UNIVERSITÀ DELLA CALABRIA

Dipartimento di Economia e Statistica
Ponte Pietro Bucci, Cubo 0/C
87036 Arcavacata di Rende (Cosenza)
Italy
http://www.ecostat.unical.it/

Working Paper n. 03 - 2012

THE CAUSAL IMPACT OF CLOSENESS ON ELECTORAL PARTICIPATION EXPLOITING THE ITALIAN DUAL BALLOT SYSTEM

Maria De Paola
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492459
Fax: +39 0984 492421
e-mail: m.depaola@unical.it

Vincenzo Scoppa
Dipartimento di Economia e Statistica
Università della Calabria
Ponte Pietro Bucci, Cubo 1/C
Tel.: +39 0984 492464
Fax: +39 0984 492421
e-mail: v.scoppa@unical.it

Febbraio 2012

The Causal Impact of Closeness on Electoral Participation

Exploiting the Italian Dual Ballot System

Maria De Paola, Vincenzo Scoppa*

This version: 23 February 2012

Using data from Italian municipal elections from 1993 to 2011, we investigate whether political competition affects electoral turnout. Taking advantage of the dual ballot system adopted for municipalities with more than 15,000 inhabitants, we measure the expected closeness in the second round through the first round electoral results. Thanks to the richness of our dataset we are able to distinguish between valid, blank and invalid ballots and to investigate the effect of closeness on each of these variables, controlling for municipalities' and candidates' characteristics and for municipal fixed effects. We also estimate a Heckman selection model to take into account for the non-randomly selected sample. It emerges that closeness strongly increases valid ballots and reduces blank ballots supporting the idea that the expected benefits of voting increase in closer competitions. The effect is much higher in magnitude than that merging when measuring closeness with ex-post electoral results, suggesting a quite relevant endogeneity bias. On the other hand, we do not find any statistically significant effect on invalid ballots.

JEL classification: D72, D78; J45
Keywords: Electoral Turnout; Closeness; Electoral Competition; Blank and Invalid Ballots.

1. Introduction

Participation at elections is crucial for well-functioning political systems. Understanding the reasons that lead people to vote in large elections is of major interest both for political scientists and for economists.

Scholars have considered different factors that may induce individuals to vote, such as the utility deriving from the election of the voter's favorite candidate, weighted by the probability of being the decisive voter ("instrumental voting"), and the benefits deriving from expressing political preferences or solidarity, or the desire to contribute to the functioning of democracy and to fulfill a civic duty ("expressive voting") (see Dhillon and Peralta, 2002; Feddersen, 2004).

* Department of Economics and Statistics, University of Calabria, 87036 Arcavacata di Rende (CS), Italy. E-mail: m.depaola@unical.it; v.scoppa@unical.it. We are grateful to Stefano Trulli of the Italian Ministry of Internal Affairs for making some data available and for helping us with their use. We thank Guglielmo Barone, Giorgio Bellettini, Marco Debenedetto, Stefano Gagliarducci, Vincenzo Galasso, Laura Mazzuca, Paolo Pinotti, Gerard Roland and seminar participants to the 1st Workshop on Economics and Politics (Bologna, October 2011) for useful comments and suggestions.

Among the factors affecting turnout, the degree of electoral competition is one of the most investigated, under the theoretical assumption that a closer electoral race stimulates greater electoral participation.

However, disentangling the effects of electoral competition on turnout and trying to investigate through which channels it works is a challenging task. Typically, electoral competition is measured using ex-post election results to proxy for electors' ex-ante expectations of closeness (Cox and Munger 1989; Matsusaka and Palda, 1993; Geys, 2006). This poses serious endogeneity problems. Suppose, for example, that a positive shock, due to an unobservable factor, affects votes obtained by a certain candidate: this has an impact on turnout and on the electoral margin, determining a correlation between the error term and the variable of interest and leading to biased estimates.

In this paper we rely on elections held using a dual ballot system and measure the expected electoral closeness at the second round with the actual closeness at the first round. To analyze the relationship between electoral closeness and turnout we consider about 1,400 electoral competitions at the municipal level taking place in Italy from 1993 to 2011. Since 1993, for cities with more than 15,000 inhabitants, elections are held using a dual ballot system. If none of the candidates obtains the majority of the votes, the two leading candidates at the first round compete in the second ballot, which takes place two weeks later. This structure allows us to proxy expected closeness at the second round with the actual closeness at the first round, given the short time span intervening among the two rounds.

Compared to existing works exploiting French and Hungarian general elections (Fauvelle-Aymar and Francois, 2006; Indridason, 2008; Simonovits, 2011), we take advantage of some institutional features of the Italian electoral system. Firstly, our study is based on independent elections (each municipality elects its own mayor and no interdependence exists among municipalities), while the literature exploiting dual ballot systems refers to political elections in which voters could be affected in the second round by the general tendency at the national level. Secondly, in Italy only two candidates compete in the second round allowing for a clear definition of the electoral margin, while in other countries typically all the candidates obtaining more than a certain threshold of votes (for example, 12.5% in France) go to the second round, making ambiguous the definition of the electoral margin. Thirdly, we are able to use a Heckman selection model to handle estimation problems deriving from a non-randomly selected sample due to the fact that the second round only takes place if no candidate has obtained the majority of votes in the first round.

A further peculiarity of our work is that, to better understand the effects of electoral competition on turnout, we look separately at the three different types of ballots casted by electors: valid, blank ad invalid ballots. While the effect of closeness on turnout is widely investigated, little is known on its effect on blank and invalid ballots, even if they are increasing over time.

We estimate several models explaining valid, invalid and blank ballots at the second round elections in relation to the expected closeness of that round based on the electoral results of the first round. In our regressions we control for the percentage of valid, invalid and blank ballots at the first

round, a number of municipal characteristics and the average characteristics of candidates in terms of education, age and gender. Furthermore, exploiting the fact that typically more than one election is held in each municipality in the period considered (18 years), we also control for municipality fixed effects. We find that closeness strongly increases valid ballots and reduces blank ballots. As far as invalid ballots are concerned, we do not find any statistically significant effect.

With our data we also show that the bias deriving from not taking into account the endogeneity problems related to the use of ex-post measures of closeness could be quite relevant. In fact, when we run our regressions to explain turnout measuring electoral closeness with the electoral results at the *same* round, we find that the effect of closeness on turnout is much smaller in magnitude compared to the unbiased estimates.

To avoid possible biases due to sample selection we also estimate a Heckman selection model: the estimates substantially confirm our OLS results. We also show that our findings are robust to the use of an estimator fitting a proportional dependent variable (frational logit model).

The paper is organized in the following way. Section 2 briefly discusses the related literature. Section 3 is devoted to the description of the institutional framework and of our dataset. In Section 4, we investigate the effects of closeness on valid ballots, controlling for municipalities' and candidates' characteristics and municipal fixed effects. Section 5 offers some robustness checks. In Section 6, we investigate the determinants of blank and invalid ballots. Section 7 explores the possibility of non-linear effects. Section 8 concludes.

2. Related Literature

The idea that electoral participation is linked to political competition goes back to Downs (1957): under the assumption of instrumental behavior, with voters undertaking a costs-benefits analysis, a closer election is expected to induce a higher turnout since it increases the probability for each voter of being decisive. This was defined by Matsusaka and Palda (1993) as the "Downsian Closeness Hypothesis".

As argued by Cox and Munger (1989) and Shachar and Nalebuff (1999), closeness may increase turnout also because of a higher mobilization effort of parties and candidates, which may persuade unconvinced voters and provide useful information that induce a number of previously uninformed subjects to cast their vote. The greater mobilization and the higher attention and discussion that typically accompanies elections characterized by highly uncertain results may also increase social pressure for taking part to the vote.

Empirical works aimed at shedding light on the impact of closeness on turnout are widespread– in his review of the literature, Geys (2006) counts more than 50 studies on the topic – especially at the national level. Findings are unfortunately rather ambiguous: more than 40% of the estimates show no effect of closeness or even effects opposite to those expected.

The unclear results are perhaps caused by a number of econometric problems mainly deriving from the fact that the closeness of the electoral competition (the explanatory variable) and the turnout at that competition (the dependent variable) are jointly determined.

Only in few works, electoral competition has been measured considering previous election results or opinion polls. However, given the temporal distance, previous election results are typically a quite imprecise measure of electoral competition, while opinion polls may not reflect the effective electoral choices and are typically not available at the local/district level.

Some recent works, relying on elections held using a dual ballot system, have measured the expected electoral competition at the second round with the actual competition at the first round. This strategy has been used by Fauvelle-Aymar and Francois (2006) and by Indridason (2008), who base their analysis on French data, and by Simonovits (2011), who focuses on the dual ballot system adopted in Hungary for general elections. From these studies, it emerges that a higher electoral competition (higher closeness) at the first round increases the turnout at the second round (both in France and Hungary).

Our work follows a similar strategy but we look separately to valid, invalid and blank ballots. Blank and invalid ballots generally represent a small share of total ballots, but their importance is increasing over time and an analysis of electoral behavior in all its components seems relevant as blank and invalid ballots are probably not random events. The percentage of blank and invalid ballots seem to be related to socio-economic and institutional factors, such as education, income and voting systems (Power and Garand, 2007). The degree of electoral competition may also affect this type of voter choices, but the relationship could be either positive or negative. On the one hand, invalid and blank ballots may decrease because of the higher expected benefits deriving from closeness, related both to the higher probability of being pivotal and to the higher costs of expressing discontent through a blank or voluntary invalid ballot (see Endersby, Galatas, and Rackaway, 2002; Aldashev and Mastrobuoni, 2011). On the other hand, a positive relationship may emerge in case of electors unwilling to support any of the candidates but wanting to avoid the stigma of non-voting deriving from higher social pressure. Indeed, when competition increases candidates' and social pressure increases leading to a higher cost of not going to the polls, but electors may still escape from supporting the candidates by casting a blank or invalid ballot.

A part from voters' behavior, in traditional paper-ballot systems, the number of invalid ballots is affected by election officers. To these subjects is, in fact, delegated the provisional decision on any dispute related to the vote count (disputes are ultimately decided by the Court). Again, the relationship between electoral closeness and invalid votes is ambiguous. In case of unbiased election officers, we would expect a positive correlation between invalid ballots and electoral closeness due to a higher effort and attention devoted in tracking the truly invalid ballots in closer electoral races. However, in case of biased elections officers (who cheat to favor their preferred candidate) the relationship is less clear. When the margin between the candidates is slim, biased election officers are, on the one hand, less likely to

invalidate the ballots supporting their preferred candidate, but, on the other hand, they are more induced to invalidate the ballots against their candidate.

Then, understanding the relationship between electoral competition and valid, invalid and blank ballots is an empirical matter. Nevertheless, while there are a number of works analyzing the effect of electoral competition on turnout, only few empirical studies have looked at the different ballot categories. Galatas (2008), using data from 1999 and 2003 provincial elections in Ontario, investigates the relationship between blank ballots and electoral competition, showing that blank ballots reduce in closer elections. Aldashev and Mastrobuoni (2011) focus instead on the relationship between the closeness of the electoral race and the number of invalid ballots. They use data on Italian parliamentary elections in years going from 1994 to 2001 and find a strong positive correlation between the closeness of the electoral race and the fraction of invalid ballots. They interpret this effect as the result of unbiased election officers who increase their effort in response to the higher closeness of the electoral race.

3. Institutional framework and data

Italy is characterized by a quite high electoral turnout compared to many European countries and to US. In municipal elections the average turnout in the period 1993-2011 has been 78.8%.[1]

Municipal administrations in Italy are responsible for a number of relevant functions such as the management of public utilities (local roads, water, sewage, garbage collection) and the provision of some services having a strong impact on the daily life of citizens (public housing and transportation, nursery schools etc.). As a consequence, voters are generally strongly interested in the choice of the mayor and in the composition of the Municipal Council.

The system currently regulating municipal elections in Italy has been introduced in 1993 (Law no. 81 of March 25, 1993). It has established the direct election of the mayor and the adoption of the plurality rule, with some differences according to the size of the city. For municipalities with a population of fewer than 15,000 inhabitants, elections are held with single ballot and plurality rule: the winning candidate is awarded a majority premium of at least two-thirds of the seats in the council. For cities with a population above 15,000, elections are held using a dual ballot system (where the second ballot is held only if none of the candidates obtains an absolute majority of votes in the first ballot).[2] Only the two leading candidates at the first round compete in the second ballot – which takes place two weeks later – and the winning candidate is awarded a majority premium of at least 60 percent of the seats in the council.

This feature of the Italian municipal system allows us to proxy voters' expected closeness at the second round with the actual closeness at the first round.

[1] A similar figure has been recorded for recent Parliamentary Elections (78%).

[2] In two regions with special legislative powers, the thresholds for the dual ballot system have been set differently. In Sicily, the threshold is set at 5,000 inhabitants. In Friuli Venezia Giulia until 2001 the threshold was at 10,000 and was set at 15,000 afterwards. We include in our sample 61 observations for the municipalities involved.

Municipalities have a registry of eligible voters (which is revised whenever there is an election): all citizens aged 18 or more on the election date are automatically registered to vote. Elections usually last two consecutive days (Sunday and Monday). Voting takes place in polling stations organized by the local authorities. Elections are organized according to a traditional paper ballot system.[3]

We base our analysis on a panel dataset (collected using the information provided by the Italian Ministry of Internal Affairs)[4] on 1,410 municipal elections held over the period 1993-2011 in 632 municipalities with more than 15,000 inhabitants (about 7% of Italian municipalities). The observations include the municipalities that over the period considered have elected their mayor through a second electoral round. We have information both for the first and second round on the number of valid, invalid and blank ballots, the number of people eligible to vote, the number of candidates at the first round and the votes they received.

In addition, we use the 1991 and 2001 Italian Census of Population to obtain data at the municipal level on the population size, the number of employed individuals, the educational attainment of the population, the percentage of people over the age of 65. Finally, from the database on local administrators, provided by the Italian Ministry of Internal Affairs, we have collected information on the identity, gender, age and highest educational attainment of candidates to the mayor position (we have some missing values for these controls and the number of municipal elections used in regressions controlling for these covariates reduces to 1,367).

We use the information on the results at the first round as a proxy for the closeness of the competition at the second round. More precisely, we build the variable *Electoral Margin* as the absolute difference between the number of votes obtained by the two leading candidates at the first round (divided by the number of eligible voters). Therefore, *Electoral Margin* represents an inverse measure of expected electoral closeness for the second round election.[5]

We define the variable *Valid (II round)* as the ratio between the number of valid ballots at the second round and the number of eligible voters.[6] Similarly, the *Blank (II round)* and *Invalid (II round)* are the ratio between, respectively, blank and invalid ballots, and eligible voters. We define analogous

[3] In each polling room there is a president, a secretary and four assistants. Each political party has the right to have a representative present in every polling room, who can check all the activities of the election officers. The count of the ballots starts as soon as the elections end. During the count each ballot paper is checked by all the six election officers and any of the party representatives. Ballots with any redundant writing, beyond the expression of a valid vote, are considered invalid. In case of disagreement about who to assign the vote to, the president decides a temporary "position" (the final decision is taken by the Court). Results of the count (valid, invalid and blank ballots) are reported in an official statement.

[4] See the webpage http://elezionistorico.interno.it/ (Archivio Storico delle Elezioni).

[5] We also experiment defining *Electoral Margin* as the absolute difference between the number of votes for the two leading candidates at the first round divided by the sum of votes obtained by them $|v_1 - v_2|/(v_1 + v_2)$. Results (not reported) are very similar to those shown in the paper.

[6] Our dependent variable *Valid (II round)* is slightly different from other measures of turnout used in the literature, since we deal separately with blank and invalid ballots, whereas they are typically included in the turnout variable. We investigate blank and invalid ballots determinants in Section 6.

measures for the first round. We also consider among our controls the number of votes obtained by the two leading candidates at the first round (as a percentage of electors), *Votes First Two Candidates*

In our main analyses, we only consider elections for which a second round was held and then elections for which at the first round none of the candidates obtained the absolute majority of the votes. In Section 5, as a robustness check, we estimate a Heckman selection model considering also the municipalities with more than 15,000 inhabitants in which no second round was held.

In Table 1 are reported some descriptive statistics. The average *Electoral Margin* is equal to 7.2%. The municipality with the closest election in our data shows a vote difference of nearly zero between the two leading candidates, while in the election with the highest margin at the first round the best performing candidate leads by 29.3%.

Table 1. Descriptive Statistics

Variables	**Mean**	**Std. Dev.**	**Min**	**Max**	**N**
Valid (II round)	0.6392	0.0785	0.3412	0.8330	1410
Blank (II round)	0.0083	0.0059	0.0010	0.0435	1410
Invalid (II round)	0.0154	0.0088	-0.0072	0.0951	1410
Electoral Margin	0.0723	0.0562	0.0001	0.2932	1410
Valid (I round)	0.7504	0.0590	0.3433	0.8934	1410
Invalid (I round)	0.0262	0.0118	-0.0033	0.1774	1410
Blank (I round)	0.0130	0.0096	0.0000	0.0796	1410
Votes First Two Cand. (I round)	0.5234	0.0977	0.1974	0.8367	1410
# Candidates	5.6752	1.9239	3	17	1410
Population (,000)	48.6894	152.0518	5.0070	2775.2500	1410
Education	7.3727	0.9320	4.9863	10.0206	1410
Employment	0.3105	0.1371	0.0549	0.8045	1410
Percentage Pop≥65	0.1452	0.0440	0.0380	0.2782	1410
Candidates' Education	16.1291	2.2881	8	18	1369
Candidates' Age	49.0448	7.5365	25	77	1384
Female Candidates	0.1348	0.3417	0	1	1387

Source: Italian Ministry of Internal Affairs and Italian Census 1991 and 2011.

The average ratio between valid ballots and electors over the period 1993-2011 is of 75% and 63.9%, at the first and second round, respectively. Blank ballots are 1.3% and 0.83% respectively at the first and second round, while invalid ballots are 2.62% at the first round and 1.54% at the second round.

The average population size is of 48,689. The average years of education in the population are 7.37, the ratio between the number of employed individuals and the total number of inhabitants is of about 31% and the share of population aged more than 65 is 14.52%. 43% of municipalities are located in the South and Islands.

The average number of years of schooling of the two leading candidates is of about 16.[7] Their average age is about 49 years. Less than 14% of second round competitions have seen a female among competing candidates.

[7] It takes 13 years to attain a High-School Degree while 17-18 years are necessary to attain a College Degree.

4. Electoral Closeness and Turnout: OLS and fixed effect estimates

In this Section we estimate an OLS model to analyze whether electoral closeness enhances turnout. More precisely, we firstly focus on the percentage of valid votes, which according to theoretical prediction should be positively affected by electoral closeness. We estimate the following model:

$$Valid\ (II\ round)_{it} = \beta_0 + \beta_1 Electoral\ Margin_{it} + \beta_2 W_{it} + \beta_3 X_{it} + \beta_4 Z_{it} + \mu_i + \eta_t + \varepsilon_{it}$$

where $Valid\ (II\ round)_{it}$ is the ratio between valid ballots at the second round and eligible voters in municipality i at election year t, $Electoral\ Margin_{it}$ is the difference (in absolute value) between the number of votes at the first round for the two leading candidates (divided by the number of eligible voters), W_{it} is a vector including the first round electoral outcomes (the percentage of votes obtained by the two leading candidates, the percentage of valid, invalid and blank ballots), X_{it} is a vector of municipal characteristics, such as the population size, the average years of education of inhabitants, the fraction of employed population, the fraction of elderly population, Z_{it} is a vector of candidates' characteristics such as education, age and gender, the η_t is a vector of election year dummies, μ_i is a vector of municipal dummies (to capture unobserved geographical heterogeneity), and ε_{it} is an error term.

In Table 2 are shown OLS estimates considering as dependent variable *Valid (II round)*. In column (1) we only consider among regressors the variable *Electoral Margin* and the percentage of valid ballots at the first round. Any observable and unobservable determinants of electoral turnout antecedent to the first round should be captured by the variable *Valid (I round)* and controlling for it should attenuate possible biases from omitted variables.

From our OLS estimates, it emerges a strong and negative effect of *Electoral Margin* on the percentage of valid ballots (significant at the 1 percent level), that is, a positive effect of the expected closeness on electoral turnout. If the electoral margin shrinks of 10 percentage points, *ceteris paribus*, the electoral turnout at the second round increases of about 2 percentage points.

As expected, *Valid (I round)* has a strong effect on the turnout at the second round. The coefficient is 0.88, implying a very high degree of correlation between the turnout at the first and second round.

In column (2) we control for the percentage of valid, blank and invalid ballots at the first round and for the votes obtained by the two leading candidates at the first round. Results are very similar to those reported in column (1). In column (3) we add among regressors a number of municipal characteristics in terms of population size, average education, employment rate and the percentage of elderly population and some candidates' characteristics (average age, average education and a dummy for

the presence of at least a female among candidates). Again the variable *Electoral Margin* at the first round produces a negative and highly statistically significant effect on the percentage of valid ballots.[8]

In columns (4)-(6) we run the same specifications described above, but we control for municipal fixed effects, since unobservable municipal traits could affect both the degree of political competition and the electoral turnout. In all these specifications, we find that the *Electoral Margin* at the first round has a strong and negative impact on the *Turnout* at the second round.

The coefficients on *Electoral Margin* are always highly statistically significant (the *t*-stat is typically greater than 6 in absolute value) and quite stable across specifications. As regards the magnitude of the effect, we find that a decrease of 10 percentage points in the electoral margin between the two leading candidates increases the percentage of valid ballots of about 1.6-1.9 percentage points according to the specification used. In terms of standard deviations, an increase of one standard deviation in the electoral margin produces a reduction of about 0.13 standard deviations in the percentage of valid votes.

Our estimated effects are in line with those found by Simonovits (2011) for general elections in Hungary, but much higher (about two-three times) than those highlighted for France by Fauvelle-Aymar (2006) and Indridason (2008).

We also find that the effect of a closer political race is particularly relevant when electoral participation at the first round was lower than the average. By interacting *Electoral Margin* with the variable *Valid (I round)* demeaned, we show that a 10 percentage points lower electoral margin leads to an increase in the turnout (at the second round) of about 3.2 percentage points in municipalities where *Valid (I round)* is 10 points lower than the average. On the other hand, we find that in municipalities where *Valid (I round)* is 10 points higher than the average this effect is only 1 percentage point (results not reported). This implies that the electoral closeness produces a strong encouraging effect particularly in municipalities where at the first round the electors were less prone to cast their votes.

[8] We have also run separate regressions for Southern and Northern municipalities to investigate whether the relationship between closeness and turnout is different in the two geographical areas but we do not find any significant difference.

Table 2. Turnout (Valid Ballots) at the Second Round and Expected Closeness. OLS estimates

	(1) Turnout	(2) Turnout	(3) Turnout	(4) Turnout	(5) Turnout	(6) Turnout
Electoral Margin	-0.1937***	-0.1801***	-0.1874***	-0.1857***	-0.1620***	-0.1725***
	(0.0249)	(0.0234)	(0.0231)	(0.0236)	(0.0213)	(0.0216)
Valid (I round)	0.8820***	0.7532***	0.8345***	0.7617***	0.6051***	0.5831***
	(0.0306)	(0.0321)	(0.0333)	(0.0456)	(0.0442)	(0.0455)
Votes First Two Cand. (I round)		0.2032***	0.1922***		0.2092***	0.2143***
		(0.0161)	(0.0146)		(0.0165)	(0.0167)
Blank (I round)		0.4682**	0.0450		0.0452	0.0903
		(0.2057)	(0.1822)		(0.3035)	(0.3178)
Invalid (I round)		0.1204	0.2193		0.1906	0.1764
		(0.1607)	(0.1519)		(0.1956)	(0.1918)
Candidates' and Municipalities' Controls	NO	NO	YES	NO	NO	YES
Municipality Fixed Effects	NO	NO	NO	YES	YES	YES
Observations	1410	1410	1367	1410	1410	1367
Adjusted R-squared	0.620	0.663	0.717	0.693	0.749	0.752
Number of clusters	632	632	620	632	632	620

Notes: The dependent variable is the ratio between valid ballots at the second round and eligible electors. Candidates' controls include Candidates' Education, Candidates' Age, Female Candidates. Municipalities' Controls include Population, Education, Employment, Percentage Pop≥65. In all regressions we control for year dummies. Standard errors (corrected for heteroskedasticity and clusterized at the municipality level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In order to evaluate the possible bias deriving from measures of electoral closeness based on ex-post electoral results (as typically done in the literature, see Geys, 2006), we have computed the electoral margin as the difference (in absolute value) between the number of votes obtained by the two candidates at the *second* round (divided by the number of eligible voters). Using this measure we have run our OLS regressions (exactly the same specifications reported in Table 2).

Results reported in Table 3 show that although the impact of the electoral margin is negative and statistically significant, the magnitude of the effect is much smaller compared to our previous estimates, in which we proxy the electoral closeness with the margin at the *first* round. A decrease of 10 percentage points in the electoral margin increases the percentage of valid ballots of only 0.5-0.7 percentage points: the effect is about one third of the effect emerging from estimates reported in Table 2. This is arguably the result of the endogeneity bias deriving from the use of an ex-post measure of closeness.

Table 3. Turnout (Valid Ballots) at the Second Round and Closeness at the Second Round. OLS estimates

	(1) Turnout	(2) Turnout	(3) Turnout	(4) Turnout	(5) Turnout	(6) Turnout
Electoral Margin (II round)	-0.0508**	-0.0347*	-0.0475**	-0.0795***	-0.0703***	-0.0707***
	(0.0211)	(0.0200)	(0.0188)	(0.0208)	(0.0184)	(0.0190)
Candidates' and Municipalities' Controls	NO	NO	YES	NO	NO	YES
Municipality Fixed Effects	NO	NO	NO	YES	YES	YES
Observations	1410	1410	1368	1410	1410	1368
Adjusted R-squared	0.603	0.648	0.701	0.675	0.735	0.735
Number of clusters	632	632	620	632	632	620

Notes: The dependent variable is the ratio between valid ballots at the second round and eligible electors. Candidates' controls include Candidates' Education, Candidates' Age, Female Candidates. Municipalities' Controls include Population, Education, Employment, Percentage Pop≥65. In all regressions we control for year dummies. Standard errors (corrected for heteroskedasticity and clusterized at the municipality level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

In sum, our estimation results strongly confirm the idea that electors are induced to go to vote when the electoral competition is high. However, we are not able to disentangle the direct effect of closeness (due to the fact that the probability of affecting results increases) from the indirect effect deriving from the mobilization effort provided by candidates and political parties and from social pressure. In fact, when candidates are in a close race, they have strong incentives to put forth effort and to invest in electoral campaigns in order to convince electors to vote for them. Whereas the effect of mobilization effort can be taken into account considering the electoral turnout in the first round, we are not able to control for the effort devoted by candidates in the two weeks intervening from the first to the second round.

5. Robustness Checks

5.1. Dealing with non-random selection: Heckman correction

The sample we have used so far is a non-randomly selected sample of municipalities with more than 15,000 inhabitants. As explained above, the second ballot is held only if none of the candidates obtains an absolute majority of votes in the first ballot. According to the data, about 42% of municipalities (1,057) elects the mayor at the first round.

As shown by Heckman (1979), OLS estimates under these conditions could in principle produce inconsistent estimates. To correct for the fact that municipalities holding a second round are a non-randomly selected sample we use the two-step Heckman selection model.

Firstly, we need to specify a selection equation in order to model the probability of going to the second round. We use the *Number of Candidates* at the first round as independent variable included in the selection equation but arguably not relevant for the outcome equation (only the two candidates obtaining more votes compete in the second round). The *Number of Candidates* presumably affects the dispersion of votes in the first round and should positively affect the probability of going to the second round, since it is more difficult that a single candidate obtains more than 50 percent of the votes. On the other hand, once we control for the percentage of votes obtained by the two leading candidates in the first round, the number of candidates in the first round should not affect the turnout in the second round.

Estimation results of the Heckman selection model are reported in Table 4. In Panel A we report the outcome equation, in which the turnout (*Valid II round*) is put in relationship with the *Electoral Margin*, while in Panel B we report the results of the selection equation. We use a standard two-step estimation method. In the selection equation, in addition the *Number of Candidates*, we control for any variable that appears as an explanatory variable in the outcome equation.

As expected, we find that the *Number of Candidates* positively affects the probability of being in the sample (significant at the 1 percent level). The inverse Mills ratio λ is always not significantly different from zero, implying that there is no evidence of a sample selection bias and OLS estimates are

consistent. This is confirmed by estimates reported in Panel A of Table 4, which are in line with the OLS estimates reported in Table 2: the coefficient on *Electoral Margin* is around -0.16/-0.20, with a *t*-stat always around 6 or greater.

We have also estimated the Heckman Selection model using full maximum likelihood obtaining results very similar to those reported in Table 4.

Table 4. Heckman Selection Model

	(1) Turnout	(2) Turnout	(3) Turnout	(4) Turnout	(5) Turnout	(6) Turnout
Panel A: Turnout						
Electoral Margin	-0.2309***	-0.1754***	-0.1877***	-0.2084***	-0.1670***	-0.1670***
	(0.0394)	(0.0261)	(0.0206)	(0.0319)	(0.0237)	(0.0241)
Panel B: Selection Equation (Second Round)						
# Candidates (I round)	0.2627***	0.1978***	0.4320***	0.2917***	0.4011***	0.5755***
	(0.0220)	(0.0519)	(0.0972)	(0.0250)	(0.0841)	(0.1211)
Electoral Margin	-16.2142***	-96.9201***	-114.6131***	-17.1901***	-107.9374***	-114.5426***
	(0.6396)	(7.7906)	(12.0157)	(0.7388)	(10.5901)	(12.8416)
Observations	2467	2464	2295	2467	2464	2294
Censored	1057	1054	926	1057	1054	925
Lambda	0.0076	0.0023	0.0005	0.0046	0.0163	0.0189
S.E. Lambda	0.0064	0.0062	0.0061	0.0054	0.0069	0.0075
Rho	0.1568	0.0420	0.0127	0.1109	0.3487	0.4066

Notes: Two-step estimation method. Columns 1-6 replicate the estimates of columns 1-6 of Table 2. In columns 4-6 we include provincial fixed effects instead of municipality fixed effects (computationally more burdensome). In the selection equation we use as explanatory variable the *Number of Candidates* and any variables that appears as an explanatory variable in the outcome equation is also included in the selection equation. Standard errors (corrected for heteroskedasticity) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

5.2. Using a Proportional Dependent Variable Estimator

As explained by Wooldridge (2002), since our dependent variable is a proportion and takes values from 0 to 1, the linear model faces the same weaknesses as the linear probability model for a binary dependent variable. In order to deal with the fact that the dependent variable is a proportion, in this section we use the fractional logit model proposed by Papke and Wooldridge (1996).[9]

We replicate the specifications of Table 2. Parameter estimations, performed by maximum likelihood, are reported in Table 5. To make easily comparable these estimates with OLS ones, we report marginal effects. We find that the *Electoral Margin* has a strong and significant impact on the percentage of valid ballots. The magnitude of the effects and the statistical significance levels are very similar to those emerging from OLS estimates.

[9] We have also tried an alternative method – the beta regression model (see Ferrari and Cribari-Neto, 2004) – obtaining very similar results.

Table 5. Fractional Logit model. Marginal Effects

	(1) Turnout	(2) Turnout	(3) Turnout	(4) Turnout	(5) Turnout	(6) Turnout
Electoral Margin	-0.1963***	-0.1810***	-0.1893***	-0.1842***	-0.1615***	-0.1708***
	(0.0253)	(0.0239)	(0.0235)	(0.0235)	(0.0211)	(0.0213)
Observations	1410	1410	1367	1410	1410	1369
Number of clusters	632	632	620	632	632	620

Notes: Columns 1-6 replicate the estimates of columns 1-6 of Table 2. In columns 4-6 we include municipality fixed effects. Standard errors (corrected for heteroskedasticity and clusterized at the municipality level) are reported in parentheses. The symbols *** indicate that the coefficients are statistically significant at the 1 percent level.

A possible alternative to the fractional logit model is to transform the dependent variable as a log-odds ratio $\ln[y/(1-y)]$, and estimate a linear regression model. In this case too, we find results that are consistent with OLS estimates (results are not reported and are available upon request).

6. Electoral Closeness and Blank and Invalid Ballots

Blank and voluntary invalid ballots represent voters' choices that are difficult to interpret. From a cost-benefit point of view, blank and voluntary invalid ballots seem irrational as a voter who shows up at the election poll incurs the cost of voting but when casting a blank or invalid ballot does not receive any benefit. However, when a civic duty component is considered, these types of vote may result consistent with a rational choice as the voter may obtain a benefit from casting a ballot and to show up on election day even if he does not obtain the benefits deriving from supporting a specific candidate or party (see Galatas, 2008).

As explained in Section 2, the relationship between closeness and blank and invalid ballots could be either positive or negative. On the one hand, as the probability of casting the decisive vote is higher, the probability of casting a blank or invalid ballot should be lower in closer elections. On the other hand, closeness may increases candidates' and social pressure, leading electors who are unconvinced of competing candidates to cast a blank or invalid ballot. Finally, invalid ballots are also affected by the behavior of election officers.

We firstly investigate the effect of closer elections on blank ballots and then turn our attention to invalid ballots. In Table 6 are reported estimation results obtained replicating the specifications estimated in Table 2, but considering as dependent variable the percentage of blank ballots in the second round. In all specifications we find that *Electoral Margin* produces a positive effect on the percentage of blank ballots (the effect is highly statistically significant), that is, when the race is closer individuals are less induced to cast a blank ballot.

Our findings support the idea that the behavior of electors casting blank ballots is not very different from the behavior of electors who abstain from going to the polls, so they react in the same way when the electoral race becomes closer.

Table 6. Blank Ballots in the second round and Closeness. OLS estimates

	(1) (Blank)	(2) (Blank)	(3) (Blank)	(4) (Blank)	(5) (Blank)	(6) (Blank)
Electoral Margin	0.0100***	0.0085***	0.0065***	0.0094***	0.0071***	0.0059**
	(0.0024)	(0.0024)	(0.0021)	(0.0024)	(0.0024)	(0.0024)
Valid (I round)	0.0226***	0.0385***	0.0393***	0.0371***	0.0486***	0.0447***
	(0.0026)	(0.0030)	(0.0030)	(0.0052)	(0.0052)	(0.0051)
Votes First Two Cand. (I round)		-0.0160***	-0.0170***		-0.0133***	-0.0137***
		(0.0016)	(0.0015)		(0.0018)	(0.0018)
Blank (I round)		0.1449***	0.1059***		0.1779***	0.1667***
		(0.0250)	(0.0191)		(0.0345)	(0.0337)
Invalid (I round)		0.0155	0.0260*		0.0539*	0.0533*
		(0.0150)	(0.0153)		(0.0282)	(0.0289)
Candidates' and Municipalities' Controls	NO	NO	YES	NO	NO	YES
Municipality Fixed Effects	NO	NO	NO	YES	YES	YES
Observations	1410	1410	1368	1410	1410	1368
Adjusted R-squared	0.385	0.465	0.557	0.566	0.622	0.618
Number of clusters	632	632	620	632	632	620

Notes: The dependent variable is the ratio between blank ballots at the second round and eligible electors. Candidates' controls include Candidates' Education, Candidates' Age, Female Candidates. Municipalities' Controls include Population, Education, Employment, Percentage Pop≥65. In all regressions we control for year dummies. Standard errors (corrected for heteroskedasticity and clusterized at the municipality level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

Let us now turn our attention to the percentage of invalid ballots. In Table 7 are reported estimation results obtained considering as dependent variable the percentage of invalid ballots at the second round. We do not find any statistically significant impact of *Electoral Margin* on invalid ballots, although in some specifications the effect turns out to be positive and with a *p*-value of 0.13.-0.16.

Therefore, electors who abstain to go to the polls and those casting a blank ballot behave similarly when the electoral competition gets harsher. Invalid ballots also show a similar pattern, but the effects are statistically insignificant. A potential higher effort provided by unbiased election officers in checking ballots in response to closeness, which would lead to a negative correlation between invalid ballots and the electoral margin – as shown in Aldashev and Mastrobuoni (2011) – does not seem to be sufficient to counterbalance the effects deriving from electors' behavior.

Table 7. Invalid Ballots in the second Round and Closeness. OLS estimates

	(1) (Invalid)	(2) (Invalid)	(3) (Invalid)	(4) (Invalid)	(5) (Invalid)	(6) (Invalid)
Electoral Margin	0.0048	0.0020	0.0010	0.0051	0.0018	0.0014
	(0.0031)	(0.0030)	(0.0030)	(0.0037)	(0.0036)	(0.0038)
Valid (I round)	0.0113**	0.0356***	0.0389***	0.0150	0.0375***	0.0344***
	(0.0055)	(0.0055)	(0.0060)	(0.0097)	(0.0093)	(0.0104)
Votes First Two Cand. (I round)		-0.0245***	-0.0253***		-0.0243***	-0.0248***
		(0.0024)	(0.0025)		(0.0029)	(0.0030)
Blank (I round)		0.1341***	0.1070***		0.0966**	0.0776
		(0.0245)	(0.0252)		(0.0470)	(0.0488)
Invalid (I round)		0.1376***	0.1474***		0.0747	0.0752
		(0.0471)	(0.0500)		(0.0500)	(0.0537)
Candidates' and Municipalities' Controls	NO	NO	YES	NO	NO	YES
Municipality Fixed Effects	NO	NO	NO	YES	YES	YES
Observations	1410	1410	1368	1410	1410	1368
Adjusted R-squared	0.375	0.462	0.472	0.481	0.533	0.531
Number of clusters	632	632	620	632	632	620

Notes: The dependent variable is the ratio between invalid ballots at the second round and eligible electors. Candidates' controls include Candidates' Education, Candidates' Age, Female Candidates. Municipalities' Controls include Population, Education, Employment, Percentage Pop≥65. In all regressions we control for year dummies. Standard errors (corrected for heteroskedasticity and clusterized at the municipality level) are reported in parentheses. The symbols ***, **, * indicate that coefficients are statistically significant, respectively, at the 1, 5, and 10 percent level.

7. Non-Linearities in the Impact of Electoral Closeness

In this section we investigate whether electoral closeness has a non-linear impact on valid, blank and invalid ballots. In fact, there are no particular reasons to expect a linear relationship. To analyze this aspect, we have classified our variable *Electoral Margin* in four categories (approximately corresponding to quartiles): the first group refers to cases in which at the first round the best performing candidate leads by less than 2%; the second one considers situations in which the electoral margin is higher than 2% but smaller than 5%; the third one refers to values of the electoral margin in the range 5% and 10%; the last case considers situations in which the best performing candidate leads by more than 10%. We build a dummy variables for each category.

In Table 8 are reported estimation results using the most complete specification used in previous estimates (column 6 in Tables 2, 6 and 7), considering as dependent variable, alternatively, the percentage of valid ballots (column 1), the percentage of blank ballots (column 2) and the percentage of invalid ballots (column 3). We leave as reference category elections characterized by an electoral margin of less than 2%.

Interestingly, as far as valid ballots are concerned (column 1), we find a statistically significant difference between the first and the second group (significant at 5% level). In addition, the percentage of valid ballots tends to furtherly reduce when the electoral margin reaches the range 5%-10% compared to electoral contexts characterized by smaller margins of victory. The negative impact of the electoral margin is even stronger when the best performing candidate leads by more than 10%.

Blank ballots (in column 2) seem to react more strongly when the electoral margin becomes particularly high. In fact, we find small differences between the first two groups, estimated rather

imprecisely, while an increase in the percentage of blank ballots emerges when the electoral margin is higher than 10%.

Estimation concerning invalid ballots show that they tend to increase when the margin increases from less than 2% to values in the range 2%-5%. However, further increases in the electoral margin do not seem to produce any statistically significant effect on invalid ballots compared to the reference category.

Table 8. Non Linear Effects of Closeness on Valid, Blank and Invalid Ballots. Municipality Fixed Effects

	(1) Valid (II round)	(2) Blank (II round)	(3) Invalid (II round)
2%<=Margin<5%	-0.0069**	0.0007**	0.0016**
	(0.0034)	(0.0003)	(0.0006)
5%<=Margin<10%	-0.0108***	0.0005	0.0009
	(0.0031)	(0.0004)	(0.0006)
Margin>=10%	-0.0262***	0.0011***	0.0008
	(0.0034)	(0.0004)	(0.0006)
Observations	1368	1368	1368
Adjusted R-squared	0.750	0.619	0.534
Number of clusters	620	620	620

Notes: We replicate the estimates of column 6 of Tables 2, 6 and 8.

8. Concluding Remarks

In this paper, relying on the dual ballot system adopted in Italy for municipal elections in cities with more than 15,000 inhabitants, we have estimated the impact of electoral closeness on valid, blank and invalid ballots, using as a proxy of expected closeness at the second round the actual closeness at the first round between the two leading candidates.

Using a very rich data set, providing information on the number of valid, invalid and blank ballots at the first and at the second round elections, on the number of people eligible to vote and on some candidates' characteristics, we estimate several models with and without municipal fixed effects. We find that closeness strongly increases valid ballots and reduces blank ballots, suggesting that voters who cast a blank ballot react to closer elections similarly to those who decide to abstain from going to the poll. A decrease of 10 percentage points in the electoral margin produces an increase of about 2 percentage points of valid votes and a decrease of about 0.1 in the percentage of blank ballots. On the other hand, we do not find any statistically significant effect on the percentage of invalid ballots, maybe because of opposing effects deriving from the behavior of electors and of election officers: the former may be less inclined to cast an invalid ballot when the electoral margin shrinks, while the latter may increase their effort in checking valid preferences so increasing the number of invalid ballots when the risk of erroneously appointing the winner increases.

Our results are robust to the use of an Heckman selection model to tackle problems deriving from the fact that the second round only takes place if no candidate has obtained the majority of votes in the

first round. The same findings emerge also when we estimate a fractional logit model to consider the nature of proportion of our dependent variables.

Unfortunately, we are not able to accurately disentangle the direct effect of closeness (due to the higher probability of affecting the electoral outcome) from the indirect effect deriving from the mobilization effort provided by candidates and political parties and from social pressure. In fact, while we can take into account mobilization campaigns taking place until the first round election, including among regressors the electoral turnout in the first round, we are not able to control for the effort devoted by candidates in the two weeks from the first to the second round.

References

Aldashev, G. and Mastrobuoni, G., (2011), "Invalid Ballots and Electoral Competition", *Carlo Alberto Notebooks*.

Cox, G., and Munger, M., (1989), "Closeness, expenditures, and turnout in the 1988 US house elections" *American Political Science Review*, 83(1), 217–231.

Dhillon, A., and Peralta, S., (2002), "Economic theories of voter turnout", *Economic Journal*, 112(june), 332–352.

Downs, A., (1957), *An Economic Theory of Democracy*. New York: Harper and Row.

Endersby, J., Galatas, S. and Rackaway, C., (2002), "Closeness counts in Canada: Voter participation in the 1993 and 1997 federal elections" *Journal of Politics*, 64, pp. 610-631.

Fauvelle-Aymar, C. and Francois, A., (2006), "The impact of closeness on turnout: An empirical relation based on a study of a two-round ballot", *Public Choice*, 127, pp. 469-491.

Feddersen, T., (2004), "Rational choice theory and the paradox of not voting", *Journal of Economic Perspective*, 18(1), 99–112.

Ferrari, S. and Cribari-Neto, F., (2004) "Beta regression for modelling rates and proportions", *Journal of Applied Statistics*, 31, pp. 799-815.

Galatas, S., (2008), "None of the Above?" Casting Blank Ballots in Ontario Provincial Elections, *Politics & Policy*, pp. 448–473.

Geys, B., (2006), "Explaining voter turnout: A review of aggregate-level research", *Electoral Studies*, 25 (2006) 637-663.

Heckman, J., (1979), "Sample selection bias as a specification error", *Econometrica* 47 (1): 153–61

Indridason, I., (2008), "Competition and turnout: the majority run-off as a natural experiment", *Electoral Studies*, 27, pp. 699-710.

Matsusaka, J., and Palda, F., (1993), "The downsian voter meets the ecological fallacy", *Public Choice*, 77, 855–878.

Papke, L. and Wooldridge, J., (1996), "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates", *Journal of Applied Econometrics* 11, pp. 619–632.

Power, T. and Garand, J., (2007), "Determinants of invalid voting in Latin America", *Electoral Studies*, pp. 432-44.

Shachar, R. and Nabeluff, B., (1999), "Follow the leader: Theory and evidence on political participation", *American Economic Review*, 89(3), 525–549.

Simonovits, G., (2011), "Closeness and Turnout in runoff elections: Evidence from Hungary", mimeo.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press.