method

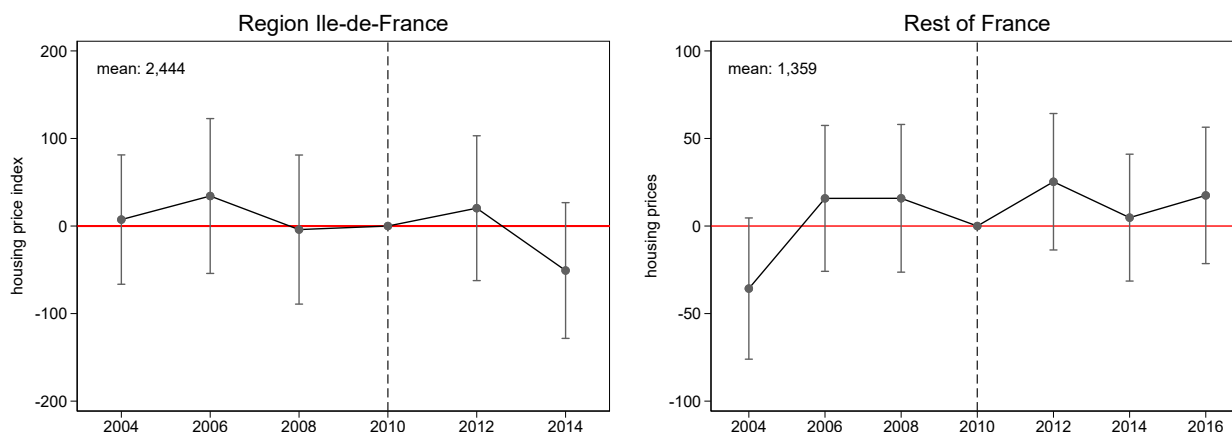
Following Combes et al. (2018), I built the housing price indexes using transaction data. These data come from an annual census conducted by the regional notary associations, which report the transactions of non-new dwellings. Although reporting is voluntary, it covers about 65 percent of all transactions. I perform the analysis separately for the Parisian region of Ile-de-France and for the rest of France, as the two databases come from two distinct notary associations and do not use the same definitions for the dwellings' characteristics. The data are made available by the Ministry of Sustainable Development for every even year since 2000. They are available until 2014 for Ile-de-France and 2016 for the rest of France.

First, following Gouriéroux and Laferrère (2009) and Musiedlak and Vignolles (2016)'s guidelines, I excluded some outliers from the transaction databases. Next, I regressed the log of the price per square meter of the dwelling on several characteristics. To build the baseline index I use in the paper, following Combes et al. (2018), I regressed the log of the price per square meter on indicator variables for the quarter of the transaction and the construction period. I next build a second index (which I refer to as the “augmented index”) using additional characteristics. For houses, I added the floorspace area, the size of the land, the number of rooms, bathrooms and floors, and whether the house had parking. For apartment, I added the floorspace area, the floor at which it was located, the number of rooms and bathrooms, whether the building had an elevator, and whether the apartment had parking and a cellar.

While the price of the transaction was never missing, the floorspace area of the dwelling was missing in 10 percent of the cases for apartments and in 36 percent of the cases for houses. To compute the price per square meter, I replaced the missing floorspace areas by the average floorspace area of an apartment or a house with the same number of rooms located in the same département. Results are the same if I instead drop the transactions for which the floorspace area was missing. Regarding the right-hand side variables (the dwelling's characteristics), I replaced the missing values by the average of the variable and added an indicator variable equal to one if the variable was missing. I also centered all explanatory variables by subtracting the means and dividing by the standard errors.

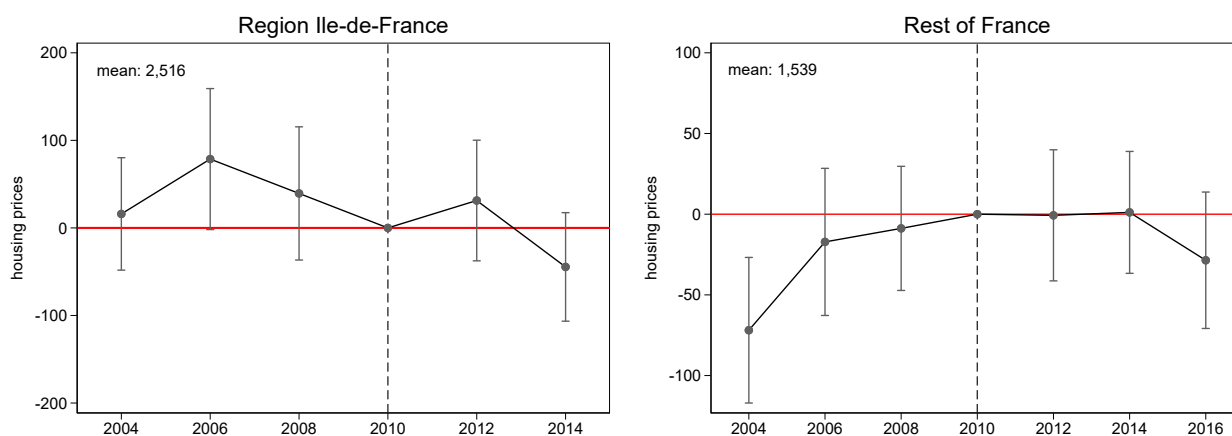
After running the hedonic regressions, I computed the mean of the residuals for each year and municipality separately, after having added the regression constant. Since the explanatory variables are centered, we can interpret the resulting indices as a price per square meter for a reference dwelling.

Figure D1: Impact on prices – unbalanced panel



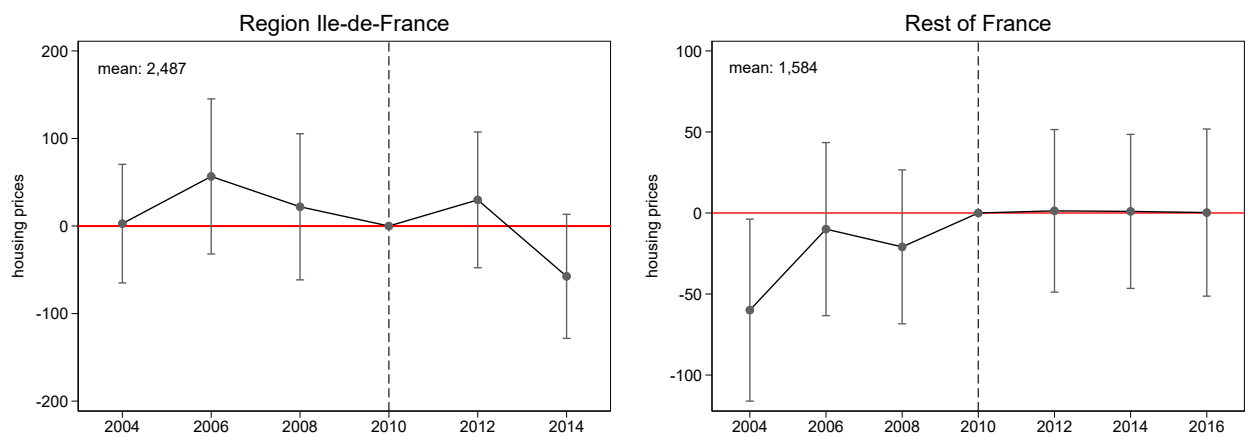
Notes: The sample includes all municipalities where at least one housing transaction took place in the period of analysis. The graph on the left-hand side includes only municipalities located in the Parisian region of Ile-de-France, while the graph on the right-hand side includes all the other municipalities. On each graph, the average value of the price index in the treatment group before 2010 is displayed on the top left corner.

Figure D2: Impact on prices – augmented index



Notes: For these graphs, I use an alternative version of the indices for which I include additional apartments' and houses' characteristics in the hedonic regressions. The sample includes only municipalities where at least one housing transaction took place each even year over the period studied. The graph on the left-hand side includes only municipalities located in the Parisian region of Ile-de-France, while the graph on the right-hand side includes all the other municipalities. On each graph, the average value of the price index in the treatment group before 2010 is displayed on the top left corner.

Figure D3: Impact on prices – urban area



Notes: The sample includes only municipalities part of an urban area and where at least one housing transaction took place each even year over the period studied. The graph on the left-hand side includes only municipalities located in the Parisian region of Ile-de-France, while the graph on the right-hand side includes all the other municipalities. On each graph, the average value of the price index in the treatment group before 2010 is displayed on the top left corner.

References

- Alesina, A. (1987). Macroeconomic policy in a two-party system as a repeated game. *Quarterly Journal of Economics*, 102(3):651–678.
- Alesina, A., Baqir, R., and Hoxby, C. (2004). Political Jurisdictions in Heterogeneous Communities. *Journal of Political Economy*, 112(2):348–396.
- Alesina, A. and Spolaore, E. (1997). On the number and size of nations. *The Quarterly Journal of Economics*, 112(4):1027–1056.
- Alesina, A. and Spolaore, E. (2003). *The size of nations*. MIT Press, Cambridge, MA.
- Alonso, W. (1964). *Location and land use. Toward a general theory of land rent*. Harvard University Press, Cambridge, MA.
- Alvarez, R. M. and Nagler, J. (2000). A New Approach for Modeling Strategic Voting in Multiparty Elections. *British Journal of Political Science*, 30(1):57–75.
- Anagol, S. and Fujiwara, T. (2016). The Runner-Up Effect. *Journal of Political Economy*, 124(4):927–991.
- Arrow, K. J. (1951). *Social Choice and Individual Values*. Wiley, New York.
- Atkeson, L. R. (1999). Sure, I Voted for the Winner! Overreport of the Primary Vote for the Party Nominee in the National Election Studies. *Political Behavior*, 21(3):197–215.
- Balinski, M. and Laraki, R. (2011). *Majority judgment: measuring, ranking, and electing*. MIT press, Cambridge, MA.
- Barankay, I. (2018). Rank incentives: Evidence from a randomized workplace experiment. *Working Paper*.
- Baron, D. P. (1996). A dynamic theory of collective goods programs. *American Political Science Review*, 90(2):316–330.
- Baron, D. P. and Ferejohn, J. A. (1989). Bargaining in legislatures. *American political science review*, 83(4):1181–1206.
- Barro, R. J. (1973). The control of politicians: an economic model. *Public Choice*, pages 19–42.
- Bartels, L. M. (1985). Expectations and Preferences in Presidential Nominating Campaigns. *American Political Science Review*, 79(3):804.

- Bates, L. J. and Santerre, R. E. (1994). The determinants of restrictive residential zoning: Some empirical findings. *Journal of Regional Science*, 34(2):253–263.
- Battaglini, M. and Coate, S. (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review*, 98(1):201–36.
- Bel, G. and Warner, M. E. (2015). Inter-municipal cooperation and costs: expectations and evidence. *Public Administration*, 93(1):52–67.
- Bel, G. and Warner, M. E. (2016). Factors Explaining Inter-municipal Cooperation in Service Delivery : A Meta-Regression Analysis. *Journal of Economic Policy Reform*, 19(2):91–115.
- Besley, T. and Coate, S. (1997). An Economic Model of Representative Democracy. *The Quarterly Journal of Economics*, 112(1):85–114.
- Blais, A., Dolez, B., and Laurent, A. (2017). Strategic voting in the second round of a two-round system: The 2014 French municipal elections. *French Politics*, 15(1):27–42.
- Blais, A., Nadeau, R., Gidengil, E., and Nevitte, N. (2001). Measuring strategic voting in multiparty plurality elections. *Electoral Studies*, 20(3):343–352.
- Blom-Hansen, J., Houlberg, K., Serritzlew, S., and Treisman, D. (2016). Jurisdiction Size and Local Government Policy Expenditure: Assessing the Effect of Municipal Amalgamation. *American Political Science Review*, 110(4):812–831.
- Bordignon, M., Nannicini, T., and Tabellini, G. (2016). Moderating Political Extremism: Single Round versus Runoff Elections under Plurality Rule. *American Economic Review*, 106(8):2349–2370.
- Borusyak, K. and Jaravel, X. (2017). Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume. *Available at SSRN 2826228*.
- Bouton, L. (2013). A theory of strategic voting in runoff elections. *American Economic Review*, 103(4):1248–1288.
- Bouton, L. and Gratton, G. (2015). Majority Runoff Elections: Strategic Voting and Duverger’s Hypothesis. *Theoretical Economics*, 10(2):283–314.
- Bouton, L., Llorente-Saguer, A., and Castanheira, M. (2015). Multicandidate Elections: Aggregate Uncertainty in the Laboratory. *Games and Economic Behavior*, 101:132–150.
- Bouton, L. and Ogden, B. (2017). Ethical Voting in Multicandidate Elections. *NBER Working Paper 23898*.
- Bowen, T. R., Chen, Y., and Eraslan, H. (2014). Mandatory versus discretionary spending: The status quo effect. *American Economic Review*, 104(10):2941–74.

- Brueckner, J. K. (1998). Testing for Strategic Interaction Among Local Governments: The Case of Growth Controls. *Journal of Urban Economics*, 44(3):438–467.
- Cagé, J. (2017). Media Competition, Information Provision and Political Participation: Evidence from French Local Newspapers and Elections, 1944-2014. *CEPR Discussion Paper 12198*.
- Callander, S. (2007). Bandwagons and momentum in sequential voting. *Review of Economic Studies*, 74(3):653–684.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2018). Regression Discontinuity Designs Using Covariates. *Forthcoming in Review of Economics and Statistics*.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.
- Castanheira, M. (2003). Why Vote for Losers? *Journal of the European Economic Association*, 1(5):1207–1238.
- Chapelle, G. and Eyméoud, J.-B. (2019). Can Big Data increase our knowledge of local rental markets? Estimating the cost of density with rents. *Working Paper*.
- Chetty, R., Hendren, N., and Katz, L. F. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902.
- Clingermayer, J. C. (1994). Electoral Representation, Zoning Politics, and the Exclusion of Group Homes. *Political Research Quarterly*, 47(4):969–984.
- Coate, S. and Conlin, M. (2004). A Group-Rule Utilitarian Approach to Voter Turnout: Theory and Evidence. *American Economic Review*, 94(5):1476–1504.
- Combes, P.-P., Duranton, G., and Gobillon, L. (2018). The Costs of Agglomeration : House and Land Prices in French Cities. *Review of Economic Studies*, 86(4):1556–1589.
- Cox, G. W. (1994). Strategic Voting Equilibria Under the Single Nontransferable Vote. *American Political Science Review*, 88(3):608–621.
- Cox, G. W. (1997). *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge University Press, Cambridge.
- Dasgupta, P. and Maskin, E. (2019). Elections and Strategic Voting: Condorcet and Borda.
- de Chaisemartin, C. and D'Haultfoeuille, X. (2019). Two-way fixed effects estimators with heterogeneous treatment effects. *NBER Working Paper No. w25904*.

- de la Cuesta, B. and Imai, K. (2016). Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections. *Annual Review of Political Science*, 19:375–396.
- Deltas, G., Herrera, H., and Polborn, M. K. (2016). Learning and Coordination in the Presidential Primary System. *Review of Economic Studies*, 83(4):1544–1578.
- Di Porto, E. and Paty, S. (2018). Cooperation among Local Governments to Deliver Public Services. *Politics & Policy*, 46(5):790–820.
- Downs, A. (1957a). *An Economic Theory of Democracy*. Harper and Row, New York.
- Downs, A. (1957b). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2):135–150.
- Duverger, M. (1954). *Political Parties*. Wiley.
- Egger, P. H., Köthenbürger, M., and Loumeau, G. (2017). Local Border Reforms and Economic Activity. *CESifo Working Paper, No. 6738*.
- Eggers, A. C. (2015). Proportionality and Turnout: Evidence From French Municipalities. *Comparative Political Studies*, 48(2):135–167.
- Eggers, A. C., Fowler, A., Hainmueller, J., Hall, A. B., and Snyder, J. M. (2015). On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races. *American Journal of Political Science*, 59(1):259–274.
- Fabre, B. (2017). Political Colleagues Matter: The Impact of Multiple Office-Holding on Intergovernmental Grants. *PSE Working Paper 2017-43*.
- Fauvelle-Aymar, C. and François, A. (2005). Campaigns, Political Preferences and Turnout: An Empirical Study of the 1997 French Legislative Elections. *French Politics*, 3(1):49–72.
- Feddersen, T. and Pesendorfer, W. (1997). Voting Behavior and Information Aggregation in Elections With Private Information. *Econometrica*, 65(5):1029–1058.
- Feddersen, T. and Sandroni, A. (2006). A theory of participation in elections. *American Economic Review*, 96(4):1271–1282.
- Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice*, pages 5–25.
- Festinger, L. (1962). Cognitive dissonance. *Scientific American*, 207(4):93–107.
- Fey, M. (1997). Stability and Coordination in Duverger's Law : A Formal Model of Preelection Polls and Strategic Voting. *American Political Science Review*, 91(1):135–147.
- Fischel, W. A. (2001). *The Homevoter Hypothesis*. Harvard University Press, Cambridge, MA.

- Fisher, S. D. and Myatt, D. P. (2017). Strategic Voting in Plurality Rule Elections. *Working Paper*.
- Fiva, J. H. and Smith, D. M. (2017). Local Candidates and Voter Mobilization : Evidence from Historical Two-Round Elections in Norway. *Electoral Studies*, 45:130–140.
- Folke, O., Persson, T., and Rickne, J. (2016). The primary effect: Preference votes and political promotions. *American Political Science Review*, 110(3):559–578.
- Forsythe, R., Myerson, R. B., Rietz, T. A., and Weber, R. J. (1993). An experiment on coordination in multi-candidate elections: The importance of polls and election histories. *Social Choice and Welfare*, 10(3):223–247.
- Foucault, M. and François, A. (2005). Le rendement des dépenses électorales en France. *Revue Economique*, 56(5):1125–1143.
- Frieden, B. (1979). A New Regulation Comes to Suburbia. *Public Interest*, 55:15–27.
- Fujiwara, T. (2011). A Regression Discontinuity Test of Strategic Voting and Duverger’s Law. *Quarterly Journal of Political Science*, 6:197–233.
- Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601.
- Glaeser, E. and Gyourko, J. (2018). The economic implications of housing supply. *Journal of Economic Perspectives*, 32(1):3–30.
- Glaeser, E. L. (2014). Land Use Restrictions and Other Barriers to Growth. *Cato Institute*.
- Glaeser, E. L. and Kahn, M. E. (2010). The greenness of cities: carbon dioxide emissions and urban development. *Journal of Urban Economics*, 67(3):404–418.
- Glaeser, E. L. and Maré, D. C. (2001). Cities and Skills. *Journal of Labor Economics*, 19(2):316–342.
- Glaeser, E. L. and Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3):265–278.
- Goodman-Bacon, A. (2018). Difference-in-Differences With Variation in Treatment Timing. *NBER Working Paper No. 25018*.
- Gordon, N. and Knight, B. (2009). A spatial merger estimator with an application to school district consolidation. *Journal of Public Economics*, 93(5-6):752–765.
- Gouriéroux, C. and Laferrère, A. (2009). Managing hedonic housing price indexes: The French experience. *Journal of Housing Economics*, 18(3):206–213.

- Gyourko, J. and Molloy, R. (2015). Regulation and Housing Supply. In Duranton, G., Henderson, V. J., and Strange, W., editors, *Handbook of Regional and Urban Economics*, volume 5b, pages 1289–1337. Amsterdam: Elsevier.
- Hall, A. B. and Snyder, J. M. (2015). Information and Wasted Votes: A Study of U.S. Primary Elections. *Quarterly Journal of Political Science*, 10(4):433–459.
- Hankinson, M. (2018). When Do Renters Behave Like Homeowners? High Rent, Price Anxiety, and NIMBYism. *American Political Science Review*, 112(3):473–493.
- Harjunen, O., Saarimaa, T., and Tukiainen, J. (2019). Political representation and effects of municipal mergers. *Political Science Research and Methods*, pages 1–17.
- Hartzmark, S. M. (2015). The Worst, the Best, Ignoring All the Rest: The Rank Effect and Trading Behavior. *Review of Financial Studies*, 28(4):1024–1059.
- Hibbs Jr, D. A. (1994). The partisan model of macroeconomic cycles: More theory and evidence for the United States. *Economics & Politics*, pages 1–23.
- Hilber, C. A. and Robert-Nicoud, F. (2013). On the origins of land use regulations: Theory and evidence from US metro areas. *Journal of Urban Economics*, 75:29–43.
- Hillygus, D. S. (2007). The Dynamics of Voter Decision Making Among Minor-Party Supporters: The 2000 Presidential Election in the United States. *British Journal of Political Science*, 37(2):225–244.
- Hinich, M. J. (1981). Voting as an act of contribution. *Public Choice*, 36(1):135–140.
- Hsieh, C. T. and Moretti, E. (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.
- Hulst, R., Montfort, A. V., Haveri, A., Airaksinen, J., and Kelly, J. (2009). Institutional Shifts In Inter-Municipal Service Delivery. *Public Organization Review*, 9:263–285.
- Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):933–959.
- Imbens, G. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.
- Jackson, K. T. (1987). *Crabgrass frontier: The suburbanization of the United States*. Oxford University Press.
- Jones, C. and Kammen, D. M. (2013). Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density. *Environmental Science & Technology*, 48(2):895–902.

- Kawai, K. and Watanabe, Y. (2013). Inferring strategic voting. *American Economic Review*, 103(2):624–662.
- Kiewiet, D. R. (2013). The Ecology of Tactical Voting in Britain. *Journal of Elections, Public Opinion and Parties*, 23(1):86–110.
- Kiss, Á. (2015). Identifying strategic voting in two-round elections. *Electoral Studies*, 40:127–135.
- Kiss, Á. and Simonovits, G. (2014). Identifying the bandwagon effect in two-round elections. *Public Choice*, 160(3-4):327–344.
- Lapointe, S. (2018). Love thy neighbour? The effect of income and language differences on votes for municipal secessions. *Regional Science and Urban Economics*, 70:229–245.
- Laslier, J.-F. (2009). The Leader Rule: A Model of Strategic Approval Voting in a Large Electorate. *Journal of Theoretical Politics*, 21(1):113–136.
- Le Pennec, C. and Pons, V. (2019). Vote Choice Formation and Minimal Effects of TV Debates : Evidence from 61 Elections in 9 OECD Countries. *NBER Working Paper 26572*.
- Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675–697.
- Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies*, 76(3):1071–1102.
- Lidström, A. (2010). The Swedish Model under Stress: The Waning of the Egalitarian, Unitary State? In *Territorial Choice. The Politics of Boundaries and Borders*, pages 61–80. Houndsmills: Palgrave.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.
- Meriläinen, J. and Tukiainen, J. (2016). Primary Effect in Open-List Elections. *Working paper*.
- Morton, R. B., Muller, D., Page, L., and Torgler, B. (2015). Exit Polls, Turnout, and Bandwagon Voting: Evidence from a Natural Experiment. *European Economic Review*, 77:65–81.
- Mouritzen, P. E. (2010). The Danish Revolution in Local Government: How and Why? In *Territorial Choice. The Politics of Boundaries and Borders*, pages 21–41. Houndsmills: Palgrave.
- Musiedlak, Y. and Vignolles, B. (2016). Les mouvements des prix immobiliers dans l'ancien au cours des années 2000 : des marchés locaux différenciés. *Document de travail du CGEDD*, (24):1–40.

- Myatt, D. P. (2007). On the Theory of Strategic Voting. *Review of Economic Studies*, 74(1):255–281.
- Myatt, D. P. and Fisher, S. D. (2002). Tactical Coordination in Plurality Electoral Systems. *Oxford Review of Economic Policy*, 18(4):504–522.
- Myerson, R. B. and Weber, R. J. (1993). A Theory of Voting Equilibria. *American Political Science Review*, 87(1):102–114.
- Nordhaus, W. D. (1975). The political business cycle. *Review of Economic Studies*, 42(2):169–190.
- Nurmi, H. (1983). Voting Procedures : A Summary Analysis. *British Journal of Political Science*, 13(2):181–208.
- Oates, W. E. (1972). *Fiscal Federalism*.
- Ortalo-Magné, F. and Prat, A. (2014). On the Political Economy of Urban Growth: Homeownership versus Affordability. *American Economic Journal: Microeconomics*, 6(1):154–181.
- Osborne, M. and Slivinski, A. (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics*, 111(1):65–96.
- Palfrey, T. R. (1989). A mathematical proof of Duverger's law. In Ordeshook, P. C., editor, *Models of Strategic Choice in Politics*, pages 69–92. University of Michigan Press, Ann Arbor.
- Palfrey, T. R. and Rosenthal, H. (1983). A Strategic Calculus of Voting. *Public Choice*, 41(1):7–53.
- Palfrey, T. R. and Rosenthal, H. (1985). Voter Participation and Strategic Uncertainty. *American Political Science Review*, 79(1):62–78.
- Piketty, T. (2000). Voting as Communicating. *Review of Economic Studies*, 67(1):169–191.
- Pons, V. and Tricaud, C. (2018). Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates. *Econometrica*, 86(5):1621–1649.
- Pope, D. G. (2009). Reacting to Rankings: Evidence from "America's Best Hospitals". *Journal of Health Economics*, 28(6):1154–1165.
- Purcell, M. (2006). Urban Democracy and the Local Trap. *Urban Studies*, 43(11):1921–1941.
- Reingewertz, Y. (2012). Do municipal amalgamations work? Evidence from municipalities in Israel. *Journal of Urban Economics*, 72(2-3):240–251.

- Rodden, J. (2002). The Dilemma of Fiscal Federalism: Grants and Fiscal Performance around the World. *American Journal of Political Science*, 46(3):670–687.
- Rodden, J. (2006). *Hamilton's paradox : the promise and peril of fiscal federalism*. Cambridge University Press.
- Rolleston, B. S. (1987). Determinants of restrictive suburban zoning: An empirical analysis. *Journal of Urban Economics*, 21(1):1–21.
- Saarimaa, T. and Tukiainen, J. (2014). I Don't Care to Belong to Any Club That Will Have Me as a Member: Empirical Analysis of Municipal Mergers. *Political Science Research and Methods*, 2(1):97–117.
- Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics*, 125(3):1253–1296.
- Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217.
- Shayo, M. and Harel, A. (2012). Non-consequentialist voting. *Journal of Economic Behavior and Organization*, 81:299–313.
- Sorensen, R. J. (2006). Local Government Consolidations: The Impact of Political Transaction Costs. *Public Choice*, 127(1):75–95.
- Spenkuch, J. (2018). Expressive vs. Strategic Voters: An Empirical Assessment. *Journal of Public Economics*, 165:73–81.
- Spenkuch, J. L. (2015). Please Don't Vote for Me : Voting in a Natural Experiment with Perverse Incentives. *Economic Journal*, 125(585):1025–1052.
- Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.
- Tocqueville, A. (1835). *De la démocratie en Amérique*.
- Ujhelyi, G., Chatterjee, S., and Szabó, A. (2017). 'None Of The Above' Votes in India and the Consumption Utility of Voting. *Working Paper*.
- Van der Straeten, K., Laslier, J.-F., Sauger, N., and Blais, A. (2010). Strategic, sincere, and heuristic voting under four election rules: An experimental study. *Social Choice and Welfare*, 35(3):435–472.
- Weese, E. (2015). Political mergers as coalition formation : An analysis of the Heisei municipal amalgamations. *Quantitative Economics*, 6:257–307.

Wright, G. C. (1993). Errors in Measuring Vote Choice in the National Election Studies, 1952-88. *American Journal of Political Science*, 37(1):291–316.

Titre : Essais en Économie Politique et Publique

Mots clés : Élections; Comportements de vote; Politiques publiques locales; Coopération intercommunale; Régression sur discontinuités; Doubles différences

Résumé : Cette thèse vise à atteindre une meilleure compréhension du comportement des citoyens et des élus, à travers l'utilisation de méthodes quasi-expérimentales et de données administratives. Les deux premiers chapitres, en collaboration avec Vincent Pons, étudient le comportement des électeurs.

Le premier chapitre évalue l'étendue et les conséquences du vote expressif. En utilisant la méthode de la régression sur discontinuités (RDD), nous comparons les résultats électoraux selon que deux ou trois candidats se présentent. Nous montrons que la présence d'un.e troisième candidat.e augmente la participation, et, dans environ 1/5^{ème} des élections, provoque la perte du candidat qui, parmi les deux premiers, était le plus proche idéologiquement. Ces résultats suggèrent qu'une large fraction des électeurs préfère voter expressivement plutôt que stratégiquement.

Le deuxième chapitre évalue l'impact du classement des candidats. En utilisant une RDD

au sein d'élections à deux tours, nous montrons qu'être classé.e marginalement premier.ère au premier tour, augmente la probabilité qu'un.e candidat.e se maintienne au second tour et gagne. Ces résultats suggèrent que le désir de voter pour le.a vainqueur.e joue un rôle clé dans la décision de vote des électeurs.

Le troisième chapitre étudie le comportement des élus: la réticence des maires à coopérer en matière de politiques publiques. J'exploite la loi de 2010, qui en France, a forcé les communes non-intégrées à rejoindre une intercommunalité. J'étudie l'impact causal de leur intégration en utilisant la méthode des doubles différences. Mes résultats montrent que les communes résistaient afin d'échapper aux conséquences économiques locales de l'intégration: les communes urbaines résistaient afin d'éviter la construction de nouveaux logements, tandis que les communes rurales résistaient afin d'éviter la perte de services publics locaux.

Title : Essays in Political Economy and Public Economics

Keywords : Elections; Voting; Local public policies; Intermunicipal cooperation; Regression-discontinuity-design; Difference-in-differences

Abstract: This thesis combines quasi-experimental designs with administrative data to better understand citizen and policymaker behavior. The first two chapters, in collaboration with Vincent Pons, study the behavior of voters in elections.

In Chapter 1, we assess the extent and consequences of expressive voting. Using a regression discontinuity design (RDD), we compare electoral results when two or three candidates compete. We find that a third candidate's presence substantially increases participation, and, in about 1/5 of the elections, causes the loss of the candidate among the top two that is ideologically closest to her. Hence, we show that many voters value voting expressively over voting strategically.

Chapter 2 assesses the impact of candidate ranking. Using an RDD in two-round elections,

we show that arriving marginally first has large effects on a candidate's likelihood of running and winning in the second round. These results suggest that the desire to vote for the winner is an important driver of voter behavior.

Chapter 3 studies the behavior of policymakers and seeks to understand mayors' reluctance to cooperate over local public policies. Exploiting a 2010 reform in France that forced non-integrated municipalities to join an intermunicipal community, I use a difference-in-differences strategy to assess the causal impact of their integration. I provide new evidence that resistance is driven by local costs of integration: urban municipalities resisted integration to prevent further housing construction, while rural municipalities resisted to avoid losing local public services.