Your Vote is (no) Secret! How Low Voter Density Harms Voter Anonymity and Biases Elections in Italy

Mauro Caselli and Paolo Falco

headed by

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

CE PS

CEPII
RESEARCH AND EXPERTISE
ON THE WORLD ECONOMY

CERGE EI

CPB Netherlands Bureau for Economic
Policy Analysis

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

KOF Swiss Economic Institute

REM
RESEARCH IN
ECONOMICS AND
MATHEMATICS

Toulouse
School
of Economics

uc3m | Universidad **Carlos III** de Madrid
Instituto Universitario de Economía

SAID
BUSINESS SCHOOL

UNIVERSITY OF
OXFORD

UNIVERSITY
OF TRENTO

VATT
VALTION TALOUDELLINEN TUTKIMUSKESKUS
VATT INSTITUTE FOR ECONOMIC RESEARCH

ZEW

Your vote is (no) secret!

How low voter density harms voter anonymity and biases elections in Italy[*]

Mauro Caselli[†] and Paolo Falco[‡]

May 29, 2019

Abstract

Italian voters are assigned to a specific polling station according to their address. After an election, candidates know how many votes they received in each polling station. When the number of voters per polling stations is low and candidates are many, this jeopardises the secrecy of voting and candidates can more easily detect deviations from pre-electoral pledges. Exploiting variation in the number of voters per polling station across cities and over time, combined with rich data on politicians in office in all Italian municipalities between 1989 and 2015, we estimate the effect of voter density on the probability of re-election for local politicians. We find that when the number of voters per polling station is lower (and secrecy is at greater risk), incumbents have a higher probability of re-election. The analysis addresses the potential endogeneity of voter density. The results are stronger in regions with lower social capital and worse institutions.

JEL codes: D02, D72, H70.
Keywords: elections, secret ballot, incumbency advantage, polling stations, voter density, institutions, social capital, Italy.

1 Introduction

The well-functioning of modern democracies rests on the anonymity of voting. Casting votes by secret ballot is commonplace around the world and is crucial to prevent illicit practices that may influence the outcome of elections, such as voter intimidation, blackmailing, and vote buying (Norris, 2014). The right to hold elections by secret ballot is enshrined in many international treaties and agreements, including the *Universal Declaration of Human Rights*, which states that "The will of the people shall be expressed in periodic and genuine elections which shall be held by secret vote or by equivalent free voting procedures." (UN, 1948). In this paper, we investigate a mechanism that may jeopardise the secrecy of voting in an advanced industrialised country.

In Italy, voters are assigned to a specific polling station according to their address, and the number of voters per polling station varies. After an election, information on the number of votes obtained by every candidate in each polling station is publicly available. When the number of voters per polling station is low, this may severely limit the secrecy of voting, as recently discussed in the Italian Parliament (Rubino, 2018). A politician who receives fewer votes than expected in a particular area knows with certainty that some of the people who pledged their vote did not keep their promise. In the limiting case, when a politician receives no votes in a given polling station, he or she can conclude with certainty that none of the people in a confined geographical area voted for him or her. As a result, he can sanction them by being less generous if elected in office, or by other systems of retaliation.[1] Such a mechanism is likely to have an impact on

---

[1]This paper uses retaliation as a generic term and does not distinguish between different forms of it. Since our study is set in a context where physical violence is relatively infrequent, the reader should think of retaliation as largely consisting of the denial of favours that politicians can grant when in office. This may span a range of areas, including health services (e.g., helping people to bypass waiting lists, influencing the hiring of health personnel in public hospitals, etc.), education (e.g., influencing the assignment of students to schools or teachers to schools), and employment in the public administration. In addition, politicians may make it more difficult for voters who do not support them to obtain services they are

the outcome of elections, especially of municipal ones, where candidates have a close relationship with voters and a detailed knowledge of their geographical distribution across the municipality. To test this claim we investigate how electoral results change as voter density (the number of voters per polling station) drops and it becomes easier for politicians to detect voters' behaviour. In particular, we hypothesise that lower voter density reduces political contestability by favouring the re-election of incumbents, who have greater retaliatory power against voters who do not support them.

The assignment of voters to specific polling stations according to their address is not unique to Italy. The peculiarity of Italian municipal elections is the use of preference voting, whereby voters can express a preference for a specific candidate from the list proposed by the party they support. Each party receives a share of council seats that depends on its share of all votes (with a possible correction for a majority premium), and the seats are assigned to candidates who have received most preferences within the list. Since voters can express a preference for a single candidate among many (each list usually has as many candidates as the number of seats available in the council, on average 14 over the period 1989-2015, and the average number of lists per election was 3.5 over the same period), the average number of votes candidates receive per polling station is extremely low (less than 20, on average, over the period analysed). This makes it easier for candidates to detect deviations from pre-electoral pledges. Such a system is largely absent in other advanced economies.[2]

---

entitled to (e.g., a construction permit).

[2]In proportional systems, preference voting is typically either not possible or extended to several candidates. In France and in some German states where preference voting is not allowed, voters cast their vote for a party (and the list of candidates it proposes), but they cannot express a preference for a specific candidate. The seats allocated to each party are distributed on the basis of a pre-determined ranking of candidates (e.g., on the basis of primary elections). In other German states where multiple preferences are allowed, voters can cast their vote for a party (and the associated list of candidates) as well as several specific candidates. In this case, the seats are allocated based on the total number of votes each candidate obtains and this number is potentially large. In countries with first-past-the-post systems, such as the United Kingdom, each party proposes a single candidate per seat available and the total number of candidates per seat is therefore significantly smaller.

Exploiting variation in the number of voters per polling station across cities and over time, combined with rich data on politicians in office in all Italian municipalities between 1989 and 2015, we estimate the effect of voter density on the probability of re-election. The analysis controls for province fixed effects and for a set of municipal-level indicators of economic activity. We test the robustness of the results using an instrumental variable (IV) strategy based on the geographical and morphological characteristics of a municipality's territory, which drive exogenous variation in voter density. We also introduce municipal fixed effects and resolve the endogeneity of changes in voter density within cities by using variation in the total number of polling stations at the macro level (regions) as an instrument for municipal-level variation. While this is a novel instrument, the use of an instrument based on a variable related to the endogenous one but aggregated at a higher level in order to keep the exogenous component of the variable of interest is not uncommon in the economic literature (see, for example, Khandelwal, 2010, and Bernini and Tomasi, 2015).

The results show that lower voter density significantly increases the probability of re-election for an incumbent in a municipal office. The effect is sizeable and robust to addressing the potential endogeneity of voter density with respect to unobserved determinants of electoral outcomes. Additional exploration also reveals that lower voter density has a stronger positive impact on the probability of re-election in cities with lower levels of social capital. Coupled with a vast literature showing a clear link between social capital and the quality of institutions (e.g., Putnam (1993)), this finding strongly suggests that the negative effects of lower density on electoral competition are larger in areas where the quality of government is lower. The explanation we advance is that when institutions are less effective, having the sympathy of politicians can help to obtain public services (e.g., a place in a certain school, a consultation in a hospital that would otherwise require a long wait, etc.), which are provided transparently in places with

better institutions.

This is the first study, to the best of our knowledge, that analyses how voter density may limit the secrecy of voting and affect the outcomes of elections. The problem has been recently debated in Italy and a proposal has been made in the Italian Parliament to increase the minimum number of voters who can be assigned to a single polling station (Rubino, 2018). Yet, Italy is not the only industrialised country in which voter anonymity and ballot secrecy have been discussed. For example, Elklit (2018) discusses electoral integrity in Sweden, one of the countries with the highest scores according to the Electoral Integrity Project (Norris et al., 2018). Gerber et al. (2013) use survey data from the United States and show that when the perception of ballot secrecy falters voting behaviour may be affected. We contribute to this literature by studying a novel mechanism that may *de facto* reduce ballot secrecy, even when the system is in principle fully anonymous, and in turn affect voters' behaviour.

More generally, our work contributes to the recent empirical literature on how different features of political systems affect elections. A number of studies, for instance, focus on the selection of politicians and show how it is affected by wage schemes (Gagliarducci et al., 2010; Gagliarducci and Nannicini, 2013), the use of closed lists (Galasso and Nannicini, 2015), the share of competitive districts (Galasso and Nannicini, 2017), and the voting systems (Gulino, 2018). Beath et al. (2016) are closer to our analysis, since they study the role of district size on electoral outcomes. They find that representatives selected by elections with a single (large) multi-member district are better educated and exhibit less extreme policy preferences. Their analysis, however, does not dwell on the effects of voter density.

Our paper also contributes to the expanding literature on the role of social capital for the well-functioning of institutions (e.g., Putnam, 1993). Recent contributions to this lit-

erature include the work by Nannicini et al. (2013), who show that, when social capital is higher, voters who share values and beliefs that foster cooperation are more likely to vote based on criteria of social welfare, rather than private gain. As a consequence, they are more likely to hold their political representatives accountable and to punish misbehaviour. In a similar vein, Batinti et al. (2017) show that incumbent Italian mayors who carry out a forward-looking and transparent fiscal agenda are more likely to be re-elected where the level of local social capital is larger. Thus, the good conduct of incumbent mayors is rewarded, but only in contexts with more social capital. We contribute to this literature by showing that the adverse effects of lower voter density on electoral competition is stronger in municipalities with lower social capital. Consistent with Nannicini et al. (2013), our findings may be explained by the fact that in areas with higher social capital voters are more likely to vote based on criteria of social welfare and they are less sensitive to the threats of retaliation that may affect their private welfare.

Thirdly, our work contributes to the literature on incumbency advantage. Early work by Krebs (1998) highlighted incumbency as a strong predictor of electoral success in the United States. More recently, Trounstine (2011) highlights an incumbency advantage of 32 percentage points at city council elections between 1915 and 1985 in four U.S. cities. In the case of Italy, De Benedetto (2014) documents the existence of an incumbency advantage by means of a regression discontinuity design that compares candidates who barely win to those who barely lose an election. While explaining the origin of such an advantage is beyond the scope of of our work, we explore a mechanism that may contribute to consolidate it.[3]

The remainder of the paper is structured as follows. Section 2 describes our concep-

---

[3]Since we are not attempting to explain the origin of incumbency advantage, our results are not inconsistent with the finding by De Benedetto (2014) that such an advantage is stronger in Northern Italy. The mechanism we uncover in this paper, which is found to be stronger in the South, may contribute to increase the incumbency advantage in those regions, but the overall level may remain lower than in the North.

tual framework. Section 3 outlines the procedure we have followed to construct our data. Section 4 presents our empirical strategy and discusses our results. Section 5 digs deeper into the role played by the quality of institutions in driving our results. Section 6 concludes.

2   A mechanism that limits the secrecy of voting

To fix ideas, we begin by reporting a short anecdote that inspired the analysis in this paper and is useful to describe the key mechanism that will be outlined in detail below.[4] The rest of the section discusses this intuition and a formal model is presented in Appendix B.

In the town of Castel del Faggio, the municipal elections are about to be held and Maria is trying to decide whom she should vote for. In the past, she has gladly followed the indications of her parents, whom she generally trusts, but she increasingly feels that the politician they have been supporting over the years, Mr. Rossi, a friend of her family who is trying to be re-elected as councillor, has not served her well. Together with a group of friends from her neighbourhood, she is inclined to choose another candidate, but they do not want to upset their parents, who have pledged their family's support to Mr. Rossi and would not want to disappoint him, as he is a rather important figure in the town. To make matters worse, the polling station where Maria typically votes is very small and each candidate running for the position of councillor can only expect to obtain a handful of votes. If she and her friends do not vote for Mr. Rossi, her parents will almost certainly know. It would be so much better, she thinks, if all votes from the entire city were pooled together. That way her vote would go unnoticed. Isn't that what anonymous voting means, she wonders?

---

[4]Please note that while the dynamics described correspond to the anecdotal evidence gathered by the authors, the name of the town and of the main characters in the story are fictional.

This simple anecdote exemplifies the mechanism we set out to test. Our central hypothesis is that by increasing the risk that a voter's decision is revealed to candidates, lower voter density skews electoral outcomes in favour of incumbent politicians, who have greater retaliatory power against voters who do not support them. In order to understand how that mechanism may play out, it is useful to begin by outlining the key features of municipal elections in Italy.

In Italian municipal elections, voters not only elect a mayor, but also city councillors. The electoral system (discussed at length in Appendix A) allows voters to cast their vote for a mayoral candidate and, in addition, to express a preference for candidates running for the role of councillors, who are grouped in lists (associated to national parties or local civic groups).[5] Since there is a minimum mandated number of candidates per list, the total number of candidates running for a post of councillor is large and, therefore, the number of votes that each one of them can expect to obtain in a given polling station is limited. For instance, in the 2015 local elections, candidates for city councillor received less than 18 votes per polling station on average.[6] Figure 1 shows the full distribution of votes and reveals that in the majority of cases those votes were very few.
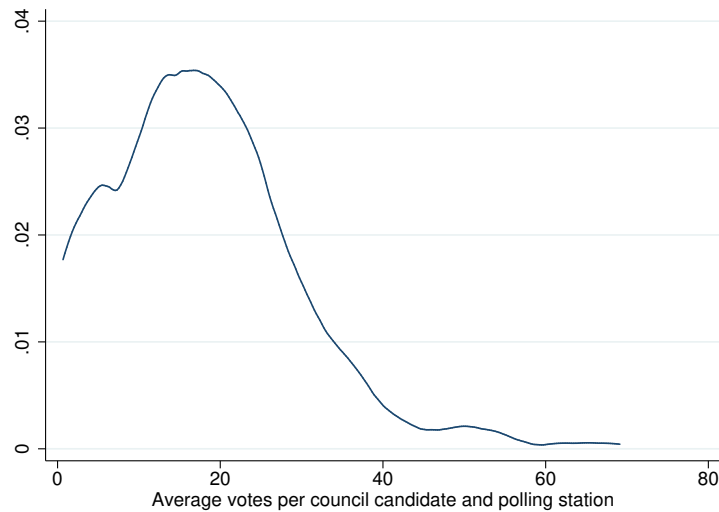
Since each polling station gathers voters from a precise part of the city, and given that local candidates are likely to have direct connections with their voters, they can form precise expectations about who should be voting for them in each polling station. Crucially, after the election, information about votes received in each polling station is available to politicians. That is because the Italian law prescribes that each list of candidates should have a person to officially represent it in every polling station (the

---

[5]The electoral system, as well as the system of government of Italian municipalities, was different before 1993, as described in Appendix A. Yet, all our main results are robust to the exclusion of all observations prior to the 1993 reform, as discussed in Appendix E.

[6]It should be noted that mayoral candidates receive on average substantially more votes in a given polling station. In the 2015 elections, mayoral candidates received on average 200 votes per polling station.

Figure 1: Average number of votes per candidate and per polling station, 2015

Notes: Number of votes received on average in a given polling station by candidates to the position of city councillors in the 2015 election. Source: Own elaborations based on data from the Italian Ministry of the Interior.

so-called *rappresentanti di lista*). The role of this person is to monitor the regularity of the count and having one representative per list ensures balance. On the other hand, this system allows politicians to have a detailed mapping of the votes they received across polling stations immediately after the election (in fact, in real-time, while the count is ongoing).[7]

Given the detailed information at their disposal, candidates who receive fewer votes than they expected in a given polling station know with certainty that some of the people they expected to vote for them did not do so. In the limiting case, when the number of votes received in a given polling station is zero, the candidate knows with certainty that all the people in that area did not vote for him or her. Such case is in fact common, as shown by the density plot in Figure 1.

---

[7]Moreover, municipalities may also publicly post the results of the latest elections for each polling station. Yet, it should be noted that local electoral results at the level of polling stations are not readily available for all municipalities in Italy.

Voters' choices are likely to be influenced by this mechanism. When the number of votes the candidate can expect in a given polling station is small, there is a significant risk that it will be zero. If the voter decides not to vote for the candidate, he/she has to face the risk that his/her choice will be effectively revealed. If candidates have retaliatory power, this may affect voters' behaviour.

Such mechanism is likely to favour incumbents in elected positions who have greater retaliatory power against voters who do not support them. This rests on the following considerations. Retaliatory power is a function of every politician's probability of being elected, which is typically higher for an incumbent than for a first-time candidate. This is a well-studied phenomenon, known as "incumbency advantage" (Krebs, 1998; Trounstine, 2011; De Benedetto, 2014). Also, incumbents are likely to be better connected and be part of relatively powerful social networks (e.g., Cruz et al., 2017), which may allow them to retaliate even if not re-elected. Recent work by Akcigit et al. (2018) documents the connections between Italian politicians and market-leading companies across a range of sectors. These connections are likely to grow stronger when a politician is in office. Appendix B outlines a simple theoretical model that formalises this intuition. It should be noted that while the model currently assumes that incumbent politicians can impose greater sanctions on voters, one could equally assume that all politicians (incumbents and challengers) can impose the same sanctions if elected. However, since the expected probability of winning an election is higher, on average, for an incumbent (a conclusion that finds support in our data), *expected* retaliatory power (i.e., sanctions weighted by the probability of re-election) is higher for an incumbent.

In what follows, our analysis will focus on the election of city councillors. That is because the mechanism we have described is likely to be most salient when the number of candidates is high relative to voters. However, the same mechanism may have similar

10

repercussions on the election of mayors (though they are likely to be weaker since in a mayoral race the number of votes received by each candidate in every polling station is substantially higher and detection of deviations from electoral pledges is more difficult). On the other hand, the effect on mayoral election may pass through the election of councillors. Since incumbent councillors are likely to be connected to incumbent mayors, their re-election is likely to contribute to the re-election of the mayor.

In sum, we hypothesise that by reducing politicians' ability to detect individual voters' choices, higher voter density (i.e., a higher number of voters per polling station) decreases the probability of re-election for incumbents, who have higher retaliatory power against those who do not vote for them. The rest of the paper is dedicated to test this hypothesis.

3   Data

The empirical analysis is based on a rich dataset containing information on all Italian politicians who are in office at the municipal level between 1989 and 2015, combined with electoral results for all the elections held over the period, and with economic and geographical information on the universe of over 8,000 Italian municipalities.

Registry of local politicians

We obtain administrative data on all municipal-level politicians in office between 1989-2015 from the website of the Italian Ministry of the Interior. For each individual in every municipality, we have demographic information, office held (i.e., council member, mayor, and member of the municipal government), and appointment date.

Electoral results

Data on electoral results and on the composition of the electorate for all municipal

Table 1: Probability of re-election by officer characteristics, councillors

	0	1	Difference
	(1)	(2)	(3)
Male	0.266	0.336	-0.071***
	(0.442)	(0.472)	(-46.79)
Age > 45	0.343	0.311	0.032***
	(0.475)	(0.463)	(29.69)
Uni degree	0.325	0.330	-0.005***
	(0.468)	(0.470)	(-4.22)

Notes: The table shows the share of city councillors re-elected in their position by officer characteristics. The variable Male distinguishes between female (corresponding to column (1)) and male (corresponding to column (2)) councillors. The variable Age > 45 distinguishes between councillors younger than 45 (corresponding to column (1)) and older than 45 (corresponding to column (2)). Variable Uni degree distinguishes between councillors who did not complete a university degree (corresponding to column (1)) and those with at least a university degree (corresponding to column (2)). Column (3) shows the difference between column (1) and column (2). Standard deviations are shown in parentheses for columns (1) and (2), while the t statistics are shown in parentheses in column (3). *** indicates coefficients significantly different from zero at 1% level.

elections held between 1989 and 2015 were also provided by the Italian Ministry of the Interior. Notably, we have information on the number of eligible voters in each municipality and year, the number of polling stations, the identities (names and demographic characteristics) of all mayoral candidates, parties and coalitions, votes obtained (by mayoral candidates and parties), and the allocation of seats in the councils.

One limitation of the data is that we do not directly observe whether incumbent councillors run for re-election, but only if they remain in office or not.[8] In the rest of the paper, we define the variable "re-election" as a dummy equal to one if a previously-elected councillor runs again and is re-elected (and, thus, remains in office), while it is equal to zero if a councillor is not re-elected or does not run for election again. Upon discussing our findings, we will clarify the implications of this approach for our conclusions.

Based on these data, Table 1 provides some descriptive statistics regarding the probability of re-election of city councillors by officer characteristics in the period 1989-2015.

---

[8]On the other hand, for the case of mayors, we know whether they run and, conditional on running for office, if they win the elections again or not.

Table 2: Descriptive statistics for voter density and number of polling stations

	Mean (1)	St. Dev. (2)	Minimum (3)	Maximum (4)
Voter density	695.98	241.72	28	1832
No polling stations	9.19	55.61	1	3688
No candidates per election	59.29	100.75	6	2124
No votes per candidate/polling station	16.82	11.06	0	139

Notes: The number of observations is 37541, equal to the number of local elections in all municipalities in Italy. The table shows the mean (column (1)), standard deviation (column (2)), minimum (column (3)) and maximum (column (4)) for voter density (i.e., number of voters per polling station), number of polling stations, number of candidates for city councillors per election and average number of votes obtained by each candidate for city councillor and by polling station.

The table shows that one third of councillors are re-elected on average. Male councillors and those younger than 45 are more likely to be re-elected. Holding a university degree also increases the probability of re-election, but by a smaller amount.

Using data for all local elections across all municipalities in Italy, Table 2 shows some descriptive statistics for voter density, i.e., the average number of voters per polling station, number of polling stations, number of candidates for city councillors per election and average number of votes obtained by each candidate for city councillor in a given polling station. Data show that there is considerable variation in voter density across Italian municipalities and we will be able to exploit such variation to estimate its effect on the probability of re-election. We can also notice that candidates for city councillors obtain on average few votes in each polling station and we argue that our results are driven by this fact.

Municipality characteristics

Data on the socio-demographic structure of municipalities and key economic indicators at the municipal level have been obtained from the Italian National Institute of Statistics (Istat). In particular, data on resident population (share of residents above 65 and below 19), educational attainment (average years of education), and labour market

conditions (share of employed and share of unemployed in the active population) are extracted from the 1991, 2001 and 2011 waves of the Population Census (CPOP).[9]

Istat also provides information on the geographical and morphological characteristics of municipalities. In particular, we focus on their size in squared kilometres, on the coefficient of variation in altitude within their territory, and on the number of localities that compose municipalities.[10] These variables are used as instruments in the IV regressions with province fixed effects.

Social Capital

Finally, in order to measure social capital at the level of municipalities, we follow Putnam (1993) and more recent work by Andini and Andini (2018), and use voter turnout in the 1987 national political elections, which pre-date the beginning of the period analysed in this paper. This deserves some discussion. Voter turnout has long been established as a measure of social capital, since it constitutes a proxy for civic engagement and participation. Two arguments are commonly used against this proxy (Andini and Andini, 2018). The first one is that voting in political elections may be a constrained choice, since it is defined by the Italian constitution as a civic duty.[11] However, since no concrete sanction is linked to abstention, it seems reasonable to assume that, in practice, the decision to vote is free from such a constraint. Second, voting may be driven by personal gain rather than social considerations, since politicians may provide benefits to their electorate in exchange for votes. However, the argument is more likely to apply to municipal elections, as argued above. An advantage of using this measure is that it

---

[9]As information for the years between censuses is missing, we attribute the values from the 1991 census to all years between 1989 and 2000, the values from the 2001 census to all years between 2001 and 2010, and the values from the 2011 census to all years between 2011 and 2015.

[10]These localities, known as *frazioni*, are smaller villages or hamlets (occasionally, just small clusters of houses) that are part of the municipality but, for historical and geographical reasons, have the characteristics of a separate town, namely their own centre, square and church.

[11]By contrast, a free choice that may serve as a better proxy for social capital would be an anonymous blood donation.

suffers from fewer distortions than turnout rates in referenda, the most commonly used alternative. Finally, turnout in national elections is also the only measure of social capital that is available at the municipal level for the period we analyse (Andini and Andini, 2018).

4 Results

We estimate the following model of re-election for incumbent j:

$$reelection_{jmpt} = \beta_D density_{mpt} + \beta_X X_{jmpt} + \beta_M M_{mpt} + \eta_p + \theta_t + \epsilon_{jmpt} \tag{1}$$

where *density* is the log of voter density, i.e., average number of voters per polling station in municipality m in province p at time t, X contains a set of candidate j's characteristics, M captures municipality characteristics. Province and year fixed effects are included.

Table 3 shows the coefficients for equation (1). In all regressions, standard errors are clustered at the municipality level. We find that the re-election probability is significantly higher in municipalities where voter density is lower (column (1)). In addition to having province fixed effects the regression controls for economic conditions at the municipal level, including the local unemployment rate.

As mentioned above, our data does not allow us to see whether an incumbent councillor decides not to seek re-election. Therefore, not being re-elected in the analysis above conflates both losing the election and not participating in the first place. If the choice of running for re-election does not change with the number of polling stations (and with voter density), this has no implications for identification. On the other hand, if elected officials have a higher propensity to run in towns with more polling stations (because they know that they are more likely to be re-elected), the coefficient on voter density will

Table 3: Probability of being re-elected, councillors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	-0.036***	-0.062***	-0.661***
	(-12.47)	(-5.37)	(-3.13)
Male	0.094***	0.093***	0.093***
	(57.38)	(57.16)	(51.88)
Age	-0.002***	-0.002***	-0.003***
	(-41.36)	(-41.41)	(-44.67)
Education	0.007***	0.007***	0.006***
	(40.73)	(40.84)	(33.86)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	496.74	490.08	420.87
Cragg-Donald F statistic		12657.14	922.04
Kleibergen-Paap F statistic		68.80	18.00
Observations	757,045	756,071	757,044

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. t statistics in parentheses. *** indicates coefficients significantly different from zero at 1% level.

de facto capture both a direct effect and an indirect effect via the higher propensity of running for office when the possibility of exerting control over the electorate is greater. While one would ideally want to separate the two effects, this adds one additional channel to our story but is not inconsistent with our conclusions.[12] As a robustness check, when we run the analysis on the election of mayors we can include a control for whether an incumbent chose not to run for re-election (the information is available in mayoral races), and we find that our results do not change.

Next, we show that our findings are robust to addressing the potential endogeneity of voter density across municipalities in a given province. To do that, we employ an IV

---

[12]To indirectly test the relative importance of the two channels, we re-estimate the model controlling for the number of re-elections the candidate has experienced. Since the probability of running for office again should decrease over the lifetime of an elected official, this should give us a tentative indication about the importance of the indirect channel. All our results are robust to the inclusion of this additional variable.

strategy that uses geographic characteristics as an instrument for voter density (column (2)) (see, among others, Romer and Frankel, 1999, for a related IV strategy based on geographic variables applied to a rather different analysis). Three variables in our data are well-suited to serve as instruments, keeping in mind that they are time invariant and can only be used in a cross-section. First, municipalities can be split in localities (smaller units that have the appearance of separate cities but are part of the same municipality). The larger is the number of localities in a given municipality, the lower is the density of voters per polling station. Such fractionalisation is typically due to geographical features of the territory (e.g., a river that splits a city in two parts) or historical reasons (e.g., two formerly separate towns that merged into one). Second, the size of a municipality's territory is inversely related to voter density per polling station, since wider areas require a larger number of polling stations for a given number of inhabitants. Third, the degree of variation in altitude within the territory of a municipality is positively related to the number of polling stations for a given population size (and thus inversely related to voter density). These three variables can be considered alternative measures of geographical dispersion of the population of a municipality. The higher the dispersion, the higher the number of polling stations required to serve the population, and the lower is the voter density.

The first stage results of the IV estimation, shown in Appendix C, as well as the Cragg-Donald and Kleibergen-Paap F statistics indicate that all three measures are highly predictive of differences in voter density across municipalities within a province. Moreover, as we have more than one instrument in this case, we can use the Hansen J statistic for overidentifying restrictions to test the validity of our instruments. The Hansen J statistic shows that the instruments are indeed valid.

Upon applying this IV strategy, the estimated effect of voter density increases in mag-

nitude and remains highly significant. The size of the coefficient is such that the probability of being re-elected is 0.6 percentage points smaller in a municipality with a voter density ten percent higher than in a similar municipality in the same province.

As an additional robustness test, we introduce municipality fixed effects and attempt identification of the effects of interest by analysing variation of voter density over time. Since this is likely to be endogenous with respect to potential omitted drivers of electoral outcomes (for instance, population density may increase in better administered municipalities), we need to devise a suitable strategy to offset the problem.[13] We construct an instrument for each municipality-year pair equal to the ratio of the number of voters to the total number of polling stations in each region and year (i.e., voter density at the regional-year level), net of the contribution of each municipality to the total, multiplied by the initial share of the municipality's population in the region's total population. The intuition behind this IV strategy is that macro-level (in our case at the regional level) variation over time in the number of polling stations is driven by macro-level trends that cannot be influenced by the year-on-year choices of individual municipalities. Hence, such variation can be deemed exogenous and used to instrument time-varying municipal-level variation in the number of polling stations. A strong justification for this approach resides in the fact that the institution or abolition of polling stations is ultimately approved by provincial authorities, the *prefetti*, who are directly appointed by the Ministry of the Interior and follow guidelines that are above the po-

---

[13]Another potential source of endogeneity is the fact that polling stations are typically located within public buildings such as schools and hospitals. Hence, public investment that may raise the probability of re-election if voters are pleased with the new infrastructure may also result in an increased number of polling stations. It should be noted, however, that such source of endogeneity should become less important after 1997. That year saw the introduction of a national law that aimed to reduce the number of polling stations across the country in order to contain the costs of elections. The law led to a 30% reduction in the average number of polling stations per municipality in Italy. It is reasonable, therefore, to assume that after 1997 it became more difficult for local politicians to propose an increase in the number of polling stations in their municipality, even when a new infrastructure that could host them (e.g., a school) became available. As a robustness check, we re-estimate all our main results with the exclusion of all observations up until 1997. As discussed in Appendix E, all the main results go through.

litical dynamics of any single municipality. Their decision is driven by indications they receive from the central government and must take into account, among other things, budgetary constraints. In light of this, it seems reasonable to argue that municipal-level circumstances are unlikely to affect the regional trend in the number of polling stations, which is overseen by the central authorities. A similar IV strategy has been applied in different contexts. Two prominent studies from the literature on international economics, for instance, are Khandelwal (2010) and Bernini and Tomasi (2015), who use instruments at a higher level of aggregation than the endogenous variable to attain identification. The first stage results shown in Appendix C together with the Cragg-Donald and Kleibergen-Paap F statistics confirm that the instrument is informative and can effectively predict changes in voter density.

Even after introducing municipality fixed effects and controlling for the endogeneity of changes in the number of polling stations over time, our results are confirmed. Higher voter density maintains its negative effect on the re-election probability and the coefficient becomes larger. The size of the coefficient is such that an increase by ten percent in voter density decreases the probability of being re-elected by 6.6 percentage points. Moreover, all the other coefficients do not vary considerably, thus confirming that the effect of voter density on the probability of re-election is stronger when we use variation within municipalities rather than across them.

Moreover, the estimated effect is non-linear. This is an important result as it constitutes a direct test of the mechanism we have outlined. The formal model presented in Appendix B predicts that the negative marginal effect of increasing worker-density on incumbents' chances of re-election should decrease as the number of voters per polling station increases. This hypothesis is confirmed by the results in Table 9 of Appendix E,

lending support to our theory.[14]

As an additional test, we run the analysis on mayoral races and the results are qualitatively similar in our preferred model with municipality fixed effects (see Appendix D). The results, however, are less stable across specifications, and statistical significance is weaker. This may be a result of lower statistical power since the number of observations available is lower. It is also plausible and consistent with our theory, however, that our mechanism of interest becomes weaker in mayoral races, since the number of voters per candidate is significantly higher and detection of voters' behaviour becomes more difficult.[15]

Finally, we would like to test whether the mechanism we have hypothesised is consistent with voting behaviour *within* a city. This could be done by investigating how support for incumbents varies across polling stations with different voter density. Such an exercise, however, cannot be conducted satisfactorily since data on the votes received by candidates in each polling station has not been compiled at the national level. Certain municipalities, however, make the results available on their website. Since our analysis indicates that results are stronger in the South of the country (as discussed below), we tried to identify a city from that region where data was available and the number of polling stations was sufficiently large to allow for the analysis. The city that best fits these criteria is Catania, a large provincial capital of about 300,000 residents, 335 polling stations and with readily available data by polling station for the 2013 municipal elec-

---

[14]Another indirect way of testing the non-linearity of the effect is to verify whether the impact of density changes with the number of candidates running for election. If the mechanism we have hypothesised works as we predict, when the number of candidates grows (and the expected number of votes each candidate can obtain in a polling station falls), detecting voters' deviations from their pre-electoral pledges becomes easier and thfe impact of voter density on electoral outcomes should become stronger. We find evidence in support of this hypothesis when we interact the estimated effect with the number of lists and, therefore, with the number of candidates running for an election (Table 10 in Appendix E).

[15]Another reason for the weaker result is that since 1993 mayors cannot be re-elected more than twice in a row.

tions. As shown by the results in Table 11 of Appendix E, incumbents in Catania tend to gain a relatively smaller share of votes in larger polling stations (where voter density is higher). This is consistent with our mechanism whereby larger polling stations allow voters to keep their anonymity and restrain incumbents' retaliatory power. Of course, this is only an illustration of the mechanism, which may vary in intensity across cities and regions, as discussed in the next section.

5 Does the quality of local institutions matter?

The mechanism we investigate may have different effects on local elections in different parts of the country. Most importantly, since it rests crucially on the ability of politicians to retaliate after an election, we hypothesise that it is stronger in areas where the quality of local institutions is lower. That is because when institutions are less effective, having the sympathy of politicians can help to obtain public services (e.g., a place in a certain school, a consultation in a hospital that would otherwise require a long wait, etc.), which would be provided transparently and independently of one's political leaning, in places with better institutions.

To test this hypothesis, we follow Andini and Andini (2018) and introduce in the analysis a measure of social capital at the municipal level, which we interact with voter density in our preferred IV estimation. While we are aware that social capital is not a direct measure of the quality of institutions, we build on a vast literature showing that in parts of Italy where social trust is lowest institutions are more likely to fail (e.g., Putnam, 1993; Nannicini et al., 2013; Batinti et al., 2017). The data by Andini and Andini (2018), therefore, constitutes a valid proxy and has the unique advantage of being disaggregated at the municipal level.

Table 4: Probability of being re-elected and social capital, councillors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	-0.038***	-0.078***	-1.205***
	(-11.05)	(-6.73)	(-3.01)
Voter density, log x	0.003	0.015***	0.579**
Voter turnout in 1987	(1.36)	(3.40)	(2.43)
Male	0.094***	0.093***	0.093***
	(57.34)	(57.07)	(51.50)
Age	-0.002***	-0.002***	-0.003***
	(-41.39)	(-41.48)	(-44.96)
Education	0.007***	0.007***	0.006***
	(40.68)	(40.48)	(33.36)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	483.44	476.76	402.71
Cragg-Donald F statistic		9728.71	413.85
Kleibergen-Paap F statistic		52.89	8.01
Observations	756,315	755,341	756,314

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The variable social capital is proxied by the municipality-level turnout at the 1987 national elections following Andini and Andini (2018). The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. *t* statistics in parentheses. *** indicates coefficients significantly different from zero at 1% level.

The results shown in Table 4 indicate that the quality of institutions does indeed amplify or mitigate the mechanism we have identified. The negative impact of low voter density on electoral competition is significantly stronger in municipalities with lower levels of social capital. Among municipalities in the bottom quarter of the social capital distribution, increasing voter density by ten percent can increase the probability of an incumbent's re-election by over seven percentage points based on the IV regression with municipality fixed effects.[16]

---

[16]As an additional check, we compute the correlation between the measure of social capital by Andini and Andini (2018) and the effect of voter density estimated at the municipal level (by means of individual regressions for each municipality). Though the regression results are much more imprecise, due to data limitations, we find a significant correlation indicating that the estimated impact is stronger in areas with lower social capital.

The regions where social capital is lowest (and the effect of voter density is therefore strongest) are predominantly located in the south of Italy, but there is substantial heterogeneity both across the country and within regions. Figure 2 in Appendix F, plots the effect of the interaction term by municipality (according to the level of social capital) and shows that it is largest in Sicily and Calabria. Even within relatively small geographical areas, however, there can be substantial heterogeneity in the estimated effect. Regions like Campania, depicted in Figure 3 in Appendix F, are a striking example.

Geographical differences in the quality of Italian institutions are the subject of a vast literature that can help us shed light on the mechanisms that may amplify the impact of voter density on electoral competition. First, areas with lower social capital have been shown to be more prone to political clientelism (the exchange of goods and services for political support) and citizens living in these areas may be more inclined to cast their vote in relation to exchange agreements (De Benedetto, 2014; Knack, 2002). This may be especially the case in contexts where the public administration is less efficient and corruption more rife. In those circumstances, the benevolence of a politician may be the only possibility to obtain services that in more efficient and less corrupt places would be accessible to all citizens independently of their political connections. This implies that the threat of retaliation from politicians is stronger (and their control over voters more effective) in regions where corruption and inefficiencies in the public administration are more common, as it is more typically the case in the south (Putnam, 1993; Nannicini et al., 2013). A second possibility is that stronger family ties in southern regions are an instrument for politicians to exert pressure on voters (Cruz et al., 2017; Davidson et al., 2017; Fafchamps and Labonne, 2019). An interesting finding in this respect is that the interaction between social capital and voter density is stronger in more inland municipalities (not only in the south), which are typically more inward looking, isolated, and characterised by tighter social networks. Third, in a poorer context where workers are,

on average, less educated and more exposed to the risk of unemployment, they may be more dependent on politicians' handouts, more hesitant to challenge an incumbent, and more prone to fall prey to the pressures that having lower work opportunities entail.[17] Finally, organised crime has historically been more rife and had a stronger grip on society in southern Italy. To the extent that the interest of organised crime intertwines with politics, this may become another mechanism to exert control over voters. Further work will be necessary to explore this possibility.

6 Conclusions

In modern democracies the well-functioning of the electoral system is based on the secrecy of voting. This paper shows that when the secrecy of voting is at risk the outcome of elections in an advanced developed country is affected: electoral competition is reduced and incumbent politicians gain an advantage that helps their re-election.

The paper focuses on Italian municipal elections. In Italy, voters are assigned to specific polling stations depending on their address and after an election candidates have access to information about the number of votes they received in each polling station. This implies that when a candidate receives no votes in a given polling station, he or she knows with certainty that none of the people in a given area voted for him or her. Therefore, from the point of view of an individual voter the risk of being caught not voting for a candidate increases when the number of voters per polling station decreases. If candidates can retaliate against voters who do not vote for them, this affects voters' choices. In particular we hypothesise that such mechanism favours incumbent politicians who have greater retaliatory power. We formalise this intuition with a simple theoretical model.

---

[17]To some extent, our analysis can exclude this channel since it controls for the socio-demographic composition of cities, but this channel cannot be entirely excluded.

The results of the empirical analysis, based on all municipal elections conducted in Italy from 1989 to 2015, show strong support for this hypothesis. When voter density increases (i.e., the number of voters per polling station increases), and hence politicians' ability to detect voters' choices decreases, the probability of re-election for an incumbent falls. The result is robust to controlling for municipal fixed effects, time-varying municipal economic conditions, and the potential residual endogeneity of voter density with respect to unobserved determinants of re-election. The result is stronger in southern regions and more isolated areas where patronage politics is more common, social networks are tighter, and politicians therefore have greater retaliatory power (since their knowledge and control over the territory is more effective).

The results have important implications for policy makers. Concerns about the adverse impacts of the mechanism documented in this paper have inspired recent parliamentary action in Italy to pass legislation that would increase the size of electoral districts. This is the first paper, however, to empirically document the phenomenon. The conclusions may extend well-beyond the Italian context, as the same mechanism may be at play in other countries with a similar voting mechanism.

References

Akcigit, U., Baslandze, S., and Lotti, F. (2018). Connecting to Power: Political Connections, Innovation, and Firm Dynamics. NBER Working Papers 25136, National Bureau of Economic Research, Inc.

Andini, C. and Andini, M. (2018). Social capital and growth: causal evidence from italian municipalities. *Journal of Economic Geography*.

Batinti, A., Filippetti, A., and Andriani, L. (2017). Why Does Social Capital Increase Government Performance? The Role of Local Elections across Italian Municipalities. Management Working Papers 13, Birkbeck Department of Management.

Beath, A., Christia, F., Egorov, G., and Enikolopov, R. (2016). Electoral Rules and Political Selection: Theory and Evidence from a Field Experiment in Afghanistan. *Review of Economic Studies*, 83(3):932–968.

Bernini, M. and Tomasi, C. (2015). Exchange rate pass-through and product heterogeneity: Does quality matter on the import side? *European Economic Review*, 77(C):117–138.

Cruz, C., Labonne, J., and Querubín, P. (2017). Politician family networks and electoral outcomes: Evidence from the philippines. *American Economic Review*, 107(10):3006–37.

Davidson, M., Hicken, A., and Ravanilla, N. (2017). Family networks, clientelism, and voter behavior: Evidence from the philippines. University of California San Diego.

De Benedetto, M. A. (2014). Incumbency Advantage at Municipal Elections in Italy: A Quasi-Experimental Approach. Birkbeck Working Papers in Economics and Finance 1408, Birkbeck, Department of Economics, Mathematics & Statistics.

Elklit, J. (2018). Is voting in sweden secret? an illustration of the challenges in reaching electoral integrity. IPSA World Congress, University of Brisbane.

Fabbrini, S. (2000). Presidentialization of italian local government? the nature and the effects of semiparliamentarism. ECPR Joint Sessions, University of Copenhagen.

Fafchamps, M. and Labonne, J. (2019). Family networks and distributive politics. *Journal of the European Economic Association*.

Gagliarducci, S. and Nannicini, T. (2013). Do Better Paid Politicians Perform Better? Disentangling Incentives From Selection. *Journal of the European Economic Association*, 11(2):369–398.

Gagliarducci, S., Nannicini, T., and Naticchioni, P. (2010). Moonlighting politicians. *Journal of Public Economics*, 94(9-10):688–699.

Galasso, V. and Nannicini, T. (2015). So closed: Political selection in proportional systems. *European Journal of Political Economy*, 40(PB):260–273.

Galasso, V. and Nannicini, T. (2017). Political selection under alternative electoral rules. *Public Choice*, 171(3):257–281.

Gerber, A. S., Huber, G. A., Doherty, D., and Dowling, C. M. (2013). Is there a secret ballot? ballot secrecy perceptions and their implications for voting behaviour. *British Journal of Political Science*, 43(1):77–102.

Gulino, G. (2018). Electoral systems, selection and re-election. evidence from italian municipalities. University of Bergamo.

Khandelwal, A. (2010). The Long and Short (of) Quality Ladders. *Review of Economic Studies*, 77(4):1450–1476.

Knack, S. (2002). Social capital and the quality of government: Evidence from the states. *American Journal of Political Science*, 46(4):772–785.

Krebs, T. B. (1998). The Determinants of Candidates' Vote Share and the Advantages of Incumbency in City Council Elections. *American Journal of Political Science*, 42(3):921–935.

Nannicini, T., Stella, A., Tabellini, G., and Troiano, U. (2013). Social Capital and Political Accountability. *American Economic Journal: Economic Policy*, 5(2):222–250.

Norris, P. (2014). *Why Electoral Integrity Matters*. Cambridge University Press.

Norris, P., Wynter, T., and Cameron, S. (2018). Corruption and coercion: The year in elections 2017. Report, The Electoral Integrity Project.

Putnam, R. D. (1993). *Making Democracy Work: Civic Traditions in modern Italy*. Princeton University Press, Princeton, New Jersey.

Romer, D. H. and Frankel, J. A. (1999). Does Trade Cause Growth? *American Economic Review*, 89(3):379–399.

Rubino, M. (2018). Elezioni pulite, la legge approda alla camera. dalle urne trasparenti alle nuove regole per i fuori sede.

Trounstine, J. (2011). Evidence of a local incumbency advantage. *Legislative Studies Quarterly*, 36(2):255–280.

UN (1948). Universal declaration of human rights. *UN General Assembly*.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, volume 1. The MIT Press, 2nd edition.

Appendix

A   The electoral system for municipal elections in Italy

Up until 1993, municipalities were characterised by a parliamentary system coupled with a majoritarian electoral system in small municipalities with up to 5,000 inhabitants and a proportional electoral system in large municipalities with more than 5,000 inhabitants. Voters could cast a vote for a list (i.e., national party or local civic group) and, within this list, for a candidate for the municipal council. Among the competing lists, the seats were distributed in a majoritarian way in small municipalities and proportionally in large municipalities, and, within each list, candidates were ranked according to the total number of votes obtained. The elected municipal councillors had then the power to elect a mayor.

The reform of 1993 changed both the system of government and the electoral system used in Italian municipalities. Currently, Italian municipalities are characterised by a so-called semi-parliamentary system, which gives the possibility of voting both for a mayoral candidate and for a list of candidates for the municipal council (Fabbrini, 2000). In small municipalities with up to 15,000 inhabitants (more than 90% of Italian municipalities), each mayoral candidate is associated with one list of candidates for posts as municipal councillors. Voters are entitled to vote for a mayoral candidate and may cast a preference vote for a candidate for the municipal council. The candidate who gains the largest number of votes is elected mayor. Two thirds of the seats available are allocated to the list associated with the winning candidate. The remaining seats are distributed proportionally among the other lists and the candidates who have received the largest number of preference votes are elected.

On the other hand, in large municipalities, the mayoral candidate is linked with several lists of candidates for the municipal council and each candidate must formalise this link by means of an official declaration. Voters can vote for any mayoral candidate and can choose to vote for a list (that may or may not be associated with that candidate). In addition, they can express a preference for one of the candidates that are part of the list they have chosen to vote for. The mayoral candidate who gains the absolute majority of votes is elected mayor. Should no candidate achieve an absolute majority, a second ballot is held between the two candidates who received the most votes in the first round. These two candidates may link with other lists apart from the ones that supported them in the first round, which gives such lists the possibility of gaining some seats if the mayor they are associated with wins the second round. With regards to the election of the municipal councillors, their fate is not entirely tied (as it is in the small municipalities) to the election of their mayor. In fact, the distribution of seats depends on the votes gained by the lists in the first round, but such distribution of seats takes place only after the election of the mayor. Once the seats are allocated to the different lists, the candidates for the position of councillor who have gained the largest number

of preference votes are elected. For more detailed information on the electoral systems used in Italian municipalities, the reader can refer, among others, to Fabbrini (2000).

B A model of voter behaviour when the secrecy of voting is at risk

This section outlines a stylised theoretical framework to formalise the intuition that drives the analysis. The model describes the choice of an individual between keeping the promise made to an incumbent to vote for him/her, versus voting for an alternative candidate.[18]

Let U_i^I be the utility voter i obtains if he/she votes for the incumbent (I) and U_i^C the utility he/she obtains from voting for the most preferred challenger (C):

$$U_i^I = f(\theta_i^I) \tag{2}$$
$$U_i^C = f(\theta_i^C) - \Pi R. \tag{3}$$

If i votes for the incumbent, the resulting utility is an increasing function of the perceived quality of the incumbent (θ_i^I).[19] The utility from voting for the challenger is a function of the perceived quality of the challenger minus the expected value of the sanction the incumbent can impose on the voter if he/she finds out i did not vote for him/her.[20] For simplicity, we assume $f(\theta_i^I) = \theta_i^I$ and $f(\theta_i^C) = \theta_i^C$. This has no implications for our conclusions.

The sanction the incumbent can impose is a function of his/her retaliatory power (R) – which can be alternatively thought of as the favours he/she can withdraw – and of the probability that a voter is caught not voting for the incumbent (Π). For simplicity, the model assumes that the challenger has no retaliatory power. This is a simple way of capturing the fact that the incumbent has higher retaliatory power than the incumbent (though in practice the challenger's retaliatory power may be non-zero). Based on the vast literature on incumbency advantage and on the observation that having spent time in power is likely to have strengthened a politician's power in the community relative to potential challengers, it seems reasonable to assume that a challenger has higher retaliatory power.

From the perspective of voter i, the probability of being caught and sanctioned de-

---

[18]While there may be several candidates running for the election, for simplicity we model the choice between an incumbent and the most preferred alternative.

[19]θ_i is subjectively assessed and may be a function of the candidate's ability, political leaning, ideology, closeness to the voter's instances, etc. For the purpose of the model, it is not necessary to specify the precise factors driving θ_i.

[20]We assume that incumbents can impose sanctions whether or not they are elected. This assumption could be relaxed and the model could allow for sanctions to occur only if a politician arrives in office (or, at least, sanctions could be assumed to be stronger upon being elected). Such an extension would not change the main implications of the model. We believe the current assumption to be more realistic because a politician who is not re-elected can also impose sanctions when out of office. This is even more likely to be the case for incumbents who are not re-elected, since their time in power may have strengthened their power in the community. Such dynamics are documented by a vast literature on incumbency advantage, lending support to our approach.

pends on the likelihood that all other voters in the same polling station choose not to vote for the incumbent. That is because if a candidate receives no votes in a polling station, he/she knows with certainty that voter i did not vote for him/her. Voter i expects other voters to vote for the challenger with probability π.[21] This implies that the probability of being caught not voting for the incumbent is:

$$\Pi = (\pi)^{N-1}. \tag{4}$$

In this simple framework, voter i chooses to vote for the incumbent if the following condition holds:

$$\theta_i^I > \theta_i^C - \Pi R \tag{5}$$
$$\theta_i^I > \theta_i^C - \pi^{N-1} R. \tag{6}$$

This condition says that even a challenger who is perceived to be of higher quality than the incumbent (such that $\theta_i^C > \theta_i^I$) may not be voted if the incumbent has sufficient retaliatory power.

Crucially, the incumbent's advantage shrinks with N

$$\frac{\partial \left[(\pi^I)^{N-1} R \right]}{\partial N} = ln(\pi^I)(\pi^I)^{N-1} R < 0 \tag{7}$$

given that $0 < \pi^I < 1$. As N grows, the expected sanction the incumbent can impose decreases and the condition for i to choose the incumbent is less likely to be satisfied. In other words, the incumbent's advantage is progressively eroded when the number of voters per polling station increases.

Finally, the incumbent's advantages decreases more and more slowly as N grows. That is captured by the positive sign of the second derivative of the incumbent's advantage with respect to voter density (which implies that the first derivative becomes less negative with N):

$$\frac{\partial^2 \left[(\pi^I)^{N-1} R \right]}{\partial N^2} = (ln(\pi^I))^2 (\pi^I)^{N-1} R > 0 \tag{8}$$

In light of this framework, we hypothesise that voter density (i.e., the number of voters per polling station) reduces the advantage of incumbent politicians and the likelihood of their re-election, and that the effect should become weaker as density grows. The empirical analysis aims to test these hypotheses.

---

[21] In this stylised model we do not explicit articulate the expectation-formation process, but this is not crucial for the analysis. This setup is a simple way to convey the idea that a voter faces a lower risk of being in a polling station where an incumbent receives no votes when the number of voters in that polling station is lower. As discussed in the text, when the number of candidates is large, as it is typically the case in Italian elections, this risk is not negligible.

C First-stage results of IV regressions

Table 5: First-stage results for the log of voter density, province fixed effects

	Baseline	Social capital	
	(1)	(2)	(3)
Number of localities	-0.003***	-0.003***	-0.007***
	(-9.69)	(-9.08)	(-11.00)
Size, squared kilometres	-0.001***	0.002**	-0.036***
	(-5.69)	(2.40)	(-14.97)
Coefficient of variation of altitude	-0.070***	-0.071***	-0.067***
	(-5.50)	(-5.49)	(-3.69)
Size, squared kilometres x		-0.003***	0.040***
Social capital		(-3.61)	(14.89)
Individual controls	yes	yes	yes
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	yes
Municipality fixed effects	no	no	no
Year fixed effects	yes	yes	yes
Sanderson-Windmeijer F statistic	68.80	73.56	100.99
Kleibergen-Paap LM statistic	170.11	172.49	
Hansen J statistic	1.25	1.20	
Observations	756071	755341	755341

Notes: The dependent variable is the log of voter density. The regression in column (1) refers to the first stage of column (2) in Table 3. The regressions in columns (2) and (3) refer to the first stage of column (2) in Table 4. The individual controls include a dummy variable for gender, age and level of education. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. The Sanderson-Windmeijer F statistic refers to the multivariate F test of excluded instruments. The Kleibergen-Paap LM statistic refers to the test for underidentification. The Hansen J statistic refers to the test for overidentifying restrictions. t statistics in parentheses. ** and *** indicate coefficients significantly different from zero at 5% and 1% level respectively.

Table 6: First-stage results for the log of voter density, municipality fixed effects

	Baseline	Social capital	
	(1)	(2)	(3)
Regional voter density, log	0.195***	-0.162	-0.595***
	(4.24)	(-1.61)	(-7.53)
Regional voter density, log x		0.393***	0.856***
Social capital		(3.78)	(10.78)
Individual controls	yes	yes	yes
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	yes
Municipality fixed effects	no	no	no
Year fixed effects	yes	yes	yes
Sanderson-Windmeijer F statistic	18.00	16.33	17.05
Kleibergen-Paap LM statistic	18.09	16.15	
Observations	757044	756314	756314

Notes: The dependent variable is the log of voter density. The regression in column (1) refers to the first stage of column (3) in Table 3. The regressions in columns (2) and (3) refer to the first stage of column (3) in Table 4. The individual controls include a dummy variable for gender, age and level of education. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. The Sanderson-Windmeijer F statistic refers to the multivariate F test of excluded instruments. The Kleibergen-Paap LM statistic refers to the test for underidentification. t statistics in parentheses. *** indicates coefficients significantly different from zero at 1% level.

D Effect of voter density on the election of mayors

Table 7: Probability of being re-elected, mayors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	0.011*	0.009	-0.811**
	(1.69)	(0.32)	(-2.06)
Male	0.072***	0.072***	0.071***
	(8.31)	(8.29)	(6.60)
Age	-0.010***	-0.002***	-0.003***
	(-44.63)	(-44.45)	(-40.21)
Education	-0.004***	-0.004***	-0.004***
	(-6.20)	(-6.10)	(-5.17)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	128.35	127.99	420.87
Cragg-Donald F statistic		883.58	77.49
Kleibergen-Paap F statistic		135.38	27.22
Observations	47,194	47,123	47,174

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. t statistics in parentheses. *, ** and *** indicate coefficients significantly different from zero at 10%, 5% and 1% level respectively.

Table 8: Probability of being re-elected and social capital, mayors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	0.012	-0.013	-1.853**
	(1.52)	(-0.45)	(-2.45)
Voter density, log x	-0.001	0.020*	1.119**
Voter turnout in 1987	(-0.23)	(1.95)	(2.49)
Male	0.072***	0.072***	0.072***
	(8.31)	(8.27)	(6.65)
Age	-0.010***	-0.010***	-0.012***
	(-44.63)	(-44.42)	(-40.30)
Education	-0.004***	-0.004***	-0.004***
	(-6.15)	(-6.05)	(-5.14)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	124.99	124.82	102.67
Cragg-Donald F statistic		685.11	34.80
Kleibergen-Paap F statistic		105.01	12.06
Observations	47,148	47,077	47,128

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The variable social capital is proxied by the municipality-level turnout at the 1987 national elections. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. Standard errors are clustered at the municipality level. t statistics in parentheses. *, ** and *** indicate coefficients significantly different from zero at 10%, 5% and 1% level respectively.

E   Additional robustness checks

This section presents a set of additional robustness checks.

Table 9 shows that the effect of voter density on the probability of re-election for councillors is non-linear.[22] While the effect is negative for all values of voter density, the effect becomes less negative (smaller in absolute value) for larger values of voter density. This is equivalent to the positive second-order derivative of the stylised model presented in Appendix B.

Next, we check whether the mechanism we have hypothesised becomes stronger as the average number of votes candidates can expect to get per polling station becomes smaller. Therefore, in Table 10, we include an interaction term between the log of voter density and the number of lists of candidates per 1000 voters. The number of lists is directly related to the number of candidates as each list usually has as many candidates as the number of councillors in each municipality. It should be noted that we restrict the sample to southern regions only as they exhibit the lowest values of social capital and, thus, the highest impact of voter density on the probability of re-election (see Figure 2). The interaction term between voter density and the number of lists is negative and statistically significant in all cases. This suggests that the negative effect of voter density on the probability of re-election for councillors becomes stronger as the number of lists and candidates increases and, thus, as candidates are expected to obtain fewer votes per polling station.

We run additional robustness checks for changes in the electoral system by restricting the years included in our sample. Since the system of municipal government and the electoral system was modified in 1993, we check that our results are robust to the exclusion of all previous years. Moreover, it should be noticed that the number of polling stations was greatly reduced (by about 30% on average) after the changes introduced in Law 27 December 1997, n. 449. This law was passed at the national level as a cost-cutting measure and the majority of municipalities implemented it by 1998. As there is the possibility that the implementation of the 1997 law by municipalities is endogenous, we check the robustness of our results by excluding all years before 1998. All our main results are robust to these checks and the additional results are available upon request.

As a final piece of analysis, we use data at the level of each polling station in a given municipality to test our mechanism. Unfortunately, electoral results at the level of individual polling stations are not compiled in a national dataset, but some municipalities post the data online. Since this analysis is purely illustrative, we decide to concentrate on the region where the impact is strongest, i.e., Sicily, and our choice of municipality is

---

[22]In order to construct valid instruments for the squared term of the log of voter density, our endogenous regressor, we follow the procedure suggested in Wooldridge (2010). The instrument is constructed by taking the square of the predicted values of the first stage of the log of voter density on all the basic instruments and their squared terms. The basic instruments change depending on whether the regression adds province or municipality fixed effects, as described in Section 4.

Table 9: Probability of being re-elected, non-linear, councillors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	-0.275***	-1.110***	-8.790***
	(-6.86)	(-6.86)	(2.88)
Voter density, log, squared	0.020***	0.086***	0.657***
	(6.10)	(6.35)	(2.84)
Male	0.093***	0.093***	0.090***
	(57.31)	(56.59)	(52.75)
Age	-0.002***	-0.002***	-0.003***
	(-41.63)	(-41.94)	(-45.18)
Education	0.007***	0.007***	0.006***
	(40.70)	(40.49)	(38.19)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	483.13	478.67	405.65
Cragg-Donald F statistic		5220.34	409.54
Kleibergen-Paap F statistic		14.99	7.54
Observations	757,045	756,071	757,044

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The municipality controls include the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. The instrument for the squared term of the log of voter density is constructed as the square of the predicted values of the first stage of the log of voter density on all the instruments and their squared terms (Wooldridge, 2010). Standard errors are clustered at the municipality level. t statistics in parentheses. *** indicates coefficients significantly different from zero at 1% level.

limited to cities where the number of polling stations is sufficiently high, i.e., sufficiently large cities. Given these constraints, we focus on Catania, a large provincial capital of about 300,000 residents, 335 polling stations and with readily available data by polling station for the 2013 municipal elections.

To test our mechanism within a municipality, we run the following OLS regression:

$$sharevotes_{cs} = \gamma_I \ incumbent_c \times voters_s + \gamma_C \ incumbent_c + \gamma_V \ voters_s + \varepsilon_{cs}, \quad (9)$$

where $sharevotes_{cs}$ is the share of votes obtained by candidate c in polling station s, $incumbent_c$ is a dummy equal to one if a candidate is an incumbent and $voters_s$ is the log of the number of voters in each polling station. We also check the robustness of this result by including candidate and polling station fixed effects. In this case, the coefficients on the incumbent dummy and the number of voters by themselves are not

Table 10: Probability of being re-elected and number of lists, councillors

	OLS - Prov FE (1)	IV - Prov FE (2)	IV - Munic FE (3)
Voter density, log	-0.035***	-0.050	0.082
	(-4.73)	(-0.96)	(0.34)
Voter density, log x	-0.004***	-0.033***	-0.142**
No of lists per 1000 voters	(-2.66)	(-3.43)	(-2.38)
Male	0.121***	0.120***	0.119***
	(37.32)	(37.06)	(35.80)
Age	-0.003***	-0.003***	-0.003***
	(-26.81)	(-27.04)	(-30.26)
Education	0.007***	0.007***	0.007***
	(24.10)	(23.43)	(23.38)
Municipality controls	yes	yes	yes
Province fixed effects	yes	yes	no
Municipality fixed effects	no	no	yes
Year fixed effects	yes	yes	yes
F statistic	164.80	160.29	158.48
Cragg-Donald F statistic		1605.81	669.20
Kleibergen-Paap F statistic		25.45	2.60
Observations	244,034	244,034	244,034

Notes: The dependent variable is a dummy equal to one if an elected official is re-elected. The municipality controls include the number of lists per 1000 voters, the log of the number of voters, the percentage of the population less than 19, the percentage of the population over 65, average years of schooling in the population, employment rate and unemployment rate. The sample is restricted to southern regions only, i.e., Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria, Sicily and Sardinia. Standard errors are clustered at the municipality level. t statistics in parentheses. ** and *** indicate coefficients significantly different from zero at 5% and 1% level respectively.

identified and, thus, we estimate the following regression:

$$sharevotes_{cs} = \gamma_I \, incumbent_c \times voters_s + \phi_c + \psi_s + \varepsilon_{cs}, \qquad (10)$$

where ϕ_c represents candidate fixed effects and ψ_s represents polling station fixed effects. We expect that in smaller polling stations, where the anonymity of voting is more likely to falter, an incumbent will obtain a relatively larger share of votes. Looking at Table 11, we find some support for this hypothesis, since the interaction term between being an incumbent and the number of voters in a polling station is negative, even when candidate and polling station fixed effects are included. As discussed above, however, this is only an illustration of the mechanism of interest, which may vary across different cities.

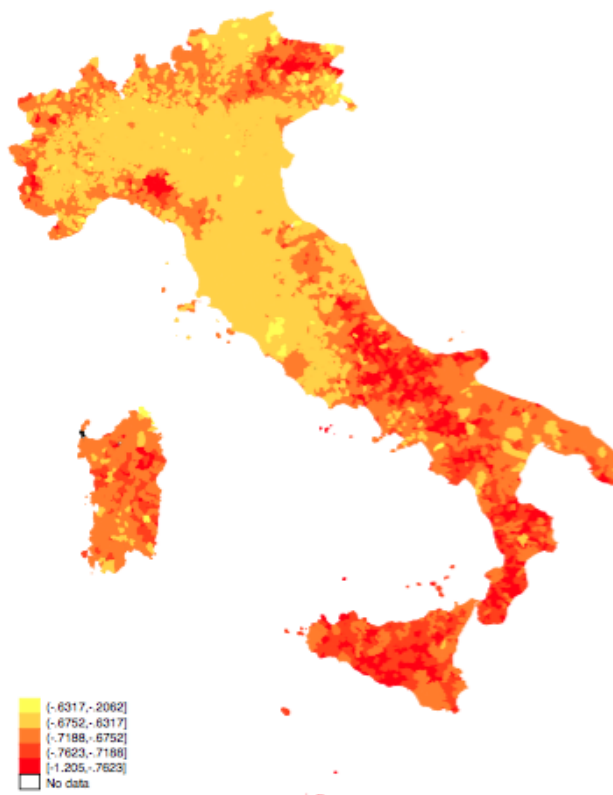**Table 11:** Probability of being re-elected in Catania, councillors

	OLS (1)	FE (2)
Incumbent dummy	-0.0072***	
	(12.56)	
Number of votes (polling station), log	-0.0000	
	(0.67)	
Incumbent dummy ×	-0.0003***	-0.0003***
Number of votes (polling station), log	(-3.44)	(-3.64)
Candidate fixed effects	no	yes
Polling station fixed effects	no	yes
F statistic	2236.81	13.22
Observations	215,070	215,070

Notes: The dependent variable is the share of votes of each candidate in each polling station. t statistics in parentheses. *** indicates coefficients significantly different from zero at 1% level.

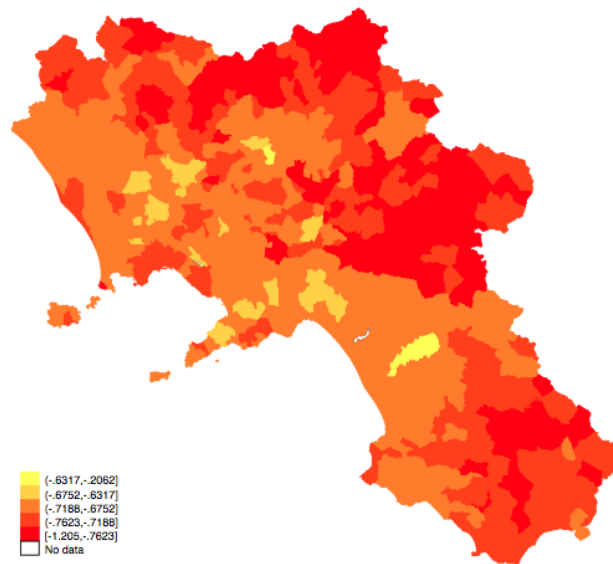# F   Heterogeneous effect by quality of institutions

The figures in this section plot the results of the regression in Table 4, where we interact voter density with a proxy for social capital at the municipal level. This allows us to show how the effect of voter density on an incumbent's probability of re-election changes across municipalities with variations in social capital.

Figure 2: Impact of voter density on probability of re-election by municipality

Notes: The figure is based on the estimation results in column (3) of Table 4, which includes an interaction between voter density and a proxy of social capital at the municipal level (measured by voter turnout in the 1987 national political elections, following Andini and Andini (2018)). The figure shows how the estimated effect of voter density varies across municipalities with changing levels of social capital. The dark red areas refer to values of the coefficient smaller than the average minus 1.2 standard deviations ($\leq$ -0.7623). The orange-red areas refer to values of the coefficient between the average minus 1.2 times the standard deviations and the average minus 0.4 times the standard deviation ($>$ -0.7623 & $\leq$ -0.7188). The orange areas refer to values of the coefficient between the average minus 0.4 times the standard deviations and the average plus 0.4 times the standard deviation ($>$ -0.7188 & $\leq$ -0.6752). The golden areas refer to values of the coefficient between the average plus 0.4 times the standard deviations and the average plus 1.2 times the standard deviation ($>$ -0.6752 & $\leq$ -0.6317). The yellow areas refer to values of the coefficient larger than the average plus 1.2 times the standard deviation ($>$ -0.6317).

Figure 3: Impact of voter density on probability of re-election, Campania

Notes: The figure is based on the estimation results in column (3) of Table 4, which includes an interaction between voter density and a proxy of social capital at the municipal level (measured by voter turnout in the 1987 national political elections, following Andini and Andini (2018)). While the estimation uses data for the whole of Italy, the map is restricted to municipalities in the region of Campania. The figure shows how the estimated effect of voter density varies across municipalities with changing levels of social capital. The dark red areas refer to values of the coefficient smaller than the average minus 1.2 standard deviations ($\leq$ -0.7623). The orange-red areas refer to values of the coefficient between the average minus 1.2 times the standard deviations and the average minus 0.4 times the standard deviation ($>$ -0.7623 & $\leq$ -0.7188). The orange areas refer to values of the coefficient between the average minus 0.4 times the standard deviations and the average plus 0.4 times the standard deviation ($>$ -0.7188 & $\leq$ -0.6752). The golden areas refer to values of the coefficient between the average plus 0.4 times the standard deviations and the average plus 1.2 times the standard deviation ($>$ -0.6752 & $\leq$ -0.6317). The yellow areas refer to values of the coefficient larger than the average plus 1.2 times the standard deviation ($>$ -0.6317).

EconPol Europe

EconPol Europe – the European network for economic and fiscal policy research – is a network of 14 policy-oriented university and non-university research institutes across 12 countries, who contribute scientific expertise to the discussion of the future design of the European Union. The network's joint interdisciplinary research covers sustainable growth and best practice, reform of EU policies and the EU budget, capital markets and the regulation of the financial sector, and governance and macroeconomic policy in the European Monetary Union.

The network was founded in spring 2017 by the ifo Institute, along with eight renowned European research institutes. A further five associate partners were added to the network in January 2019.

Our mission is to contribute our research findings to help solve the pressing economic and fiscal policy issues facing the European Union, and to anchor more deeply the idea of a united Europe within member states.

With our cross-border cooperation on fiscal and economic issues, EconPol Europe promotes growth, prosperity and social cohesion in Europe. In particular, we provide research-based contributions to the successful development of the European Economic and Monetary Union (EMU).

Our joint interdisciplinary research covers:

– Sustainable growth and best practice
– Reform of EU policies and the EU budget
– Capital markets and the regulation of the financial sector
– Governance and macroeconomic policy in the European Monetary Union

We will also transfer our research results to the relevant target groups in government, business and research, as well as to the general public.