Voter Registration Costs and Disenfranchisement: Experimental Evidence from France*

Céline Braconnier, Jean-Yves Dormagen and Vincent Pons

Paper in progress - this version: March 19th, 2013

Abstract

We provide experimental evidence on distortions introduced by non-automatic voter registration processes in terms of registration, turnout, and electoral outcomes. Before the 2012 French Presidential and General elections, 4,118 buildings hosting 38,000 citizens were randomly allocated to a control group or one of six treatment groups varying by the timing, number, and content of door-to-door canvassing visits: simple encouragement and information, or offer to register at home instead of having to go to the town hall. The visits did not affect the participation of citizens already registered, but increased the number of new registrations by 30%, bringing in citizens almost equally likely to vote and with different socioeconomic characteristics and political preferences than the average. This suggests that facilitating the registration process would enfranchise some groups, including the youth and immigrants otherwise less likely to register due to higher costs, and improve the representativeness of electoral outcomes. This could also theoretically disengage active citizens, a hypothesis for which our intricated design provides the first existing test: we evaluate the counterfactual participation of those who got registered at home, if they had registered on their own, and reject the hypothesis.

*We are grateful to Esther Duflo, Ben Olken, Abhijit Banerjee, Daniel Posner, James Snyder, Adrien Auclert and seminar participants at MIT and the Tobin Project Graduate Students Forum for suggestions that have improved the paper. We thank Caroline Le Pennec and Ghislain Gabalda for the outstanding research assistance they provided throughout the entire project and Aude Soubiron for her assistance in the administration of the interventions in the cities surrounding Bordeaux. We thank the town hall administration of each of the twelve cities included in the experiment for their generous collaboration and are indebted to all canvassers who administered the interventions, including local units of the Socialist Party in Cergy, Sevran, Carcassonne and Pessac, the local unit of the Front de Gauche in the 20th arrondissement of Paris, the NGO of retired workers of the MGEN, the NGO "Tous Citoyens", the NGO "RAJ-LR", and students of the universities of Montpellier 1, Cergy-Pontoise, the Ecole Normale Supérieure and the IEP of Bordeaux. We gratefully acknowledge funding by the Abdul Latif Jameel Poverty Action Lab, MIT France, the Russell Sage Foudation, the Tobin Project, the University of Montpellier 1, the University of Cergy-Pontoise and the cities of Montpellier, Blanquefort, Le Taillan, Pessac, Eysines and Lormont.

1 Introduction

Voter registration serves competing purposes such as preventing electoral fraud and ensuring that all eligible citizens can vote. In most countries, it falls under the responsibility of the state and all citizens are automatically registered to vote. On the contrary, in countries such as the United States and France, it is the citizens' responsibility to register. The registration process can then make voting costly (Rosenstone and Hansen 1993, Timpone 1998). In addition to the nonregistered citizens, who cannot vote (5% in France (Bréchon 2009), 29% in United States (US Census Bureau 2012)), a large but unknown number are "misregistered": they have moved out of their previous address but stay registered there. They have to travel or vote by mail or by proxy (if possible) if they want to participate in the elections. Braconnier and Dormagen (2007, 2012) track the registration status, electoral participation and moving out of inhabitants of a small French "banlieue" and observe dramatic differences in turnout between misregistered and well-registered citizens. In this paper, we explore to what extent non-automatic registration processes introduce important distortions in terms of participation and electoral outcomes, using the case of France.

Such distortions cannot be taken for granted: the registration procedure, identical for all, might simply select citizens who are more willing to participate in the elections. In this view, failure to register or reregister after moving signals a low propensity to vote or, for the misregistered, decreased intensity of the mobilizing influence of family, friends and neighbors related to their arrival to a new city or neighborhood (Franklin 2005). Even if they were added to the voter rolls, very few nonregistered and misregistered citizens would actually vote. An alternative view is that people relate differently to the registration procedure: some are unaware of it or fail to project themselves to the elections before the registration deadline; others have difficulties gathering the appropriate information or find it costly to complete the required administrative steps. Failure to register then signals a higher cost rather than a lower propensity to vote and the registration process significantly reduces turnout. A last view is that the nonregistered and misregistered might have a lower propensity to vote than other citizens but, conditional on being registered, their actual participation would be enhanced by the mobilization effect of political campaigns and they would display particularly large turnout differences between high- and low- saliency elections. In addition to affecting overall turnout, and to the extent that the cost of registering is correlated with specific socioeconomic characteristics and political preferences, the non-automatic procedure can further contribute to marginalize some categories of citizens, and distort electoral outcomes. Finally, beyond these negative effects, the early involvement of the citizens in the electoral process might play a positive role: based on theoretical and empirical results established by the self-determination theory (Ryan and Deci 2000, Tirole **XXX**) one could expect that non-automatic registration procedures enhance the feeling, among registered citizens, that their political engagement has an internal rather than external "locus of causality", and thus increase their subsequent turnout and involvement in the elections.

It is key to find the answers to these questions, as the political recommendations are orthogonal. But this is also challenging. First, the nonregistered and misregistered are hard to localize. Unlike studies of voter turnout, which can use voter lists as their sample, there is usually no preexisting list of nonregistered and misregistered citizens and it is even more difficult to distinguish these two groups. Second, it is difficult to measure these citizens' propensity to vote or the cost that the registration process means to them directly, with survey questions. Ideally, these would be inferred by comparing electoral outcomes obtained in the same area observed at the same time under a non-automatic and automatic registration procedure. This is of course impossible. Instead, the existing literature has relied on difference-in-difference strategies to identify how variations in the registration procedure affect registration and participation (Martinez and Hill 1999). This methodology relies on the strong assumption that changes in registration rules are independent from any other trend that could also affect electoral outcomes. Other papers estimate determinants of registration and turnout separately (Erikson 1981) and, based on the estimated parameters, predict the likelihood that non-registrants would go to the polls, conditional on being registered (Timpone 1998). This prediction can however be biased by unobserved variables which affect the decisions to register or vote. A last and more

	Early visit	Late visit
Control		
Early Canvassing	*Canvassing*	
Late Canvassing		*Canvassing*
Early Home registration	*Home registration*	
Late Home registration		*Home registration*
Early Canvassing + Late Home registration	*Canvassing*	*Home registration*
Early Home registration + Late home registration	*Home registration*	*Home registration*

Figure 1: Experimental design

recent strand of the literature draws on the wave of randomized controlled trials launched by Gerber and Green (2000) and now conducted in an increasing number of countries (e.g. e.g. John and Brannan 2008; Banerjee et al. 2012; Aker, Collier and Vicente 2011; Pons and Liegey 2013; Pons 2013), to estimate the impact of interventions facilitating registration. In line with results obtained by interventions encouraging turnout, Bennion and Nickerson (2009) find that face-to-face encouragement significantly increases registration among college students and is more effective than mail and email. Beyond registration, Nickerson (2010) measures the turnout of citizens registered thanks to registration drives and finds that they substantially increase registration but that only a small fraction of the additional newly registered voters participate in the following election. To some extent, this subsequent turnout indicates how likely these citizens would be to vote, conditional on being automatically registered. However, their turnout might also be directly affected by the registration visits, which inevitably give information about the elections and can contribute to make it more salient and create commitment to vote towards the canvasser, or alternatively disengage them as the self-determination theory would predict.

In this paper, we evaluate the impact of two different types of partisan and non-partisan registration visits on actual registration and participation. In some addresses, door-to-door canvassers provided information and encouragement to register. In others, they brought the procedure as close as possible to automatic registration: while French citizens usually register by going to the town hall, canvassers offered to fill the application form with them at their place, took pictures of their ID and a proof of address to complete the file and brought it to the town hall themselves. Twelve French cities and a total of 4,118 addresses hosting approximately 38,000 citizens were included in this experiment. Prior to the interventions, enumerators identified all apartments of the addresses likely to host nonregistered or misregistered citizens by comparing names found on the voter rolls and on the mailboxes. The addresses were then randomized between one control and six treatment groups, each defined by the timing, content and number of visits. The visits took place two to three months (for the early visits) or during the last month (for the late visits) before the December 31st 2011 registration deadline.

We find that all interventions but the least intensive (group 1) significantly increased the number of votes cast by initially nonregistered and misregistered citizens at the subsequent Presidential and General elections. Averaged on all rounds, the increase was by 45% for group 6 which received the most intensive intervention. This result is the product of the increased number of new registrations and differences in turnout rates of the different groups, which we then disentangle. Simple door-to-door canvassing visits (groups 1 and 2) increased new registrations by 14% on average, compared to 26% for visits offering home registration (groups 3 and 4). This shows that both lack of information and the time to go and register at the town hall enter in the cost of registering. In addition, late visits (groups 2 and 4) were more effective than early visits (groups 1 and

3) although these left more time to register, suggesting that the decision to register is subject to important procrastination, which is more effectively addressed by visits made close to the deadline.

Official turnout data further show that, on average, the turnout of newly registered citizens in the treatment groups is nearly as high as in the control group, and higher than the turnout of citizens already on the voter rolls prior to 2011. This could theoretically reflect different factors. First, the canvassers' visits might have affected turnout directly, by providing information about the forthcoming elections and encouragement to vote, in addition to reducing the cost of registering. This direct impact was assumed away by previous papers. We estimate it on the subsample of citizens who had registered before our visits, and could only have been affected by this direct impact, and do not find any evidence of an effect. Second, in addresses that received home registration visits, getting registered at home might have affected the participation of those who accepted this procedure, either by disengaging them or, on the contrary, by creating a feeling of indebtedness towards the canvassers. Groups 5 and 6 of our design were built to provide the first existing test for these effects, using a strategy inspired from Karlan and Zinman (2009). In both groups, each address was visited twice. We focus on the citizens who opened their door during the second visit. All of them received the most intensive treatment, a late visit of home registration. As expected, the fraction of newly registered is thus the same in both groups. The only difference is that a higher fraction was registered at home in group 6, since citizens were offered this opportunity earlier and had thus less time to register on their own, at the town hall. However, we do not find any significant turnout difference between the two groups, showing that registering people at home rather than asking them to register at the town hall did not affect in itself their participation. These possible explanations ruled out, we can attribute differences in turnout between the control and treatment groups to differences between the profile of always-takers, who register even absent the intervention, and compliers, who got registered only thanks to the intervention. The latter were only slightly less likely to vote than the former, a difference that is significant only in addresses that were offered home registration.

This does not necessarily mean that the interest in participating in the elections was equally high among always-takers and compliers, at the time when they had to take the decision to register. Indeed, by comparing turnout figures at the Presidential and less salient General elections (INSEE 2012), we find some evidence that the compliers are relatively more affected than always-takers by the mobilization effect of the Presidential elections: their turnout drops further at the General elections. A comprehensive postelectoral survey conducted after the elections on a sample of 1,500 citizens corroborates this view: in the treatment groups, the compliers express an increased interest in the recent political campaigns as well as higher competence. This survey further shows that newly registered compliers and always-takers were more likely to vote for left candidates than the average registered citizen of the same polling stations. Finally, the compliers and never-takers, who fail to vote partly because of the registration procedure, have different socioeconomic profile than the average citizens in their neighborhoods: they are younger and a higher share of them are immigrants. In short, although identical for all, the non-automatic French registration procedure creates important distortions: it decreases turnout by imposing a higher cost to some subgroups of the population that are already economically and socially marginalized, and biases electoral outcomes.

The reminder of the paper is organized as follows. In Section 2, we provide more details about the experiment and the research design. In Section 3, we describe the data used in the paper and provide a broad picture of the structure of the population in the sample. Section 4 presents a simple model of the two-step process of registering and voting, which maps unobserved distributions of benefits of voting and registration cost with empirical predictions about the impacts of our interventions. Section 5 outlines the empirical analysis, conducted in Sections 6 to 9. Section 10 concludes with a discussion of the results: how do they compare with results obtained by experiments conducted in the US, and what do they teach us about the impact of other procedures that can still make voting costly, in a context where many traditional costs, including distance to the polling stations, poll taxes or literacy tests, have dramatically decreased.

2 Research design

2.1 The context of the experiment

The French registration process

Unlike in most other countries, where voter registration falls under the responsibility of governments and all citizens beyond age of majority and with full civil rights can vote, in a few countries, including France, the United States, Australia, Great Britain or Portugal, it is the citizens' responsibility to complete a preliminary registration step if they want their name to be added to the voters' list. In France, registering is compulsory, but failure to register is not punished in any way, except than the impossibility to vote.[1]

The registration deadline is early: one has to register before December 31st to be able to partake in the elections of the following year, which typically do not take place until April. Registering requires to fill in a form and provide an ID and a proof of address, such as a recent electricity or gas bill or the latest tax return. Most people register in person at the town hall, although the registration file, once signed by the applicant, can be brought to the town hall by a third party, mailed, or, in some cities, completed online. The documents accepted as a valid ID or proof of address differ from one city to another, which can make the registration process more or less cumbersome.[2]Once submitted, the application is entered on a software, examined by an electoral committee, and is subject to be rejected if one of the pieces is missing or considered invalid.

A registered voter does not have to update her status until she moves. If she moves to a new address or a new city, she has to file a new application for her name to be erased from the list of voters allocated to her previous polling station and added to the list related to the polling station closest to her new place. Voters who fail to update their registration status can continue to vote at their previous polling station, unless they get struck off from the list. The town hall's administration is responsible for striking off voters to whom the political propaganda and voter's ID is repeatedly sent and sent back (signalling that the recipients have changed address) and who fail to come to the polling station and vote[3]. This procedure is however applied more or less systematically, depending on the city.

Any registration intervention can potentially target two different groups: unregistered citizens and misregistered citizens. The first group includes naturalized citizens who were not offered to register, previously registered citizens who failed to update their registration status after moving to a new address and were struck off from the list, and younger citizens. Since a law passed in November 1997, they automatically get registered when they turn 18. However, the law is still sometimes imperfectly applied, and the procedure systematically fails to register teenagers who changed address between 16 and 18[4]. In 2012, unregistered citizens accounted for 7% of all citizens living in metropolitan France[5] and they are not taken into account in French abstention rate computations. Although still high, this rate is a historical low: it had been close to 10% since the 1980's and until the massive registration wave which took place before the 2007 presidential elections.

[1] Automatic registration goes together with demanding citizens to inform their town hall when they move to a new place and having the town halls update municipal population registers accordingly. In France, this would be considered an infringement upon their freedom to move without notifying public authorities. Countries where registration is automatic and based on citizens' mandatory declaration of a change of address include Germany, Belgium, Danemark, Spain, Italy and Netherlands. In a third group of countries, registration is semi-automatic: in Canada, for instance, it is based on income tax returns, provided citizens have given their consent.

[2] In our sample too, there were differences from one city to another: driving licenses were accepted as a valid ID in some cities only. As another example, a proof of address of a relative or roommate, combined to a letter in which she recognizes that the applicant lives at her place was accepted as a valid proof of address in most cities, but not in Montpellier.

[3] It is not unusual for citizens who moved a long time ago and did not vote for a long time to show up at the polling station on a more salient election and find out that their name was dropped from the voters' list.

[4] Indeed, they are registered at the address at which they used to live when they were 16 and partook in the National Defence Preparation Day which, since 1998, replaces the French military service. If they moved away, their voter's ID gets sent back to the town hall, which does not register them.

[5] Niel and Lincot (2012)

registration rate increases with age: below 90% for those younger than 18. increases with age, and 95% for more than 50 years old. Registration rate increases with education: 85% for people with no diploma, 96% for those with higher education. higher for employed than unemployed people and increases with social category: lower for working class and employees. Lower for urban poles.

The group of misregistered citizens includes some citizens who are happily misregistered, because of some attachment to their previous city, and many others who just forgot, did not know that they had to update their registration status, or knew it but did not bother to do it. Whatever its reason, misregistration significantly lowers turnout by increasing the cost of voting, in particular at low salience elections: for a misregistered citizen, voting requires travelling or applying for proxy voting. In the absence of any systematic and representative study, it is generally estimated that misregistered citizens account for 20 to 25% of all citizens[6][7].

The quality of information that people have on the registration process varies but is low on average: many are unaware of the early deadline and assume they can register up to a few days before the election; few know the detailed list of pieces that enter in the registration file until they go to the town hall, which asks a substantial fraction of applicants to come back with a missing document; and some believe that they have to deregister from their previous place before registering in their new city, making the perceived cost of the registration process bigger than it really is. Among several reasons explaining this low quality information, the fact that one has to register only rarely certainly plays an important role. Even when people have the appropriate information, actually going through the registration process can be relatively costly: one has to find the appropriate documents, go to the town hall during opening hours and sometime wait in long queues, in particular during the last days of December.

In 2011, the INSEE [8]recorded about 714,000 registrations (excluding the automatic registrations mentioned before), most of which occured in the last weeks or days before the deadline. This number is relatively high compared to other years, due to the upcoming of the presidential elections, and close to the historic peak of 800,000 attained in 2006. But it only accounted for a small fraction of the unregistered and misregistered citizens.

The French registration process differs from that of other countries which make registration the responsibility of individuals on two important dimensions.

First, the deadline to register is closer to the elections in the other countries. In Portugal, for instance, the voters' list is updated up to 60 days before the election. In the United States, as a result of the 1970 Voting Rights Act Amendements and the Dunn vs. Blumstein Supreme Court Decision, all states have closing dates of 31 days or less. Some states even allow registration on election day.

Second, the registration process is less costly on average in the other countries. In Great Britain, registration forms are sent each year to each household in the Fall: the registration only requires the head of the household to list the household members and send the form back to the town hall. In the United States, a series of provisions decreased the registration cost and, so, the disenfranchisement of the young and the residentially mobile. In particular, "motor voter" laws, first adopted in some states in the 1970s, and generalized after the National Voter Registration Act of 1993, allow citizens to register at motor vehicle agencies, where they also apply for or renew their driver's licenses or state identification cards. As a result of these reforms, some scholars argue that, in the United States, "registration reform has reached its limits of enhancing turnout" (Highton, 2004). However, while most middle class households certainly benefit from the motor voter laws,

[6]In a 2007 survey conducted by the Cevipof, 12% of the registered citizens declared to be registered in a place different from their own town. This does not take into account misregistered citizens registered at a different address but in the same city, and might be subject to reporting bias.

[7]While the number of unregistered citizens can be precisely estimated at the national level, as the difference between the number of citizens entitled to be registered (which excludes some subgroups deprived of their civil rights) and the number of citizens actually registered, no similar systematic method exists to compute the number of misregistered citizens.

[8]Institut National de la Statistique et de l'Etude economique, www.insee.fr

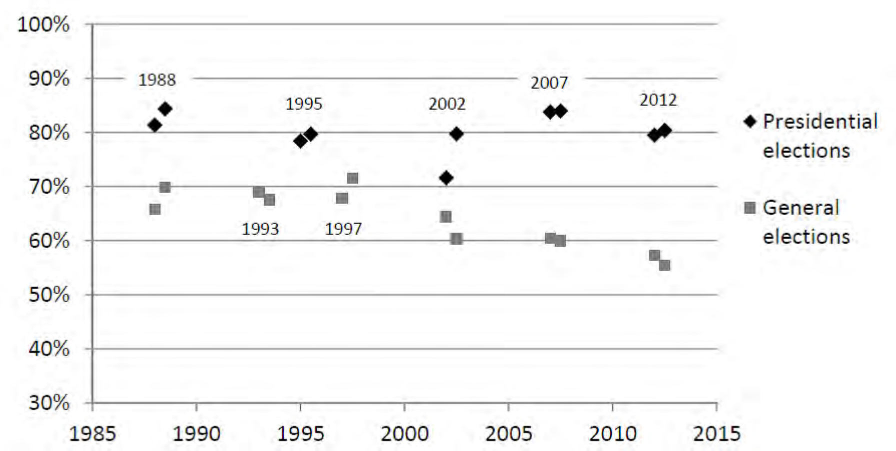


Figure 2: Turnout at French Presidential and General elections since 1988

the same is not true of the citizens, often the poorest, who do not have a car, and the very existence of a registration step, however easy to complete it is, might create a barrier for some.

In short, the results of our study, although not directly transposable to other countries, should still be informative for other contexts.

The 2012 French presidential and general elections

Every five years, French registered citizens choose the President of the Republic and the representative of their electoral district at the National Assembly. Both elections are led in two rounds unless a candidate gets more than 50% of the votes at the first round[9] and, since 2002, they take place on the same year.

The two rounds of the Presidential elections took place on April 22nd and May 6th. They were characterized by a high participation (79% at the first round and 80% at the second round). François Hollande and Nicolas Sarkozy qualified for the second round. François Hollande was finally elected President of France with 51.6% of the votes, a share lower than predicted by most polls.

The two rounds of the General elections took place on June 10th and 17th. They were characterized by a low participation (57% at the first round and 55% at the second round), relative to the 2012 Presidential elections and the previous General elections: this confirms both the presidential nature of the French regime, and the declining turnout observed at General elections and all other types of polls, except for the Presidential elections (see Graph 2 below). The Socialist Party, also the political party of the recently elected President, won 57% of the 577 seats.

2.2 The interventions

We measure the impact of interventions relying on two types of visits, conducted from July to late December 2011 : "simple" door-to-door canvassing visits and more intensive "home registration" visits.

In addresses selected to receive the first type of visits, canvassers knocked at people's doors, introduced themselves and the organization they belonged to, and encouraged French citizens who were not registered or had moved in without updating their registration status to do so. They gave them information about the registration process, the localization and opening hours of the city's town hall, reminded them of the early deadline, emphasized the importance of the 2012 elections and answered their questions. After the

[9]This happened for the General elections in one city in the sample, Pessac, where the winner got elected at the first round.

3 to 5 minutes conversation, they left a leaflet customized with the logo of their organization (an example can be found in Appendix 1) and containing general information about the registration process as well as city-specific information.

This first type of visits was not atypical: French political parties increasingly use door-to-door canvassing campaigns in electoral periods (Liegey and Pons, 2012), and, in 2011, in other cities than the ones included in our sample, a few local units organized door-to-door canvassings to foster registration. Our home registration visits were more innovative and experimental: no similar campaign had been led at such a large scale in France before.[10]It added the following key component to the visits: canvassers carried official registration forms with them, and offered their interlocutors (or other unregistered or misregistered members of their household) to get registered at home. When someone was willing to register with them, canvassers would help them fill in the registration form and offered to take pictures of the additional documents (ID and proof of address)[11]. In each city, all applications were then centralized, completed with prints of the pictures, and brought to the town hall. The French law does not forbid such registration campaigns, as long as they are not organized by a public institution. Nonetheless, to avoid any complaint, canvassers asked home applicants to sign an official authorisation letter allowing them to transfer their file to the town hall. Town halls were informed about the experiment and most of them accepted to check the applications that we transferred them before the official electoral committees took place, so that we could inform the home applicants in advance if their application would be rejected and help them update it when possible. Less than 5% of the applications were rejected.

This intervention was designed to greatly reduce the cost of registering for its beneficiaries, but it also required a great deal of trust from them. The belonging of the canvassers to well-identified groups, as well as the professional-looking leaflets they were carrying certainly helped generate this trust but some people refused to register at home even though they would eventually register at the town hall, by lack of trust, or because they considered the registration process as a personal duty[12]. People who refused to register at home were given the same information and the same leaflet as in the simple door-to-door canvassing visits. When an application was filled out, the interaction took a longer time. Oftentime, canvassers had to make appointments and come back when a document was missing or if a citizen potentially willing to register was absent at the time of the visit.

The two types of visits were carried out by a total of 230 canvassers belonging to different groups or institutions. Three broad types can be distinguished: students, NGO members, and party activists. Students took part in the experiment as part of a graduate or undergraduate political science class at the Ecole Normale Supérieure, the University Cergy-Pontoise, the IEP of Bordeaux and the University of Montpellier 1. In the three former schools, this class and the participation in the experiment were optional. In the latter, it was made mandatory for students enrolled in the last year of the political science bachelor. Three NGOs took part in the experiment: the RAJ Languedoc-Roussillon, an NGO specialized in actions towards political mobilization of young adults; an association of students studying at Science Po Paris; and an NGO of retired workers. Beyond their civic interest in the experiment, some political activists found an additional partisan interest in it, as they hoped that the new voters that would register thanks to the interventions would majoritarily vote for their candidate in the subsequent elections. In five cities, they belonged to the Socialist Party; in one city, to the Front de Gauche, another left-wing party.[13]

[10]We heard about only one similar very small-scale campaign led by an NGO based in Nantes in one neighborhood of the city. Its unusual character secured its organizors important media coverage, including in the national media.

[11]In some cases, people already had copies of these documents, or could print them right away from a printer at their home. In a few other cases, when people were reluctant to canvassers taking a picture of these documents, canvassers asked them to make a copy themselves and to come back a few days later to collect it.

[12]Several town halls reported that they received phone calls from people who had received the visit of canvassers and wanted to make sure that it was not a con. In one case (only), people alerted the police, which interrogated two canvassers.

[13]Contacts had been established with local units of other political parties as well: in Sevran, activists from the Green party were supposed to take part in the experiment before local political tensions with the local unit of the Socialist Party, also involved in the experiment, made them decide not to. In Montpellier, activists belonging to the UMP (the main right-wing party) started

All canvassers were trained before the start of the interventions. At the training, they were reminded of all the details of the registration process and received practical information about the interventions and the related monitoring sheets they would have to fill in. They were engaged in role plays, whereby two participants would play the role of canvassers, and a third one would play an unregistered citizen, to increase the confidence of those who had no previous similar experience, and teach them how to react to a few typical situations. Finally, they were asked to draw a sharp line between the two types of interventions, i.e. refrain from offering home registration in the first group[14] and systematically offer it in the second group.

The great diversity of the canvassers' profiles increases the external validity of the study. Although we could not randomly allocate the addresses to different groups of canvassers, the relative registration impact obtained in addresses covered by one or the other group remains informative.

2.3 Sampling frame

The interventions were conducted in 12 cities and 48 polling stations[15]. In addition to the 20th arrondissement of Paris, four cities are located in Ile-de-France, the region surrounding Paris (Cergy, Saint-Denis, Gonesse and Sevran). Two are located in Languedoc-Roussillon (Montpellier and Carcassonne) and five in Aquitaine (Blanquefort, Eysines, Le Taillan, Lormont and pessac). All cities are localized on a map included in Appendix 2. Their size ranges from less than 10,000 inhabitants (Le Taillan) to more than 200,000 (Montpellier and Paris). Some of them enjoy a good reputation and attract many visitors and tourists (Carcassonne, Paris) while others belong to the "French banlieues" (Sevran, Saint Denis, Gonesse). Some of them have existed for ages whereas others started developing over the last decades only (the cities in Aquitaine and Cergy, a planned city created in the 1960s). In all twelve cities, however, the mobility rate, arguably an important explanatory factor of misregistration, is surprisingly close to the national average: in 2008, in each of them, between 20 and 30% of the population had arrived in the past 5 years. Overall, in the the cities included in the experiment, 25.8% of the population had arrived in the city in the past 5 years, for a national average of 24.3%.[16].

The main criterions for including cities in our experiment were the availability of canvassers willing to partake in the experiment, and the logistical and financial support some municipalities were willing to offer. While all cities supported the experiment in some extent, financial contributions were made by the five cities located in Aquitaine and Montpellier.

In all cities, we chose to include polling stations characterized by relatively lower turnout rates at previous elections. Indeed, low participation rates are good proxies for high rates of unregistered and misregistered citizens: misregistration is in itself a factor generating low turnout and, to some extent, other factors predicting lower turnout also predict high rates of unregistered voters. The areas included in the sample are therefore not representative of the whole French population (we notably lack any data point in rural areas) but they are relatively representative of areas that would be targeted by registration campaigns or would mostly be affected by country-wide changes in the registration process.

To further define our sample, a preparatory stage took place between May and September 2011 in the selected areas: 20 surveyors went from building to building to compare the last names of citizens listed as registered at this address to the names actually found on mailboxes (or on intercoms). After the preparatory work, we

covering one polling station but stopped halfway because they got the impression that the voters they were encouraging to register would majoritarily vote for other parties afterward.

[14] Some later reported that it was sometime very difficult, in particular when people living in addresses covered by the simple door-to-door canvassing intervention asked them whether they could help them to register.

[15] By "polling station", we refer to the area associated to a specific voting booth. Polling stations include 1,000 registered citizens on average, but their size varies around this mean across cities or even within a city.

[16] Source: Institut National de la Statistique et de l'Administration, http://www.recensement-2008.insee.fr/basesFluxMobilite.action

decided to exclude from the sample those addresses in which surveyors had not been able to enter (due to the presence of multiple barriers, including doors with digital locks), as we were missing the information for these addresses and we could expect canvassers to encounter similar difficulties. We further excluded a few other buildings that were about to be brought down (and their inhabitants relocated). Among the remaining addresses, only those addresses in which at least one name found on the mailbox did not appear on the 2011 voters' list were included in the final sample as we knew that at least one citizen who was either unregistered, misregistered or a foreigner lived there. In "targetable" addresses where the matching between apartments and mailboxes was possible, only those apartments corresponding to mailboxes showing one such name at least were then targeted by the interventions. In other "non-targetable" addresses, it was impossible to link apartments to mailboxes, as either the first or the second showed neither a number nor any other type of identification.

Our final sample includes 4,118 addresses in which we were able to identify at least one potential unregistered or misregistered citizen during the preparatory work. In these addresses, 20,502 apartments were included in the sample. [17]. According to the 2011 voters' lists, the addresses in our sample hosted a total of 32,399 registered voters at the start of 2011. During the year 2011, 5,486 new voters got registered, and 3,888 were struck off the list, for a new total of 33,997 registered voters on January 1st 2012[18].

2.4 Experimental design

The addresses included in our sample were randomly allocated to a control group or one of six treatment groups. The control group is twice as big as any of the treatment groups, which have all equal size. Randomization was made at the address level as there is a fixed cost to locating an address, reaching it, and passing its main door (one sometime has to wait for someone to respond at the intercom). Moreover, randomizing at the level of the apartment would have generated a greater risk of blurring between the different treatment groups. Thus, all unregistered and misregistered citizens living at a given address were assigned to the same group and administered the same treatment.

Prior to randomization, addresses were stratified by polling station and address size, to maximize the comparability between the different treatment groups and the control group.

Figure 1 shows which visits each group received, and their timing. Early visits took place between July and November 2011 and late visits in December 2011.

Of the 48 polling stations included in the sample, 4 were excluded from the final analysis, because canvassers who covered these polling stations did not follow the experimental design.[19]

[17] In the 3,417 targetable addresses, only the apartments with at least one potential unregistered or misregistered citizen were included (16,567). In the non-targetable addresses, all the apartments were included (3,935).

[18] The corresponding numbers are 26,454; 4,403; 3,094 and 27,763 for targetable addresses

[19]

In one polling station in Gonesse, canvassers offered home registration in all treated addresses as well as in some control addresses. In two polling stations in Pessac, canvassers never offered home registration and put the leaflets in mailboxes corresponding to apartments that had not opened. Finally, in one polling station in Montpellier, students left blank registration files to each household (independently on their group) and mentioned we would come back to complete their registration. These four polling stations are not included in the sample size figures provided in the previous section.

3 Data collection

3.1 Localization of unregistered and misregistered citizens: the preparatory work

The localization of unregistered and misregistered citizens requires to compare the list of all citizens registered at a given address with the list of all citizens entitled to be registered who actually live at this address. While the first list is often available, the second usually does not exist, or is not made publicly available[20].

This transforms any attempt to localize or count the unregistered and misregistered citizens into a difficult challenge, well identified by previous scholars who studied registration, and circumvented in some experiments taking place on college campuses, where student directories provide the desired second list (Bennion and Nickerson (2009), Nickerson (2010)).

We first addressed it by completing a preliminary step of preparatory field work, prior to the first visits: last names found on mailboxes were systematically noted down, along with the corresponding apartment numbers, and compared to the last names of the citizens registered at this address. Although inherently imperfect, this strategy enabled us to identify apartments likely to host unregistered or misregistered citizens, which would then be targeted by the canvassers.[21]

The preparatory work also served to determine whether an address would be "targetable" or not (i.e. whether the information displayed on mailboxes and on the apartment doors would enable canvassers to identify pre-selected apartments), count the number of mailboxes, which we use as a proxy for the address' size, and identify the apartment number of those well-registered citizens whose name appeared on a mailbox.

3.2 Determining the initial number of unregistered and misregistered citizens: the tracking sheets filled out by canvassers

To assess which fractions of the initially unregistered and misregistered citizens registered on their own or thanks to our interventions, we first have to estimate their total number. Our strategy relies on the monitoring sheets which listed the apartments the canvassers had to cover, and were completed by the canvassers during their visits. In each apartment in which there was an interaction, the canvassers tried to identify the number of well-registered, misregistered and unregistered citizens, as well as the number of foreigners.

Among targetable addresses, we were able to collect monitoring sheets for addresses accounting for 89% of all targeted apartements. Canvassers were able to identify the types of citizens (registered, misregistered, unregistered or foreigner) for 73% of the apartments that opened their door, and the number of individuals of each type in 67% of them.

We infer the initial number of unregistered and misregistered citizens from this data after addressing two types of issues. First, the apartments which opened their door are not necessarily representative of all pre-identified apartments, which we can address by comparing the household composition in apartments which opened their door at the first visit only, the second visit only, or at both visits in addresses that were targeted for two visits. Second, the identification information recorded by the canvassers is potentially subject to systematic biases. Appendix 3 explains the strategies that we use to address these two issues.

Among all households included in the sample, 18.9% included at least one unregistered citizen, 20.1% one misregistered citizen, 45% one registered citizen and 29.9% one foreigner. Overall, 37.6% included at least

[20] In France, citizens are not required to inform any public institution when they move to a new place and although a rolling census takes place every year, only aggregate census data are publicly available, to protect individual privacy.

[21] Names found on a mailbox do not always perfectly match with the names of the people actually living in the apartment, so that we included some apartments hosting only registered citizens. Moreover, even if the match had been perfect, we would have left out some apartments hosting unregistered of misregistered citizens, but belonging to the same family as registered voters. Finally, we included many apartments which did not host any unregistered or misregistered citizen, but only foreigners.

one unregistered or misregistered citizen. As discussed in Appendix 3, this estimate is likely to be a lower bound on the true fraction.

On average, targeted households hosted 0.45 registered citizens, 0.48 misregistered citizens, 0.25 unregistered citizens and 0.47 foreigners, for a total of 1.65 adult members. Overall, we estimate that the 4,118 addresses included in our sample hosted initially 38,375 citizens, among which 55% were "well-registered", 30% misregistered and 15% unregistered. Approximately 31% of the initially misregistered and unregistered citizens (4,403) got registered in the pre-presidential year of 2011.

Finally, for each apartment that was visited, we know from the monitoring sheets whether the door remained closed or whether there was an interaction. Combining this information with the identification information, we estimate the fraction of households with at least one unregistered or misregistered citizen that were actually treated. We later use this estimate to infer the treatment on the treated impact of our interventions on the number of registrations in each group from the intention to treat estimates.

47.5% of the apartments targeted by the canvassers in the first period and 44.9% of the apartments targeted in the second period opened their door[22]: on average, 46.2% of the apartments visited only once opened their door[23]. In addresses that were randomly selected to receive two visits, 65.1% of the apartments opened their door at least once. Among households that opened their door the first time, 59% opened it the second time. Conversely, among households that opened their door the second time, 62% had opened it the first time.

Households hosting unregistered or misregistered citizens that were covered in the first period were slightly more likely to open their door to canvassers: we estimate that, among addresses covered in the first period, 49.5% of the apartments with at least one unregistered or misregistered citizen initially were treated. On the contrary, in the second period, households with unregistered or misregistered citizens were less likely to enter in a discussion: in this period, only 41.3% of the apartments with at least one unregistered or misregistered citizen initially were treated. In addresses that received 57.7% of the apartments hosting at least one unregistered or misregistered citizen initially were treated

Pooling all treated addresses together, 50.7% of all apartments and 51.9% of the apartments with at least one unregistered or misregistered citizens opened their door at least once.

3.3 Administrative data: voters' lists and turnout

Voters' lists

The 2011 and 2012 voters' lists are the most important data used in this study. For each city, these two files draw the list of all citizens registered respectively on January 1st of 2011 and January 1st of 2012. We compare the two lists to identify newly registered citizens (present on the second list only), citizens who were struck off the list in year 2011 (present on the first list only) and citizens listed in both years.

We were able to obtain lists containing the following information for any citizen registered in 2011 and/or 2012: first name, last name, polling station[24], exact administrative address (street name and number), gender, date and place of birth (country, city, and "department", if born in France)[25].

In some of the empirical analysis presented below, we use the apartment as our unit of analysis. As a preliminary step, we allocated the citizens registered in 2011 and/or 2012 to the apartments identified during

[22]This small difference probably results from the fact that the last second period visits took place during the Christmas vacations.

[23]Most frequently, the reason why the door does not open is that no household member is at home at the time of the visit, or that the person at home distrusts strangers.

[24]The limits of polling stations change regularly, following important changes in the relative size of the population living in the geographical areas they cover. A few such changes occured in our sample, which we had to take into account when merging the lists together.

[25]We merge the two lists based on names, addresses and date of birth, and systematically check for spelling changes between 2011 and 2012, to make sure that we are not identifying artificial registrations and striking off the list.

the preparatory work by comparing their last names and spouse names (when listed) to the last names found on the mailboxes [26]. To leave as few well-registered citizens unallocated as possible, we also used the complementary address available for some citizens, when it included their apartment number[27].

We were unable to identify the apartment number of 17% of the newly registered citizens, although we can take for granted that almost all of them actually live at the address listed on the voters' file. For that reason, regressions done at the individual level include all individuals living in the addresses included in the sample (and not only those mapped to the apartments targeted by the interventions).

Additional information on the newly registered voters

In all 12 cities, we were able to obtain very useful additional information on the newly registered citizens that is recorded when they register, but usually not publicly available: the initial registration status of the applicant (never registered before, registered at another address within the same city, or in another city), a piece of information central to our analysis, as it enables us to differentiate between former unregistered and misregistered citizens; the former place of registration, for citizens previously registered in another city; the registration date[28].

Turnout at the Presidential and General elections

Attendance sheets signed by voters who cast a ballot on Election Day are the second most important set of data used in this study. They enable us to measure the actual voting behavior of all registered citizens in our sample without any mistake, differently from survey reports which are often unreliable when it comes to voters' turnout, constantly overreported (Ansolabehere and Hersh, 2011).

In France, attendance sheets are available for consultation by any registered French citizen until 10 days after each poll. They show whether each registered citizen voted or not and, in the former case, whether they voted in person, by proxy, or in a consulate abroad. We were allowed to take pictures of all these sheets and entered the data for the 33,997 citizens registered in 2012 and included in our sample. For each of them, we have 4 data points, except for Pessac where the general election consisted in only one round, and for a few missing sheets. Overall, we use a total of 135,583 turnout data observations.

3.4 Surveys

Survey of canvassers

All canvassers who took part in the experiment were encouraged to answer a short online survey which included questions about their gender, age, activity, education, date of arrival in the city, registration status, interest in politics and political knowledge. They were further asked to provide feedback about the experiment, in particular: how difficult it was to interact with their interlocutors; the way their visit was perceived by their interlocutors; the relative interest in politics of people targeted by the interventions, those who registered after simple door-to-door canvassing, and those who registered at home; the most frequent motives

[26] We assume that all citizens bearing the same name and living at the same address live in the same apartment. Matching names together was relatively tedious, due to the high proportion of identical names written with two different spellings.

[27] The richness of the complementary information varies from one city to the other, but it is always recorded in a non-systematic way and available for a fraction of the registered voters only. We systematically clean this information "by hand" for all observations and use it for a second purpose as well: some administrative addresses encompass several actual buildings. When the building name is available for a large enough fraction of registered citizens, to maximize our statistical power, we reconstruct the complete address (street name and number AND building name) and randomize at this level.

[28] In most cities, only the date when the newly registered citizen was entered in the system is available. It measures the actual date of registration with a delay that varies, since applications are often handled in stacks.

mentioned by unregistered and misregistered citizens who refused to register at home; the best method to foster new registrations.

We were able to get the answers from 75% of the 230 canvassers.

Survey of newly registered citizens at the time of registration

Canvassers offered citizens who registered at their place to answer a short questionnaire about their socio-economic status, their interest in politics and the reasons why they chose to register. The same questionnaire was also available at the town halls, but the extent to which citizens who registered were offered to answer the questionnaire varied considerably from one city to another and the fraction of newly registered citizens who actually answered the questionnaire was low on average. In Montpellier, however, the city accounting for 37% of the sample, surveyors were posted at the town hall during opening hours, from November 15th to December 31st, so that the answer rate is much higher.

Post electoral suvey

From June 18th (the day following the General elections) to July 15th, 50 surveyors administered a post-electoral survey to 1,500 respondants. All surveyors were students in political science, economics, social sciences or law coming from different universities. The survey was administered at the respondants' place. To facilitate the coordination of the surveyors, it took place in only four cities, Saint-Denis, Cergy, Sevran and Montpellier, which account for 84% of the entire sample.

The questionnaire includes questions about socio-demographic characteristics, registration status, political opinions, participation in the last elections, interest in the campaign, and personal views on the future. Administering the questionnaire required 15 to 20 minutes on average. Only 2% of respondants who had started answering the questionnaire refused to go to the end.

Since some questions were related to the General elections, the survey had to be conducted afterward. Its timing was further constrained by the Summer vacations: we wanted the survey to be completed as early as possible, before the massive vacation departures.

The sample included apartments targeted by the interventions in the targetable addresses (both in the control and the treatment groups) and all apartments in non-targetable addresses. Surveyors were asked to survey no more than one person in each apartment and only French citizens who were not registered on the 2011 voters' list, and thus had been eligible to register before December 31st. We further excluded citizens who had just reached 18 and had been automatically registered[29].

We expected an important fraction of the initially unregistered and misregistered citizens to have a limited interested in politics, if they did not reject institutional politics overall. We were particularly eager to get answers from such citizens also, and thus asked canvassers to introduce the survey in broad terms, as a survey related to people's day-to-day life and opinions, without mentioning the word "politics". For the same reason, questions regarding the respondants' sociodemographic status were asked first, and questions related to registration and politics only later. Finally, respondants were instructed not to mention anything related

[29] Canvassers were given a list of people they should NOT survey and were asked to identify respondants corresponding to the above-defined criteria as follows: after introducing themselves and explaining the purpose of their visit, they asked the person who had openened the door whether she was a French citizen. If yes, they asked her whether she accepted to respond, wrote down her first and last name and rapidly checked that she was not listed on their list. If not, they went on administering the questionnaire. If their interlocutor was not French, not willing to answer, or if her name appeared on the list, they asked whether they could survey another member of the household.

to registration, to minimize the risk of getting answers from different types of citizens in the treatment groups and in the control group[30].

To further control for differential response rate in the control and treatment groups and increase the representativity of our respondants' answers, we further randomly allocated all addresses of any polling station to two surveyors, and asked them to cover twice half of the addresses, also randomly selected: in these addresses, surveyors knocked again at all doors that had remained closed the first time.

In addition to the questionnaires, the surveyors filled out monitoring sheets that tell us, for each apartment, whether the door opened; if so, whether a questionnaire was administered; and if not, for what reason.

3.5 Additional data

Housing price data

For a subset of the addresses of the cities included in Ile-de-France, we obtained housing price data at the address level from www.meilleursagents.com, a real estate company, which we use as a proxy for voters' social status.

Polling station-wise electoral results

For all 12 cities, we obtained polling station-wise electoral results for both the Presidential and General elections: turnout and shares of the votes obtained by each candidate.

City-wise data from the French Institute for Statistics (INSEE)

City-wise data give us the localization and characteristics of the registered citizens' city of birth and the newly registered citizens' previous place of registration, if they had been registered before. We use the following characteristics: size of the population, urban / rural, unemployment rate and median revenus.

3.6 Verifying randomization

Table 1 presents summary statistics for addresses in the sample. We also identify significant differences between the different treatment groups and the control group for a variety of characteristics. In the regressions reported in odd-number columns, we measure differences in the baseline characteristics between the control group and the treatment groups taken altogether. In the regressions reported in the even-number columns, we measure differences in the baseline characteristics between the control group and each treatment group. We run joint T tests of the joint significance of the six treatment dummies.

Panel A takes the building as the unit of observation. On average, 83% of the buildings were "targetable" and each building counted 8 apartments. In the average building, 7 last names found on the mailboxes and written down during the preparatory work did not match with any name in the voters' list. The differences between the control group and the treatment groups, taken altogether or individually, are significant neither for these variables, nor for the housing price.

Panel B takes the individual as the unit of observation and includes all previously registered citizens. These people were not targeted by our interventions but, given the likely correlation between their sociodemographic

[30] If our treatments increased the interest in politics of the additional newly registered citizens, we feared that we would get more answers from respondants in the treatment groups. Similarly, we feared that unregistered or misregistered citizens would exclude themselves from the survey if we mentioned the word "registration".

characteristics with those of the unregistered and misregistered citizens, the extent to which they are homogeneously distributed between the control and treatment groups can be used as a proxy for the success of randomization regarding the distribution of unregistered and misregistered citizens in the different groups as well. 46% of the previously registered citizens of the sample were men and the average previously registered citizen was 45 years old. 20% of the previously registered citizens were born in their city of residence, 17% in another city in the "departement", 13% in another "departement" in the region, 27% in another region, and 23% abroad. 96% were born in a city (instead of a village), in a populated city on average (slightly less than 300,000 inhabitants).

Previously registered citizens living in the treated addresses considered altogether did not differ significantly from the control group. Once again, considering the treatment groups separately, we cannot reject the hypothesis that their joint significance is null, although a few treatment dummies (2 over 30) are significant at the 5% level, as should be expected.

4 Model

4.1 2 stages: registration, and vote

Each citizen that is initially unregistered goes through two steps: registration and vote.[31] For simplicity, we assume that there is only one electoral round.

In the *second stage*, individual i can cast a vote only if she registered in the first stage. She actually votes if $u_i + \varepsilon_i \geq 0$, where u_i represents the net benefits of voting[32] and ε_i is a utility shock realized after registering and characterized by the density function and the distribution functions f_ε and F_ε, identical across individuals. We set $E[\varepsilon_i] = 0$ and write $P(V_i = 1)$ i's propensity to vote.

Conditional on registering in the first stage, $P(V_i = 1) = P(u_i + \varepsilon_i \geq 0) = 1 - F_\varepsilon(-u_i)$.

In the *first stage*, when i decides whether or not to register, she expects to get utility $g(u_i) \equiv \int_{-u_i}^{\infty} (u_i + \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon$ in the second stage.[33]

i registers even without receiving the treatment if $\widetilde{c_i} \leq \beta_i g(u_i) \Leftrightarrow c_i \equiv \frac{\widetilde{c_i}}{\beta_i} \leq g(u_i)$, where $\widetilde{c_i}$ represents i's net present registration cost[34] and β_i her intertemporal actualization rate.

The treatment, a door-to-door canvassing or home registration visit, decreases the registration cost to λc_i for some $\lambda \in [0, 1[$, assumed to be identical for all i's for simplicity. The closer λ is from 0, the more efficient the intervention.

Henceforth, we call u_i and c_i i's "degree of politicization" and "registration cost". The distribution of types over the entire population of unregistered citizens is described by the the continuous bivariate random vector (U, C), with joint density function $f(u, c)$ and marginal density functions $f_U(u)$ and $f_C(c)$.[35]

The always-takers, who register even if they do not benefit from an intervention, are $\{i \mid i \text{ is always-taker}\} = \{i \mid c_i \leq g(u_i)\}$. Their share is $P(i \text{ is always-taker}) = \int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} f(u, c)\, dc\, du$. The compliers, who reg-

[31] We focus here on unregistered citizens. The case of misregistered citizens is discussed in section 4.6.

[32] u_i is positive only if the benefits of voting outweigh its cost

[33] g is strictly increasing in u. It is therefore a bijection, and g^{-1} is defined.

[34] $\widetilde{c_i}$ is the difference between the gross cost of registering (which includes gathering information about the registration process and actually going through the process) and the benefits derived from registering (e.g. the benefits derived from complying with a norm). For simplicity, we assume that it is always positive. Otherwise, the registration condition should be written as $c_i \leq max\{g(u_i), 0\}$.

[35] As usual, for any u such that $f_U(u) > 0$, we write $f(c \mid u) \equiv \frac{f(u,c)}{f_U(u)}$ the conditional density of C given that $U = u$. And for any c such that $f_C(c) > 0$, we write $f(u \mid c) \equiv \frac{f(u,c)}{f_C(c)}$ the conditional density of U given that $C = c$.

ister only if they receive the treatment, are $\{i \mid i \text{ is complier}\} = \{i \mid g(u_i) < c_i \leq {}^{g(u_i)}/\lambda\}$. Their share is $P(i \text{ is complier}) = \int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)\, dc\, du$.

4.2 Differences between compliers and always-takers

Compliers have a lower degree of politicization u on average than always-takers

Claim 1: $E[u_i \mid i \text{ is complier}, c_i = \bar{c}] \leq E[u_i \mid i \text{ is always-taker}, c_i = \bar{c}]$ for any $\bar{c}$: compliers characterized by a given registration cost $\bar{c}$ have a lower expected degree of politicization u than always-takers facing the same $\bar{c}$.[36]

If "large enough" fractions of always-takers have a low c and "large enough" fractions of compliers a high c, this result is nonetheless compatible with $E[u_i \mid i \text{ is complier}] > E[u_i \mid i \text{ is always-taker}]$.[37]

However, joint density functions characterized by these patterns are unrealistic. *Conditions ID and R1* are sufficient (but not necessary) conditions under which $E[u_i \mid i \text{ is complier}] \leq E[u_i \mid i \text{ is always-taker}]$.

Condition ID: $-f(u,c)$ satisifies log-increasing differences in u and c : $\frac{f(u',c')}{f(u',c)} < \frac{f(u,c')}{f(u,c)}$ for any $u' > u$ and $c' > c$.

The condition can be interpreted as follows: there are relatively less citizens with a higher c among citizens with a higher u. It directly implies, for instance, that people with a higher u have a lower c, on average. This assumption corresponds to the expectation that the degree of politicization is an increasing function and the individual registration cost a decreasing function of some variables, such as the level of education.

Claim 2: *Condition ID* is satisfied for instance by any bivariate normal density (the type bivariate density most commonly used) with negative correlation.

Condition R1 (regularity condition): For any u, and any $u" \geq u'$ with $u' \epsilon [g(u), {}^{g(u)}/\lambda]$, $\frac{u" f(u"|u)}{F(u"|u)} \leq \frac{u' f(u'|u)}{F(u'|u)}$.

Claim 3: If *Conditions ID* and *R1* hold, then $E[u_i \mid i \text{ is complier}] \leq E[u_i \mid i \text{ is always-taker}]$

Compliers have a higher registration cost c on average than always-takers

A symmetric reasoning goes for the comparison of c between the always-takers and compliers.

Claim 4: $E[c_i \mid i \text{ is complier}, u_i = \bar{u}] \geq E[c_i \mid i \text{ is always-taker}, u_i = \bar{u}]$ for any $\bar{u}$: compliers characterized by a given degree of politicization $\bar{u}$ have a higher expected registration cost c than always-takers facing the same $\bar{u}$.

Claim 5: If *Conditions ID* and *R1* hold, then $E[c_i \mid i \text{ is complier}] \geq E[c_i \mid i \text{ is always-taker}]$.

Compliers are less likely to vote on average than always-takers

We write the expected participation of an individual with degree of politicization u_i as $v(u_i) \equiv P(u_i + \varepsilon_i \geq 0) = 1 - F_\varepsilon(-u_i)$.

Claim 6: If *Conditions ID* and *R1* hold, then $E[v(u_i) \mid i \text{ is complier}] \leq E[v(u_i) \mid i \text{ is always-taker}]$

[36] The proof of this and the other claims are included in Appendix 4.

[37] This can easily be seen when one writes the two latter objects as the weighted averages $E[u_i \mid i \text{ is complier}] = \int_{-\infty}^{\infty} E[u_i \mid i \text{ is complier}, c_i = c] \left(\frac{\int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c) du}{\int_{-\infty}^{\infty} \int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c) du\, dc} \right) dc$ and $E[u_i \mid i \text{ is always-taker}] = \int_{-\infty}^{\infty} E[u_i \mid i \text{ is always-taker}, c_i = c] \left(\frac{\int_{g^{-1}(c)}^{\infty} f(u,c) du}{\int_{-\infty}^{\infty} \int_{g^{-1}(c)}^{\infty} f(u,c) du\, dc} \right) dc$.

Compliers who vote have a lower degree of politicization u on average than always-takers who vote

Claim 7: If *Conditions ID* and *R1* hold, then $E\left[u_i \mid i \text{ is complier}, i \text{ votes}\right] \leq E\left[u_i \mid i \text{ is always-taker}, i \text{ votes}\right]$

Interpretation: The intervention does not only select unregistered citizens who are less likely to vote to the subsequent elections: among the compliers, those who vote have a lower degree of politicization. On average, they experienced utility shocks that were higher than those experienced by the always-takers: they were more susceptible to vote as a consequence of a recent societal or personal event for instance. The efficiency of the intervention is mitigated twice: it includes citizens who are relatively less likely to participate afterward, and, when they participate, express short-term preferences rather than long-term interest in politics.

4.3 Comparative statics

How do compliers differ when selected by a more or less efficient intervention?

Claim 8: If *Conditions ID* and *R1* hold, a more efficient intervention, characterized by $\lambda' < \lambda < 1$, selects additional compliers characterized by a lower degree of politicization, a higher registration cost, a lower turnout and a lower degree of politicization conditional on voting than those selected by the less efficient intervention.

Relationship between the compliers' (observed) propensity to vote and their (unobserved) degree of politicization, degree of politicization conditional on voting and registration cost

In Sections 6 and 7 below, we provide estimates of the share of compliers and their propensity to vote. We observe neither their degree of politicization nor their registration cost or degree of politicization conditional on voting.

Claim 9 draws the theoretical link between these observed and unobserved quantities so that we can draw empirical inferences from the first ones to the second ones.

Condition R2 (regularity condition): $z(u) \equiv E_f\left[c_i \mid i \text{ is complier}, u_i = u\right] = \frac{\int_{g(u)}^{g(u)/\lambda} c f(u,c) dc}{\int_{g(u)}^{g(u)/\lambda} f(u,c) dc}$ increases in u.[38]

Claim 9: For a given share of compliers and unchanged conditional densities $f(c \mid u)$, if *Conditions ID, R1* and *R2* hold, an increase in the compliers' propensity to vote, generated by an increase in the relative number of compliers with a higher u, is concomitant to an increase in their degree of politicization, registration cost and degree of politicization conditional on voting.

Interpretation: The higher the (observed) compliers' turnout, the higher we should expect their (unobserved) degree of politicization, degree of politicization conditional on voting and registration cost to be. *Claim 9* identifies the conditions under which the following intuition can be stated rigorously: in a world in which the compliers' degree of politicization is relatively higher (and, so, less binding), their registration cost has to be relatively higher (and, so, more binding) for their share to be left unchanged.

This is of great importance for several reasons: first, if there are reasons to think that the registration cost is relatively more binding, then one might be relatively more unsatisfied with the registration process, as it excludes citizens who did not self-select out of registration due to too little interest in the elections.

[38] Absent the condition, we have that, for any u, $z(u') \geq z(u)$ for any u' and u such that $g(u') \geq g(u)/\lambda$. Indeed, the conditional expectation is then taken on a separated support, "higher" for u' than u. For u' and u such that $g(u') < g(u)/\lambda$, $z(u') - z(u)$ is driven by two opposite effects. The support is still "higher" for u' which tends to make the difference positive. However, for any $\left(\widetilde{c}, \widetilde{\widetilde{c}}\right) \in [g(u'), g(u)/\lambda[$ with $\widetilde{\widetilde{c}} > \widetilde{c}$ the relative weight of compliers facing the higher cost $\widetilde{\widetilde{c}}$ is higher for those who have the lower degree of politicization u, which tends to make the difference negative: $\frac{f\left(u', \widetilde{\widetilde{c}}\right)}{f(u', \widetilde{c})} \leq \frac{f\left(u, \widetilde{\widetilde{c}}\right)}{f(u, \widetilde{c})}$ since $-f(u,c)$ satisfies log-increasing differences. The closer u' to u, the bigger the importance of this second effect relative to the first one. *Condition R2* ensures that the second effect never outweighs the first one so that $z(u') \geq z(u)$ even for u' close to u.

Second, this would make actions and interventions designed to decrease the registration cost or the cost of voting overall relatively more desirable compared to interventions designed to increase people's interest in politics, for governments, parties or NGOs looking forward to increasing turnout. Third, to the extent that the registration cost is higher for citizens with specific socio-demographic characteristics and political preferences, this would finally make us worry that the registration process decreases the representativity of the electoral outcomes.

4.4 Beyond the degree of politicization, two other determinants of the propensity to vote

In the model presented thus far, each citizen's propensity to vote is entirely determined by her degree of politicization, u_i. Overall, the lower the degree of politicization of the compliers, compared to the always-takers, the lower their propensity to vote.

However, other factors might determine newly registered citizens' propensity to vote. Assessing their importance is interesting in itself, and necessary to draw accurate inferences from compliers' turnout to their relative degree of politicization and registration cost.

The relative mobilization effect of the electoral campaign on the compliers

In addition to the degree of politicization, the propensity to vote can be affected by the mobilization that goes together with the political campaign, after the registration stage and independently from the interventions .

To account for this possibility, we maintain the assumption that, when she registers, i anticipates that she will vote if $u_i + \varepsilon_i \geq 0$, with a likelihood of $v(u_i) = P(u_i + \varepsilon_i \geq 0)$: the mobilization effect of the campaign is not anticipated (at least not entirely) at the time of the registration. However, once registered, and thanks to the media coverage of the campaign and partisan and nonpartisan mobilization efforts, her actual propensity to vote becomes $P(V_i = 1) = w(u_i)$.

We investigate the case in which the mobilization effect of the campaign is higher for citizens with a lower degree of politicization, although they keep a lower propensity to vote: $w(u') - v(u') \leq w(u) - v(u)$ and $w(u') \geq w(u)$ for any $u' \geq u$. .[39]The previous subsections can be seen as a special subcase of this extended version of the model, in which $w(u) = v(u)$ for all u .

Claim 10: All previous results hold in the extended version of the model, where a registered citizen's actual propensity to vote is $w(u_i)$, with $w(u') - v(u') \leq w(u) - v(u)$ and $w(u') \geq w(u)$ for any $u' \geq u$.

Claim 11: The difference between compliers and always-takers' turnout is lower if the propensity to vote of a registered citizen with utility u is given by $w(u)$ rather than $v(u)$.

Claim 12: The difference between the propensity to vote of compliers and always-takers can be written as the sum of two terms. The first one, negative, and predominant, is bigger, the bigger the difference between the degree of politicization of the compliers and always-takers. The second one, positive, comes from the fact that the mobilization effect of the campaign is lower for citizens with a higher degree of politicization.

Neglecting the mobilization effect of the campaign (and, so, the second term) would lead us to underestimate the difference in the degree of politicization between the compliers and the always-takers.

[39]In the subsection "Microfounding the assumption that $w(u) - v(u)$ decreases with u" of Appendix 2, we discuss how this assumption can be grounded in a more fundamental assumption about the way the campaign affects the perceived benefits of voting u_i.

The treatment impact of home registration

Home registration differs from door-to-door canvassing along two important dimensions. First, it further reduces the cost of registering. This selects additional compliers characterized by a lower degree of politicization, a higher registration cost, a lower turnout and a lower degree of politicization conditional on voting than those selected by door-to-door canvassing, as predicted by *Claim 8*. The additional compliers selected by home registration are also relatively more affected by the mobilization effect of the campaign and the elections, provided this mobilization effect is higher for citizens with a lower degree of politicization.

Second, registering people at their place instead of having them register at the town hall might affect their subsequent propensity to vote: beyond the *fact* of getting registered, the *way* in which one gets registered might itself matter. By decreasing the registration cost, home registration might decrease the propensity to vote of those who choose this way of registering, if it decreases their involvement. Alternatively, it can increase it if it makes them feel indebted towards the canvassers. We write $\widetilde{w}(u_i)$ the propensity to vote of i when i gets registered at home. For citizens who accept to get registered at home, the propensity to vote changes by $E\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right]$ on average: we call this quantity the **"treatment impact of home registration"** on the propensity to vote.

Importantly, citizens who accept to register at home are not a subset of the compliers: some always-takers who would have registered at the town hall shortly before the deadline otherwise, accept to register at home provided they receive the canvassers' visit early enough. Conversely, some compliers register at the town hall after the visit of the canvassers, thanks to the information and encouragement that they received, even though they refuse to register at home with them.

4.5 The case of initially misregistered citizens

Let us now consider one last extension to the model presented in section 4.1: while we have so far restricted the analysis to the case of unregistered citizens, our interventions also targeted misregistered citizens. How do the results obtained for unregistered citizens extend to the misregistered citizens?

k, the additional cost of voting for misregistered citizens

Each misregistered citizen's type can be characterized by her registration cost and degree of politicization c_i and u_i, as well as a third parameter, k_i. k_i is the additional cost of voting that i faces if she votes in her previous city or polling station rather than at the polling station closest to her new address: it reflects the time and/or financial cost required to reach this previous place.[40]

The distribution of types over the entire population of misregistered citizens is described by the the continuous multivariate random vector (U, C, K), with density function $f(u, c, k)$.

Similarly to unregistered citizens, misregistered citizens expect to get utility $g(u_i) \equiv \int_{-u_i}^{\infty} (u_i + \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon$ if they update their registration status. However, if they fail to do so, they can expect to get utility $g(u_i - k_i) = \int_{-u_i+k}^{\infty} (u_i - k_i + \varepsilon) f_\varepsilon(\varepsilon) d\varepsilon$ rather than 0.

The always-takers, who register even if they do not benefit from an intervention, are $\{i \mid c_i \leq g(u_i) - g(u_i - k_i)\}$.
The compliers, who register only if they are in the treatment group, are $\left\{i \mid g(u_i) - g(u_i - k_i) < c_i \leq \frac{g(u_i) - g(u_i - k_i)}{\lambda}\right\}$.

Misregistered citizens facing the additional cost $\bar{k}$ of voting at their previous address

Holding $\bar{k}$ constant, we call $f_{\bar{k}}(u, c)$ the distribution of types of misregistered citizens who face the additional cost $\bar{k}$ of voting at their previous address and define $g_{\bar{k}}(u) \equiv g(u) - g\left(u - \bar{k}\right)$.

[40] k_i can be < 0, for instance if i deliberately remains registered at her old address as a commitment device to maintain ties with family members still living there. Evidently, misregistered citizens who have a negative k_i will not register even if the registration cost is null. Interested in the compliers and always-takers, we thus focus on misregistered citizens for whom $k_i > 0$.

We define three conditions, similarly as for unregistered citizens:

Condition $ID_{\bar{k}}$: $-f_{\bar{k}}(u,c)$ satisifies log-increasing differences in u and c.

Condition $R1_{\bar{k}}$: For any u, and any $u" \geq u'$ with $u' \epsilon \left[g_{\bar{k}}(u), g_{\bar{k}}(u)/\lambda\right]$, $\frac{u" f_{\bar{k}}(u"|u)}{F_{\bar{k}}(u"|u)} \leq \frac{u' f_{\bar{k}}(u'|u)}{F_{\bar{k}}(u'|u)}$.

Condition $R2_{\bar{k}}$: $z_{\bar{k}}(u) \equiv E_{f_{\bar{k}}}\left[c_i \mid i \text{ is complier}, u_i = u\right]$ increases in u.

Claim 13: If *Conditions* $ID_{\bar{k}}$, $R1_{\bar{k}}$ and $R2_{\bar{k}}$ hold for any $\bar{k}$, all results established for unregistered citizens hold for misregistered citizens facing an additional cost $\bar{k}$ of voting at their previous address, for any $\bar{k}$.

All misregistered citizens pooled together

Absent any further assumption on the general shape of $f(u,c,k)$, the results obtained for misregistered citizens facing a given cost $\bar{k}$ of voting at their previous address do not necessarily hold when all misregistered citizens pooled together.

For instance, if the share of compliers is larger for lower values of $\bar{k}$ and if misregistered citizens facing a lower $\bar{k}$ have a higher u, on average, it is possible that, among misregistered citizens, compliers have a higher u on average than always-takers, even though the reverse is true when the sample is restricted to any given $\bar{k}$.

Even if this is unlikely, at the very least, comparing the propensity to vote between previously misregistered compliers and always-takers will be less informative when they are all pooled together: although interesting politically, it will not allow us to draw backward inferences about the relative cost of registration faced by the compliers, which we can do when separating previously misregistered citizens by their $\bar{k}$.

We will thus conduct both types of comparisons: in some regressions, previously unregistered and misregistered citizens will all be pooled together. In others, we will allow for heterogeneity in the results obtained for unregistered and misregistered citizens, as well as for misregistered citizens facing different additional costs of voting at their previous address. Administrative data tell us who had been registered before and provide us with a strong predictor of k_i for the misregistered citizens : where they were previously registered, from which we infer the distance between this place and their actual address.

5 From the model to the data - Outline of the empirical analysis

Our empirical analysis, guided by the theoretical results obtained in Section 4, seeks to answer two broad questions. First, what was the effectiveness of the interventions? Second, who are the compliers selected by the interventions and for what reason(s) do their counterparts in the control group fail to register?

The impact of the interventions on the number of new registrations and votes cast by newly registered citizens: Section 6

We assess the effectiveness of the interventions on two outcomes: the number of new registrations and votes cast by newly registerered citizens.

The impact of the interventions on the number of new registrations is best described by the ratio $\frac{p_C}{p_A} \equiv \frac{P(i \text{ is complier})}{P(i \text{ is always-taker})} = \frac{\#\{\text{compliers}\}}{\#\{\text{always-takers}\}} = \frac{\#\{\text{registrations, treatment gr.}\} - \#\{\text{registrations, control gr.}\}}{\#\{\text{registrations, control gr.}\}}$, which compares the number of newly registered citizens in the control and treatment addresses or apartments.

We differentiate the overall impact of the six interventions from the individual impact of each of them. Pooling interventions together based on the timing and content of the visits, we measure the specific contribution of visiting the addresses Late (i.e. close to the deadline) vs. Early, or offering Home registration vs. information and encouragement through a simple Canvassing visit.

Similarly to the impact of the interventions on the number of new registrations, their impact on the number of votes cast by newly registered citizens is given by the ratio $\frac{\#\{\text{votes, treatment gr.}\}-\#\{\text{votes, control gr.}\}}{\#\{\text{votes, control gr.}\}}$. This outcome, a reduced form policy number of interest, is determined both by the impact of the interventions on the number of new registrations and the relative propensity to vote of the additional citizens registered thanks to the interventions, the compliers: to fully understand the relative cost-effectiveness of the different interventions, our first question, we must turn to our second question.

Selection and treatment impacts of the interventions on the propensity to vote: Section 7

We call the **"selection impact of door-to-door canvassing"** on the propensity to vote the difference between the propensity to vote of compliers selected by door-to-door canvassing and always-takers: using the notation of the model, it is $P(V_i = 1 \mid i \text{ is canvassing complier}) - P(V_i = 1 \mid i \text{ is always-taker})$ or $E\left[w\left(u_i\right) \mid i \text{ is canvassing complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]$.

We call the **"selection impact of home registration"** on the propensity to vote the difference between the propensity to vote of compliers selected by home registration and always-takers that we would observe if all had been registered in the same way, at the town hall: $E\left[w\left(u_i\right) \mid i \text{ is home registration complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]$.

As defined in Section 4.4, the **"treatment impact of home registration"** on the propensity to vote is the difference between the propensity to vote of those who register at home and their propensity to vote if they had registered at the town hall:$E\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right]$.

These three objects have to be inferred from preliminary estimates of the difference between the propensity to vote of newly registered citizens in the different treatment and control groups: $P(V_i = 1 \mid \text{door-to-door}) - P(V_i = 1 \mid \text{control})$ and $P(V_i = 1 \mid \text{home registration}) - P(V_i = 1 \mid \text{control})$.

Claim 14[41]:

$P(V_i = 1 \mid \text{door-to-door}) - P(V_i = 1 \mid \text{control})$

$= \left[E\left[w\left(u_i\right) \mid i \text{ is canvassing complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A}$

, where p_C is the proportion of compliers selected by door-to-door canvassing.

Claim 15:

$P(V_i = 1 \mid \text{home registration}) - P(V_i = 1 \mid \text{control})$

$= \left[E\left[w\left(u_i\right) \mid i \text{ is home registration complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A}$

$+p \times E\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right]$

,where p_C is the proportion of compliers selected by home registration and $p = P(i \text{ gets registered at home} \mid i \text{ registers, home registration})$.

Our treatment design enables us to estimate the treatment impact of home registration and, so, infer its selection impact. Comparing it to the selection impact of door-to-door canvassing visits, we describe the way the marginal registrant changes as the registration process is facilitated on a first dimension, the propensity to vote.

Decomposing the selection impacts on turnout: The relative long-term interest in politics and short-term mobilization by the campaign of the compliers vs. always-takers: Section 8

Estimating the selection impacts of the interventions on the propensity to vote is only a first and incomplete step towards assessing the role played by self-selection or, more precisely, lack of interest in politics and the elections, in compliers' failure to register, absent any intervention. Indeed, as we argued in Section 4.4, these

[41]The proofs of *Claims 14* and *15* are included in Appendix 4.

selection impacts sum two terms, which reflect respectively the difference between the degree of politicization of the compliers and always-takers and the fact that the mobilization effect of the campaign might differ for the first and the second.

Fortunately, the lower saliency of the General elections help us disentangle the two: arguably, the campaign for the General elections had a much smaller mobilization effect, so that differences in propensity to vote measured at these elections should better reflect differences in the long-term interest in politics and the elections. Data from the post-electoral survey further help us to disentangle the two terms which enter in the interventions' selection impact on turnout.

Based on our assessment of the relative long-term interest in politics of the compliers, compared to the always-takers, we can finally evaluate the role of the different factors which explain compliers' failure to register, absent the interventions. The discussion is based on *Claim 9* of the model, which draws a connection between the relative degree of politicization of the compliers and the relative cost the registration process means to them.

Selection impacts of the interventions on other dimensions than the propensity to vote: Section 9

Understanding the way the compliers differ from the always-takers in terms of their idiosyncratic cost of registering and interest in politics is important to appraise the registration process from an individual perspective: does it exclude some citizens from participating, who would not otherwise self-select out of the political game?

To assess the process from a social welfare point of view (does it contributes to the marginalization of some specific subgroups of the population? does it undermine the representativity of the electoral outcomes?), one has to evaluate the way the marginal registrant changes as the intervention becomes more intensive on dimensions other than the propensity to vote: we estimate the selection impacts of the interventions on sociodemographic characteristics and political attitudes and preferences of the newly registered citizens.

Note regarding this preliminary version of the paper

We are yet to receive the data about the previous registration status of the newly registered citizens by the city of Montpellier, and thus only show and discuss results obtained when pooling all newly registered citizens together, whether they were previously unregistered or misregistered. Moreover, we have not used the data of the postelectoral survey, the survey of newly registered citizens at the time of registration and the survey of canvassers yet. Analyses based on this data will complete Sections 7 and 8 in particular.

6 The impact of the interventions on the number of new registrations and votes cast by newly registered citizens

6.1 The impact of the interventions on the gross number of new registrations

We first estimate the overall impact of our interventions on the gross number of new registrations by running

$$NR_{i,b} = \alpha + \beta T_b + \gamma N_{i,b} + X_b^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b} (1)$$

where subscript i refers to the unit of observation (the building or the apartment) and b to the building (the unit of randomization), the outcome NR is the number of new registrations, T_b is a dummy equal to 1 if the building belongs to one of the six treatment groups, N is the number of last names found on mailboxes that did not appear on the 2011 voters' list[42], used as a proxy for the initial number of unregistered and misregistered voters, X is a vector of building characteristics and δ^s are strata fixed effects.

X includes the building size, measured by the number of mailboxes and used as an (imperfect) proxy for social housing; the initial registration rate in this building, proxied by the ratio between the initial number of registrations and the number of mailboxes; and the average turnout of previously registered voters at this address at the 2012 elections, as a proxy for the degree of politization in this address.

Running equation (1) with the building at the unit of observation, we find that our interventions had a positive, large and significant impact on the gross number of new registrations: on average, they increased the number of new registrations by 0.3 (27%) in each treated building, as shown in Table 2, Panel A, column 1. This impact is large also when related to the overall population in the addresses included in the sample. They hosted 7.9 registered citizens on average in 2011[43] so that our interventions increased the number of registered citizens by 3.8% on average. This impact is robust to the inclusion of strata fixed effects (column 2) and building characteristics (column 3) and significant at the 1% level. As expected, the number of registrations was significantly higher in addresses with a higher N.

The individual impact of each intervention is estimated by running

$$NR_{i,b} = \alpha + \sum_{t=1}^{6} \beta_t T_b^t + \gamma N_{i,b} + X_b^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b}(2)$$

where the T_b^t's are dummies indicating treatment status and the β_t's measure the individual impact of each intervention. The results are displayed in columns 4 to 6 of Table 2: each of our six interventions significantly increased the number of new registrations. The impact of an Early Canvassing visit is the lowest, and only significant at the 10% level. The impact of any other intervention is significant at the 1% level. The increase in the number of new registrations is the highest in the addresses which received both an Early and a Late visit of Home registration: 0.62 (55%) in our preferred specification (column 6), which amounts to an overall increase of the number of registered citizens by 8%. A series of Wald tests show that the impact of this intervention was significantly higher than the impact of any other intervention. Conversely, the impact of an Early Canvassing visit is lower than the impact of any other intervention, except for a Late Canvassing visit.

We then use the apartment as the unit of analysis (Table 2, Panel B) and restrict the sample to apartments that were targeted by the interventions. The upside of using this smaller unit of analysis is that it should increase our statistical power; the downside is that we have to leave out some of the newly registered citizens for whom we know the address but not the apartment (we were not able to allocate them to the apartments identified during the preliminary work). We adjust the standard errors for clustering at the building level since the randomization was conducted at this level.

Among the targeted apartments, our interventions increased the number of registrations by 0.05 (30%) on average. Again, this impact is robust to the inclusion of strata fixed effects and building characteristics, as shown in columns 2 and 3, and significant at the 1% level. Scaled-up by the fraction of apartments with at least one unregistered or misregistered citizen that opened their doors, this intention to treat impact translates into a treatment on the treated estimate of $\frac{0.05}{0.519} = 0.10$ (58%). This corresponds to 11% of the initially unregistered and misregistered citizens living in apartments that opened their doors.[44]

[42] N is demeaned from the mean in the control group so that α can be read as the mean of R in the control group in regressions which do not include X and δ^s.

[43] We directly derive this number as the ratio between 32,399 registered citizens and 4,118 addresses

[44] We compute this fraction by relating 0.1 to the initial 0.27 unregistered citizens and 0.6 misregistered citizens.

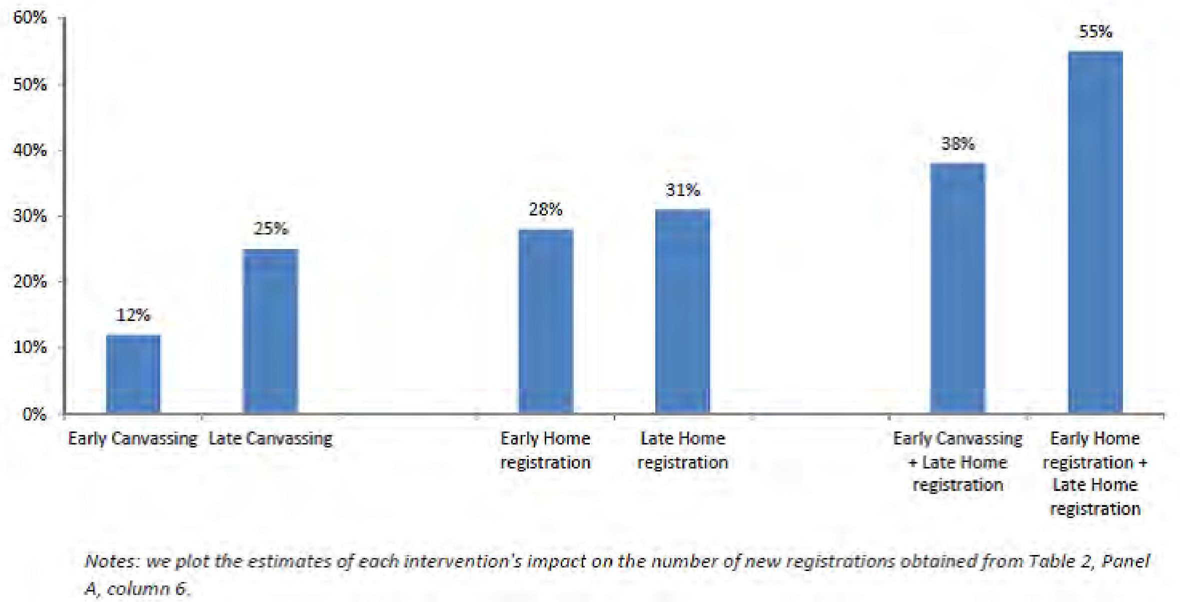


Figure 3: Impact of the interventions on the gross number of registrations

Except for the Early Canvassing visit, the impact of each individual intervention is significant at the 1% or 5% level. The impact of the combination of an Early and Late visit of Home registration is again the highest: this intervention increased the gross number of registrations by 0.09 (55%). Scaled up by $\frac{1}{0.577}$, this ITT estimate translates into a ToT registration impact of 0.16 (95%) new registrations, which corresponds to 19.5% of the initially unregistered and misregistered citizens living in the apartments that were included in this group and opened their doors: although the impact of two visits of Home registration is very high, approximately 60% of the initially unregistered and misregistered citizens remained unregistered, even after benefitting from the intervention.

6.2 The relative impact of a Late vs. Early visit on the gross number of new registrations

We run the rest of the regressions of the Section 6 at the household level, and investigate which component(s) of the different interventions best explain their relative impact. We first measure the specific contribution of visiting the addresses Late vs. Early by running equation (2), but with three treatment dummies only, which indicate whether the address received an Early visit (of Canvassing or Home registration), a Late visit (of Canvassing or Home registration) or two visits (Early Canvassing or Home registration + Late Home registration).

The comparison between Early and Late visit is the most meaningful: in both pooled treatment groups, half of the buildings benefitted from a simple door-to-door canvassing, and half were offered home registration, so that the groups differ only in the timing of the visit.

Columns 1 to 4 of Table 3 are identical to columns 2, 3, 5 and 6 of Table 2, Panel B, and included for reference.

In our preferred specification (column 6 of Table 3), we estimate that Early visits generated 0.02 (14%) more new registrations and Late visits 0.04 (27%) more new registrations, compared to the control group: this latter impact is 91% higher than the former, a difference significant at the 5% level.

This result can be interpreted in different ways: for instance, it could signal a complementarity between the interventions and the mediatic coverage of registration which increased as the deadline was getting closer. However, different from 2006, the 2011 media and public information campaign was centered on the very last days before the deadline, after even the late visit had taken place in most addresses. Although a series of other explanations can probably be suggested, the most likely is, according to us, the following one: the decision to register is subject to important procrastination and our visits were more effective at addressing it when they were made closer to the deadline. First, procrastinators who had postponed the decision to register had done it for a longer time then and were thus more likely to become sophisticated by talking to the canvassers. Second, when urged to register, it was more difficult for them to convince themselves that they still had lots of time to register when the visit was made closer to the deadline, thus making procrastination less likely to deaden the urging stimulus they had received. In short, we interpret the higher impact of a late visit as the sign that procrastination (also evidenced for instance by the large share of citizens who register in the last days and hours before the deadline) contributes to a large extent to the registration cost and that a late visit is an effective way to address it.

6.3 The relative impact of a Home registration vs. Canvassing visit on the gross number of new registrations

We now measure the specific contribution of offering Home registration vs. simple information and encouragement (in the Canvassing visits) in explaining the relative impacts of our different interventions. Again, we run equation (2) with three treatment dummies, which indicate whether the address received a Canvassing visit (either Early or Late), a Home registration visit (either Early or Late) or two visits (Early Canvassing or Home registration + Late Home registration).

The comparison between Canvassing and Home registration visits is the most meaningful: in both pooled treatment groups, half of the buildings were visited Early and half were visited Late, so that the groups differ only in the content of the visit.

In our preferred specification (column 8 of Table 3), we estimate that Canvassing visits generated 0.02 (14%) more new registrations and Home registration visits 0.04 (26%) more new registrations, compared to the control group: this latter impact is 91% higher than the former, a difference significant at the 10% level.

Moreover, two visits generated 0.08 (46%) more new registrations: this impact is 75% higher than the impact of Home registration visits, a difference significant at the 1% level.

The economically and statistically significant impact of simple door-to-door Canvassing visits shows that this intervention effectively addresses some components of the registration cost which are binding for important fractions of unregistered and misregistered citizens. Possible mechanisms underlying the registration impact of these visits are all related to the "uphill" side of the registration cost, the cost of deciding to register and gathering the information to do so. Door-to-door canvassing visits might have increased the saliency of the forthcoming elections and created a sense of urgency; they might have created a commitment to register towards the canvassers; or simply improved information about the registration process.

Differently, the additional registration impact of Home registration comes from the fact that it reduces the cost of actually going through the registration process, ie preparing the documents, filling out the application form, and bringing the file to the town hall. Its important size shows that "uphill" barriers to registration are not the only ones that matter: for equally important fractions of unregistered and misregistered citizens, it is rather the actual cost of going through the process which is binding. The success of Home registration, which effectively drops the registration cost to zero, was nevertheless not granted, as it could have been seen as generating a new cost, related to the risk that the assistance offered was hiding a fraud. So, the second take-away of its high impact is that a large fraction of unregistered citizens were trustful enough to let strangers take copies of sensitive individual documents, which opens the door for experimenting similar assistance interventions on other issues.

The relative size of the registration impact of the two types of interventions allows us to confidently conclude to the higher cost-effectiveness of Home registration visits, in terms of number of new registrations generated by a given amount of canvassers' time: although these visits took more time, the difference was not nearly as big as the difference in registration impact. This said, offering Home registration is clearly more sensitive than simply providing information and encouragement about the registration process: canvassers need to be trained and monitored, which might generate important fixed costs.

The higher impact of two visits compared to a single Home registration visit is more difficult to interpret as it is theoretically driven by several effects: first, a higher fraction of apartments that were visited twice opened their door at least once. Moreover, the apartments which opened their door to canvassers twice will show a higher number of registrations provided the impact on the number of new registrations is an increasing function of the number of interactions with canvassers.

6.4 The relative impact of a single vs. two visits on the gross number of new registrations

To test this hypothesis and estimate the additional impact of receiving both an Early and a Late visit instead of a Late visit only, we restrict the sample to the apartments in groups "Late Home registration", "Early canvassing + Late Home registration" and "Early Home registration + Late Home registration" which opened their door at the second visit and run equation (2) with only one treatment dummy, which indicates whether the address belongs to the groups "Early canvassing + Late Home registration" or "Early Home registration + Late Home registration", i.e. whether it received two visits.

All apartments in this subsample received a second visit of Home registration. Absent any control variables, α now estimates the average number of registrations in the treatment group "Late Home registration". The additional impact of receiving both an Early and a Late visit instead of a Late visit only is estimated by β scaled up by the inverse of the fraction of apartments (62%) which had opened their door at the first visit, among those in the groups "Early canvassing + Late Home registration" and "Early Home registration + Late Home registration" which opened their door at the second visit.
[45]

We do not find any significant additional impact of receiving both an Early visit (be it simple Canvassing or Home registration) and a Late visit (of Home registration), compared to receiving a Late visit only (Table 4, columns 1 to 3). The size of the sample used here is smaller than for the other regressions (2171 apartments in 371 addresses)[46], but the point estimate is small enough to make us confident that we are not missing an important impact by lack of power. Scaled up by the inverse of the fraction of apartments which had opened their door at the first visit, the 0.01 (3%) intention to treat estimate translates into a treatment on the treated estimate of 0.01 (5%).

The complementarity between the Early and Late visits might depend on the content of the first visit: Canvassing or Home regitration. To test this hypothesis, we run the same equation, but with two treatment dummies, indicating whether the address belongs to the groups "Early canvassing + Late Home registration" and "Early Home registration + Late Home registration".

[47]. We don't find any significant difference between these two treatments: the two point estimates are small and not significant, and we cannot reject the hypothesis that the two parameters are equal (colums 4 to 6).

[45] This does in two steps what a regression of $NR_{i,b}$ on a variable indicating whether the apartment opened its door at the first visit, instrumented by the treatment dummy, would do in one step. The rationale for using two steps is that we miss the information regarding whether the apartment opened its door at the first visit for some targetable addresses for which we could not collect the monitoring sheet and for all non-targetable addresses.

[46] For this reason, strata fixed effects are replaced by city fixed effects.

[47] Theoretically, we should expect the impact of "Early Home registration + Late Home registration" to be bigger if households who were offered to register at home a first time and refused because they thought they would register on their own are more

These results were expected as all households included in the sample received a Home registration visit, identified above as the most efficient, in the last period, identified as the one in which they were most receptive to the interventions. Although coming back a second time was not useless, as it enabled canvassers to interact with people who were absent the first time, which did generate additional registrations, it was certainly not cost-effective since the interactions with households which had already opened their door the first time did not generate additional registrations. The lack of impact of repeated canvassing enables us to identify the specific treatment impact of getting registered at home, as we discuss in section 7.3.

6.5 Robustness check: The impact of the interventions on the net number of new registrations

We test the robustness of the results obtained with the *gross* number of registrations as the outcome of interest to replacing it by the *net* number of new registrations after substracting the gross number of deregistrations.

The number of deregistrations was significantly lower in the treatment groups: using the household as the unit of analysis, and controlling for the number of voters registered at each address in 2011, we observe a difference of 15%, which is significant at the 5% level[48]. This difference is robust to the inclusion of strata fixed effects and building characteristics (number of mailboxes and ratio between the initial number of registrations and the number of mailboxes).

This difference should not be understood as an impact of the interventions. First, it is difficult to identify a channel through which our interventions might have caused this impact. Indeed, the main reasons why voters get deregistered from a voters' list are death, registration in a new city and striking off from the register if the town hall notices that the voter has moved away. We can thus only think of one channel through which our interventions might have impacted the number of deregistrations: households who received a canvasser's visit might have been more likely to alert previous household members who had moved away (or would do it between the visit and the registration deadline) and encourage them to update their registration status. But in this case, we should observe a *positive* impact of the interventions on the number of deregistrations. Second, the number of deregistrations does not seem to follow any clear pattern across our different treatment groups: it is lowest for the treatment group "Early Home registration," which did not receive the most intensive intervention. In the other groups, the difference with the control group is not significant.

Therefore, we interpret the higher number of deregistrations in the control group as a baseline difference which occured despite randomization. It could for instance reflect a relatively higher turnover of inhabitants in this group. In this case, the initial pool of incoming unregistered citizens was also higher in this group, and our estimate of the impact of canvassing on the gross number of new registrations is a lower bound on the true impact of canvassing. Under the assumption that, on average, each deregistration was mirrored by a new registration (to the extent that each person who moves away is replaced by an incomer), the number of deregistrations offers a counterfactual for the number of automatic registrations which would have happened in each building, absent the experiment. We can thus use the impact of the interventions on the net number of new registrations as an upper bound of their true impact of our interventions.

Controlling for the variable N, we estimate that our interventions increased the net number of new registrations by 0.06 (66%) in each treated building. This difference is significant at the 1% level and robust to the inclusion of strata fixed effects and building characteristics (Table 5, columns 1 to 3). The patterns identified

likely to accept it the second time, as they realize they did not register on their own, and might thus also fail to do it before the deadline. Alternatively, we should expect the impact of "Early canvassing + Late Home registration" to be bigger if the first visit anchors the cost of registering, and the second Home registration visit thus surprises those who first received simple Canvassing by revealing an actual lower cost, thus generating a higher number of registrations.

[48] For brevity, the related table is not included, but available upon request

with the gross number of registrations as the outcome of interest still hold: each of our six interventions significantly increased the number of new registrations (columns 4 to 6). The impact of an Early visit of simple Canvassing is still the lowest and significant only at 10%. The impact of any other intervention is significant at 1%. The increase in the net number of new registrations is the strongest in the group of addresses which received both an Early and a Late visit of Home registration: 0.11 (113%).

The impact of Late visits is no longer significantly higher than the impact of Early visits. Differently, Canvassing visits generated 0.04 (37%) more new net registrations and Home registration visits 0.07 (67%) more new net registrations: this latter impact is 80% higher than the former, a difference significant at the 5% level. Moreover, two visits generated 0.09 (95%) more new net registrations: this impact is 41% higher than the impact of Home registration visits, a difference significant at the 5% level.

6.6 The impact of the interventions on the number of votes cast by newly registered citizens

The impacts of our interventions on the number of new registrations are very high, but this would mean little if they did not translate in significant impacts on the number of votes cast by newly registered citizens.

To measure the impact of our interventions on the number of votes cast by newly registered citizens, we use the same empirical strategy as above, taking the number of votes cast by newly registered voters as the new outcome. We run

$$NV_{i,b} = \alpha + \beta T_b + \gamma N_{i,b} + X_b^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b}(3)$$

and

$$NV_{i,b} = \alpha + \sum_{t=1}^{6} \beta_t T_b^t + \gamma N_{i,b} + X_b^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b}(4)$$

where NV is the number of votes cast by newly registered citizens and the other variables have the same meaning as earlier.

We first run these two regressions for each round separately (Columns 1 to 8 of Table 6) and then for all rounds pooled together (Columns 9 and 10)[49].

On average, our interventions increased the number of votes cast by newly registered voters by 0.04 (28%) and 0.04 (26%) for the first and second rounds of the Presidential elections, by 0.01 (14%) and 0.02 (24%) for the first and second rounds of the General elections and by 0.03 (24%) when averaging over all rounds. All these estimates are statistically significant at 1% or 5%.

Averaging over all rounds, the impact was lowest and is not significant in the group of addresses that received an "Early Visit" only. It is significant for all other interventions, and highest in the group "Early Home registration + Late Home registration".

7 Selection and treatment impacts of the interventions on the propensity to vote

The impact of the interventions on the number of votes cast by newly registered citizens is determined both by their impact on the number of new registrations and by the relative propensity to vote of the additional citizens registered thanks to the interventions.

[49]The correlation between the observations for the same apartment is taken care of since we already adjust standard errors for clustering at the building level.

In this section, we estimate the selection and treatment impacts of the interventions on the propensity to vote of the beneficiaries. Our estimates of the difference between the propensity to vote of compliers and always-takers, and between compliers selected by simple door-to-door canvassing visits and home registration are a first central piece of evidence to understand the way the marginal registrant changes as the intervention becomes more intensive, and assess the role played by self-selection and lack of interest in the elections in explaining the failure of some citizens to register, absent any intervention.

7.1 The propensity to vote of newly registered citizens in the control and treatment groups

We first compare the propensity to vote of newly registered citizens in the control and the treatment groups, without distinguishing between the always-takers (who would have registered even absent any visit) and the compliers (who registered thanks to the interventions): we run

$$V_{i,b,r} = \alpha_r + \sum_{t=1}^{6} \beta_{t,r} T_b^t + \epsilon_{i,b,r} (5)$$

where $V_{i,b,r}$ is a dummy indicating the individual participation of voter i in building b at electoral round r, the T_b^t's are dummies indicating treatment status, and we take the individual participation at some electoral round as the unit of observation. We include neither strata fixed effects, nor any other control variable in this regression. Indeed, if the impact of our interventions varied across different strata and depending on these control variables, the latter would capture part of the difference between the participation of newly registered voters in the control and in the treatment groups.

The results obtained running regression (5) are displayed in Table 7. The coefficient on the constant shows the average turnout among newly registered citizens in the control group. It was very high for the Presidential elections (87% and 90% at the first and the second rounds) and much lower for the General elections (53% and 49% at the first and the second rounds). This difference is 14 percentage points (60%) larger than the difference in turnout measured at the national level: on average, the turnout was of 79% and 80% at the Presidential, and 57% and 55% at the General elections.

The turnout of newly registered citizens in the treatment groups follows a similar pattern. Overall, pooling all treatment groups together, it is lower than the turnout of newly registered citizens in the control group at all rounds, but significantly so (at 5% level) for the second round of the Presidential elections and the first round of the General elections only (Columns 1, 4, 7 and 10). On average, over the four rounds, the propensity to vote of newly registered citizens was lower in the treatment groups than in the control group by 2.3 percentage points (3%), a difference signifiant at the 10% level (Column 13).

The difference between the participation of newly registered citizens in a specific treatment group and in the control group at any given round is usually not significant (Columns 2, 5, 8 and 11). However, over the four rounds, the propensity to vote of newly registered citizens was lower in the two groups which received a single visit of Home registration (Column 14).

7.2 The selection impact of door-to-door canvassing

To test whether the propensity to vote of newly registered citizens was actually significantly lower in the groups which received a more intensive Home registration visit, we run equation (5) with three (instead of six) treatment dummies, which indicate whether the address received a Canvassing visit (either Early or Late), a Home registration visit (either Early or Late) or Two visits (Early Canvassing or Home registration + Late Home registration).

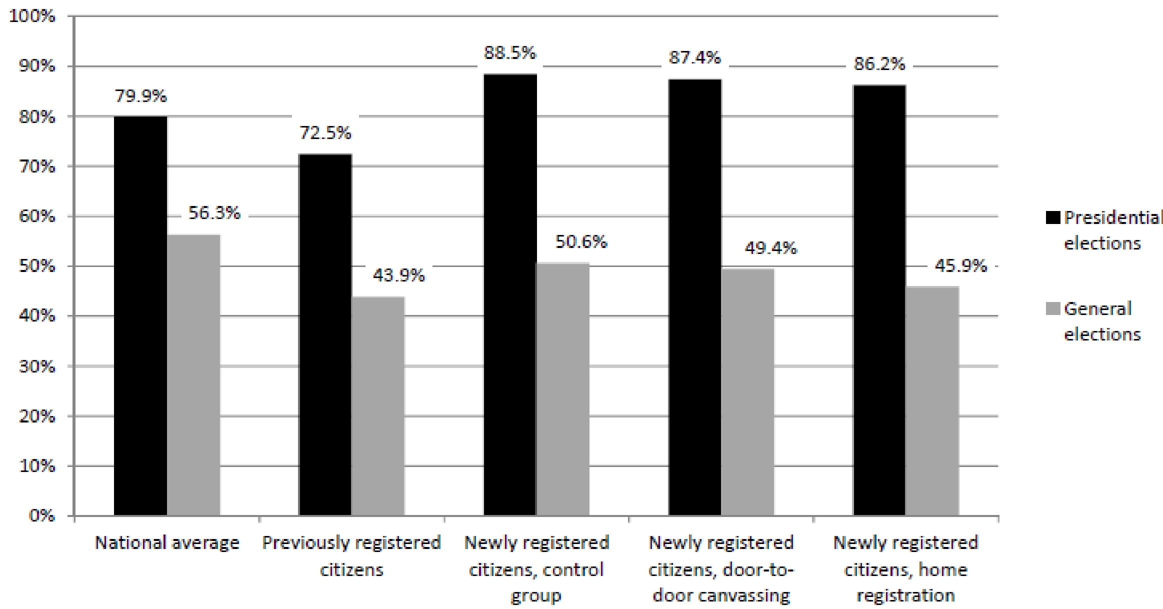


Figure 4: Propensity to vote of previously and newly registered citizens, by treatment group

$$V_{i,b,r} = \alpha_r + \beta_{\text{Canvassing},r} T_b^{\text{Canvassing}} + \beta_{\text{Home registration},r} T_b^{\text{Home registration}} + \beta_{\text{Two visits},r} T_b^{\text{Two visits}} + \epsilon_{i,b,r} \quad (6)$$

On average, over the four rounds, we estimate that the propensity to vote of newly registered citizens was lower by only 1 percentage point (2%) in the two groups which had received a Canvassing visit, compared to the control group. This difference is not significant at any standard level: we can thus directly infer that the selection impact of door-to-door canvassing itself is not significantly different from 0. In other words, on average, over the four rounds, the propensity to vote of the compliers selected by Canvassing visits was not significantly different from the propensity to vote of the always-takers.

Further, we estimate that, on average, the propensity to vote of newly registered citizens was lower by 3 percentage points (5%) in the two groups which had received a Home registration visit, compared to the control group. This difference is significant (at 1%). It is more than thrice as important as the former difference, a difference significant at 10% (Column 15 of Table 7). What does this tell us about the selection impact of home registration?

7.3 The treatment impact of home registration

We know from Section 5 (*Claims 14* and *15*) that, differently from the selection impact of Canvassing, the selection impact of Home registration cannot be directly inferred from the difference in propensity to vote between the newly registered citizens in the groups which received a Home registration visit and the control group. Indeed, this also depends on the treatment impact of home registration.

We can estimate this treatment impact using a strategy inspired from Karlan and Zinman (2009) and built in our experimental design. In the group "Early Canvassing + Late Home registration", people first received a less intensive visit (door-to-door canvassing) than in the group "Early Home Registration + Late Home Registration". However, a few weeks later, both groups received the most intensive home registration treatment in a second visit[50]. We showed in section 6.4 that the increase in the number of new registrations was almost

[50] Our strategy is reversed to the one used by Karlan and Zinman: they isolate a treatment impact (the moral hazard associated

identical in the two groups, among apartments which had opened their door at the second visit. This gives us strong reasons to believe that the corresponding newly registered citizens have the same socio-demographic characteristics and interest in politics, so that, their propensity to vote would be identical if they had been registered in the same way[51]

. We further check that there is no significant difference between the two groups of newly registered citizens in terms of observed characteristics (age, gender, a dummy indicating whether the citizen was born abroad, turnout of previously registered voters living at the same address and baseline registration rate), as shown in the odd-number columns of Table 8.

So, the groups of newly registered voters in both groups who had opened their door at the second visit differ only on one dimension: the fraction of them that got registered at home vs. at the town hall. Indeed, as expected, this fraction is more than twice as large in the group "Early Home registration + Late Home registration" (42%) than in the group "Early Canvassing + Late Home registration" (20%), a difference significant at 1% (Table 9, Column 1): in the former group, people were offered to register at home earlier, and were thus given less time to register at the town hall on their own.

The treatment group is thus a valid instrument for the type of registration, and we can run

$$V_{i,b,r} = \alpha + \xi H_i + \gamma N_{i,b} + X_b^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b,r} (12)$$

where H_i is a dummy equal to 1 if i was registered at home, and instrumented by the treatment dummy "Early Home registration + Late Home registration", and we have one observation per individual per round.

The results are reported in Table 10, columns 1 and 2: the treatment impact of home registration is small and not significant. This result is robust to the inclusion of individual and building control variables.

Unfortunately, the sample on which we estimate this treatment impact is small, since it is restricted to two treatment groups only, and to a fraction of newly registered citizens in these subgroups: those from whom we know that they opened their door at the second visit. This excludes all newly-registered citizens in non-targetable addresses and those that we were unable to allocate to an apartment.

To increase our power, and test the robustness of our finding, we include the treatment group "Late Home registration" in the analysis. Although addresses in this group were visited only once, so that the content of the intervention administered to this group differs to a larger extent from the interventions administered to the groups "Early Home registration + Late Home registration" and "Early Canvassing + Late Home registration" than they differ from each other, the increase in the number of new registrations was not statistically different in the former group than in the two latters ones. Moreover, we check that there is no significant difference between newly registered citizens in this group and the two other groups in terms of observed characteristics (even-number columns of Table 8). In short, we can be confident that the "exogeneity assumption" holds for this group assignment as well.

The "instrument relevance" itself holds: In this group, people were given more time to register at the town hall on their own than in the group "Early Home registration + Late Home registration", but they were not

with a relatively higher interest rate) from the related selection impact (adverse selection) by first selecting recipients through an identifical offer, and then surprising some with a lower interest rate. In our case, it was not possible to identify the group of citizens willing to register at the town hall before they actually registered. Once registered, their registration cost had been incurred and could no longer be changed. The advantage of our design it that it captures a dimension of the treatment effect that Kaplan and Zinman and other papers using a similar strategy (such as Cohen and Dupas (2010)) cannot capture: the signal about one's type that one seeks and sends to oneself when choosing a relatively more costly treatment. Indeed, individuals who benefit from a lower interest rate had self-selected for a higher interest rate and are thus affected by this dimension of the treatment effect equally as much as those who have to stick with the initial interest rate initially offered. In our study maybe more than in other contexts, this signalling process could be expected to be important, in so far as voting can be seen as contributing to a public good: as theorized by Bénabou and Tirole (2006), individuals' contributions to such goods might depend in part from a concern for self-respect.

[51] In the notation of the model: $E\left[w\left(u_i\right) \mid i \text{ is in group "Early Canvassing + Late Home registration"}\right] - E\left[w\left(u_i\right) \mid i \text{ is in group "Early Home registration + Late Home registration"}\right] = 0$

encouraged and informed early about the registration process as were people in the group "Early Canvassing + Late Home registration". Therefore, fewer got registered at the town hall before being offered to register at home at the second visit and, overall, a higher fraction of newly registered citizens who had opened their door at the second visit were registered at home than in the latter group (Table 9, Column 2).

We thus run equation (12) on a sample including newly registered citizens of this third group, and instrument H_i by the treatment dummies "Early Home registration + Late Home registration" and "Late Home registration". The results are robust to the inclusion of this subgroup, as we see in columns 3 and 4 of Table 10.

7.4 The selection impact of home registration

Although this exercise does not allow us to precisely estimate the treatment impact of home registration, it provides some evidence that it is small, so that the difference in propensity to vote between the newly registered citizens in the control group and the groups which received a Home-registration visit is mostly driven by the selection impact of home registration.

From *Claim 15*, we thus get that, for any round, the selection impact of home registration is approximately equal to

$$\beta_{\text{Home registration},r} \times \frac{1 + {}^{p_C}/_{p_A}}{{}^{p_C}/_{p_A}}$$

where p_C is the proportion of compliers selected by home registration, p_A is the proportion of always-takers, and $\beta_{\text{Home registration},r}$ is obtained from the results of equation (6).

From Section 6.3, we know that Home registration visits generated 0.04 (26%) more new registrations, compared to the control group. Further, from Section 7.1, we know that, on average, over the four rounds, $\beta_{\text{Home registration}} = -0.035$. Therefore, on average, over the four rounds, the selection impact of home registration was approximately $-0.035 \times \frac{1+0.262}{0.262} = -0.035 \times 4.817 = -0.169$.

On average, the compliers selected by home registration had a propensity to vote lower by 16.9 percentage points (24%) than the always-takers who got registered in the control group, absent any intervention.

7.5 Can we conclude about the reasons why the citizens registered thanks to the interventions would have failed to do so otherwise yet?

The analysis above shows that the propensity to vote of the compliers selected by door-to-door canvassing was not significantly different than that of the always-takers. Moreover, the propensity to vote of the compliers selected by the more intensive home registration visits was lower, granted, but only modestly so. It is tempting to infer from these results that the lack of interest in the elections had little importance in explaining the failure of their counterparts to register in the control group.

However, as we argue in Section 4.6 of the model, the selection impacts of the interventions might reflect not only differences in the long-term interest in politics and the elections more particularly but also a different mobilization effect of the campaign on compliers and always-takers. In other words, at this stage, we cannot completely rule out the hypothesis that, before and around the registration deadline, compliers were finding much lower benefits to the perspective of voting, which explains to a great extent their failure to register absent the interventions. And that it is only thanks to a relatively higher mobilization effect of the campaign on the compliers that such important fractions ended up going to the polls.

Fortunately, the much lower saliency of the General elections together with data from the post-electoral survey can help us disentangle the two terms which determine the selection impact.

8 Decomposing the selection impacts on turnout: The relative long-term interest in politics and short-term mobilization by the campaign of the compliers vs. always-takers

8.1 The relative propensity to vote of the compliers at the Presidential and General elections

Arguably, if it is likely that the Presidential campaign had an important mobilization effect, and more so for the less politicized citizens, including those selected by our interventions, we would expect the campaign for the General elections to have had a much smaller mobilization effect. Differences in propensity to vote measured at the two rounds of these elections should thus better reflect differences in the long-term interest in politics and the elections.

We investigate whether the difference between the propensity to vote of newly registered citizens in the treatment groups and the control group was larger at the less salient General elections by running

$$V_{i,b,r} = \alpha + G_r + \sum_{t=1}^{3} \beta_t^P T_b^t \times P_r + \sum_{t=1}^{3} \beta_t^G T_b^t \times G_r + \epsilon_{i,b,r} (13)$$

with one observation per person per round, where $P_r = 1$ for Presidential elections turnout data and $G_r = 1$ for General elections turnout data, and the three treatment dummies indicate whether the address received a Canvassing visit (either Early or Late), a Home registration visit (either Early or Late) or Two visits (Early Canvassing or Home registration + Late Home registration)

The results are displayed in Table 11. Columns 1 and 3 are identical to the columns 13 and 15 of Table 7, and included for reference. When all treatment groups are pooled together (Column 2), we find that the relative turnout of newly registered citizens in the treatment groups was lower by 1 percentage point (2%) at the Presidential elections and 3 percentage points (6%) at the General elections, compared to the control group. This latter difference only is significant, at the 10% level. It is twice as big as the former, but this difference is not statistically significant.

In Column 4, we let the groups which received a Canvassing visit, a Home registration visit or Two visits enter separately in the regression. The results obtained for the two treatment groups in which a Home registration visit was administered mirror those obtained for all treatment groups pooled together: the relative turnout of newly registered citizens in these groups was lower by 2 percentage point (3%) at the Presidential elections and 5 percentage points (9%) at the General elections, compared to the control group. Both differences are significant at the 5% level. The latter is twice as big as the former, but this difference is not statistically significant.

The difference between the relative turnout of newly registered citizens in the groups which received a Canvassing visit or Two visits and the control group is significant neither for the Presidential nor for the General elections. In these cases also, the difference was larger for the General elections, but this is not statistically significant

Although we cannot reject the hypothesis that the difference between the turnout of newly registered citizens in the treatment groups and the control group was identical at the Presidential and General elections, by lack of power, the above results suggest that it might have been higher for the General elections, at least for the addresses which received a visit of home registration. Thus, the long-term differences in interest in politics and the elections between the always-takers and the compliers selected by these visits might have been alleviated by the mobilization effect of the Presidential campaign, and might be only partially reflected in turnout differences.

Nonetheless, we can quite confidently conclude that door-to-door canvassing visits selected additional registrants whose turnout was close, if not equal to the turnout of the always-takers. Moreover, even when we take into account a possible mobilization effect of the Presidential campaign on the propensity to vote of compliers selected by the home regitration visits, what we can infer about their relative long-term interest in politics and the elections from their turnout is still higher than what the authors of the paper expected.

Applying the theoretical result stated in *Claim 9* of the model, readers who shared our initial expectations and find the compliers' willigness to participate higher than they expected will thus have to conclude with us that this was less binding and, therefore, their idiosyncratic registration cost higher and more binding at the time of registration. The cost that registering means to the unregistered and misregistered plays an important role in explaining their failure to register.

8.2 Data from the post-electoral survey

Data from the post-electoral survey further help us to disentangle the two terms which determine the selection impact: differences in the long-term interest in politics and the elections more particularly, and different mobilization effect of the campaign on compliers and always-takers.

TO BE COMPLETED

8.3 Lack of spillovers on the participation of previously registered voters

Evidence that participation is higher among people living in couple than people living alone (Niel and Lincot, 2012). Because of domino effet?

focus on people with same age

The extent to which the registration and turnout of compliers had spillover effects on the participation of previously registered citizens living in the same apartments and addresses brings further indirect evidence on the relative mobilization effect of the campaign on compliers, provided that at least part of this effect was conditional on being registered (and, thus, knowing that one would have the possibility to vote).

Indeed, the existence of significant spillover effects would be the sign of such a conditional mobilization effect, as enhanced interest of the compliers in the campaign would be their most likely channel: we do not expect our interventions to have directly affected the participation of previously registered citizens. Canvassers emphasized the close registration deadline rather than the upcoming elections. Moreover, early get-out-the-vote interventions have been repeatedly shown to not have any impact on turnout[52].

To test for the existence of spillovers, we run

$$V_{i,b} = \alpha + \beta T_b + X_b^{'}\lambda + Y_{i,b}^{'}\lambda + \sum_s \delta_b^s + \epsilon_{i,b}(14)$$

on a sample including all previously registered citizens living in buildings included in our sample, with one observation per citizen per round. X, the vector of building characteristics, includes the size of the building and the usual proxy for the baseline registration rate in this building. Y is a vector of individual characteristics that includes age, gender, and the number of other previously registered voters in the household.

The results are displayed in Table 12. We first observe that the participation of voters who were registered before 2011 is lower by 16 percentage points for the Presidential elections and 7 percentage points for the

52

Evaluating the extent to which there were such spillovers is interesting on its own: in some sense, the setting of this experiment creates an ideal context to test for peer effects in the choice to participate in an election.

General elections than the participation of newly registered citizens in the control group. This should be attributed at least in part to the fact that some of these voters have moved away. Among them, some are now registered in a new city and forgot to mention their previous registration address when they filled out their application. The others are the counterpart of the misregistered citizens who live in the same addresses: voting is more difficult and costly for them than for citizens registered at the address and actually living there, a group that includes the newly registered. Differently, their participation is higher by

Second, on average, our interventions did not significantly affect the participation of previously registered voters, as shown in columns 9 and 10 of Panel A. In Panel B, we restrict the sample to the previously registered citizens whose names had been identified on the mailboxes during the preparatory work and who should thus be on average more likely to actually live at the listed address. Their turnout is relatively higher (76.5% for the first round and 77% for the second round of the presidential elections) but still unaffected by the interventions[53].

9 Selection impacts of the interventions on other dimensions than the propensity to vote

In the previous section, we have shown that the compliers selected by the interventions differ little from the always-takers in terms of their propensity to vote and long-term interest in politics and the elections, and relatively more in terms of the specific cost that the registration process means to them.

The fact that the registration process prevents citizens willing to vote to register is annoying in itself, but it is even more so if these citizens are not random draws from the population, so that the process marginalizes some specific subgroups of the population and, if these subgroups have political preferences which differ from the rest of the population, undermines the representativity of the electoral outcomes.

We now turn to estimating the selection impacts of the interventions on sociodemographic characteristics and political attitudes and preferences of the newly registered citizens.

9.1 Characteristics available from the voters' lists

We first consider a set of characteristics available on the voters' lists, i.e. for all registered citizens, whether registered previously to 2011 or newly registered. We look for systematic differences between the compliers and always-takers, as we did for the propensity to vote, but also compare the newly registered citizens altogether (whether always-takers or compliers) to the previously registered citizens. For characteristics for which the former type of difference is small or not significant, so that the marginal registrant does not change on this dimension as registration becomes easier, it is still important to know if the latter difference is important: in this case also, we should conclude that the interventions increased the representation of some types of citizens otherwise underrepresented among the registered.

For any characteristic of interest X, we run

$$X_{i,b} = \alpha + \text{New}_i + \sum_{t=1}^{3} \beta_t^P T_b^t \times \text{New}_i + \epsilon_{i,b} \quad (15)$$

on the sample of all registered citizens, where New_i is a dummy equal to 1 if i is a newly registered citizen, and the three treatment dummies indicate whether the address received a Canvassing visit (either Early or

[53] Previously registered citizens who actually live at the address at which they are registered are still a subsample of the sample used in Panel B, and they were probably more likely to turn out. Indeed, the group of voters whose last names were found on the mailboxes includes people who have moved away and share their last name with someone still living there or but whose name was not erased from the mailbox. This is indirectly testified by the fact that 24% of the citizens who were deregistered had been identified as living there during the preparatory work; this fraction is too high to be entirely accounted for by moves that happened between the preparatory work and the interventions.

Late), a Home registration visit (either Early or Late) or Two visits (Early Canvassing or Home registration + Late Home registration).

As for regressions predicting voters' turnout we include neither strata fixed effects nor other control variables in this regression. The β's evaluate the difference between newly registered citizens in the treatment groups and the control group. We scale them up by $\frac{1+p_C/p_A}{p_C/p_A}$ to interpret them as the difference between compliers and always-takers: $\frac{1+0.262}{0.262} = 4.817$ for the Home registration visits; $\frac{1+0.137}{0.137} = 8.299$ for the Door-to-Door canvassing visits; and $\frac{1+0.286}{0.286} = 4.497$ when pooling all treatment groups together.

We first run this regression for a set of characteristics regarding the place of birth and report the results in Table 13, Panel A. We find that, in the control group, newly registered citizens are twice as less likely to be born in the city where they live than previously registered citizens. They are also 27% less likely to be born in another city in the department, and 40% more likely to be born in another region or abroad. Finally, they are 4% more likely to be born abroad. Similarly, compliers are less likely to be born in another city in the department, more likely to be born abroad and more likely to be born in a city.

We then consider other types of characteristics (Panel B) and find that the share of women is nearly identical among previously registered citizens, always-takers, and compliers. Differently, always-takers are on average 9 years younger than the previously registered citizens and compliers are even younger, a difference which is however not statistically significant. Unsurprisingly the newly registered citizens live in buildings in which the baseline registration rate was initially 25% lower. Differences between the housing price in addresses of the different categories of citizens are not significant.

Finally, and perhaps most interestingly, newly registered citizens in the control group live in buildings in which the degree of politicization (proxied by the propensity to vote of previously registered citizens) is 3% higher on average than in the addresses of previously registered citizens. On the contrary, the degree of politicization is 4% lower on average in the buildings where newly registered citizens in the treatment groups live, compared to the control group. Scaling-up this difference by 4.5, we find that the degree of politicization was 17% lower in the addresses of the compliers, compared to the always-takers: the interventions helped counterbalance an environment otherwise relatively less inducive to increasing one's political participation.

9.2 Postelectoral survey

TO BE COMPLETED

10 Conclusion

In this paper, we study the impact of door-to-door canvassing interventions implemented in 2011 on the registration rate and participation of newly registered citizens at the 2012 French Presidential and General elections. 4,118 buildings hosting 38,000 citizens were randomly allocated to a control group or one of six treatment groups varying by the timing and number of visits as well as their content: simple door-to-door canvassing (providing encouragement and information), or home registration.

Our interventions were highly effective: they increased the number of new registrations by 30% on average. This impact was significant in the groups of addresses which received door-to-door canvassing or ealy visits, and significantly higher in addresses which received the more intensive home registration visits and when the visit was closer to the registration deadline. This shows that, next to lack of information, procrastination and the cost of actually going through the process are important factors explaining some citizens' failure to register which were addressed by the interventions.

Overall, although the propensity to vote of citizens registered thanks to the interventions was significantly lower on average than the participation of the newly registered citizens in the control group, this difference is small and more than 80% of the newly registered voters selected thanks to the interventions participated in the Presidential elections.

Home registration visits selected additional citizens characterized by a lower propensity to vote than those selected by simple door-to-door canvassing, a difference which results from a higher selection impact rather than a negative treatment impact of getting registered at home on the beneficiaries' motivation and involvement. The comparison between the difference in propensity to vote of the citizens selected by home registration and in the control group at the Presidential and less salient General elections provides some evidence that the high turnout of the former at the Presidential elections was driven in part by a mobilization effect of the campaign. Nonetheless, the much higher than expected turnout of the additional citizens selected by the interventions at the subsequent elections, interpreted at the light of our model, suggests that the higher idiosyncratic cost that the registration process means to them plays a more important role than what could have been expected (given that the process is identical for all) in explaining the failure of their counterparts in the control group to register.

Finally, unregistered and misregistered citizens that the registration process prevents from participating in the elections systematically differ from others on other dimensions than the cost that the registration process means to them, as evidenced by the fact that the interventions increased the share of registered citizens who were born abroad and selected younger people, compared to previously registered citizens. They further selected citizens living in addresses in which the average propensity to vote was lower than for the newly registered citizens in the control group: although identical for all, the registration process imposes a higher cost to some subgroups of the population, it reinforces preexisting exclusions and undermines the representativity of the electoral outcomes.

Our results differ to an important extent from those obtained in experiments conducted in the United States, which find a much larger difference between the participation of citizens registered thanks to interventions and other voters (Nickerson, 2010). Our results show that the marginal registrant changes as the registration process is made easier. This suggests that the difference between our results and Nickerson's could be partly explained by the fact that the French registration process is more costly than the American one, so that it excludes unregistered and misregistered citizens with higher benefits of voting on average. But also that, replicating our experiment in the United States, we might find in this different context as well that the citizens selected by relatively less intensive interventions have a subsequent propensity to vote much closer to that of the citizens who register absent any intervention.

Now turning to political recommendations, our findings make a strong case for the generalization of interventions similar to the ones evaluated here and a general reform of the registration process and, more generally, transposed to other contexts characterized heavy administrative procedures, call for enhanced vigilance: such administrative procedures not only impose a cost on everyone. Although in principle identical for all, they might actually deter some specific subgroups, already marginalized on other dimensions, from accessing a good, service or status that they desire or deserve as much as others.

Should political parties willing to run door-to-door canvassing campaigns generalize interventions similar to the ones we evaluated and privilege the targetting of unregistered citizens compared to the registered citizens? In contexts close to the one from the experiment, the answer is a clear yes: when door-to-door canvassing campaigns that target abstentionists generate one extra vote for 14 non voters on average and fail to produce long-term effects, in this study, a similar campaign generated one additional newly registered voter and, with her, a total of 2.4 extra votes every 10 doors. This adds the ballots cast at the four subsequent rounds, a lower bound on all ballots cast in the future. Moreover, targeting areas with high fractions of unregistered and misregistered citizens should in principle be equally easy as targeting areas with high fractions of abstentionists and predicting their voting preferences should not be much harder[54].

Second, our results suggest that the French registration system should be made less costly, by increasing people's information, postponing the registration deadline and simplifying and standardizing the actual process of registering.

Postponing the registration deadline would address procrastination and lack of perceived saliency of the election at the time of registration. The example of the United States, where, depending on the state, registration is possible up to a few weeks before the election, and, in a few states, on Election Day itself, shows that this is logistically feasible.

In addition to this, there are a number of ways in which the actual process of registering could be standardized from one city to the other, in particular in terms of the type of documents accepted as a proof of address. Today, important differences between cities make any information collected in one city partly obsolete when one moves to a new place and has to update one's registration status. Further, our results suggest that the enfranchising impact of a simplification of the process (for instance by accepting a sworn statement as a proof of address) would greatly outweigh possible costs, including strategic registrations in swing cities or fraud. In North Dakota, the only American state that does not require registration, lists of voters are maintained from the previous election and citizens not on the list who show up to vote on Election Day are asked to sign a document swearing to the fact that they are qualified voters of the precinct. There has been no incidence of widespread fraud (Jaeger, 2002).

Bibliography

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association* 91: 444-455.

Ansolabehere, Stephen and Eitan Hersh (2011). Pants On Fire: Misreporting, Sample Selection, and Participation. Mimeo.

Ansolabehere, Stephen, Eithan Hersh and Kenneth Shepsle (2011). Movers, Stayers, and Registration: Why Age is Correlated with Registration in the US. Mimeo.

Arceneaux, Kevin (2005). Using Cluster Randomized Field Experiments to Study Voting Behavior. *The Annals of the American Academy of Political and Social Science* 601:169-179.

Ariely, Dan and Klau Wertenbroch (2002). Procrastination, Deadlines, and Performance: Self-Control by Precommitment. 13(3): 219-224.

Arkes, Hal R. and Catherine Blumer (1985). The Psychology of Sunk Cost. *Organizational Behavior and Human Decision Processes* 35: 124-140.

Banerjee, Abhijit V., Selvan Kumar, Rohini Pande and Felix Su (2011). Do Informed Voters Make Better Choices? Experimental Evidence from Urban India. *Working Paper*.

Banerjee, Abhijit V., Esther Duflo, Clément Imbert and Rohini Pande (2012). Fostering Political Competition: Evidence from India. *Work in Progress*.

Bergan, Daniel E., Alan S. Gerber, Donald P. Green and Costas Panagopoulos (2005). Grassroots mobilization and voter turnout in 2004. *Public Opinion Quarterly* 69:760–777.

Bénabou, Rolland and Jean Tirole (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies* 70: 489-520.

Bénabou, Rolland and Jean Tirole (2006). Incentives and Prosocial Behavior. *The American Economic Review* 96(5): 1652-1678.

[54]This last statement might be truer in France than in the US where an important fraction of registered citizens, including abstentionists, declare an affiliation with a political party, an information which is not available for any unregistered citizen.

Bennion, Elizabeth A. and David W. Nickerson (2009). I Will Register, if You Teach Me How: Results from Voter Registration Field Experiments on College Campuses. Mimeo.

Bennion, Elizabeth A. and David W. Nickerson (2010). The Cost of Convenience: An Experiment showing E-Mail Outreach Decreases Voter Registration. *Political Research Quarterly.* 1-12

Blais, André (2006). What affects voter turnout? *Annual Review of Political Science* 9: 111-125.

Blais, André, Elisabeth Gidengil, Neil Nevitte and Richard Nadeau (2004). Where does turnout decline come from? *European Journal of Political Research* 43: 221–236.

Bobo Lawrence and Franklin D. Gilliam (1990). Race, Sociopolitical Participation, and Black Empowerment. *The American Political Science Review* 84:377-393.

Bon Frédéric and Bernard Denni (1978). Population électorale, population électorale potentielle, population totale de la région Rhône- Alpes. *Revue Française de Science Politique* 28(6): 1055-1066.

Braconnier, Céline and Jean-Yves Dormagen (2007). La démocratie de l'abstention. Gallimard.

Braconnier, Céline and Jean-Yves Dormagen (2012). Logiques de Mobilisation et Inégalités Sociales de Participation Électorale en France, 2002-2012. *French Politics, Culture and Society.* 30(3) 20-44

Brady, Henry E., Sidney Verba and Kay Lehman Schlozman (1995). Beyon SES: A resource model of political participation. *American Political Science Review* 89(2): 271-94.

Bréchon Pierre and Bruno Cautrès (1987). L'inscription sur les listes électorales : indicateur de socialisation ou de politisation ? *Revue Française de Science Politique* 37(3): 502-528.

Bréchon Pierre (2009). La France aux urnes, Paris, La documentation française.

Brouard, Sylvain and Vincent Tiberj (2011). As French as everyone else ? A Survey of French Citizens of Maghrebin, African and Turkish French. Philadephia, Temple University Press.

Campbell, Philip E. and Roy Pierce (1986). Political Representation in France. Harvard University Press.

Cohen, Jessica and Pascaline Dupas (2010). Free Distribution or Cost-Sharing? Evidence From a Randomized Malaria Prevention Experiment. *Quarterly Journal of Economics* 125(1): 1-45

Devoto Florencia, Esther Duflo, Pascaline Dupas, William Pariente and Vincent Pons (2012). Happiness on Tap: Piped Water Adoption in Urban Morocco. *American Economic Journal: Economic Policy* forthcoming

Eldersveld, Samuel J. (1956). Experimental Propaganda Techniques and Voting Behavior. *American Political Science Review* 50:154-165.

Erikson, Robert S. (1981). Why Do People Vote? Because They Are Registered. *American Politics Quarterly* 9(3): 259-276.

Feddersen Timothy J. and Pesendorfer, Wolfgang (1996). The swing voter's curse. *The American Economic Review* 86,3: 408-424.

Filer, John E., Lawrence W. Kenny and Rebecca B. Morton (1991). Voting Laws, Educational Policies, and Minority Turnout. *Journal of Law and Economics* 34: 371-393.

Franklin, Marc (2005). You Want to Vote Where Everybody Knows Your Name: Anonymity, Expressive Engagement, and Turnout Among Young Adults. Paper presented at the 2005 Annual Meeting of the American Political Science Association (Washington DC, 1-4 September 2005).

Garrigou, Alain (1992). Le Vote et la vertu, comment les Français sont devenus électeurs. Paris, Presses De Sciences Po.

Gerber, Alan S. and Donald P. Green (1999). Does Canvassing Increase Turnout? A Field Experiment. *Proc. Natl. Acad. Sci.* USA. 96: 10939–10942.

Gerber, Alan S. and Donald P. Green (2000). The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *The American Political Science Review* 94:653-663.

Gerber, Alan S. and Donald P. Green (2005). Recent Advances in the Science of Voter Mobilization. *The Annals of the American Academy of Political and Social Science* 601:6-9.

Gerber, Alan S., Donald P. Green and David W. Nickerson (2006). Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments. *The Journal of Politics.* 65:1083–1096.

Gerber, Alan S. and Donald P. Green (2008). Get Out The Vote. Brookings Institution Press.

Highton Benjamin and Raymond E. Wolfinger (1998). Estimating the Effects of the National Voter Registration Act of 1993. *Political Behavior* 20(2): 79-104.

Highton Benjamin (1997). Easy Registration and Voter Turnout. *Journal of Politics* 59: 565-575.

Highton Benjamin (2004). Voter Registration and Turnout in the United States. *Perspectives on Politics* 2: 507-515.

Jackson, Robert A. (1996). A reassessment of voter mobilization. *Political Research Quarterly* 55:331-349.

Jackson, Robert A. (2002). Gubernatorial and Senatorial Campaign Mobilization of Voters. *Political Research Quarterly* 55:825-844.

Jaeger, Alvin A. (2002). North Dakota... The Only State Without Voter Registration. North Dakota Secretary of State, Elections Division. http://www.state.nd.us/sec/forms/pdf/votereg.pdf.

Karlan, Dean and Jonathan Zinman (2009). Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment. *Econometrica.* 77(6): 1993-2008.

Kelley S. Junior, R .E. Ayres and W.G Bowen (1967). Registration and Voting : Putting First Things First. *American Political Science Review.*

Knack Stephen and James White (2000). Election-Day Registration and Turnout Inequality. *Political Behavior* 22(1): 29-44.

Kramer, Gerald H. (1970). The Effects of Precinct-Level Canvassing on Voting Behavior. *Public Opinion Quarterly* 34: 560-572.

Kymlicka W. (2001). Politics in the Vernacular: Nationalism, Multiculturalism and Citizenship. Oxford: Oxford Univ. Press.

Leighley, Jan E. and Arnold Vedlitz (1999). Race, Ethnicity and Political Participation : Competing Models and Contrasting Explanations. *The Journal of Politics* 61 :1092-1114.

Levine, Peter and Mark Hugo Lopez (2005). What We Should Know about the Effectiveness of Campaigns but Don't. *The Annals of the American Academy of Political and Social Science* 601:180-191.

Martinez, Michael D., and David Hill (1999). Did Motor Voter Work?. *American Politics Quarterly* 27: 296-315.

Nagler J. (1991). The Effect of Registration Laws and Education on US Voter Turnout. *American Political Science Review* 85(4): 1393-1405.

Nannicini, Tommaso and Francesco Trebbi (2012). How do Voters Respond to Information? Evidence from a Randomized Campaign. *Work in progress*

Nickerson, David W. (2006). Forget me Not? The Importance of Timing and Frequency in Voter Mobilization. Mimeo

Nickerson, David W. (2006). Hunting the Elusive Young Voter. *Journal of Political Marketing* 5:47-69.

Nickerson, David W. (2010). Do Voter Registration Drives Increase Participation. Mimeo

Niel Xavier and Liliane Lincot (2012). L'inscription et la participation électorales en 2012. Insee Premiere n° 1411

Norris Pippa (2004). Electoral Engineering. Voting Rules and Political Behavior. Cambridge, Cambridge University Press.

Pacek , Alexander and Benjamin Radcliff (1995). Turnout and the Vote for Left-of-Centre Parties: A Cross-National Analysis. *British Journal of Political Science* 25:137-143

Pan Ké Shon, Jean-Louis (2004). Determinants of Electoral Non-Registration and Sensitive Neighbourhoods in France. *Population* 59(1): 143-156.

Pan Ké Shon, Jean-Louis (2010). The ambivalent nature of ethnic segregation in France's disadvantaged neighbourhoods. *Urban Studies* 47: 1603-1623.

Percheron Annick, Françoise Subileau and Marie-France Toinet (1987). Non inscription, abstention et vote blanc et nul en France. *Espace, Population, Sociétés* 3: 511-521.

Pons, Vincent and Guillaume Liegey (2012). Increasing the political participation of immigrants - Evidence from a field randomized experiment in France. Mimeo.

Rhine, Staci L. (1996). An Analysis of the Impact of Registration Factors on Turnout in 1992. *Political Behavior* 18(2): 171-198.

Riker, William H., and Peter C. Ordeshook (1968). A theory of the calculus of voting. *American Political Science Review* 62(1): 25-43.

Rohrschneider, Robert (2002). Mobilization versus chasing: how do parties target voters in election campaigns? *Electoral Studies* 21:367–382.

Rosenstone, Steven J., and John Mark Hansen (1993). Mobilization, Participation, and Democracy in America. New York: Macmillan

Rusk, Jerrold G., and John J. Stucker (1978). The effect of the southern system of election laws on voting participation. In *The history of American electoral behavior*, ed. Joel H. Silbey, Allan G. Bogue, and William H. Flanigan: 198-250. Princeton: Princeton University Press.

Ryan, Richard, and Edward L. Deci (2000). Self-determiantion theory and the facilitation of intrinsic motivation, social development and well-being. *American Psychologist.* 55(1): 68-78.

Ryan, Richard (1982). Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. *Journal of Personality and Social Psychology* 43(3): 450-461.

Sénat (2006). L'Inscription sur les listes électorales. *Etude de législation comparée* 161.

Teixeira, Ruy A. (1992). *The Disappearing American Voter.* Washington, CD: Brookings Institution.

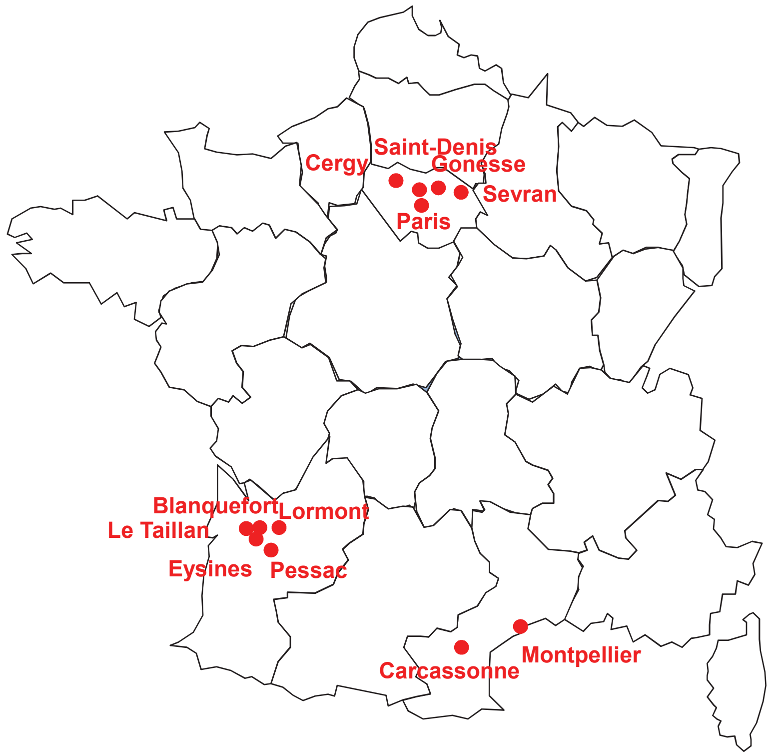
Timpone, Richard J. (1998). Structure, Behavior, and Voter Turnout in the United States. *American Political Science Review* 92(1): 145-158.

Wolfinger, Raymond E., and Steven J. Rosenstone (1980). *Who votes?* New Haven: Yale University Press.

Appendix

Appendix 1: Example of leaflets handed out by the canvassers

Appendix 2: Localization of the 12 cities included in the experiment



Appendix 3: Strategies used to infer the number of unregistered and misregistered citizens from the information collected by canvassers

We analyze the identification data recorded by the canvassers on the monitoring sheets with the two following caveats in mind: the apartments which opened their door are not necessarily representative of all pre-identified apartments; and the identification information recorded by the canvassers is potentially subject to systematic biases.

1st caveat : From the identification done by canvassers to a picture of the initial population

The first source of concern can be illustrated by the following example: on average, we would expect the likelihood to open the door to be higher for households with relatively more members, in which case the identification information collected by the canvassers would induce us to overestimate the size of the average household/apartment. Other systematic differences between apartments which opened their door and the others could lead us to systematically under- or overestimate the fraction of one or several types of citizens.

We use the groups of apartments which were targeted for two visits (and are located in addresses in which individual apartments could be targeted) to address this issue and draw correct inferences from the identification information collected by canvassers to the structure of the underlying population.

The strategy is as follows: each apartment characteristic we are interested in can be expressed as a single dummy or as a weighted sum of dummies. Whether the apartment counts an unregistered or misregistered citizen or not is a characteristic of the first type. The number of unregistered citizens in each apartment is a characteristic of the second type. Its average can be written as $A = \sum_{n=1}^{\infty} (n.\lambda_n)$ where λ_n is the fraction of apartments with i unregistered citizens.

For any dummy we are interested in (either in itself, or as a term of a sum), f_1, the proportion of apartments of type 1 (for which the dummy is equal to 1) can be estimated as

$$f_1 \equiv \frac{N_1}{N_1 + N_0} = \frac{X_1^1 X_1^2 X_0^{1,2}}{X_1^1 X_1^2 X_0^{1,2} + X_0^1 X_0^2 X_1^{1,2}}$$

where $N_i$is the (unknown) number of apartments of type i ($i = 0$ or 1) in the sample and X_i^j is the (known) number of apartments of type i which opened their door in phase j (j = 1 or 2 or $1,2$ if the apartment opened its door in both phases).

The first equality is by definition. The second equality comes from recognizing that $X_i^j = N_i p_i^j$ for i = (0 or 1) and j = (1 or 2)and $X_i^{1,2} = N_i p_i^1 p_i^2$ for i = (0 or 1) where p_i^j is the probability that a type i apartment opens its door in phase j. This gives us a system of six linear equations for six unknowns: N_0, N_1, p_0^1, p_1^1, p_0^2 and p_1^2. Solving it gives $N_0 = \frac{X_0^1 X_0^2}{X_0^{1,2}}$ and $N_1 = \frac{X_1^1 X_1^2}{X_1^{1,2}}$ from which we derive the desired result. [55]

2nd caveat: Systematic biases in the identification information

While the aforementioned strategy enables us to accurately infer the proportion of different types of individuals in our sample from the identification information collected by the canvassers, it does not correct for possible biases in the identification itself.

First, many citizens are confused about their registration status: some think that they are still registered at an old address when they were actually struck off from the register; others ignore that registration is address-specific and thus claim that they are registered at the right address even though they moved away without updating their registration status.

Second, although the canvassers were asked both to encourage or help people to register and to collect information about their registration status, the first goal was the most important one. So, canvassers were instructed to only ask respondants questions that were actually useful to advise the respondant and her household members and would thus seem natural. In this context, the information they could obtain was often only partial. For instance, they were often able to identify that a household only hosted adult foreigners, without being able to assess their exact number. In other cases, they were not able to entirely clarify whether the respondant was registered or misregistered: asking too many questions would have been unnatural. Canvassers might then have drawn inferences that were actually inaccurate.

Third, reemphasizing the norm of participation in politics was part of the intervention: not all canvassers insisted explicitly on this norm. But their very presence was making the point. Therefore, we cannot rule out that some respondants consciously lied about their registration status, to avoid being perceived as deviants.

Altogether, we can expect the identification information to be susceptible to three systematic biases: underestimation of the number of household members; overestimation of the fraction of unregistered citizens, to the extent that some foreigners were mistakenly included in this group; overestimation of the fraction of registered citizens, to the extent that misregistered citizens were mistakenly included in this group.

Although we cannot directly assess the importance of these biases, we can evaluate the overall quality of the identification by comparing its outcome, for apartments that were visited twice, and in which the identification

[55] f_1 can be estimated using this method only if $X_0^{1,2} \neq 0$, $X_1^{1,2} \neq 0$ and $\left((X_1^1 \neq 0 \text{ and } X_1^2 \neq 0) \text{ or } (X_0^1 \neq 0 \text{ and } X_0^2 \neq 0)\right)$. We use the following approximations for characteristics for which these conditions are not satisfied: $f_1 \sim 1$ if $X_0^{1,2} = 0$; $f_1 \sim 0$ if $X_1^{1,2} = 0$; $f_1 \sim 0$ if $X_0^{1,2} \neq 0$ and $X_1^{1,2} \neq 0$ but $X_1^1 = 0$ or $X_1^2 = 0$; $f_1 \sim 1$ if $X_0^{1,2} \neq 0$ and $X_1^{1,2} \neq 0$ but $X_0^1 = 0$ or $X_0^2 = 0$.

was possible twice: in 69% of these apartments, the two identifications give the same conclusion about the presence or absence of a misregistered citizen or unregistered citizen in the apartment.

To address the above-mentioned biases, we treat our estimates about the average household size, the fraction of households with at least one unregistered or misregistered citizen and the average number of unregistered and misregistered citizens as lower bounds. The rationale for the two latter assumptions is that it is likely that our third bias is greater than our second one, ie that we underestimate the fraction of misregistered voters to a much greater extent than we overestimate the fraction of unregistered voters.

There does not seem to be any easy way to build an upper bound of the average household size and the fraction of households with at least one unregistered or misregistered citizen. Fortunately, we can build an upper bound of the number of unregistered and misregistered citizens, which turns out to be the most important object for the rest of the analysis: we estimate the average number of well-registered citizens in each household based on the names found on the mailboxes and reallocate the other self-reported well-registered citizens to the group of misregistered citizens.[56]

We take the average between the upper bound and the lower bound as our preferred estimate of the average number of unregistered and misregistered citizens[57].

According to the identification information recorded by the canvassers, on average, apartments targeted by our interventions in targetable addresses included 0.68 registered citizens, 0.25 misregistered citizens, 0.25 unregistered citizens and 0.47 foreigners.

We construct an upper bound using the above-mentioned strategy: based on the names found on the mailboxes, we estimate that households targeted by our interventions hosted 0.21 registered citizens on average. Reallocating the rest of the reported well-registered citizens to the group of misregistered citizens, we get upper bounds estimates of the average number of misregistered citizens and the total number of unregistered and misregistered citizens in the sample: respectively $0.72 = 0.25 + (0.68 - 0.21)$ and $0.97 = 0.5 + (0.68 - 0.21)$.

Taking the average between the upper bound and the lower bound, we estimate that on average, targeted households hosted 0.45 registered citizens, 0.48 misregistered citizens, 0.25 unregistered citizens and 0.47 foreigners. In the 16,567 apartments targeted by our interventions, we thus estimate the numbers of unregistered and misregistered citizens as 4,142 and 7,952 respectively.

From the composition of targeted households in targetable addresses to the composition of the total sample population

A few more steps are required to go from the composition of targeted households in targetable addresses to the composition of the total sample population.

First, we need to account for the fact that not all apartments hosting unregistered or misregistered citizens were identified as such and included in the sample. The (weighted) fraction of apartments which were misidentified as hosting only registered citizens when they actually hosted at least one unregistered or misregistered citizen can be estimated as the number of newly registered citizens that we can allocate to such apartments over the number of all newly registered citizens that we were able to allocate to an apartment : 85%.[58]. Assuming that the composition of households living in these misidentified apartments is similar to that of households living in the targeted apartments, we estimate that the targetable addresses included in the sample hosted a total of 4,873 unregistered citizens and 9,355 misregistered citizens.

[56]This gives an upper bound to the extent that we then underestimate the average number of well-registered citizens and thus overestimate the average number of unregistered and misregistered citizens.

[57]This amounts to assuming that exatly half of the citizens identified as well-registered are actually misregistered.

[58]This identification is fully valid only under the assumption that the likelihood to register was identical among these initially unregistered or misregistered citizens and the others. This assumption is more likely to hold in the control group: in the treatment groups, our interventions increased the likelihood to register for targeted apartments only.

Second, while 26,454 citizens were officially registered in the targetable addresses in 2011, not all of them actually lived there: the misregistered citizens targeted by our interventions, who live there but are registered at another address, are mirrored by misregistered citizens of the opposite type, who used to live here, but moved away without being struck off from the list. The two groups should be of equal size on average, so that we can estimate that the targetable addresses hosted $26454 - 9355 = 17099$ well-registered citizens.

We finally assume that the structure of the underlying population was identical in non-targetable addresses and in targetable addresses.

We end up with the following estimate: the 4,118 addresses included in our sample hosted initially 38,375 citizens, among which 55% were "well-registered", 30% misregistered and 15% unregistered.

Appendix 4: Proofs of claims stated in Sections 4 and 5

To prove the claims of section 4, we use the following definitions and theorems.

First-order stochastic dominance

The distribution function F first-order stochastically dominates G if, for every weakly increasing $z : \mathbb{R} \to \mathbb{R}$, $\int_{-\infty}^{\infty} z(u)f(u)du \geq \int_{-\infty}^{\infty} z(u)g(u)du$, where f and g are the density functions corresponding to F and G.

Monotone Likelihood Ratio dominance

F dominates G in the Monotone Likelihood Ratio sense if $l(u) \equiv \frac{g(u)}{f(u)}$ is weakly decreasing.

Theorem 1

F first-order stochastically dominates G if and only if $F(u) \leq G(u)$ for all u.

Proof of Theorem 1

Define $H(u) = F(u) - G(u)$ for all u.

Proof that F first-order stochastically dominates $G \Rightarrow F(u) \leq G(u)$ for all u. Suppose towards contradiction that $\exists u*$ such that $H(u*) > 0$. Define $z(u) = \mathbf{1}_{\{u \geq u*\}}$. Then, $\int_{-\infty}^{\infty} z(u)h(u)du = \int_{u*}^{\infty} h(u)du = -H(u*) < 0$, from the definition of H and the assumption that $H(u*) > 0$, and $\int_{-\infty}^{\infty} z(u)h(u)du > 0$, from the fact that $z(u)$ is weakly increasing and the assumption that F first-order stochastically dominates G. This finishes the proof by contradiction.

Proof that $F(u) \leq G(u)$ for all $u \Rightarrow F$ first-order stochastically dominates G . Take any weakly increasing z that is differentiable everywhere. Then, by integration by parts, $\int_{-\infty}^{\infty} z(u)h(u)du = [z(u)H(u)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} z'(u)H(u)du = -\int_{-\infty}^{\infty} z'(u)H(u)du \geq 0$ since z is weakly increasing and $H(u) \leq 0$ for all u. This shows that F first-order stochastically dominates G.

Theorem 2

If F dominates G in the Monotone Likelihood Ratio sense, then F also first-order stochastically dominates G.

Proof of Theorem 2

If F dominates G in the Monotone Likelihood Ratio sense, then $\frac{g(u')}{g(u)} \leq \frac{f(u')}{f(u)}$ for any $u' > u$.

Since f and g are density functions, $\int_{-\infty}^{\infty} f(u)du = \int_{-\infty}^{\infty} g(u)du = 1$ and $\int_{-\infty}^{\infty} (f-g)(u)du = 0$. Thus, there exists $u*$ such that $f(u*) = g(u*)$.[59]
$\frac{g(u')}{g(u)} < \frac{f(u')}{f(u)}$ for any $u' > u$ implies $\int_{-\infty}^{x} g(u)du \geq \frac{g(x)}{f(x)} \int_{-\infty}^{x} f(u)du$ for any x.
We can further show that $\int_{-\infty}^{x} g(u)du \geq \int_{-\infty}^{x} f(u)du$ for any x:

- for any $x \leq u*$, $\frac{g(x)}{f(x)} \geq \frac{g(u*)}{f(u*)} = 1$ so that $\int_{-\infty}^{x} g(u)du \geq \frac{g(x)}{f(x)} \int_{-\infty}^{x} f(u)du \geq \int_{-\infty}^{x} f(u)du$
- for any $x > u*$, $\frac{g(x)}{f(x)} \leq \frac{g(u*)}{f(u*)} = 1$. $g(x) \leq f(x)$ for any $x > u*$ implies $\int_{x}^{\infty} g(u)du < \int_{x}^{\infty} f(u)du$ for any $x > u*$. Since $\int_{-\infty}^{\infty} f(u)du = \int_{-\infty}^{\infty} g(u)du = 1$, this implies $\int_{-\infty}^{x} g(u)du \geq \int_{-\infty}^{x} f(u)du$ for any $x > u*$

So, $G(x) \geq F(x)$ for any x: using *Theorem 1*, this shows that F first-order stochastically dominates G.

Claim 1

$E[u_i \mid i \text{ is complier}, c_i = \bar{c}] < E[u_i \mid i \text{ is always-taker}, c_i = \bar{c}]$ for any $\bar{c}$: compliers characterized by a given $\bar{c}$ have a lower expected u than always-takers facing the same $\bar{c}$.

Proof of Claim 1

$E[u_i \mid i \text{ is complier}, c_i = \bar{c}] = \int_{g^{-1}(\lambda\bar{c})}^{g^{-1}(\bar{c})} u \frac{f(u,\bar{c})}{\int_{g^{-1}(\lambda\bar{c})}^{g^{-1}(\bar{c})} f(u,\bar{c})du} du$ and
$E[u_i \mid i \text{ is always-taker}, c_i = \bar{c}] = \int_{g^{-1}(\bar{c})}^{\infty} u \frac{f(u,\bar{c})}{\int_{g^{-1}(\bar{c})}^{\infty} f(u,\bar{c})du} du$ for any $\bar{c}$.
$\int_{g^{-1}(\lambda\bar{c})}^{g^{-1}(\bar{c})} u \frac{f(u,\bar{c})}{\int_{g^{-1}(\lambda\bar{c})}^{g^{-1}(\bar{c})} f(u,\bar{c})du} du < \int_{g^{-1}(\bar{c})}^{\infty} u \frac{f(u,\bar{c})}{\int_{g^{-1}(\bar{c})}^{\infty} f(u,\bar{c})du} du$ is immediate.

Claim 2

Condition ID ($-f(u,c)$ satisifies log-increasing differences in u and c) is satisfied for instance by any bivariate normal density (the type bivariate density most commonly used) with negative correlation.

Proof of Claim 2

Let's consider any bivariate normal density $f(u,c)$ with correlation $\rho < 0$. The bivariate density is fully characterized by ρ, μ_u, μ_c, σ_u and σ_c:
$f(u,c) = \left(2\pi\sigma_u\sigma_c\sqrt{1-\rho^2}\right)^{-1} \times \exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{u-\mu_u}{\sigma_u}\right)^2 + \left(\frac{c-\mu_c}{\sigma_c}\right)^2 - 2\rho\left(\frac{u-\mu_u}{\sigma_u}\right)\left(\frac{c-\mu_c}{\sigma_c}\right)\right)\right)$ for any $u \in]-\infty,\infty[$ and $c \in]-\infty,\infty[$.
Now take any u and any $c' > c$.
$\frac{f(u,c')}{f(u,c)} = \exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{c'-\mu_c}{\sigma_c}\right)^2 - \left(\frac{c-\mu_c}{\sigma_c}\right)^2 - 2\rho\left(\frac{u-\mu_u}{\sigma_u}\right)\left(\frac{c'-c}{\sigma_c}\right)\right)\right)$
Now taking any $u' > u$ and using the fact that $\exp(x)$ is strictly increasing for any x and our assumption that $\rho < 0$, we get
$\frac{f(u',c')}{f(u',c)} < \frac{f(u,c')}{f(u,c)}$
$\Leftrightarrow -\frac{1}{2(1-\rho^2)}\left(\left(\frac{c'-\mu_c}{\sigma_c}\right)^2 - \left(\frac{c-\mu_c}{\sigma_c}\right)^2 - 2\rho\left(\frac{u'-\mu_u}{\sigma_u}\right)\left(\frac{c'-c}{\sigma_c}\right)\right) < -\frac{1}{2(1-\rho^2)}\left(\left(\frac{c'-\mu_c}{\sigma_c}\right)^2 - \left(\frac{c-\mu_c}{\sigma_c}\right)^2 - 2\rho\left(\frac{u-\mu_u}{\sigma_u}\right)\left(\frac{c'-c}{\sigma_c}\right)\right)$
$\Leftrightarrow -2\rho\left(\frac{u'-\mu_u}{\sigma_u}\right)\left(\frac{c'-c}{\sigma_c}\right) > -2\rho\left(\frac{u-\mu_u}{\sigma_u}\right)\left(\frac{c'-c}{\sigma_c}\right)$
$\Leftrightarrow (u'-u)(c'-c) > 0$
Thus,$-f(u,c)$ satisifies log-increasing differences in u and c.

[59] This implicitly assumes the continuity of f and g. However, the proof holds even without this assumption.

Claim 3

If *Condition ID* and *R1* hold, then $E[u_i \mid i \text{ is complier}] < E[u_i \mid i \text{ is always-taker}]$

Proof of Claim 3

$E[u_i \mid i \text{ is complier}] = \frac{\int_{-\infty}^{\infty} u \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$ and $E[u_i \mid i \text{ is always-taker}] = \frac{\int_{-\infty}^{\infty} u \int_{-\infty}^{g(u)} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} f(u,c)dc\ du}$.

$E[u_i \mid i \text{ is complier}]$ and $E[u_i \mid i \text{ is always-taker}]$ can be rewritten as $E[u_i \mid i \text{ is complier}] = \int_{-\infty}^{\infty} uh(u)du$ and $E[u_i \mid i \text{ is always-taker}] = \int_{-\infty}^{\infty} uk(u)du$ with $h(u) = \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$ and $k(u) = \frac{\int_{-\infty}^{g(u)} f(u,c)dc}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} f(u,c)dc\ du}$.

h and k are two density functions: since $f(u,c) > 0$ for any u and c, $h(u) > 0$ and $k(u) > 0$ for any $u \in]-\infty,\infty[$. Moreover, $\int_{-\infty}^{\infty} h(u)du = \int_{-\infty}^{\infty} k(u)du = 1$. We call H and K the distribution functions corresponding to h and k.

We now show that $\frac{h(u')}{h(u)} \leq \frac{k(u')}{k(u)}$ for any $u' > u$.

$\frac{h(u')}{h(u)} \leq \frac{k(u')}{k(u)}$ for any $u' > u$ $\Leftrightarrow \frac{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc}{\int_{-\infty}^{g(u')} f(u',c)dc} \leq \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u)} f(u,c)dc}$ for any $u' > u$, which we show in two steps

First, we show that $\frac{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc}{\int_{-\infty}^{g(u')} f(u',c)dc} \leq \frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u')} f(u,c)dc}$ for any $u' > u$ using *Condition ID*.

Take $u' > u$. Since$-f(u,c)$ satisfies log-increasing differences: $\frac{f(u',c')}{f(u',c)} \leq \frac{f(u,c')}{f(u,c)}$ for any $c \in]-\infty, g(u')]$ and $c' \in [g(u'), {}^{g(u')}/_{\lambda}]$. Therefore, $\frac{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc}{f(u',c)} \leq \frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{f(u,c)}$ or $\frac{f(u',c)}{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc} \geq \frac{f(u,c)}{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}$ for any $c \in]-\infty, g(u')]$ and $\frac{\int_{-\infty}^{g(u')} f(u',c)dc}{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc} \geq \frac{\int_{-\infty}^{g(u')} f(u,c)dc}{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}$ or $\frac{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc}{\int_{-\infty}^{g(u')} f(u',c)dc} \leq \frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u')} f(u,c)dc}$.

Second, we show that $\frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u')} f(u,c)dc} \leq \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u)} f(u,c)dc}$ for any u and $u' > u$ using *Condition 2*.

Consider any $u' > u$. We show $\frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u')} f(u,c)dc} \leq \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u)} f(u,c)dc}$ or $\frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{F(g(u')|u)} \leq \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{F(g(u)|u)}$ by showing that $z\left(\widetilde{\lambda}\right) = \frac{\int_{g(u')}^{g(u')\widetilde{\lambda}} f(c|u)dc}{F(g(u')|u)} - \frac{\int_{g(u)}^{g(u)\widetilde{\lambda}} f(c|u)dc}{F(g(u)|u)}$ decreases in $\widetilde{\lambda}$ for any $\widetilde{\lambda} \in [1, {}^1/_{\lambda}]$ and that $z(1) = 0$.

$z(1) = 0$ is immediate.

We prove that $z\left(\widetilde{\lambda}\right)$ decreases in $\widetilde{\lambda}$ for any $\widetilde{\lambda} \in [1, {}^1/_{\lambda}]$ by iteration.

First, $z'(1) \leq 0$. Indeed, $z'(1) = \lim_{\varepsilon > 0, \varepsilon \to 0} \frac{z(1+\varepsilon) - z(1)}{\varepsilon}$.

For any $\varepsilon > 0$, $\frac{z(1+\varepsilon)-z(1)}{\varepsilon} \leq 0 \Leftrightarrow \frac{1}{\varepsilon}\left(\frac{\int_{g(u')}^{g(u')(1+\varepsilon)} f(c|u)dc}{F(g(u')|u)} - \frac{\int_{g(u)}^{g(u)(1+\varepsilon)} f(c|u)dc}{F(g(u)|u)}\right) \leq 0$.

As $\varepsilon \to 0$, $\frac{1}{\varepsilon}\left(\frac{\int_{g(u')}^{g(u')(1+\varepsilon)} f(c|u)dc}{F(g(u')|u)} - \frac{\int_{g(u)}^{g(u)(1+\varepsilon)} f(c|u)dc}{F(g(u)|u)}\right) \to \frac{g(u')f(g(u')|u)}{F(g(u')|u)} - \frac{g(u)f(g(u)|u)}{F(g(u)|u)}$.

From *Condition R1*, we have that $\frac{g(u')f(g(u')|u)}{F(g(u')|u)} < \frac{g(u)f(g(u)|u)}{F(g(u)|u)}$. Therefore, $z'(1) \leq 0$

We now show that if $z\left(\widetilde{\lambda}\right)$decreases in $\widetilde{\lambda}$ for any $\widetilde{\lambda} \in \left[1, \widetilde{\widetilde{\lambda}}\right[$, we also have $z'\left(\widetilde{\widetilde{\lambda}}\right) < 0$, where $\widetilde{\widetilde{\lambda}} \leq {}^1/_{\lambda}$.

Since $z\left(\widetilde{\lambda}\right)$ decreases in $\widetilde{\lambda}$ for any $\widetilde{\lambda} \in \left[1, \widetilde{\widetilde{\lambda}}\right[$, $z\left(\widetilde{\widetilde{\lambda}}\right) \leq 0$: $\frac{\int_{g(u')}^{g(u')\widetilde{\widetilde{\lambda}}} f(c|u)dc}{F(g(u')|u)} \leq \frac{\int_{g(u)}^{g(u)\widetilde{\widetilde{\lambda}}} f(c|u)dc}{F(g(u)|u)}$, which implies $\frac{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u')|u)} = \frac{F\left(g(u')|u\right)+\int_{g(u')}^{g(u')\widetilde{\widetilde{\lambda}}} f(c|u)dc}{F(g(u')|u)} \leq \frac{F(g(u)|u)+\int_{g(u)}^{g(u)\widetilde{\widetilde{\lambda}}} f(c|u)dc}{F(g(u)|u)} = \frac{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u)|u)}$.

By definition, $z'\left(\widetilde{\widetilde{\lambda}}\right) = \lim_{\varepsilon>0,\varepsilon\to 0} \frac{z\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)-z\left(\widetilde{\widetilde{\lambda}}\right)}{\varepsilon}$.

For any $\varepsilon > 0$, $\frac{z\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)-z\left(\widetilde{\widetilde{\lambda}}\right)}{\varepsilon} = \frac{1}{\varepsilon}\frac{\int_{g(u')\widetilde{\widetilde{\lambda}}}^{g(u')\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F(g(u')|u)} - \frac{1}{\varepsilon}\frac{\int_{g(u)\widetilde{\widetilde{\lambda}}}^{g(u)\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F(g(u)|u)} = \frac{1}{\varepsilon}\frac{\int_{g(u')\widetilde{\widetilde{\lambda}}}^{g(u')\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}\frac{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u')|u)} - \frac{1}{\varepsilon}\frac{\int_{g(u)\widetilde{\widetilde{\lambda}}}^{g(u)\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}\frac{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u)|u)}$.

Since $\frac{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u')|u)} \leq \frac{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}{F(g(u)|u)}$, it is enough to show that $\frac{1}{\varepsilon}\frac{\int_{g(u')\widetilde{\widetilde{\lambda}}}^{g(u')\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)} - \frac{1}{\varepsilon}\frac{\int_{g(u)\widetilde{\widetilde{\lambda}}}^{g(u)\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)} \leq 0$ to show that $\frac{z\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)-z\left(\widetilde{\widetilde{\lambda}}\right)}{\varepsilon} \leq 0$.

As $\varepsilon \to 0$, $\frac{1}{\varepsilon}\frac{\int_{g(u')\widetilde{\widetilde{\lambda}}}^{g(u')\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)} - \frac{1}{\varepsilon}\frac{\int_{g(u)\widetilde{\widetilde{\lambda}}}^{g(u)\left(\widetilde{\widetilde{\lambda}}+\varepsilon\right)} f(c|u)dc}{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)} \to \frac{g(u')f\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)} - \frac{g(u)f\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}{F\left(ug(u)\widetilde{\widetilde{\lambda}}|u\right)}$.

From *Condition R1*, we have that $\frac{g(u')f\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)}{F\left(g(u')\widetilde{\widetilde{\lambda}}|u\right)} < \frac{g(u)f\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}{F\left(g(u)\widetilde{\widetilde{\lambda}}|u\right)}$. Therefore, $z'\left(\widetilde{\widetilde{\lambda}}\right) \leq 0$.

This finishes the proof that $z\left(\widetilde{\lambda}\right)$ decreases in $\widetilde{\lambda}$ for any $\widetilde{\lambda} \in [1, 1/\lambda]$ and, combined with $z(1) = 0$, that $\frac{\int_{g(u')}^{g(u')/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u')} f(u,c)dc} \leq \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{g(u)} f(u,c)dc}$.

Since $\frac{h(u')}{h(u)} \leq \frac{k(u')}{k(u)}$ for any $u' > u$, we can apply *Theorem 2* to h and k: K first-order stochastically dominates H. By definition of the first-order stochastic dominance, this implies that for every weakly increasing function $z(u)$, $\int_{-\infty}^{\infty} z(u)h(u)du \leq \int_{-\infty}^{\infty} z(u)k(u)du$. In particular, for the identity function $z(u) = u$, we get $E\left[u_i \mid i \text{ is complier}\right] \leq E\left[u_i \mid i \text{ is always-taker}\right]$ Q.E.D.

Claim 4

$E\left[c_i \mid i \text{ is complier}, u_i = \bar{u}\right] > E\left[c_i \mid i \text{ is always-taker}, u_i = \bar{u}\right]$ for any $\bar{u}$: compliers characterized by a given $\bar{u}$ have a higher expected c than always-takers facing the same $\bar{u}$.

Proof of Claim 4

$E\left[c_i \mid i \text{ is complier}, u_i = \bar{u}\right] = \int_{g(\bar{u})}^{g(\bar{u})/\lambda} c\frac{f(u,c)}{\int_{g(\bar{u})}^{g(\bar{u})/\lambda} f(u,c)dc}dc$ and $E\left[c_i \mid i \text{ is always-taker}, u_i = \bar{u}\right] = \int_{-\infty}^{g(\bar{u})} c\frac{f(u,c)}{\int_{-\infty}^{g(\bar{u})} f(u,c)du}dc$ for any $\bar{u}$.

$\int_{g(\bar{u})}^{g(\bar{u})/\lambda} c\frac{f(u,c)}{\int_{g(\bar{u})}^{g(\bar{u})/\lambda} f(u,c)dc}dc > \int_{-\infty}^{g(\bar{u})} c\frac{f(u,c)}{\int_{-\infty}^{g(\bar{u})} f(u,c)du}dc$ is immediate.

Claim 5

If *Condition ID* and *R1* hold, then $E\left[c_i \mid i \text{ is complier}\right] > E\left[c_i \mid i \text{ is always-taker}\right]$.

Proof of Claim 5

$E\left[c_i \mid i \text{ is always-taker}\right]=\frac{\int_{-\infty}^{\infty} c \int_{g^{-1}(c)}^{\infty} f(u,c) du\ dc}{\int_{-\infty}^{\infty} \int_{g^{-1}(c)}^{\infty} f(u,c) du\ dc}$ and $E\left[c_i \mid i \text{ is complier}\right]=\frac{\int_{-\infty}^{\infty} c \int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c) du\ dc}{\int_{-\infty}^{\infty} \int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c) du\ dc}$.

The proof follows the same steps and is symmetric to the proof of *Claim 3*.

First, we write $E\left[c_i \mid i \text{ is complier}\right]=\int_{-\infty}^{\infty} ch(c)dc$ and $E\left[c_i \mid i \text{ is always-taker}\right]=\int_{-\infty}^{\infty} ck(c)dc$.

We then show that H first-order stochastically dominates K, using *Theorem 2*, which concludes.

Claim 6

If *Condition ID* and *R1* hold, then $E\left[v\left(u_i\right) \mid i \text{ is complier}\right]<E\left[v\left(u_i\right) \mid i \text{ is always-taker}\right]$

Proof of Claim 6

$E\left[v\left(u_i\right) \mid i \text{ is always-taker}\right]=\frac{\int_{-\infty}^{\infty} v(u) \int_{-\infty}^{g(u)} f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} f(u,c) dc\ du}=\int_{-\infty}^{\infty} v(u)k(u)du$

and $E\left[v\left(u_i\right) \mid i \text{ is complier}\right]=\frac{\int_{-\infty}^{\infty} v(u) \int_{g(u)}^{g(u)/\lambda} f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c) dc\ du}=\int_{-\infty}^{\infty} v(u)h(u)du,$

where $h(u)$ and $k(u)$ are defined as in the proof of *Claim 2*:

$h(u)=\frac{\int_{g(u)}^{g(u)/\lambda} f(u,c) dc}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c) dc\ du}$ and $k(u)=\frac{\int_{-\infty}^{g(u)} f(u,c) dc}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} f(u,c) dc\ du}$.

$v(u)=1-F_{\varepsilon}(-u)$ is increasing since $F_{\varepsilon}(.)$ is increasing.

Since K first-order stochastically dominates H (as proved in the proof of *Claim 2*), we thus get $\int_{-\infty}^{\infty} v(u)h(u)du \leq \int_{-\infty}^{\infty} v(u)k(u)du$, ie $E\left[v\left(u_i\right) \mid i \text{ is complier}\right]<E\left[v\left(u_i\right) \mid i \text{ is always-taker}\right]$.

Claim 7

If *Condition ID* and *R1* hold, then $E\left[u_i \mid i \text{ is complier}, i \text{ votes}\right]<E\left[u_i \mid i \text{ is always-taker}, i \text{ votes}\right]$

Proof of Claim 7

$E\left[u_i \mid i \text{ is always-taker}, i \text{ votes}\right]=\frac{\int_{-\infty}^{\infty} u \int_{-\infty}^{g(u)} v(u) f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} v(u) f(u,c) dc\ du}$ and

$E\left[u_i \mid i \text{ is complier}, i \text{ votes}\right]=\frac{\int_{-\infty}^{\infty} u \int_{g(u)}^{g(u)/\lambda} v(u) f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} v(u) f(u,c) dc\ du}$.

We write $w(u,c) \equiv v(u) f(u,c)$. Since$-f(u,c)$ satisifies log-increasing differences in u and c, $-w(u,c)$ satisfies log-increasing differences as well : $\frac{w\left(u',c'\right)}{w\left(u',c\right)}=\frac{f\left(u',c'\right)}{f\left(u',c\right)}<\frac{f\left(u,c'\right)}{f(u,c)}=\frac{w\left(u,c'\right)}{w(u,c)}$ for any $u'>u$ and $c'>c$.

Therefore, substituting $w(u,c)$ to $f(u,c)$ in the proof of *Claim 2*, we get $\frac{\int_{-\infty}^{\infty} u \int_{g(u)}^{g(u)/\lambda} v(u) f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} v(u) f(u,c) dc\ du}<\frac{\int_{-\infty}^{\infty} u \int_{-\infty}^{g(u)} v(u) f(u,c) dc\ du}{\int_{-\infty}^{\infty} \int_{-\infty}^{g(u)} v(u) f(u,c) dc\ du}$, ie $E\left[u_i \mid i \text{ is complier}, i \text{ votes}\right]<E\left[u_i \mid i \text{ is always-taker}, i \text{ votes}\right]$.

Claim 8

If *Condition ID* and *R1* hold, a more efficient intervention, characterized by $\lambda'<\lambda<1$, selects additional compliers characterized by a lower second stage utility, a higher registration cost, a lower turnout and a lower second stage utility conditional on voting than those selected by the less efficient intervention.

Proof of Claim 8

The more efficient intervention selects additional compliers characterized by ${}^{g(u_i)}/\lambda \leq c_i < {}^{g(u_i)}/\lambda'$.

The additional compliers have a lower second stage utility: $\frac{\int_{-\infty}^{\infty} u \int_{g(u)/\lambda}^{g(u)/\lambda'} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{u/\lambda}^{u/\lambda'} f(u,c)dc\ du} < \frac{\int_{-\infty}^{\infty} u \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$.
The proof is derived in the same way as the proof of *Claim 3*.

They have a higher registration cost: $\frac{\int_{-\infty}^{\infty} c \int_{g^{-1}(\lambda' c)}^{g^{-1}(\lambda c)} f(u,c)du\ dc}{\int_{-\infty}^{\infty} \int_{g^{-1}(\lambda' c)}^{g^{-1}(\lambda c)} f(u,c)du\ dc} > \frac{\int_{-\infty}^{\infty} c \int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c)du\ dc}{\int_{-\infty}^{\infty} \int_{g^{-1}(\lambda c)}^{g^{-1}(c)} f(u,c)du\ dc}$. The proof is derived in the same way as the proof of *Claim 5*.

They have a lower turnout: $\frac{\int_{-\infty}^{\infty} v(u) \int_{g(u)/\lambda}^{g(u)/\lambda'} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)/\lambda}^{g(u)/\lambda'} f(u,c)dc\ du} < \frac{\int_{-\infty}^{\infty} v(u) \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$. The proof is derived in the same way as the proof of *Claim 6*.

Finally, they have a lower second stage utility conditional on voting:
$\frac{\int_{-\infty}^{\infty} u \int_{g(u)/\lambda}^{g(u)/\lambda'} v(u)f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)/\lambda}^{g(u)/\lambda'} v(u)f(u,c)dc\ du} < \frac{\int_{-\infty}^{\infty} u \int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc\ du}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc\ du}$. The proof is derived in the same way as the proof of *Claim 7*.

Claim 9

For a given share of compliers and unchanged conditional densities $f\left(c \mid u\right)$, if *Conditions ID, R1 and R2* hold, an increase in the compliers' likelihood to vote, generated by an increase in the relative number of compliers with a higher u, is concomitant to an increase in their degree of politicization, registration cost and degree of politicization conditional on voting.

Proof of Claim 9

We construct a new density $f_2\left(u,c\right)$ based on the density $f\left(u,c\right)$ and such that, among compliers characterized by a given u, the shape of the conditional density of C given $U = u$ is unchanged: for any u and any $(c,c') \in [u, {}^{u}/\lambda]^2$: $\frac{f_2\left(c'|u\right)}{f_2(c|u)} = \frac{f\left(c'|u\right)}{f(c|u)}$.

$\frac{f_2\left(c'|u\right)}{f_2(c|u)} = \frac{f\left(c'|u\right)}{f(c|u)}$ for any u and any $(c,c') \in [u, {}^{u}/\lambda]^2$ is equivalent to $\frac{f_2\left(u,c'\right)}{f_2(u,c)} = \frac{f\left(u,c'\right)}{f(u,c)}$ for any u and any $(c,c') \in [u, {}^{u}/\lambda]^2$.

This requires $f_2\left(u,c\right) = f\left(u,c\right)h(u)$ for any u and $c \in [g(u), {}^{g(u)}/\lambda]$, for some function $h(u)$ positive.

For u and c such that $c \notin [u, {}^{u}/\lambda]$, we set $f_2\left(u,c\right) = f\left(u,c\right)$ otherwise.

h must be positive, so that $f_2\left(u,c\right) \geq 0$ for any u and c, a condition to qualify as a density.

Further, h must satisfy $\int_{-\infty}^{\infty} h(u) \left(\int_{g(u)}^{g(u)/\lambda} f\left(u,c\right)dc\right) du = \int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f\left(u,c\right)dc\ du$ for the fraction of compliers to be unchanged.

This also satisfies the second condition for $f_2\left(u,c\right)$ to qualify as a density: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_2\left(u,c\right)dc\ du = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f\left(u,c\right)dc\ du + \int_{-\infty}^{\infty} h(u) \left(\int_{g(u)}^{g(u)/\lambda} f\left(u,c\right)dc\right) du - \int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f\left(u,c\right)dc\ du = 1$

Finally, to obtain that the expected participation of the compliers is higher under the joint density $f_2\left(u,c\right)$ than the joint density $f\left(u,c\right)$, we impose that $h(u)$ be increasing as a sufficient (but not necessary) condition. We show below that this condition is indeed sufficient:

$E_{f_2}\left[v\left(u_i\right) \mid i \text{ is complier}\right] = \int_{-\infty}^{\infty} v(u) \frac{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc\ du} du = \int_{-\infty}^{\infty} v(u)k(u)du$ and $E_f\left[v\left(u_i\right) \mid i \text{ is complier}\right] = \int_{-\infty}^{\infty} v(u) \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty} \int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du} du = \int_{-\infty}^{\infty} v(u)l(u)du$

With h increasing in u, we get that for any $u' > u$, $\frac{k(u')}{k(u)} = \frac{h(u')}{h(u)} \frac{\int_{g(u')}^{g(u')/\lambda} f(u',c)dc}{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc} = \frac{h(u')}{h(u)} \frac{l(u')}{l(u)} \geq \frac{l(u')}{l(u)}$

Applying *Theorem 2* to the density functions k and l, we thus get that K first-order stochastically dominates L, and $E_{f_2}\left[v\left(u_i\right) \mid i \text{ is complier}\right] \geq E_f\left[v\left(u_i\right) \mid i \text{ is complier}\right]$

Now, for any $h(u)$ satisfying the conditions listed above, we sign the difference between the compliers' second stage utility, second stage utility conditional on voting and registration cost when the joint density is $f\left(u,c\right)$ or $f_2\left(u,c\right)$.

First, $E_{f_2}\left[u_i \mid i \text{ is complier}\right] = \int_{-\infty}^{\infty} u \frac{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc\ du} du \geq \int_{-\infty}^{\infty} u \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du} du = E_f\left[u_i \mid i \text{ is complier}\right]$ comes directly from the fact that K first-order stochastically dominates L.

Second, $E_{f_2}\left[u_i \mid i \text{ is complier}, i \text{ votes}\right] = \int_{-\infty}^{\infty} u \frac{\int_{g(u)}^{g(u)/\lambda} v(u)f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} v(u)f_2(u,c)dc\ du} du \geq \int_{-\infty}^{\infty} u \frac{\int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc\ du} du = E_f\left[u_i \mid i \text{ is complier}, i \text{ votes}\right]$.

The proof of this is identical to the proof above, rewriting $k(u) = \frac{\int_{g(u)}^{g(u)/\lambda} v(u)f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} v(u)f_2(u,c)dc\ du}$ and $l(u) = \frac{\int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} v(u)f(u,c)dc\ du}$.

Third, $E_{f_2}\left[c_i \mid i \text{ is complier}\right] \geq E_f\left[c_i \mid i \text{ is complier}\right]$.

These two objects can be written as:

$E_f\left[c_i \mid i \text{ is complier}\right] = \int_{-\infty}^{\infty} E_f\left[c_i \mid i \text{ is complier}, u_i = u\right] \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du} du$

and $E_{f_2}\left[c_i \mid i \text{ is complier}\right] = \int_{-\infty}^{\infty} E_{f_2}\left[c_i \mid i \text{ is complier}, u_i = u\right] \frac{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc\ du} du$.

But $E_{f_2}\left[c_i \mid i \text{ is complier}, u_i = u\right] = \frac{\int_{g(u)}^{g(u)/\lambda} cf_2(u,c)dc}{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc} = \frac{h(u)}{h(u)} \frac{\int_{g(u)}^{g(u)/\lambda} cf(u,c)dc}{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc} = E_f\left[c_i \mid i \text{ is complier}, u_i = u\right]$.

Moreover, $z(u) \equiv E_f\left[c_i \mid i \text{ is complier}, u_i = u\right]$ is increasing in u by assumption (*Condition R2*).

Writing again, $l(u) = \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$ and $k(u) = \frac{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc\ du}$, we have $\frac{k(u')}{k(u)} \geq \frac{l(u')}{l(u)}$ for any $u' > u$: K dominates L in the MLR sense. Thus, by *Theorem 2*, K first-order stochastically dominates L. Since $z(u)$ is increasing in u, we get $\int_{-\infty}^{\infty} z(u) \frac{\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f_2(u,c)dc\ du} du \geq \int_{-\infty}^{\infty} z(u) \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du} du$

Q.E.D.

Claim 10

All previous claims hold in the extended version of the model, where a registered citizen's actual propensity to vote is $w(u_i)$, with $w(u') - v(u') \leq w(u) - v(u)$ and $w(u') \geq w(u)$ for any $u' \geq u$.

Proof of Claim 10

Claims 1, 3, 4, 5 and *8* are unaffected, since the selection process of compliers and always-takers is unchanged: we assume that at the registration stage, individual i still anticipates that she will vote if $u_i + \varepsilon_i \geq 0$.

The proofs of *Claims 6, 7* and *9* can be redone, substituting $w(u)$ to $v(u)$. They rely on relations of first-order stochastic dominance between distribution functions. and thus hold for any weakly increasing function of u, be it v or w.

Claim 11

The difference between compliers and always-takers' turnout is lower if the propensity to vote of a registered citizen with utility u is given by $w(u)$ rather than $v(u)$.

Proof of Claim 11

Claim 11 can be restated as
$\int_{-\infty}^{\infty} w(u)k(u)du - \int_{-\infty}^{\infty} w(u)h(u)du \leq \int_{-\infty}^{\infty} v(u)k(u)du - \int_{-\infty}^{\infty} v(u)h(u)du$
$\Leftrightarrow \int_{-\infty}^{\infty} [v(u) - w(u)]\, k(u)du \geq \int_{-\infty}^{\infty} [v(u) - w(u)]\, h(u)du$
where, as before, $h(u) = \frac{\int_{g(u)}^{g(u)/\lambda} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{g(u)}^{g(u)/\lambda} f(u,c)dc\ du}$ and $k(u) = \frac{\int_{-\infty}^{g(u)} f(u,c)dc}{\int_{-\infty}^{\infty}\int_{-\infty}^{g(u)} f(u,c)dc\ du}$.
The proof comes immediately from the fact that K first-order stochastically dominates H and $v(u) - w(u)$ increases in u (by assumption).

Claim 12

The difference between the propensity to vote of compliers and always-takers can be written as the sum of two terms. The first one, negative, and predominant, is bigger, the bigger the difference between the degree of politicization of the compliers and always-takers. The second one, positive, comes from the fact that the mobilization effect of the campaign is lower for citizens with a higher degree of politicization.

Proof of Claim 12

We can decompose $\int_{-\infty}^{\infty} w(u)k(u)du - \int_{-\infty}^{\infty} w(u)h(u)du = \left[\int_{-\infty}^{\infty} v(u)k(u)du - \int_{-\infty}^{\infty} v(u)h(u)du\right] +$
$\left[\int_{-\infty}^{\infty} w(u)k(u)du - \int_{-\infty}^{\infty} w(u)h(u)du - \left(\int_{-\infty}^{\infty} v(u)k(u)du - \int_{-\infty}^{\infty} v(u)h(u)du\right)\right]$
The first term is negative, as shown in *Claim 6*, and bigger, the bigger $\int_{-\infty}^{\infty} uk(u)du \geq \int_{-\infty}^{\infty} uh(u)du$, as can directly be inferred from *Claim 9*. The second term is negative, as shown in *Claim 11*.

Microfounding the assumption that $w(u) - v(u)$ decreases with u

We discuss how this assumption can be grounded in a more fundamental assumption about the way the campaign affects the perceived benefits of voting u_i: suppose that a registered citizen votes if $m(u_i)+\varepsilon_i \geq 0$. Then, the propensity to vote of an individual with politicization u is $w(u) = v\,(m(u))$. Under what condition on m do we have $z(u) \equiv w(u) - v(u)$ decrease in u?
Since $v(u) = 1 - F_\varepsilon\,(-u)$, we have $v'(u) = f_\varepsilon(-u)$. Therefore,
$z'(u) \leq 0 \Leftrightarrow v'\,(m(u))\, m'(u) - v'(u) \leq 0 \Leftrightarrow m'(u) \leq \frac{f_\varepsilon(-u)}{f_\varepsilon(-m(u))}$
If f_ε is increasing on $]-\infty, 0]$ (a condition fulfilled by many usual density functions, including the normal density), this condition is satisfied for any $u \geq 0$ by any function m such that $m(u) \geq u$ and $x(u) \equiv \frac{m(u)}{u}$ decreases: the mobilization increases each citizen's perceived benefits of voting, but less so for citizens with a higher u. Indeed, then, we have $x'(u) \leq 0$ and $m(u) \geq u \Rightarrow m'(u) \leq \frac{m(u)}{u} \leq 1$ and $m(u) \geq u$ and f_ε increasing on $]-\infty, 0] \Rightarrow 1 \leq \frac{f_\varepsilon(-u)}{f_\varepsilon(-m(u))}$.

Claim 13

If *Conditions* $ID_{\bar{k}}$, $R1_{\bar{k}}$ and $R2_{\bar{k}}$ hold for any $\bar{k}$, all results established for unregistered citizens hold for misregistered citizens facing an additional cost $\bar{k}$ of voting at their previous address, for any $\bar{k}$.

Proof of Claim 13

We first prove that $g_{\bar{k}}(u)$ is strictly increasing in u for any $\bar{k}$.
$g_{\bar{k}}(u) = g\left(u\right) - g\left(u - \bar{k}\right) = \int_{-u}^{\infty} \left(u + \varepsilon\right) f_{\varepsilon}(\varepsilon) d\varepsilon - \int_{-u+\bar{k}}^{\infty} \left(u - \bar{k} + \varepsilon\right) f_{\varepsilon}(\varepsilon) d\varepsilon = \int_{-u}^{-u+\bar{k}} \left(u + \varepsilon\right) f_{\varepsilon}(\varepsilon) d\varepsilon + \bar{k} \int_{-u+\bar{k}}^{\infty} f_{\varepsilon}(\varepsilon) d\varepsilon$.
$g'_{\bar{k}}(u) = -\bar{k} f_{\varepsilon}\left(-u + \bar{k}\right) + \int_{-u}^{-u+\bar{k}} f_{\varepsilon}(\varepsilon) d\varepsilon + \bar{k} f_{\varepsilon}\left(-u + \bar{k}\right) = \int_{-u}^{-u+\bar{k}} f_{\varepsilon}(\varepsilon) d\varepsilon > 0$: $g_{\bar{k}}(u)$ is strictly increasing in u for any $\bar{k}$.
Considering any $\bar{k}$, since $g_{\bar{k}}(u)$ is strictly increasing, it can be substituted to $g(u)$ in the proofs above.

Claim 14

$P(V_i = 1 \mid \text{door-to-door}) - P(V_i = 1 \mid \text{control})$
$= \left[E\left[w\left(u_i\right) \mid i \text{ is canvassing complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A}$
, where p_C is the proportion of compliers selected by door-to-door canvassing.

Proof of Claim 14

$P(V_i = 1 \mid \text{door-to-door}) = \frac{p_C}{p_C+p_A} E\left[w\left(u_i\right) \mid i \text{ is canvassing complier}\right] + \frac{p_A}{p_C+p_A} E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]$ and $P(V_i = 1 \mid \text{control}) = E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]$.
Thus, $P(V_i = 1 \mid \text{door-to-door}) - P(V_i = 1 \mid \text{control})$
$= \left[E\left[w\left(u_i\right) \mid i \text{ is canvassing complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A}$.

Claim 15

$P(V_i = 1 \mid \text{home registration}) - P(V_i = 1 \mid \text{control})$
$= \left[E\left[w\left(u_i\right) \mid i \text{ is home registration complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A} +$
$p \times E\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right]$
,where p_C is the proportion of compliers selected by home registration and
$p = P(i \text{ gets registered at home} \mid i \text{ registers, home registration})$.

Proof of Claim 15

$P(V_i = 1 \mid \text{home registration}) = pP(V_i = 1 \mid \text{home registration, } i \text{ gets registered at home}) +$
$(1-p)P(V_i = 1 \mid \text{home registration, } i \text{ gets registered at town hall})$.
Thus,
$P(V_i = 1 \mid \text{home registration}) = pE\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right] +$
$P(V_i = 1 \mid \text{home registration, } i \text{ gets registered at town hall})$.
As in the proof of *Claim 13*,
$P(V_i = 1 \mid \text{home registration, } i \text{ gets registered at town hall}) - P(V_i = 1 \mid \text{control}) =$
$\left[E\left[w\left(u_i\right) \mid i \text{ is home registration complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A}$, where p_C is now the proportion of compliers selected by home registration. Thus,
$P(V_i = 1 \mid \text{home registration}) - P(V_i = 1 \mid \text{control})$
$= \left[E\left[w\left(u_i\right) \mid i \text{ is home registration complier}\right] - E\left[w\left(u_i\right) \mid i \text{ is always-taker}\right]\right] \times \frac{p_C/p_A}{1+p_C/p_A} +$
$p \times E\left[\widetilde{w}(u_i) - w(u_i) \mid i \text{ gets registered at home}\right]$

Table 1: Verifying randomization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A.	Building characteristics							
	Targetable address		Number of mailboxes		Number of additional names		Housing price	
Any treatment	0.003		-0.166		-0.26		47.1	
	(0.014)		(0.378)		(0.365)		(66.2)	
Early Canvassing		0.004		-0.103		-0.312		22.1
		(0.020)		(0.566)		(0.547)		(99.1)
Late Canvassing		0.013		-0.167		-0.37		15.4
		(0.020)		(0.566)		(0.547)		(101.7)
Early Home registration		0		0.051		0.054		60.6
		(0.020)		(0.568)		(0.548)		(97.7)
Late Home registration		0.003		-0.131		-0.199		50.5
		(0.020)		(0.565)		(0.546)		(99.1)
Early Canvassing + Late Home registration		-0.003		-0.275		-0.329		34.9
		(0.020)		(0.565)		(0.545)		(97.2)
Early Home registration + Late Home registration		0.001		-0.368		-0.4		97.4
		(0.020)		(0.567)		(0.547)		(99.4)
Constant	0.827	0.827	7.942	7.942	6.803	6.803	3103.4	3103.4
	(0.012)***	(0.012)***	(0.327)***	(0.327)***	(0.316)***	(0.316)***	(57.5)***	(57.6)***
Wald test: Joint significance of the six treatment dummies								
Test statistic		0.1		0.11		0.2		0.2
p-value		0.996		0.995		0.976		0.978
Observations	4118	4118	4118	4118	4118	4118	941	941
R-squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Panel A takes the building as the unit of observation. Panel B takes the individual as the unit of observation and includes all previously registered citizens.*
In the regressions reported in odd-number columns, we measure differences in the baseline characteristics between the control group and the treatment groups taken altogether. In the regressions reported in the even-number columns, we measure differences in the baseline characteristics between the control group and each treatment group. We run joint T tests of the joint significance of the six treatment dummies.

Table 1 (continued): Verifying randomization

Panel B.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Individual characteristics									
	Gender		Age		Born in the city of residence		Born in another city in the department		Borth in another department in the region	
Any treatment	-0.006		0.101		0.009		-0.004		-0.003	
	(0.008)		(0.368)		(0.012)		(0.011)		(0.007)	
Early Canvassing		-0.001		1.315		-0.008		0.01		0
		(0.012)		(0.625)**		(0.020)		(0.018)		(0.012)
Late Canvassing		-0.023		-0.351		0.016		-0.01		-0.011
		(0.011)**		(0.539)		(0.019)		(0.016)		(0.011)
Early Home registration		-0.006		-0.114		0.016		-0.011		-0.006
		(0.013)		(0.541)		(0.015)		(0.015)		(0.010)
Late Home registration		-0.014		0.513		0.017		0		0.001
		(0.014)		(0.630)		(0.020)		(0.018)		(0.011)
Early Canvassing + Late Home		0.005		-0.172		0.007		-0.012		-0.002
registration		(0.012)		(0.567)		(0.019)		(0.016)		(0.011)
Early Home registration + Late		0.003		-0.644		0.009		0.001		0.001
Home registration		(0.010)		(0.536)		(0.018)		(0.015)		(0.011)
Constant	0.457	0.457	44.771	44.771	0.195	0.195	0.165	0.165	0.134	0.134
	(0.007)***	(0.007)***	(0.374)***	(0.374)***	(0.014)***	(0.014)***	(0.012)***	(0.012)***	(0.009)***	(0.009)***
Wald test: Joint significance of the six treatment dummies										
Test statistic		1.44		1.6		0.43		0.42		0.3
p-value		0.197		0.143		0.856		0.863		0.937
Observations	17201	17201	32234	32234	28175	28175	28175	28175	28175	28175
R-squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Born in another region		Born abroad		Born in a city		Size of the population in home city	
Any treatment	-0.017		0.014		-0.004		7786	
	(0.010)*		(0.009)		(0.003)		########	
Early Canvassing		-0.006		0.002		-0.011		21863
		(0.014)		(0.013)		(0.006)*		########
Late Canvassing		-0.029		0.033		-0.004		6644
		(0.015)*		(0.014)**		(0.005)		########
Early Home registration		-0.025		0.026		-0.007		-5561
		(0.014)*		(0.013)**		(0.005)		########
Late Home registration		-0.019		0.001		-0.002		9956
		(0.015)		(0.013)		(0.005)		########
Early Canvassing + Late Home		-0.002		0.009		0		3101
registration		(0.015)		(0.013)		(0.005)		########
Early Home registration + Late		-0.023		0.012		0.004		10334
Home registration		(0.015)		(0.014)		(0.005)		########
Constant	0.275	0.275	0.232	0.232	0.958	0.958	293451	293451
	(0.011)***	(0.011)***	(0.009)***	(0.009)***	(0.003)***	(0.003)***	(15,830)***	(15,831)***
Wald test: Joint significance of the six treatment dummies								
Test statistic		1.19		1.47		1.31		0.18
p-value		0.31		0.186		0.251		0.982
Observations	28175	28175	28175	28175	21270	21270	21270	21270
R-squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Panel A takes the building as the unit of observation. Panel B takes the individual as the unit of observation and includes all previsously registered citizens.*
In the regressions reported in odd-number columns, we measure differences in the baseline characteristics between the control group and the treatment groups taken altogether. In the regressions reported in the even-number columns, we measure differences in the baseline characteristics between the control group and each treatment group. We run joint T tests of the joint significance of the six treatment dummies.

Table 2: Impact of the interventions on the gross number of new registrations

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A.	At the building level					
Any treatment	0.31	0.3	0.36			
	(0.05)***	(0.05)***	(0.06)***			
Early Canvassing				0.15	0.12	0.14
				(0.08)*	(0.08)*	(0.090)
Late Canvassing				0.25	0.21	0.28
				(0.08)***	(0.08)***	(0.09)***
Early Home registration				0.27	0.27	0.32
				(0.08)***	(0.08)***	(0.09)***
Late Home registration				0.29	0.3	0.35
				(0.08)***	(0.08)***	(0.09)***
Early Canvassing + Late Home				0.35	0.34	0.43
registration				(0.08)***	(0.08)***	(0.09)***
Early Home registration + Late				0.54	0.53	0.62
Home registration				(0.08)***	(0.08)***	(0.09)***
Number of extra names	0.16	0.13	0.12	0.16	0.13	0.12
	(0.00)***	(0.00)***	(0.01)***	(0.00)***	(0.00)***	(0.01)***
Constant	1.13	0.96	1.17	1.13	0.93	1.16
	(0.05)***	(0.47)**	(0.62)*	(0.05)***	(0.46)**	(0.62)*
Strata fixed effects		x	x		x	x
Building controls			x			x
Observations	4118	4105	3344	4118	4105	3344
R-squared	0.53	0.61	0.61	0.53	0.61	0.61
Mean in Control Group	1.13	1.13	1.13	1.13	1.13	1.13
Panel B.	At the apartment level					
Any treatment	0.05	0.05	0.048			
	(0.009)***	(0.009)***	(0.009)***			
Early Canvassing				0.019	0.017	0.013
				(0.01)	(0.01)	(0.01)
Late Canvassing				0.038	0.034	0.034
				(0.013)***	(0.012)***	(0.012)***
Early Home registration				0.036	0.033	0.032
				(0.015)**	(0.014)**	(0.015)**
Late Home registration				0.05	0.054	0.056
				(0.015)***	(0.014)***	(0.013)***
Early Canvassing + Late Home				0.062	0.063	0.063
registration				(0.014)***	(0.012)***	(0.012)***
Early Home registration + Late				0.096	0.098	0.092
Home registration				(0.015)***	(0.014)***	(0.014)***
Number of extra names	0.06	0.075	0.069	0.06	0.075	0.069
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Constant	0.168	0.148	0.244	0.168	0.145	0.243
	(0.008)***	-0.1	(0.123)**	(0.008)***	-0.098	(0.118)**
Strata fixed effects		x	x		x	x
Building controls			x			x
Observations	20502	20458	19461	20502	20458	19461
R-squared	0.01	0.03	0.03	0.01	0.03	0.03
Mean in Control Group	0.17	0.17	0.17	0.17	0.17	0.17

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. Panel A takes the building as the unit of observation and includes all newly registered citizens in the addresses in the sample. Panel B takes the apartment as the unit of observation and includes all newly registered citizens in the targeted apartments.*

In all regressions, we control for the number of last names found on mailboxes that did not appear on the 2011 voters' lists, as a proxy for the initial number of unregistered and misregistered citizens.

Table 3: Relative impact of a Late vs. Early visit and a Home registration vs. Canvassing visit on the gross number of new registrations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any treatment	0.05	0.048						
	(0.009)***	(0.009)***						
Early Canvassing			0.017	0.013				
			(0.013)	(0.013)				
Late Canvassing			0.034	0.034				
			(0.012)***	(0.012)***				
Early Home registration			0.033	0.032				
			(0.014)**	(0.015)**				
Late Home registration			0.054	0.056				
			(0.014)***	(0.013)***				
Early Canvassing + Late Home			0.063	0.063				
registration			(0.012)***	(0.012)***				
Early Home registration + Late			0.098	0.092				
Home registration			(0.014)***	(0.014)***				
Early visit					0.025	0.023		
					(0.011)**	(0.011)**		
Late visit					0.044	0.045		
					(0.011)***	(0.011)***		
Canvassing visit							0.025	0.023
							(0.010)**	(0.011)**
Home registration visit							0.044	0.044
							(0.011)***	(0.011)***
Two visits					0.08	0.078	0.08	0.077
					(0.011)***	(0.011)***	(0.011)***	(0.011)***
Number of extra names	0.075	0.069	0.075	0.069	0.075	0.069	0.075	0.069
	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Constant	0.148	0.244	0.145	0.243	0.143	0.237	0.143	0.237
	(0.100)	(0.123)**	(0.098)	(0.118)**	(0.096)	(0.117)**	(0.096)	(0.117)**
Strata fixed effects	x	x	x	x	x	x	x	x
Building controls		x		x		x		x
Wald tests:								
Late visit = Early visit								
Test statistic					3.28	4.92		
p-value					0.071*	0.027**		
Home registration visit = Canvassing visit								
Test statistic							2.74	3.46
p-value							0.098*	0.063*
Two visits = Home registration visit								
Test statistic							9.97	8.35
p-value							0.002***	0.004***
Observations	20458	19461	20458	19461	20458	19461	20458	19461
R-squared	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Mean in Control Group	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the apartment as the unit of observation and include all newly registered citizens in the targeted apartments.*
Columns 1 to 4 are identical to columns 2, 3, 5 and 6 of Table 2, Panel B, and included for reference. In columns 5 and 6, we measure the specific contribution of visiting the addresses Late vs. Early in explaining the relative impact of our different interventions. In columns 7 and 8, we measure the relative contribution of the Home registration vs. simple Canvassing visits.
In all regressions, we control for the number of last names found on mailboxes that did not appear on the 2011 voters' lists, as a proxy for the initial number of unregistered and misregistered citizens

Table 4: Relative impact of a single vs. two visits on the gross number of new registrations

	(1)	(2)	(3)	(4)	(5)	(6)
Two visits	0.029	0.021	0.01			
	(0.029)	(0.028)	(0.028)			
Early Canvassing + Late Home registration				0.02	0.012	0.01
				(0.034)	(0.033)	(0.033)
Early Home registration + Late Home registration				0.038	0.031	0.011
				(0.035)	(0.034)	(0.033)
Number of extra names	0.063	0.065	0.071	0.063	0.065	0.071
	(0.021)***	(0.023)***	(0.024)***	(0.021)***	(0.023)***	(0.024)***
Constant	0.3	0.285	0.395	0.3	0.284	0.395
	(0.024)***	(0.058)***	(0.094)***	(0.024)***	(0.059)***	(0.094)***
City fixed effects		x	x		x	x
Building controls			x			x
Wald test: Early Canvassing + Late Home registration = Early Home registration + Late Home registration						
Test statistic				0.22	0.29	0
p-value				0.636	0.591	0.974
Observations	2171	2171	2066	2171	2171	2066
R-squared	0	0.01	0.02	0	0.01	0.02
Mean in Treatment Group "Late Home registration"	0.3	0.3	0.3	0.3	0.3	0.3

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the apartment as the unit of observation and include all newly registered citizens in the targeted apartments of the treatment groups "Late Home registration", "Early Canvassing + Late Home registration" and "Early Home registration + Late Home registration" which opened their door at the second visit. Columns 1 to 3 measure the relative impact of a single vs. two visits, and columns 4 to 6 allow this impact to differ depending on the content of the first visit: Canvassing or Home registration.*

In all regressions, we control for the number of last names found on mailboxes that did not appear on the 2011 voters' lists, as a proxy for the initial number of unregistered and misregistered citizens.

Table 5: Impact of the interventions on the NET number of new registrations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Any treatment	0.064	0.065	0.062							
	(0.010)***	(0.010)***	(0.010)***							
Early Canvassing				0.03	0.029	0.026				
				(0.015)**	(0.015)*	(0.015)*				
Late Canvassing				0.049	0.047	0.043				
				(0.015)***	(0.014)***	(0.014)***				
Early Home registration				0.064	0.062	0.062				
				(0.017)***	(0.016)***	(0.016)***				
Late Home registration				0.062	0.067	0.063				
				(0.017)***	(0.016)***	(0.016)***				
Early Canvassing + Late Home registration				0.071	0.075	0.071				
				(0.014)***	(0.014)***	(0.014)***				
Early Home registration + Late Home registration				0.108	0.11	0.106				
				(0.016)***	(0.016)***	(0.016)***				
Early visit							0.046	0.044		
							(0.013)***	(0.013)***		
Late visit							0.057	0.053		
							(0.012)***	(0.012)***		
Canvassing visit									0.038	0.035
									(0.012)***	(0.012)***
Home registration visit									0.065	0.063
									(0.013)***	(0.013)***
Two visits							0.092	0.089	0.092	0.089
							(0.012)***	(0.012)***	(0.012)***	(0.012)***
Number of extra names	0.078	0.072	0.069	0.078	0.071	0.069	0.072	0.069	0.072	0.069
	(0.007)***	(0.008)***	(0.008)***	(0.007)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***
Constant	0.094	0.053	0.143	0.094	0.051	0.142	0.048	0.138	0.048	0.139
	(0.009)***	(0.140)	(0.139)	(0.009)***	(0.141)	(0.139)	(0.139)	(0.137)	(0.139)	(0.136)
Strata fixed effects		x	x		x	x	x	x	x	x
Building controls			x			x		x		x
Wald tests:										
Late visit = Early visit										
Test statistic							0.89	0.6		
p-value							0.346	0.44		
Home registration visit = Canvassing visit										
Test statistic									4.59	5.18
p-value									0.032**	0.023**
Two visits = Home registration visit										
Test statistic									4.95	4.38
p-value									0.026**	0.037**
Observations	20502	20458	20458	20502	20458	20458	20458	20458	20458	20458
R-squared	0.01	0.03	0.03	0.01	0.03	0.03	0.03	0.03	0.03	0.03
Mean in Control Group	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the apartment as the unit of observation and include all newly registered citizens in the targeted apartments. This table is similar to Table 3, replacing the gross number of new registrations by the net number of new registrations as the outcome of interest.*

In columns 7 and 8, we measure the specific contribution of visiting the addresses Late vs. Early in explaining the relative impact of our different interventions. In columns 9 and 10, we measure the relative contribution of the Home registration vs. simple Canvassing visits.

In all regressions, we control for the number of last names found on mailboxes that did not appear on the 2011 voters' lists, as a proxy for the initial number of unregistered and misregistered citizens.

Table 6: Impact of the interventions on the number of votes cast by newly registered citizens

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Presidential Elections, 1st round			Presidential Elections, 2nd round			General Elections, 1st round			General Elections, 2nd round			All 4 rounds		
Any treatment	0.042	0.041		0.040	0.039		0.018	0.013		0.021	0.020		0.030	0.028	
	(0.008)***	(0.008)***		(0.008)***	(0.008)***		(0.006)***	(0.006)**		(0.006)***	(0.005)***		(0.007)***	(0.006)***	
Early Canvassing			0.012			0.012			(0.003)			0.009			0.007
			(0.012)			(0.012)			(0.008)			(0.008)			(0.009)
Late Canvassing			0.028			0.026			0.013			0.016			0.021
			(0.012)**			(0.012)**			(0.010)			(0.009)*			(0.010)**
Early Home registration			0.027			0.019			0.012			0.017			0.019
			(0.014)*			(0.013)			(0.010)			(0.008)**			(0.010)*
Late Home registration			0.046			0.045			0.012			0.015			0.029
			(0.012)***			(0.012)***			(0.009)			(0.008)*			(0.009)***
Early Canvassing + Late Home			0.053			0.054			0.024			0.026			0.039
registration			(0.012)***			(0.012)***			(0.009)***			(0.008)***			(0.009)***
Early Home registration + Late			0.078			0.079			0.022			0.039			0.054
Home registration			(0.013)***			(0.013)***			(0.009)**			(0.010)***			(0.010)***
Number of extra names	0.055	0.062	0.062	0.055	0.064	0.064	0.035	0.038	0.038	0.030	0.033	0.033	0.044	0.049	0.049
	(0.006)***	(0.007)***	(0.007)***	(0.006)***	(0.007)***	(0.007)***	(0.005)***	(0.005)***	(0.005)***	(0.004)***	(0.005)***	(0.005)***	(0.005)***	(0.005)***	(0.005)***
Constant	0.148	0.247	0.247	0.151	0.241	0.239	0.090	0.240	0.239	0.082	0.246	0.245	0.118	0.244	0.242
	(0.007)***	(0.120)**	(0.117)**	(0.007)***	(0.121)**	(0.117)**	(0.005)***	(0.114)**	(0.112)**	(0.005)***	(0.115)**	(0.114)**	(0.006)***	(0.117)**	(0.114)**
Strata fixed effects		x	x		x	x		x	x		x	x		x	x
Building controls		x	x		x	x		x	x		x	x		x	x
Observations	20502	19461	19461	20502	19461	19461	20502	19461	19461	20502	19461	19461	82008	77844	77844
R-squared	0.01	0.03	0.03	0.01	0.03	0.03	0.00	0.02	0.02	0.00	0.02	0.02	0.01	0.02	0.02
Mean in Control Group	0.15	0.15	0.15	0.15	0.15	0.15	0.09	0.09	0.09	0.08	0.08	0.08	0.12	0.12	0.12

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the apartment as the unit of observation and include all newly registered citizens in the targeted apartments. We estimate the impact of the interventions on the number of votes cast by newly registered citizens for each electoral round separately (columns 1 to 12) and for all rounds taken together (columns 14 and 15). In all regressions, we control for the number of last names found on mailboxes that did not appear on the 2011 voters' lists, as a proxy for the initial number of unregistered and misregistered citizens.*

Table 7: Propensity to vote of the newly registered citizens in the control and the treatment groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Presidential Elections, 1st round			Presidential Elections, 2nd round			General Elections, 1st round			General Elections, 2nd round			All 4 rounds		
Any treatment	-0.006			-0.025			-0.042			-0.019			-0.023		
	(0.012)			(0.011)**			(0.020)**			(0.020)			(0.012)*		
Early Canvassing		-0.009			-0.010			-0.026			0.009			-0.009	
		(0.018)			(0.014)			(0.028)			(0.028)			(0.017)	
Late Canvassing		-0.002			-0.024			-0.022			-0.008			-0.014	
		(0.018)			(0.017)			(0.034)			(0.030)			(0.018)	
Early Home registration		0.006			-0.058			-0.040			-0.024			-0.029	
		(0.017)			(0.018)***			(0.027)			(0.027)			(0.016)*	
Late Home registration		-0.011			-0.030			-0.065			-0.059			-0.041	
		(0.018)			(0.017)*			(0.026)**			(0.026)**			(0.015)***	
Early Canvassing + Late Home registration		-0.018			-0.013			-0.033			-0.025			-0.023	
		(0.019)			(0.017)			(0.027)			(0.027)			(0.016)	
Early Home registration + Late Home registration		-0.002			-0.012			-0.060			-0.003			-0.019	
		(0.017)			(0.015)			(0.026)**			(0.027)			(0.016)	
Canvassing visit			-0.006			-0.017			-0.024			0.000			-0.011
			(0.015)			(0.013)			(0.026)			(0.025)			(0.015)
Home registration visit			-0.003			-0.044			-0.053			-0.041			-0.035
			(0.014)			(0.014)***			(0.022)**			(0.022)*			(0.013)***
Two visits			-0.009			-0.013			-0.047			-0.013			-0.021
			(0.015)			(0.013)			(0.023)**			(0.023)			(0.014)
Constant	0.874	0.874	0.874	0.896	0.896	0.896	0.526	0.526	0.526	0.486	0.486	0.486	0.695	0.695	0.695
	(0.011)***	(0.011)***	(0.011)***	(0.009)***	(0.009)***	(0.009)***	(0.018)***	(0.018)***	(0.018)***	(0.018)***	(0.018)***	(0.018)***	(0.011)***	(0.011)***	(0.011)***
Wald test:															
Home registration visit = Canvassing visit															
Test statistic			0.04			3.47			1.58			3.92			3.19
p-value			0.838			0.063*			0.21			0.048**			0.075*
Observations	5456	5456	5456	5456	5456	5456	5478	5478	5478	5471	5471	5471	21861	21861	21861
R-squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the individual participation at a given electoral round as the unit of observation and include all newly registered citizens. We estimate differences in the propensity to vote of newly registered citizens in the control and the treatment groups for each electoral round separately (columns 1 to 12) and for all 4 rounds taken together (columns 13 to 15).*

Table 8: Characteristics of newly registered citizens in apartments which opened their door at the late visit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Individual characteristics						Building characteristics			
	Gender		Age		Born abroad		Baseline registration rate		Turnout of previously registered citizens	
Early Home registration + Late Home registration	-0.033	-0.033	-0.243	-0.243	0.095	0.095	-0.079	-0.079	0.007	0.007
	(0.054)	(0.054)	(1.481)	(1.480)	(0.059)	(0.059)	(0.107)	(0.107)	(0.023)	(0.023)
Late Home registration		-0.015		-0.372		0.067		-0.077		0.005
		(0.049)		(1.347)		(0.052)		(0.108)		(0.029)
Constant	0.460	0.460	37.438	37.438	0.300	0.300	1.165	1.165	0.552	0.552
	(0.036)***	(0.036)***	(0.911)***	(0.911)***	(0.037)***	(0.037)***	(0.087)***	(0.087)***	(0.018)***	(0.018)***
Observations	300	437	460	692	460	692	460	692	423	643
R-squared	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. The sample includes all newly registered citizens living in apartments which opened their door at the late visit in the treatment groups "Early Canvassing + Late Home registration", "Early Home registration + Late Home registration" and, for regressions reported in the even-number columns, "Late Home registration".*
We consider individual characteristics (columns 1 to 6) as well as characteristics of the addresses in which the newly registered citizens live (columns 7 to 10).

Table 9: Fraction of citizens registered at home among newly registered citizens in apartments which opened their door at the late visit

	(1)	(2)
Early Home registration + Late Home registration	0.219	0.219
	(0.051)***	(0.051)***
Late Home registration		0.176
		(0.053)***
Constant	0.2	0.2
	(0.031)***	(0.031)***
Observations	452	681
R-squared	0.06	0.04

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. The sample includes all newly registered citizens living in apartments which opened their door at the late visit in the treatment groups "Early Canvassing + Late Home registration", "Early Home registration + Late Home registration" and, for regressions reported in the even-number columns, "Late Home registration".*

Table 10: Treatment impact of home registration on the propensity to vote

	(1)	(2)	(3)	(4)
	Without the treatment groupe "Late Home registration"		With the treatment groupe "Late Home registration"	
Registered at home	0.009	-0.005	-0.062	-0.068
	(0.181)	(0.175)	(0.172)	(0.164)
Constant	0.600	0.301	0.593	0.358
	(0.058)***	(0.113)***	(0.043)***	(0.103)***
City fixed effects	x	x	x	x
Building controls		x		x
Individual controls		x		x
Observations	1801	1659	2710	2520
R-squared	0.01	0.04	0.01	0.04

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. The sample includes all newly registered citizens living in apartments which opened their door at the late visit in the treatment groups "Early Canvassing + Late Home registration", "Early Home registration + Late Home registration" and, for regressions reported in columns 3 and 4, "Late Home registration".*
The dummy "Registered at home" is instrumented by the treatment dummy "Early Home registration + Late Home registration" (columns 1 and 2) or the two treatment dummies "Early Home registration + Late Home registration" and "Late Home registration" (columns 3 and 4).

Table 11: Propensity to vote of the newly registered citizens in the control and the treatment groups, at the Presidential and General elections

	(1)	(2)	(3)	(4)
Any treatment	-0.023			
	(0.012)*			
Any treatment x Pres. Elections		-0.015		
		(0.010)		
Any treatment x Gen. Elections		-0.030		
		(0.018)*		
Canvassing visit			-0.011	
			(0.015)	
Home registration visit			-0.035	
			(0.013)***	
Two visits			-0.021	
			(0.014)	
Canvassing visit x Pres. Elections				-0.011
				(0.012)
Canvassing visit x Gen. Elections				-0.012
				(0.023)
Home registration visit x Pres. Elections				-0.023
				(0.012)**
Home registration visit x Gen. Elections				-0.047
				(0.019)**
Two visits x Pres. Elections				-0.011
				(0.012)
Two visits x Gen. Elections				-0.030
				(0.020)
Gen. Elections		-0.379		-0.379
		(0.015)***		(0.015)***
Constant	0.695	0.885	0.695	0.885
	(0.011)***	(0.009)***	(0.011)***	(0.009)***
Wald tests:				
Any treatment x Pres. Elections = Any treatment x Gen. Elections				
Test statistic		0.83		
p-value		0.364		
Canvassing visit x Pres. Elections = Canvassing visit x Gen. Elections				
Test statistic				0
p-value				0.987
Home registration visit x Pres. Elections = Home registration visit x Gen. Elections				
Test statistic				1.54
p-value				0.216
Two visits x Pres. Elections = Two visits x Gen. Elections				
Test statistic				1.04
p-value				0.309
Observations	21861	21861	21861	21861
R-squared	0	0	0	0

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the individual participation at a given electoral round as the unit of observation and include one observation per round and newly registered citizen in the sample.*
Column 1 and 3 are identical to Columns 13 and 15 of Table 7, and included for reference. In Columns 2 and 4, we allow the propensity to vote of newly registered citizens to differ at the Presidential and General elections, and, in Column 4, depending on their pooled treatment group.

Table 12: Propensity to vote of the previously registered citizens in the control and the treatment groups (Spillovers)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Presidential Elections, 1st round		Presidential Elections, 2nd round		General Elections, 1st round		General Elections, 2nd round		All rounds	
Panel A.	All previously registered citizens									
Any treatment	-0.016	-0.018	-0.011	-0.012	-0.001	0.001	0.001	0.001	-0.007	-0.007
	(0.008)**	(0.007)**	(0.008)	(0.007)*	(0.010)	(0.008)	(0.010)	(0.008)	(0.007)	(0.006)
Constant	0.715	0.882	0.734	0.926	0.448	0.750	0.429	0.596	0.582	0.788
	(0.007)***	(0.028)***	(0.007)***	(0.014)***	(0.009)***	(0.020)***	(0.009)***	(0.033)***	(0.007)***	(0.014)***
Strata fixed effects		x		x		x		x		x
Building controls		x		x		x		x		x
Individual controls		x		x		x		x		x
Observations	28441	28383	28440	28382	28434	28376	28407	28350	113722	113491
R-squared	0.00	0.03	0.00	0.03	0.00	0.07	0.00	0.07	0.00	0.04
Panel B.	Previously registered citizens whose names were identified on the mailboxes									
Any treatment	-0.022	-0.019	-0.011	-0.008	-0.005	0.000	-0.006	-0.002	-0.011	-0.007
	(0.008)***	(0.008)**	-0.009	-0.008	-0.012	-0.010	-0.012	-0.010	-0.008	-0.007
Constant	0.780	0.889	0.790	0.940	0.489	0.742	0.470	0.581	0.632	0.788
	(0.008)***	(0.033)***	(0.008)***	(0.015)***	(0.011)***	(0.022)***	(0.011)***	(0.035)***	(0.008)***	(0.016)***
Strata fixed effects		x		x		x		x		x
Building controls		x		x		x		x		x
Individual controls		x		x		x		x		x
Observations	19096	19061	19095	19060	19086	19051	19067	19033	76344	76205
R-squared	0.00	0.03	0.00	0.03	0.00	0.08	0.00	0.08	0.00	0.04

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We take the individual participation at a given electoral round as the unit of observation and include all previoulsy registered citizens (Panel A) before restricting the sample to the previously registered citizens whose names were identified on the mailboxes during the preparatory work (and thus more likely to actually live at the address listed on the voters' list (Panel B).*

We estimate differences in the propensity to vote of newly registered citizens in the control and the treatment groups for each electoral round separately (columns 1 to 8) and for all 4 rounds taken together (columns 9 and 10).

Table 13: Individual characteristics of the newly registered citizens in the control and the treatment groups and the previously registered citizens in all groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A.	Characteristics regarding the place of birth													
	Born in the city of residence		Born in another city in the department		Borth in another department in the region		Born in another region		Born abroad		Born in a city		Size of the population in home city	
Newly registered	-0.102	-0.102	-0.043	-0.043	-0.001	-0.001	0.104	0.104	0.041	0.041	0.034	0.034	-30,126	-30,126
	(0.013)***	(0.013)***	(0.013)***	(0.013)***	(0.011)	(0.011)	(0.018)***	(0.018)***	(0.018)**	(0.018)**	(0.004)***	(0.004)***	(22,387)	(22,388)
Newly registered x Any treatment	0.023		0.014		-0.005		-0.045		0.013		-0.003		2,426	
	(0.012)*		(0.015)		(0.013)		(0.020)**		(0.019)		(0.004)		(23,432)	
Newly registered x Canvassing visit		0.035		-0.001		-0.001		-0.045		0.012		-0.002		-2,565
		(0.015)**		(0.017)		(0.017)		(0.026)*		(0.022)		(0.005)		(29,328)
Newly registered x Home registration visit		0.019		0.023		-0.003		-0.047		0.007		0.001		16,879
		(0.014)		(0.017)		(0.015)		(0.025)*		(0.022)		(0.005)		(28,526)
Newly registered x Two visits		0.015		0.019		-0.010		-0.044		0.020		-0.008		-6,835
		(0.015)		(0.018)		(0.015)		(0.023)*		(0.023)		(0.006)		(27,181)
Constant	0.202	0.202	0.162	0.162	0.132	0.132	0.262	0.262	0.242	0.242	0.955	0.955	299,280	299,280
	(0.012)***	(0.012)***	(0.009)***	(0.009)***	(0.007)***	(0.007)***	(0.008)***	(0.008)***	(0.005)***	(0.005)***	(0.002)***	(0.002)***	(11,059.943)***	(11,060.383)***
Observations	33621	33621	33621	33621	33621	33621	33621	33621	33621	33621	25102	25102	25102	25102
R-squared	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel B.	Other characteristics									
	Individual characteristics						Building characteristics			
	Gender		Age		Baseline registration rate		Turnout of previously registered citizens in the building		Housing price	
Newly registered	0.001	0.001	-8.832	-8.832	-0.347	-0.347	0.017	0.017	-13.6	-13.6
	(0.02)	(0.02)	(0.556)***	(0.556)***	(0.038)***	(0.038)***	(0.010)*	(0.010)*	(195.17)	(195.19)
Newly registered x Any treatment	0.003		-0.537		-0.016		-0.022		-59.9	
	(0.02)		(0.55)		(0.04)		(0.012)*		(228.24)	
Newly registered x Canvassing visit		0.024		-0.623		-0.014		-0.023		-104.6
		(0.03)		(0.70)		(0.05)		(0.02)		(262.19)
Newly registered x Home registration visit		-0.013		-0.654		-0.036		-0.028		-47.3
		(0.02)		(0.62)		(0.05)		(0.014)**		(267.31)
Newly registered x Two visits		0.000		-0.357		0.000		-0.016		-39.6
		(0.02)		(0.64)		(0.05)		(0.01)		(215.08)
Constant	0.453	0.453	45.108	45.108	1.399	1.399	0.575	0.575	3382.0	3382.0
	(0.003)***	(0.003)***	(0.291)***	(0.291)***	(0.024)***	(0.024)***	(0.004)***	(0.004)***	122.090)**	122.103)***
Observations	19705	19705	33991	33991	33997	33997	33693	33693	9840	9840
R-squared	0.00	0.00	0.04	0.04	0.03	0.03	0.00	0.00	0.00	0.00

*Notes: Clustered standard errors are in parentheses. ***, **, * indicate significance at 1, 5 and 10%. We include all registered citizens in the control and treatment groups and estimate differences between the previously and newly registered citizens, as well as the newly registered citizens in the control group and the treatment groups.*
We consider characteristics regarding the place of birth (Panel A) and other individual and building characteristics (Panel B).