Invalid Ballots and Electoral Competition*

Gani Aldashev† and Giovanni Mastrobuoni‡

August 2, 2016

Abstract

In close elections, a sufficiently high share of invalid ballots - if driven by voter mistakes or electoral fraud - can jeopardize the electoral outcome. We study how the closeness of electoral race relates to the share of invalid ballots, under the traditional paper-ballot hand-counted voting technology. Using a large micro-level dataset from the Italian parliamentary elections in 1994-2001, we find a strong robust negative relationship between the margin of victory of the leading candidate over the nearest rival and the share of invalid ballots. We argue that this relationship is not driven by voter mistakes, protest, or electoral fraud. The explanation that garners most support is that of rational allocation of effort by election officers and party representatives, with higher rates of *detection* of invalid ballots in close elections.

Keywords: vote counting; invalid ballots; election officers; party representatives.
JEL classification codes: D72; D72 - Political Processes: Rent-Seeking, Lobbying, Elections, Legislatures, and Voting Behavior D73. Bureaucracy; Administrative Processes in Public Organizations; Corruption;

*We thank Enriqueta Aragones, Georg Kirchsteiger, Jim Snyder, and seminar participants at Collegio Carlo Alberto and the University of Padua for useful comments.

†Corresponding author. Universitè libre de Bruxelles (ULB), ECARES, and CRED, University of Namur. Mailing address: 50 Avenue F Roosevelt, CP 114, 1050 Brussels, Belgium. Email: gani.aldashev@ulb.ac.be .

‡Department of Economics, University of Essex and Collegio Carlo Alberto. Mailing address: Wivenhoe Park, CO43SQ Colchester, UK. Email: gmastrob@essex.ac.uk.

"It's not the voting that's democracy, it's the counting." (Stoppard, 1972)

The heated debate in policy-making and academic circles following vote-counting problems at the 2000 U.S. Presidential elections (Ansolabehere and Stewart III, 2005, Caltech and MIT Voting Technology Project, 2001, Card and Moretti, 2007, Knack and Kropf, 2003, Wand et al., 2001) indicates that the organization of ballot casting and counting is a fundamental issue at the heart of democratic elections. During the last decades, numerous countries (including the United States, Brazil, Estonia, France, and Switzerland) carried out important policy changes regarding how voting process and vote counting is organized. Under the Help America Vote Act (HAVA), for instance, the United States invested around 3 billion USD into improving and replacing voting technologies (Stewart III, 2011), replacing punchcard- and lever-based systems with electronic voting machines. Contrarily, the Netherlands chose to abolish electronic voting technologies and to move back to paper ballots, following the technical problems with voting machines in 2006 and the subsequent public debate (Jacobs and Pieters, 2009).

In general, in the majority of large elections, there is a relatively small fraction of votes that is counted as invalid. When electoral race is tight, even a small number of votes can make a difference for the electoral outcome, and thus the importance of invalid ballots increases disproportionately. Intuitively, if the share of invalid ballots is sufficiently high as compared to the margin of victory of the winning candidate, and if the presence of invalid ballots is driven by voter mistakes or electoral fraud, then such ballots might seriously undermine the correct functioning of the electoral system.

The validity of this common-sense intuition crucially depends on the origin of invalid ballots and the relationship between the share of invalid ballots and electoral competition. In this paper, we study how the closeness of electoral race relates to the share of invalid ballots under the traditional paper-ballot hand-counted voting technology. We rely on a large micro-level dataset from the Italian parliamentary elections in 1994-2001 (the period during which three-quarters of the lower chamber of the Italian Parliament was elected using first-past-the-post system). Italian elections in this period present a natural setting for studying the above question, both because of the specific details of the organization of ballot counting (as explained below) and the availability of highly disaggregated electoral data containing rich variation in political behavior and socioeconomic variables, which

allows for testing between alternative theoretical explanations.

We document a strong robust negative correlation between the margin of victory of the leading candidate over the nearest rival and the share of invalid ballots. We then investigate the possible theoretical explanations for this relationship and argue, on the basis of econometric evidence, that this relationship is unlikely to be driven by voter mistakes, protest, or electoral fraud. The explanation that garners most support is that election officers and party representatives rationally allocate more effort in detecting invalid ballots when the stakes are highest, i.e. when the electoral race is closer. In other words, the relationship that we document corresponds to higher rates of *detection* of invalid ballots in closer elections.

To the best of our knowledge, this is the first paper that documents and analyzes the relation between electoral competition and invalid ballots. Numerous papers in economics and political science (e.g. Ansolabehere and Stewart III, 2005, Card and Moretti, 2007, Dee, 2007, Fujiwara, 2015, Shue and Luttmer, 2009) have studied the electoral outcomes, including the number of invalid or residual votes, under different voting technologies.[1] However, what matters for electoral outcomes is not so much the average level of ballots counted as invalid under different technologies, but whether the number of (truly) invalid ballots increases or decreases when the electoral race becomes tighter. Having a substantial fraction of invalid ballots or misvotes in a landslide election clearly matters less (as noted by Dee (2007) and Shue and Luttmer (2009) for the California recall election in 2003) than a much smaller fraction of invalid ballots in a close election (as, for instance, in the case of misvotes in Palm Beach County, Florida, during the 2000 Presidential election). This issue is exactly the focus of our paper.

Another related strand of literature studies vote buying and ballot rigging (see the collection of papers in Schaffer 2007 and Lehoucq 2003 for a good survey). In a study of Chilean elections before 1965, Baland and Robinson (2008) find that the introduction of the secret ballot in 1958 had effectively destroyed the "market" for votes that existed between landed aristocracy (which controlled the votes of its agricultural workers) and the right-wing parties, thus sharply decreasing the votes for the right-wing parties. Lehoucq and Molina (2002) study the accusation of ballot-rigging filed in Costa Rica, and find that such accusations where substantially more numerous in close-race districts. This

[1] See Stewart III (2011) for an excellent survey.

finding does not, however, imply that fraud is more frequent in close-race districts: it might as well be that, holding the number of rigged ballots constant, the higher incentives for parties to file an accusation of ballot-rigging were stronger in close-race districts. Thus, understanding whether the correlation between the fraction of invalid ballots and the closeness of electoral race is driven by fraud is a key question. The additional contribution of this paper lies in providing an answer to this query, using the case of Italian parliamentary elections.

1 Invalid Ballots: What Are They?

In any election based on a paper ballot system, all ballots cast by voters belong to one of the three categories: valid, blank (the voter did not express any preference), or invalid (the election officers consider that the voter did not express her preference correctly). Under the traditional hand-counted paper ballot system, after the vote count is completed, the election administration reports the number of ballots belonging to each category. A ballot can be considered invalid for different reasons: for instance, the voter over-votes (i.e. casts more than one preference when only one preference is allowed) or takes an action that undermines the secrecy of the vote (e.g. signs the ballot). The duty of an election officer is to invalidate any ballot on which the voter does not uniquely and anonymously identify her preference. These rules apply to most, if not all, democratic elections using the paper-ballot system. The stated objective of this procedure is to avoid antidemocratic and illegal voting behavior.

The Institute for Democracy and Electoral Assistance, which maintains a database on parliamentary and presidential elections across most countries in the world, reports the share of invalid votes for about 100 countries over the last 10 years. The average share of invalid ballots is around 3 percent. However, looking across countries (see Figure 1), one sees a large variation in this measure. In all the developed countries the share of invalid ballots is a single-digit number, typically below 5 percent. The number is much higher for the developing countries, with double-digit numbers in several developing countries, in particular in Latin America and Western Africa.

Political scientists have tried to link the variation in the share of invalid ballots to some principal characteristics of the political system. Power and Garand (2007) analyze,

using an aggregate-level panel-data analysis from 80 legislative elections held in 18 Latin American democracies between 1980 and 2000, the influence of three sets of factors on the number of invalid ballots: socio-demographic (literacy, education, wealth), institutional (electoral system and ballot structure), and political (alienation and protest). They find some support for all the three sets of factors: socioeconomic factors (urbanization and income inequality) correlate with the number of invalid votes, institutional factors (compulsory voting, electoral disproportionality, and the combination of high district magnitude with personalized voting) increase the number of blank and spoiled ballots, whereas political factors such as political violence and the level and direction of democratic change also correlate with the share of invalid votes. Uggla (2008) also conducts an aggregate study by looking at over 200 elections in Western Europe, Australia, New Zealand, and the Americas in the 1980-2000 period. He finds support for the hypothesis that the variation in the share of invalid ballots reflect the voters' reaction to the perceived absence of political choice.

The key problem with these aggregate studies is that one cannot rule out the influence of some unobserved factors (e.g. political culture) that influences simultaneously the number of invalid ballots and the institutional factors. We are able to overcome this problem by using highly disaggregated data from a setting with homogeneous formal political institutions, but with a sufficiently large variation (both across time and space) in political behavior.[2]

2 Context and Data

2.1 Organization of Italian Parliamentary Elections

In this section, we describe the political and institutional context from which our data comes. We analyze electoral data from the three Italian parliamentary elections (1994, 1996, and 2001), during which three-quarters of the Chamber of Deputies (the lower chamber) was elected through the first-past-the-post majoritarian system, in 475 uni-

[2]Two other papers (McAllister and Makkai, 1993, Power and Roberts, 1995) study within-country variation in invalid ballots (in Australia and Brazil, respectively). Only the former paper documents some relationship between the closeness of electoral race and the share of invalid ballots, which the authors of the study attribute to the differences in the socio-economic composition of close-race versus safe districts in Australia. Neither study attempts to distinguish between alternative explanations for the relationship between electoral competition and invalid ballots.

nominal electoral districts (explained in detail below). We restrict our dataset to this period because these elections exhibit a natural measure of the closeness of the electoral competition: the margin of victory between the candidates with the highest and the second-highest number of votes.

Our dataset has several advantages over similar data from other settings. First, our data is highly disaggregated: the unit of observation is a municipality (for the electoral districts that contain more than one municipality) or an electoral district (for large municipalities that contain more than one district); this substantially increases the power of statistical analysis. Second, it counts separately blank ballots and invalid ballots. This is important because under the traditional paper-ballot voting, a blank ballot clearly indicates intentional abstention by the voter, whereas in settings that use other technologies (such as, for instance, in most elections in the United States), it is difficult to separate clearly a voter's intention from the malfunctioning of the voting technology, in case of an empty ballot (see Ansolabehere and Stewart III, 2005). In such settings, the election data usually lumps together blank and invalid ballots into the category of "residual votes." Third, in Italian elections, invalid ballots represent a relatively small but non-negligible fraction of total votes. However, the quality of democratic institutions and voter literacy are high, and therefore it is unlikely that the majority of invalid ballots are driven by protest or voter mistakes (as, for instance, in several Latin American countries). Finally, while the electoral system is uniform throughout the country, there is rich geographic variation in the measures of electoral behavior and outcomes, both across large (regions and provinces) and small (municipalities and within-city electoral districts) administrative units.

For three legislatures - that started in 1994, 1996, and 2001 - Italian citizens elected their representatives using a two-tier system (75 percent of representatives via the majoritarian system and the remaining 25 percent via the proportional-representation system). Before 1994 and after 2001 the entire Italian electoral system was based on proportional representation.

On the election day, each voter received two ballots: one to cast a vote for a candidate in her single-member district, and another to cast a vote for a party list in her larger proportional district. Figure 2 shows a typical ballot, for the majoritarian and proportional-representation systems in Italy. In the districts with the majoritarian sys-

tem, the voter has to put a cross on the name of the candidate of her preference, whereas in the districts under the PR system, she has to put a cross on the party/coalition symbol.

475 out of the 630 House members were elected in single-member majoritarian-election districts, while the rest was elected from closed party lists in 26 multiple-member districts (with 2 to 12 seats per district). In our analysis, we focus on these 475 majoritarian-election single-member districts.

The polling stations during the elections operated in the following way. Parliamentary elections in Italy take place on a Sunday between 8 AM and 10 PM, and on the following Monday between 7 AM and 3 PM. As soon as the elections end (i.e. on Monday afternoon), election officers start counting the ballots. The counts typically last uninterrupted until late Monday night or, sometimes, up to Tuesday morning.

Each polling station has three types of election officers: the president of the polling station, a secretary, and three canvassers. Party list representatives (at most two individuals for each party list) can also assist the vote count at the polling station. At least three election officers, including the president or the vice-president (chosen by the president among the canvassers), have to be present through the entire count. The president of the polling station decides, after consulting with the canvassers, on the outcome of any disputes related to the vote count, including those about the validity of any particular ballot. She then registers her provisional decision (while the Parliament has the last word about official protests). The secretary keeps the official record about all the activities during the count. At the end of the counting all members of the polling stations sign the official record. Both the election officers and party list representatives can contest ballots, i.e. question the decision about the validity of any given vote.

Each election officer receives a monetary compensation of about 100 euros for her work at the polling station. In addition, both the election officers and party representatives are compensated by their employers with (at most) 3 days of paid leave.

Each ballot is scrutinized by all the canvassers, the president, and - if present - by party representatives. The number of valid votes for each candidate is marked (typically on a board) and gets regularly updated as the count proceeds. This implies that people present at the count can observe the evolution of the number of votes for each candidate and have a perception of the margin (at their polling station) between the leading candidate and the nearest rival.

Given that voters are instructed to put just one sign ("x") on the ballot (and no other mark is allowed), detection of any visible irregularity on the ballot (e.g. more than one preference expressed, a signature, an additional mark made by mistake, etc.) implies that the president declares the ballot as invalid. However, given the large number of ballots that have to be scrutinized and the fact that often the count continues late into the Monday night, guaranteeing that each ballot containing an irregularity gets detected is difficult.

According to the Electoral Law, the president is a public official and has the authority to arrest those who disturb the voting process, including the party representatives and the other canvassers. The president, the secretary, and the canvassers are all obliged to denounce any criminal act related to the voting and the vote counting. If the president is involved in such acts, the secretary and the canvassers are supposed to contact the judicial authorities. Public officials who deliberately alter the ballots or the final counting face the risk of being punished with up to 8 years of jail.

2.2 Data

We extracted the information on three majoritarian parliamentary elections from the Historical Electoral Atlas of Italy (Corbetta and Piretti, 2008). An observation in our dataset represents the smallest level of aggregation of polling stations available, that is the smallest unit between a municipality and a district (hereafter, we refer to this as 'the electoral unit'). Larger cities have several districts: for instance, the municipality of Rome has 24 and the municipality of Milan has 11. These cities are rather exceptional and represent only 1.4 percent of our observations. We can thus think of an electoral unit in our dataset as a municipality.

Next, we matched each municipality to its (province-level) total crime rate, corruption crime rate, organized crime rate, and the share of city councils in the province that were dissolved for Mafia infiltration.[3]

Finally, we complemented the municipality-level data with province-level measures of education, turnout at national referenda, labor activity rate, unemployment rate, GDP

[3]These data are only available for the years 2004 to 2009. Since the variation across municipalities explains 86 percent of the total variation in crime rate, we use the average crime rates across all years as a measure of crime rate for the three election years 1994, 1996, and 2001.

per capita, and the rate of urbanization.

Each polling station is responsible for 500 to 1,200 eligible voters, though there can be exceptions to this rule for isolated areas that are difficult to reach. While in general we do not know whether an electoral unit corresponds to a polling station, electoral units with an electorate size smaller than 1,200 (which represent 32 percent of electoral units) are likely to be single polling stations, because of the rule noted above. We perform most of our analysis using all the electoral units (so as to avoid any selection bias coming from the size of the electorate). However, we also present results of the analyses that uses only the electoral units that are likely to be single polling stations (i.e. municipalities with less than 1,200 voters).

Table 1 presents the summary statistics for the key variables used in our empirical analysis. Overall, our dataset contains 8,224 electoral units for the three election years, which gives a (slightly unbalanced) panel of 23,109 observations. Our main variable of interest, i.e. ballots reported as invalid, represent a non-negligible fraction of votes. On average, in a typical district or municipality, 3.9 percent of all the ballots is reported as invalid. We also see that there is substantial variation in this measure: the standard deviation is 2.2 percent. A slightly higher fraction of ballots (4.6 percent) is cast as blank. Turnout rate is relatively high (which is a traditional characteristic of Italian elections): in a typical district, 82 percent of all eligible voters participate. The average leading margin (i.e. the difference in the share of votes between the candidate with the highest number of votes and the nearest rival) is substantial; at the electoral unit level it is 18.4 percent; however, the variation is large (the standard deviation is 14.8 percentage points). The electoral unit with the smallest leading margin in our dataset exhibits a vote difference of zero percent, whereas in the one with the largest margin, the first candidate leads by 96.1 percent.[4]

The main two party coalitions (center-left and center-right) lead the electoral competition in most electoral units. The center-left coalition leads in 36.2 percent of electoral units, whereas the center-right coalition leads in 38 percent of cases. We observe a strong party incumbency effect: in 94.4 per cent of cases, the candidate from the incumbent

[4]Ideally, we would like to conduct our analysis using the measures of ex-ante or expected margins of victory. Unfortunately, such measures are not available. Nevertheless, the measures of ex-post or realized margins of victory are good proxies of ex-ante margins, under the assumption that relevant political actors (voters, election officers, party representatives, etc.) hold rational expectations.

coalition is the one with the highest number of votes. The incumbent politician effect is, however, much smaller: in 46.5 percent of cases, the voting ballot contains the name of the incumbent politician, and in slightly more than half of these cases (or in 26.4 percent of the total), the incumbent politician leads. The number of candidates also varies across districts. On average, a voter is confronted with a ballot containing 4.14 candidates.

Figure 3 indicates that the share of invalid ballots varies substantially across different regions and provinces. Southern regions exhibit the highest levels of invalid ballots, followed by the North-West. The Northern and Central regions exhibit the lowest levels of invalid ballots. While North-South divide is large, there is also substantial within-region (province-level) variation, both in the North and the South of Italy.

To match our data against the hypothesis of invalid ballots being driven by electoral fraud, we will also use measures of the rates of prevalence of different types of crime. Figure 4 shows the crime patterns across Italian provinces. Interestingly, it indicates that there no clear North-South divide in crime rates. The corruption-related and organized crime is more common in the Southern and North-Western provinces. Another measure of Mafia infiltration in public offices is the share of city councils dissolved: these are essentially concentrated in Sicily, Calabria, and Campania regions. Finally, the total crime rate (most likely being driven by economic crimes) is higher in the North and the Center-South of Italy.

3 Invalid Ballots and Electoral Competition

The left panel of Figure 5 presents graphically the relationship between the leading margin (the difference in the share of votes between the candidate with the highest number of votes and the nearest rival) and the fraction of ballots reported as invalid (out of the total number of ballots). Each dot represents the fraction of invalid ballots for a given percentile of the margin of victory. For the levels of leading margin that are close to zero, invalid ballots represent almost 4.5 percent of all votes, whereas to the largest percentiles of the leading margin correspond the lowest fractions of invalid ballots (around 2.5 percent). Overall, there is a clear negative correlation between the two variables: the larger is the leading margin, the smaller is the fraction of invalid ballots. This is also confirmed in column 1 of Table 2: the regression coefficient on the leading margin is negative and highly

statistically significant.[5] The effect is quantitatively important: one standard deviation increase in the leading margin (14.8 percentage points) corresponds to a reduction in the fraction of invalid ballots of about one-sixth of a standard deviation (0.35 percent of total votes). Restricting the analysis to the 2,704 electoral units with less than 1,200 eligible voters leads to very similar results (see the right panel of Figure 5 and Table 3).

4 Competing Explanations

What are the possible explanations for the empirical relationship that we have established above? Although direct tests of theories of political behavior that aim at explaining this pattern are not feasible (because they would require obtaining measures of individual behavior such as, for instance, the attention of voters, the effort of election officers and party representatives, etc.), we can rely on indirect tests that allow us to rule out some theories in favor of others.

The prime suspect is the behavior of voters. Consider a simple cost-benefit calculation of an individual voter. Suppose that filling out the ballot requires concentration, and filling it out correctly implies some attention cost. Moreover, suppose that the probability of making a mistake (and, therefore, submitting an invalid ballot) decreases with the attention allocated by the voter. On the benefit side, if the voter prefers one candidate over the other, she might perceive a benefit from feeling that her vote helped to increase the chances of victory of her preferred candidate. This might be justified by either the fact that a voter considers her probability of being pivotal (see Ch. 14 in Mueller 2003), or - more realistically - by the fact that the voter might feel the moral duty to "do her part" in helping her preferred party to win (as in the models by Feddersen and Sandroni 2006a,b).

The higher is the expected margin of victory of one of the candidates in the district, the lower is the voter's expected benefit of casting a valid vote. Given that the margin of victory does not affect the cost side, the higher is the margin, the lower is the attention that the voter devotes to casting a valid vote, and thus the higher is the probability of submitting an invalid ballot.

[5]This result is robust to using the logarithms of the variables instead of their levels on one or both sides of the regression equation.

Under this simple opportunity-cost theory of voter behavior, we would obtain a prediction that a higher margin of victory should be *positively* correlated with the fraction of invalid ballots. However, as we have seen above, the relationship is negative. Thus, this correlation cannot be explained by voters' attention.

Another explanation is based on voter protest. Voters might have (negative) feelings about the (lack of) choice that they are facing, and may act in the voting booth in reaction to these feelings. For instance, Brighenti (2003) analyzes a selection of invalid ballots in a regional election in Italy, and finds that a part of the invalid ballots report emotionally-charged (typically, negative) messages written by voters on their ballots. These are examples of voluntary invalidation.

If the expected margin of victory is sufficiently large, some of the voters that support the losing candidate might feel that their electoral choice is unfairly constrained. If this triggers negative emotions in them, some of the voters might voluntarily invalidate their ballots. A related possibility is that of expressive voting (Brennan and Lomasky, 1993, Schuessler, 2000). If voters want to express their general discontent about the political system, they might want to cast an invalid ballot as a protest. At the same time, each voter might have a political preference for some party. The closer is the electoral race, the higher is the opportunity cost of invalidating the ballot to express one's discontent. Under this theory as well, we should observe a positive correlation between the margin of victory and the fraction of invalid ballots (which goes against what we find in the data).

We can further refine our analysis, if we consider that ballot invalidation is not the only way of expressing one's protest. Some of the voters might express their feelings by leaving their ballots blank. Then, the fraction of invalid ballots and that of blank ballots should be correlated. Based on this intuition, we can use the fraction of blank ballots as a regressor, so as to capture, at least in part, the voters' protest.

Our empirical results show that this explanation based on voter protest does not fit the data. First, in columns 1 and 2 of Table 2 and of Table 3, the coefficient on the leading margin is negative (which is the opposite to the prediction of this theory). Second, while blank ballots correlate positively with the fraction of invalid ballots, even when we include the blank ballots and the number of candidates in the regression (in column 2 of both tables), the negative coefficient on the leading margin remains highly statistically significant.

In column 3, we verify how robust our main finding is to the inclusion of a series of local (province-level) characteristics that can affect the benefits and costs for voters of casting valid ballots. These characteristics include the level of education (measured by the fractions of the population with university and high-school degrees), social capital (measured by the average turnout at national referenda), income (measured by the labor activity rate, the unemployment rate, and GDP per capita), crime (the incidence of corruption, of organized crime, and total crime cases per 100 inhabitants), as well as the rate of urbanization. While some of these variables capture a part of the variation in invalid ballots, the coefficient on the leading margin remains highly significant. In column 4, we perform an even more stringent test, by adding year- and electoral-unit fixed effects. This means that the remaining variation in the share of invalid ballots is within the same electoral unit across the three elections. The coefficient on the leading margin remains basically unchanged (both in terms of size and significance). In other words, if the leading margin increases from one election to the other within the same electoral unit, the fraction of invalid ballots reported at this unit significantly decreases. Given that the coefficient on leading margin is unchanged while the R-squared increases to about 70 percent there is fairly little scope for unobservable characteristics to explain our results (Altonji et al., 2005, see).

Another possibility is that closer races lead to an increase in voter turnout. If the marginal voters are less educated and more prone to invalidate the ballot the effect that we observe could be driven by voter selection. While all the regressions control for the linear term of turnout, the selection might easily generate non-linear effects. To control for selection in Column 1 of Table 4 we add the cubic term in turnout and in Column 2 we interact the cubic term with the fraction of the population with high school and with a university degree. The coefficients on the leading margin are almost unchanged (whereas if the voter selection was the only driver behind our observed effect, we would expect the coefficients on leading margin in Columns 1 and 2 to go to zero). Columns 3 and 4 show that while the margin of victory effects are larger in the South they also present in the North. We see this as indirect evidence that fraud, which is more widespread in the South, is unlikely to be the major driver, though later we are going to devise more stringent tests.

If the explanation for the negative correlation that we observe does not come from

voter behavior, it plausibly comes from the behavior of those who count the votes, i.e. election officers. Let's suppose that election officers act rationally. Given that they are called to act as public officials to ensure that all the ballots cast are counted correctly (which, in particular, includes detecting ballots that are cast incorrectly), we can formalize the problem of an election officer as follows.

Suppose that each officer considers all the ballots that have to be counted, one by one. Each ballot that she scrutinizes can be either valid or invalid. The objective of the election officer is to minimize the likelihood that the victory is incorrectly adjudicated to the candidate that, in reality, has fewer valid ballots in her favor. However, the officers might make mistakes. There are two types of error that the officer might commit. Type I error consists in invalidating a truly valid ballot. Type II error consists in counting as valid a ballot which is in reality invalid. Given that the type I error is very unlikely to happen (it is virtually impossible to detect a non-existent irregularity in a ballot which has been correctly filled out by the voter), we can assume such errors away. Instead, the type II error - missing an existing irregularity - is much more important. Moreover, the likelihood of this error is affected by the effort that the officer exerts. These type-II errors might jeopardize the true outcome of the elections if they are sufficiently numerous as compared to the difference in the number of valid votes between the two candidates.[6]

In other words, the officer exerts the effort of attention to minimize the number of type-II errors. However, the effort is costly, and the higher is the number of ballots to scrutinize, the higher is the marginal cost of effort. On the other hand, the risk of jeopardizing the election outcome depends on the expected margin of victory: the larger is this margin, the less it is likely that a given number of type-II errors influence the election outcome.

The rational officer chooses the level of effort that equates the marginal cost of effort to its marginal benefit. If the expected margin of victory increases, the benefit of effort falls, and thus the officer puts lower effort. This, in turn, implies a lower number of truly invalid ballots that are counted as invalid. Thus, this theory can explain the negative correlation between the margin of victory and the fraction of invalid ballots: election officers rationally

[6]This implicitly assumes that the election officer acts taking into account the worst-case scenario: that, if let pass, all the invalid ballots are counted as votes for the same candidate. While this assumption is unrealistic in its pure form, our reasoning remains valid as far as the election officer considers as possible a scenario sufficiently close to the worst case.

allocate effort, depending on the (expected) closeness of elections, which implies higher rates of *detection* of invalid ballots in relatively closer races.

Let's now consider the effect of a variation in turnout. Given that the number of election officers is fixed (i.e. it is not adjusted on the basis of turnout), a higher number of voters showing up at the polls implies a higher number of ballots that each officer has to scrutinize. This means (under the standard convexity assumption on the cost function) that the incremental cost of effort increases. The officer then finds it optimal to reduce the effort that she puts in scrutinizing each individual ballot, which, in turn, leads to a lower fraction of invalid ballots. Moreover, this effort-reducing effect is stronger when the electoral competition is weaker. Thus, we should observe that the negative correlation between invalid ballots and margin of victory is stronger for higher-turnout electoral units.

Columns 5 to 7 of Table 4 report the estimates with this interaction term included in the regression. Consistent with the above hypothesis, we indeed find that the coefficient on the interaction between turnout and leading margin is negative and highly statistically significant (for the whole sample and for Southern Italy).

The explanations above disregard the role of political parties. However, as we discussed in the previous section, party representatives can attend the vote count, and it is highly plausible that parties act strategically in this respect and try to actively use their resources to influence the intensity of vote counting, depending on the incentives. In particular, if the total number of representatives that each party can allocate to any given election is limited (parties either have to pay the representatives to do this job, or to mobilize volunteers), parties are likely to allocate their representatives in units that give them the highest expected return.

The institutional organization of the electoral system allows us to obtain a set of additional predictions. For candidates, what matters for each party is not winning the race at each municipality, but obtaining the highest number of votes at the district level. However, the closeness of electoral race at the district level does not necessarily coincide with that in each electoral unit within that district. Election officers do not have access to information about the evolution of the vote count in other units (and thus at the district level). It is plausible that they use the tightness of the race in their own units as the best predictors of the electoral competition at the district level. Parties, however, have a clear informational advantage in this respect. Obviously, allocating representatives

(even in locally competitive races) in districts that are won or lost almost for sure (i.e. those with wide expected margins of victory at the district level) makes little strategic sense. Therefore, parties allocate disproportionately more representatives - in both locally competitive and non-competitive units - into the districts where the (district-level) competition is stronger, and where even a few votes counted mistakenly might imply winning or losing the parliamentary seat. Then, in a two-party election, if both parties act similarly, we should expect that ballots are scrutinized much more closely in the districts with lower district-level victory margins. This would imply a higher fraction of ballots invalidated in such districts, and this should occur independently of the closeness of the electoral race at the unit level.

We explore the role of parties in Table 5. In Column 1 we add as regressors the leading margin at the district level, as well as the dummies for the identity of the leading coalition. We see that the coefficient on the leading margin at the unit level remains highly statistically significant. However, when the leading coalition is the center-right one, the fraction of invalid ballots is significantly smaller. This could be because center-right coalition has more resources to devote to observing the vote counts (note that the specification includes electoral-unit fixed effects; therefore, the remaining variation is within units, so this right-wing effect cannot be driven by fixed characteristic of voters at the unit level).

Column 2 includes the measure of electoral competition at both the unit and the district levels, as well as the interaction term between the two leading margins. The coefficient on the interaction term is negative and highly statistically significant. What does this mean? Figure 6 resumes this finding in the graphical form. Let's define the electoral race to be competitive at the district level when the margin of victory is below the overall median margin of victory (calculated across all districts and elections). We can see that the slope of the relationship between the share of invalid ballots and the (unit-level) leading margin is very different in competitive versus non-competitive parliamentary districts. In particular, in the non-competitive districts, larger leading margin at the electoral unit level implies smaller fraction of the invalid ballots. This is consistent with the explanation of the rational allocation of detection effort by election officers, who do not observe what happens in other units and thus consider their own units as representative of the district-level race. Contrarily, in competitive districts, larger leading margins at

the electoral unit level imply virtually no reduction in the share of invalid ballots. This is likely to be because of the attention that party representatives devote to all (or most) electoral units in such districts, regardless of the intensity of competition locally. Given that the race is very close at the district level, party representatives keep putting pressure on election officers so that these latter detect all the invalid ballots possible, even in units with large local margins.

In the above discussion, we have implicitly assumed that election officers are motivated by duty. A plausible alternative is that they have preferences over candidates or parties, and - given the difficulty to monitor their behavior during the count - try to help their preferred candidates or parties to win by invalidating some of the valid ballots that are in favor of their less-preferred candidates. Researchers in political science have been aware of this possibility for a long time (Harris (1934)).

Given that electoral fraud is an illegal activity, it implies the risk of punishment if the illegal action of the election officer is discovered. Unless the three canvassers, the secretary, the president, as well as the party representatives that are present in the polling station agree to forge the ballot count, the decision problem of a law-breaking officer can be described as follows. Each incremental valid ballot that the officer invalidates increases the risk of getting caught. Moreover, it is plausible to assume that this risk increases more than proportionally with each additional ballot. If the officer invalidates just a few valid ballots, the likelihood that the authorities discover this misbehavior are very low. However, if she invalidates a few more ballots, this likelihoods starts to increase relatively quickly, as - for instance - the discrepancy of the election outcomes with exit polls starts to increase.

On the benefit side, the biased officer wanting to increase the likelihood that her preferred party wins the election understands that this likelihood is large when the expected margin of victory of one candidate over the other is slim. Contrarily, when the expected margin of victory is wide (either in favor of her preferred candidate or against), additional invalidated ballots contribute negligibly to increasing this likelihood. There are two possible theoretical cases. One is that the risk of getting caught for invalidation does not depend on the expected margin of victory. Then, clearly, a wider expected margin of victory implies lower number of invalidations by a rational biased officer. In such case, one obtains the prediction that higher margin of victory should be negatively correlated

with the fraction of invalid ballots, just like under the explanation based on duty-driven election officers that we have described above.

Alternatively, the other - non-biased - actors (other election officers and party representatives) allocate attention not only to counting ballots but also to monitoring their colleagues. In such case, they would rationally allocate more attention to monitoring when the expected margin of victory is small, taking into account that the incentives to misbehave of biased officers are then higher. Such rational behavior would (negatively) link the probability of capture of misbehaving officers to the expected margin of victory. If this link is sufficiently strong, then theoretically it is possible that a biased officer would actually invalidate at the same rate when the expected margin of victory is small and when it is large (because his marginal costs and benefits rise at the same rate when the margin of victory increases). This second case fails to generate a prediction that corresponds to the empirical patterns that we have documented above.

Suppose, however, that the first of the two cases described above is true. In other words, the theory of rational duty-oriented election officer and that of a biased officer generate the same main prediction concerning the correlation between the expected margin of victory and the reported share of invalid ballots. Then, how can one discern empirically between these two alternative explanations? If the electoral fraud is an important driver of the variation in invalid ballots, then - under an auxiliary assumption that the extent of electoral fraud is correlated with other measures of crime - we should normally observe a stronger effect of electoral competition on invalid ballots in areas with higher rates of crime. Here we can exploit the large variation in different measures of crime across Italian provinces: while there are some regions that have, on average, substantially higher crime rates than others, there is still a lot of within-region variation in crime. We report in Table 6 the results of the estimations in which we add as regressors different measures of crime (at province level), as well as the interaction terms between crime rates (standardized for the ease of comparison across columns) and the margin of victory. If the electoral-fraud mechanism is empirically important, we should find a significant negative coefficient on the interaction term. However, in neither of the three specifications (in which crime is measured with the rate of corruption, organized-crime rate, and the share of city councils dissolved for Mafia infiltration) the coefficient on the interaction term is statistically different from zero. We can see this clearly also on Figure

7, in which we plot the relationship between electoral competition and invalid ballots in high-crime versus low-crime areas. Independently of how we measure the crime rates, the slopes of the relationship between electoral competition and invalid ballots in two types of provinces are very similar. This implies that there is no evidence that electoral fraud explains the empirical relationship that we have established earlier.

5 Conclusion

Invalid ballots have been considered a problem for democratic elections, as they might jeopardize the electoral outcome when the electoral race is close. Our paper challenges this view. Using a detailed micro-level dataset from the Italian parliamentary elections in 1994-2001 (which use the traditional paper-ballot hand-counted voting technology), we find a strong robust negative correlation between the margin of victory of the leading candidate over the nearest rival and the share of invalid ballots. We then show that this relationship is unlikely to be driven by voter mistakes, protest, or electoral fraud. The explanation that garners most support is that election officers and party representatives rationally allocate more effort in *detecting* invalid ballots when the stakes are higher, i.e. when the electoral race in the district is closer. In other words, under hand-counted voting technology, in closer elections there are higher rates of *detection* of invalid ballots.

What are the implications of our findings for the current debates on voting technologies? Our analysis shows that the traditional paper-based hand-counted ballot system seems to have a correction mechanism that adjusts the likelihood that the winner is announced correctly to the closeness of electoral race. This mechanism functions thanks to the increased attention that the election officers and parties allocate to making sure that invalid ballots are not counted as valid ones when the electoral race becomes closer. Now, suppose a government considers whether to abandon the paper-ballot system in parliamentary elections in favor of an electronic voting (DRE) system that does not have paper trail (which makes it impossible to recount the votes in case of need). Plausibly, such a government has several considerations in its social-welfare function: (i) minimizing the risk that the winner in each district is identified/announced incorrectly, (ii) the economic cost (including the time cost of delay in announcing the election outcomes), and (iii) maximizing the correct aggregation of voter preferences. The electronic voting

system is not error-free (see, for instance, (Stewart III, 2011) for a detailed discussion); moreover, it is unlikely that the likelihood of machine failure under this system is correlated with the expected margin of victory. On the other hand, the DRE system can present some cost advantage over the paper-ballot system (especially in terms of rapid announcement of election results). Finally, in terms of the aggregation of voter preferences, (Fujiwara, 2015) shows that in an electorate with relatively low level of education (in his case, Brazil), the adoption of electronic voting clearly improves such aggregation by enfranchising uneducated voters which frequently cast invalid votes (under the paper ballot system). Such advantage of the electronic voting is plausibly less important in a polity with relatively high level of literacy/education. Thus, if electronic voting is unlikely to strongly affect the aggregation of voter preferences and the cost advantage of machine counting is fairly small, the paper-ballot system could still be a better alternative, especially if the intensity of electoral competition is fairly high in many districts. If, contrarily, most parliamentary seats are decided in lopsided elections, the electorate has an important share of low-education voters, and the cost advantage of the DRE system is large, abandoning the paper-ballot system would be a welfare improvement.

Our findings suggest an interesting future research direction. As discussed in the introduction, in the last several decades numerous countries implemented reforms in their voting technologies. Moreover, in some countries (for example, the United States) such reforms occurred gradually and at varying speed within the country. Our theory predicts that the correlation between the closeness of electoral race and the share of invalid ballots should be present in hand-counted paper-ballot systems but not in machine-counted ones. Thus, using an electoral district-level panel (with several countries or across the United States), one can test whether (i) this difference in correlation exists across countries/systems, and (ii) whether the switch from paper-ballot to machine-counting system leads to the disappearance of such correlation. Such analysis would contribute to verifying the applicability of the theory developed in this paper to other settings and polities.

References

Joseph G Altonji, Todd E Elder, and Christopher R Taber. Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184, 2005.

Stephen Ansolabehere and Charles Stewart III. Residual Votes Attributable to Technology. *Journal of Politics*, 67(2):365–389, May 2005.

Jean-Marie Baland and James Robinson. Land and Power: Theory and Evidence from Chile. *American Economic Review*, 98(5):1737–65, December 2008.

Geoffrey Brennan and Loren Lomasky. *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge University Press, Cambridge, UK, 1993.

Andrea Brighenti. L'impatto del Voto Elettronico sull'Elettorato Trentino: Un'Indagine sulle Schede Nulle. University of Trento, mimeo, 2003.

Caltech and MIT Voting Technology Project. Voting: What Is, What Could Be. Pasadena and Cambridge: Caltech and MIT, 2001.

David Card and Enrico Moretti. Does voting technology affect election outcomes? touch-screen voting and the 2004 presidential election. *The Review of Economics and Statistics*, 89(4):660–673, 2007.

Piergiorgio Corbetta and Maria Serena Piretti. *Atlante Storico-Elettorale D'Italia. (1861-2008)*. Zanichelli, Bologna, Italy, January 2008.

Thomas Dee. Technology and voter intent: Evidence from the california recall election. *The Review of Economics and Statistics*, 89(4):674–683, 2007.

Timothy Feddersen and Alvaro Sandroni. A Theory of Participation in Elections. *American Economic Review*, 96(4):1271–1282, September 2006a.

Timothy Feddersen and Alvaro Sandroni. Ethical Voters and Costly Information Acquisition. *Quarterly Journal of Political Science*, 1(3):287–311, July 2006b.

Thomas Fujiwara. Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464, 2015.

Joseph P. Harris. *Electoral Administration in the United States*. The Brookings Institution, Washington, 1934.

Bart Jacobs and Wolter Pieters. *Foundations of Security Analysis and Design V: FOSAD 2007/2008/2009 Tutorial Lectures*, chapter Electronic Voting in the Netherlands: From Early Adoption to Early Abolishment, pages 121–144. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

Stephen Knack and Martha Kropf. Voided Ballots in the 1996 Presidential Election: A County-level Analysis. *Journal of Politics*, 65(3):881–897, August 2003.

Fabrice Lehoucq. Electoral Fraud: Causes, Types,and Consequences. *Annual Review of Political Science*, 6:233–256, June 2003.

F.E. Lehoucq and I. Molina. *Stuffing the ballot box: fraud, electoral reform, and democratization in Costa Rica*. Cambridge University Press, 2002.

Ian McAllister and Toni Makkai. Institutions, society or protest? explaining invalid votes in australian elections. *Electoral Studies*, 12(1):23–40, 1993.

Dennis Mueller. *Public Choice III*. Cambridge University Press, Cambridge, MA, 2003.

Timothy Power and James Garand. Determinants of Invalid Voting in Latin America. *Electoral Studies*, 26(2):432–444, June 2007.

Timothy Power and Timmons Roberts. Compulsory Voting, Invalid Ballots, and Abstention in Brazil. *Political Research Quarterly*, 48(4):795–826, December 1995.

Fredric Schaffer. *Elections For Sale: The Causes and Consequences of Vote Buying*. Lynne Rienner Publishers, Boulder, CO, 2007. ISBN 9781588264343.

Alexander A. Schuessler. *A Logic of Expressive Choice*. Princeton University Press, Princeton, NJ, 2000.

Kelly Shue and Erzo F. P. Luttmer. Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes. *American Economic Journal: Economic Policy*, 1(1): 229–57, February 2009.

Charles Stewart III. Voting technologies. *Annual Review of Political Science*, 14:353–378, 2011.

Tom Stoppard. *Jumpers*. Grove Press, New york, 1972.

Fredric Uggla. Incompetence, Alienation, or Calculation? Explaining Levels of Invalid Ballots and Extra-Parliamentary Votes. *Comparative Political Studies*, 41(8):1141–1164, August 2008.

Jonathan Wand, Kenneth Shotts, Jasjeet Sekhon, Walter Mebane, Michael Herron, and Henry Brady. The butterfly did it: The aberrant vote for buchanan in palm beach county, florida. *American Political Science Review*, 95(4):793–810, 2001.

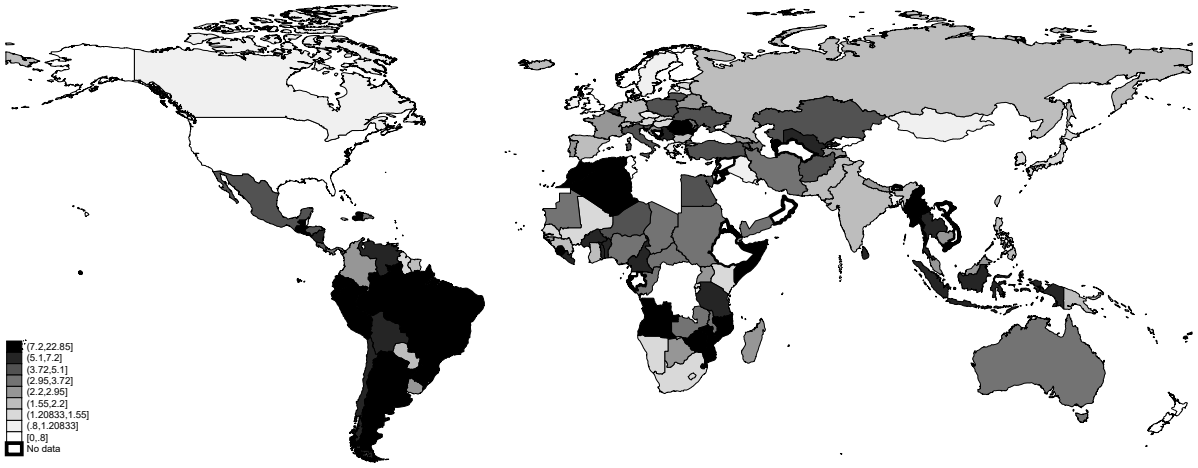

Figure 1: Invalid ballots (as a fraction of total ballots) in the parliamentary elections around the world

Source: Institute for Democracy and Electoral Assistance (www.idea.int)

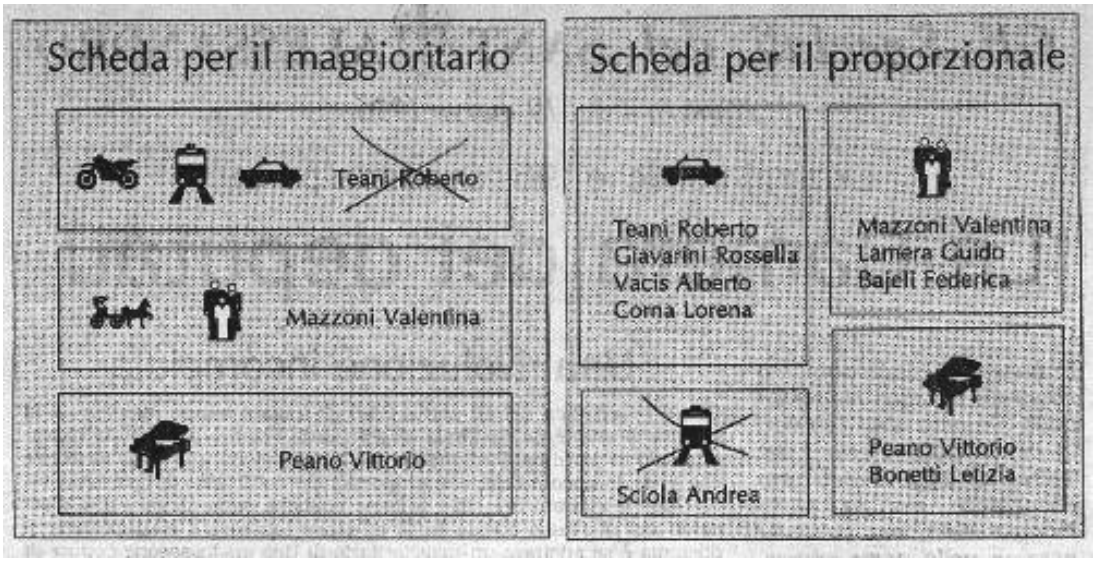

Figure 2: A typical ballot in Italian parliamentary elections

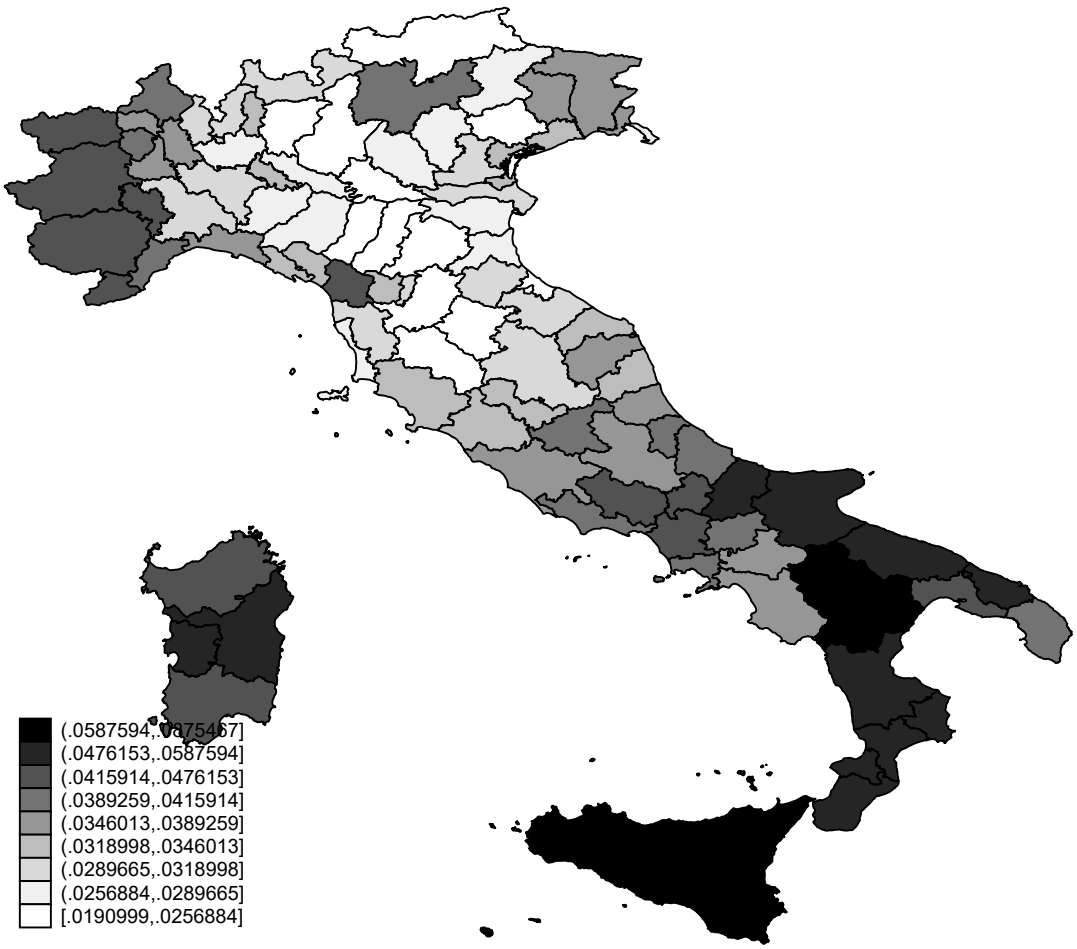

Figure 3: Invalid ballots (as a fraction of total ballots) in Italian parliamentary elections, 1994-2001 (majoritarian districts)

Notes: Author's calculation based on Italian national elections data (Corbetta and Piretti, 2008).

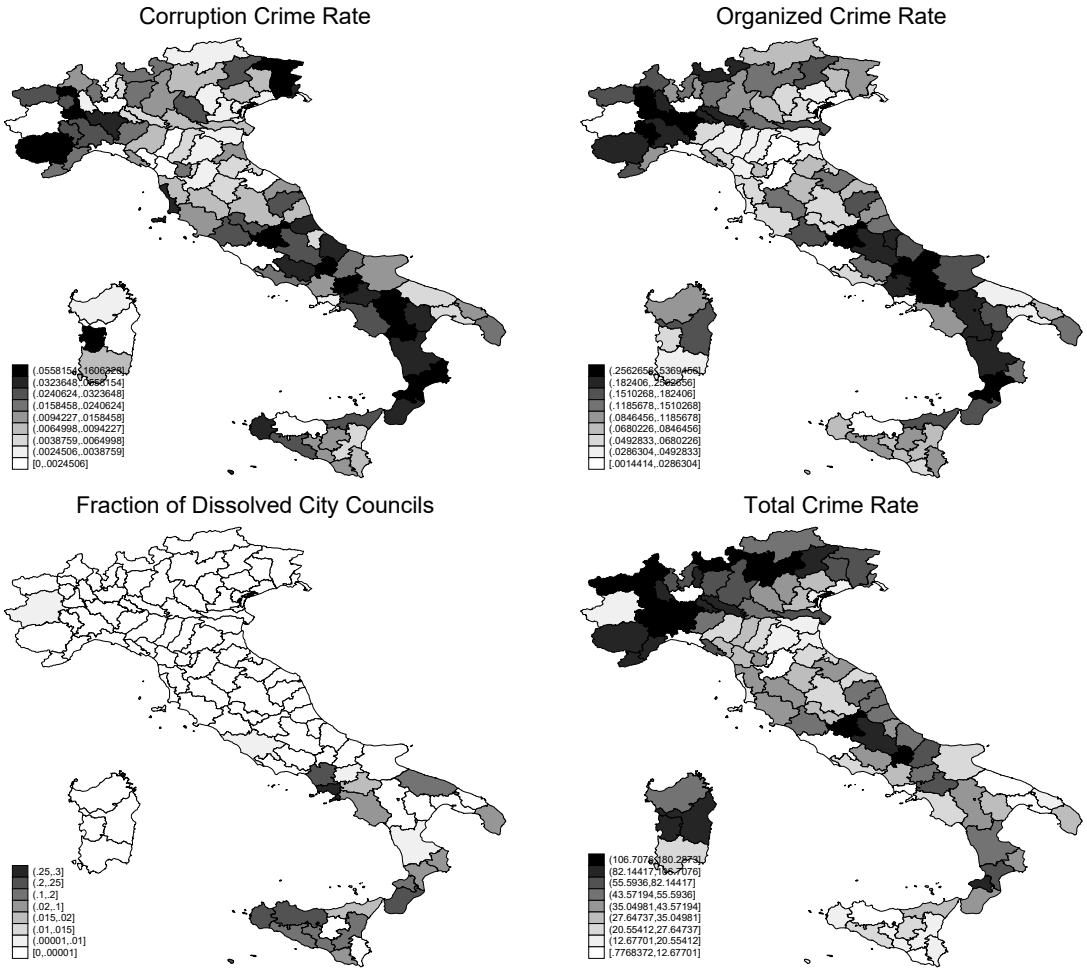

Figure 4: Crimes and Dissolution Rates of City Councils by Provinces

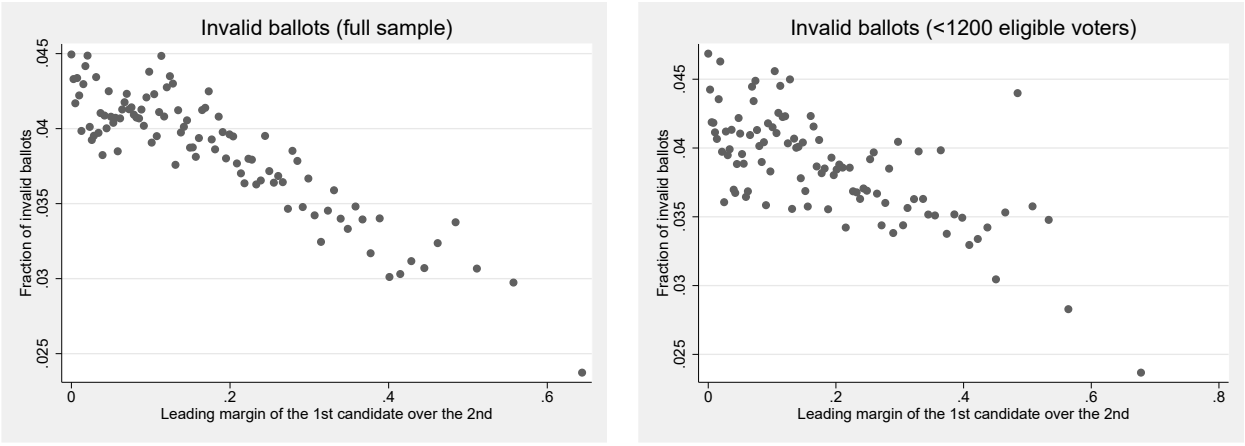

Figure 5: Fraction of invalid ballots by percentile of leading margin

Notes: The right panel restricts the data to municipalities with less than 1200 eligible voters, which is the maximal size of a polling station. Author's calculation based on Italian national elections data (Corbetta and Piretti, 2008).

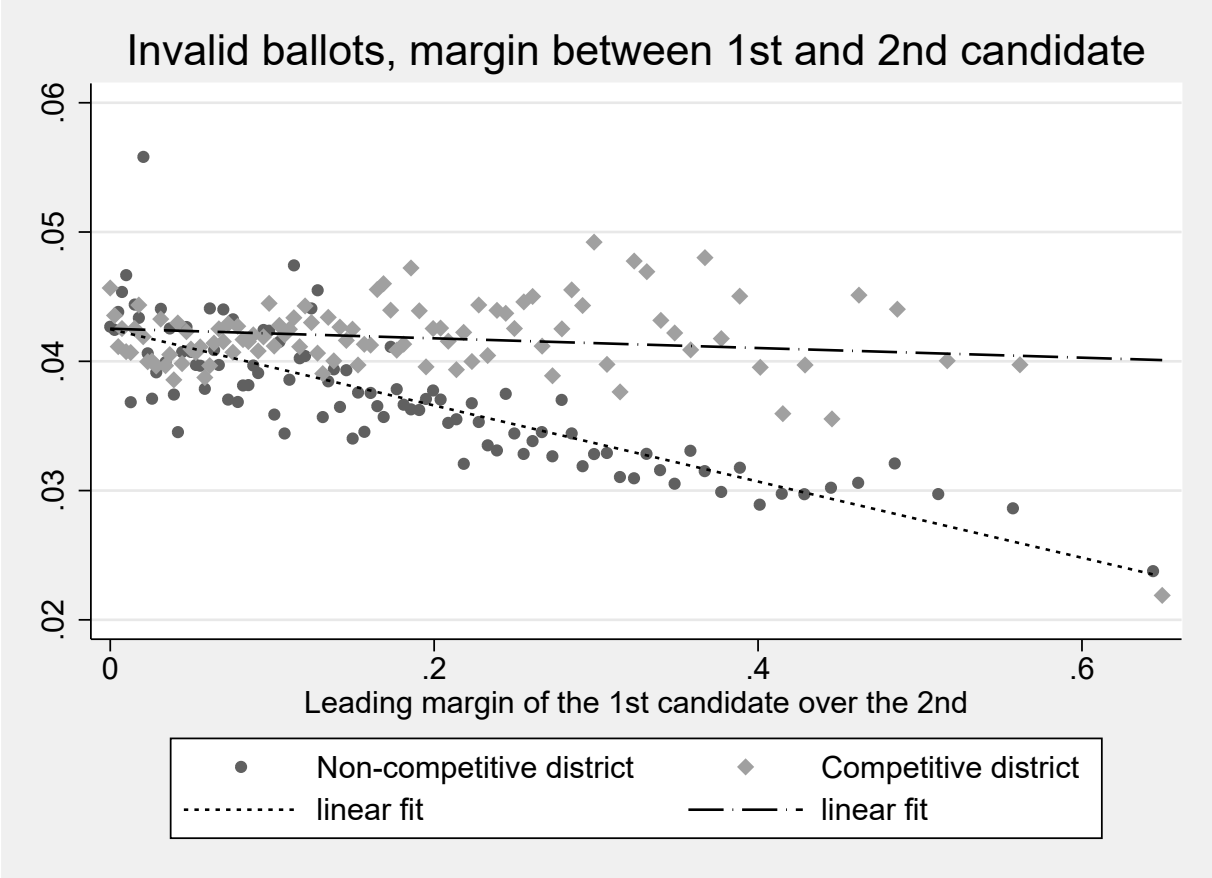

Figure 6: Fraction of invalid ballots by percentile of leading margin and district-level competition

Notes: Author's calculation based on Italian national elections data (Corbetta and Piretti, 2008).

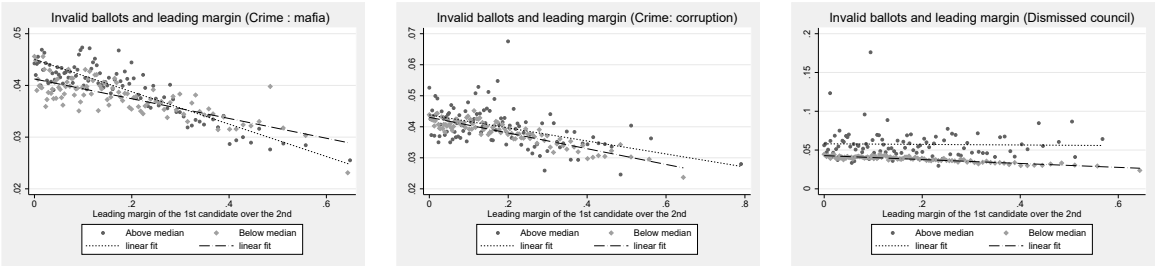

Figure 7: Fraction of invalid ballots by percentile of leading margin, in high-crime versus low-crime provinces

Notes: Author's calculation based on Italian national elections data (Corbetta and Piretti, 2008).

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Invalid ballots	0.039	0.022	0	0.573	23019
Blank ballots	0.046	0.022	0	0.245	23019
Turnout	0.820	0.108	0.049	1	23019
Leading margin at electoral unit level	0.184	0.148	0	0.961	23019
Leading margin at district level	0.139	0.124	0	0.756	23019
Right coalition leads	0.38	0.485	0	1	23019
Left coalition leads	0.362	0.481	0	1	23019
Incumbent party leads	0.945	0.229	0	1	14965
Incumbent leads	0.264	0.441	0	1	14965
Number of candidates	4.136	1.005	2	9	23019
Corruption Crime Rate (per 100 inh.)	0.001	0.007	0	0.239	23019
Organized Crime Rate (per 100 inh.)	0.007	0.012	0	0.21	23019
Dissolution of a city council between 1999-2009	0.017	0.13	0	1	23019
Fraction of pop. with university degree	0.067	0.009	0.044	0.094	22979
Fraction of pop. with high school degree	0.3	0.027	0.257	0.392	22979
Turnout at national referenda	0.684	0.086	0.48	0.811	22979
Labor activity rate	0.487	0.04	0.355	0.581	22979
Unemployment rate	0.093	0.077	0.017	0.305	22979
GDP per capita (in €10,000)	1.897	0.464	1.028	3.028	22979
Rate of urbanization	0.235	0.125	0.088	0.872	22979

Table 2: Basic regression results

	(1)	(2)	(3)	(4)
	Fraction of invalid ballots			
Leading margin at electoral unit level	-0.0242***	-0.0181***	-0.0125***	-0.0137***
	(0.002)	(0.002)	(0.002)	(0.002)
Turnout		-0.0474***	-0.0042	0.0164
		(0.006)	(0.008)	(0.091)
Blank ballots		0.2085***	0.1457***	0.1347***
		(0.020)	(0.019)	(0.046)
Number of candidates		-0.0006	-0.0022***	-0.0005
		(0.000)	(0.000)	(0.001)
Corruption Crime Rate (per 100 inh.)			0.0166	
			(0.020)	
Organized Crime Rate (per 100 inh.)			-0.0251	
			(0.020)	
Total Crime Rate (per 100 inh.)			-0.0004***	
			(0.000)	
Fraction of pop. with university degree			-0.0390	
			(0.175)	
Fraction of pop. with high school degree			-0.1025**	
			(0.052)	
Turnout at national referenda			-0.0358**	
			(0.016)	
Labor activity rate			-0.0500*	
			(0.027)	
Unemployment rate			0.0607***	
			(0.016)	
GDP per capita (in 10,000)			0.0007	
			(0.002)	
Rate of urbanization			0.0119**	
			(0.005)	
Year FE				√
Electoral Unit FE				√
Observations	23,019	23,019	22,979	23,019
R-squared	0.028	0.158	0.263	0.702

NOTE.– Standard errors, clustered at the district level, are reported in parentheses (there are 465 majoritarian districts). There are 8,042 electoral-unit fixed effects. *,**, and *** denote statistical significance at 10%, 5%, and 1% level.

Table 3: Basic regression results for electoral units with less than 1,200 eligible voters

	(1)	(2)	(3)	(4)
	Fraction of invalid ballots			
Leading margin at electoral unit level	-0.0177***	-0.0138***	-0.0122***	-0.0095***
	(0.003)	(0.003)	(0.003)	(0.002)
Turnout		-0.0312***	-0.0074	-0.0397*
		(0.007)	(0.007)	(0.021)
Blank ballots		0.1653***	0.1453***	0.1321***
		(0.022)	(0.019)	(0.027)
Number of candidates		-0.0006	-0.0020***	-0.0007
		(0.001)	(0.001)	(0.001)
Corruption Crime Rate (per 100 inh.)			-0.0135	
			(0.019)	
Organized Crime Rate (per 100 inh.)			-0.0532**	
			(0.024)	
Total Crime Rate (per 100 inh.)			-0.0002	
			(0.000)	
Fraction of pop. with university degree			-0.1971	
			(0.247)	
Fraction of pop. with high school degree			-0.0468	
			(0.077)	
Turnout at national referenda			-0.0441**	
			(0.022)	
Labor activity rate			0.0196	
			(0.042)	
Unemployment rate			0.0483**	
			(0.021)	
GDP per capita (in 10,000)			0.0016	
			(0.005)	
Rate of urbanization			0.0141**	
			(0.007)	
Year FE				√
Electoral Unit FE				√
Observations	7,275	7,275	7,272	7,275
R-squared	0.017	0.085	0.147	0.689

NOTE.– Standard errors, clustered at the district level, are reported in parentheses (there are 465 majoritarian districts). There are 2,607 electoral-unit fixed effects. *,**, and *** denote statistical significance at 10%, 5%, and 1% level.

Table 4: Regressions on Sub-Samples and with Interaction Terms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Fraction of invalid votes						
Sample	Whole	Whole	North	South	Whole	North	South
Leading margin at electoral unit level	-0.0123***	-0.0098***	-0.0060***	-0.0175***	0.0353*	0.0130	0.0469*
	(0.004)	(0.003)	(0.002)	(0.004)	(0.020)	(0.031)	(0.024)
Turnout	0.0153	-0.1326	-0.0070	0.1272	0.0159	-0.0073	0.1254
	(0.151)	(0.193)	(0.013)	(0.174)	(0.091)	(0.013)	(0.173)
Turnout × Leading margin					-0.0449**	-0.0281	-0.0804**
					(0.018)	(0.031)	(0.031)
N. of candidates × Leading margin					-0.0029*	0.0013	-0.0014
					(0.002)	(0.002)	(0.002)
Other controls (as in Table 2)	√	√	√	√	√	√	√
Cubic term in turnout	√	√					
Cubic term in turnout interacted with education		√					
Observations	23,019	22,105	13,261	9,718	23,019	13,261	9,718
R-squared	0.703	0.715	0.709	0.726	0.703	0.710	0.727

NOTE.– Standard errors, clustered at the district level, are reported in parentheses (there are 465 majoritarian districts). Controls include also year and electoral unit fixed effects. *,**, and *** denote statistical significance at 10%, 5%, and 1% level.

Table 5: Invalid ballots and competition at electoral-unit and district levels

	(1)	(2)	(3)	(4)
	Fraction of invalid votes			
	Whole Sample		Less than 1200 Voters	
Leading margin at electoral unit level	-0.0094***	-0.0083***	-0.0059**	-0.0048
	(0.002)	(0.002)	(0.003)	(0.003)
Leading margin of candidate	-0.0098*	-0.0061	-0.0107*	-0.0077
	(0.006)	(0.006)	(0.006)	(0.007)
Interaction between the margins		-0.0554***		-0.0432***
		(0.016)		(0.016)
Turnout	0.0188	0.0202	-0.0402*	-0.0389*
	(0.091)	(0.091)	(0.020)	(0.021)
Blank ballots	0.1318***	0.1291***	0.1300***	0.1262***
	(0.047)	(0.047)	(0.027)	(0.027)
Number of candidates	-0.0007	-0.0008	-0.0009	-0.0010
	(0.001)	(0.001)	(0.001)	(0.001)
Right coalition leads	-0.0047***	-0.0043***	-0.0051***	-0.0049***
	(0.001)	(0.001)	(0.002)	(0.002)
Left coalition leads	-0.0016	-0.0010	-0.0015	-0.0010
	(0.001)	(0.001)	(0.001)	(0.001)
Electoral Unit and Year FE	√	√	√	√
Observations	23,019	23,019	7,275	7,275
R-squared	0.706	0.707	0.693	0.694

NOTE.– Standard errors, clustered at the district level, are reported in parentheses (there are 465 majoritarian districts). *,**, and *** denote statistical significance at 10%, 5%, and 1% level. There are 8,042 electoral-unit fixed effects.

Table 6: Invalid Ballots, Electoral Competition, and Crime

	(1)	(2)	(3)	(4)	(5)	(6)
			Fraction invalid ballots			
Leading margin at electoral unit level	-0.0175***	-0.0137***	-0.0175***	-0.0137***	-0.0175***	-0.0137***
	(0.0020)	(0.0019)	(0.0020)	(0.0019)	(0.002)	(0.002)
Corruption Crime Rate (per 100 inh.)	0.0543**					
	(0.0260)					
Organized Crime Rate (per 100 inh.)			0.0734**			
			(0.0296)			
Fraction of dismissed city councils for Mafia					0.0138***	
					(0.003)	
Interaction term (in standard deviation)	0.0013	-0.0018	-0.0014	-0.0002	0.0016	-0.0044
	(0.0013)	(0.0015)	(0.0015)	(0.0014)	(0.002)	(0.005)
Turnout	-0.0478***	0.0164	-0.0476***	0.0164	-0.0464***	0.0164
	(0.0068)	(0.0909)	(0.0068)	(0.0909)	(0.007)	(0.091)
Blank ballots	0.2065***	0.1345***	0.2060***	0.1347***	0.2071***	0.1348***
	(0.0195)	(0.0462)	(0.0197)	(0.0463)	(0.020)	(0.046)
Number of candidates	0.0002	-0.0005	0.0001	-0.0005	0.0001	-0.0005
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.000)	(0.001)
Year FE	√	√	√	√	√	√
Electoral Unit FE		√		√		√
Observations	23019	23019	23019	23019	23019	23019
R-squared	0.169	0.702	0.170	0.702	0.175	0.703

NOTE.– Standard errors, clustered at the district level, are reported in parentheses (there are 465 majoritarian districts). *,**, and *** denote statistical significance at 10%, 5%, and 1% level. There are 8,042 electoral-unit fixed effects.