CLOSENESS AND TURNOUT IN RUNOFF ELECTIONS

EVIDENCE FROM HUNGARY

By Gábor Simonovits

Submitted to Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Master of Economics

Supervisor: Professor Gábor Kézdi

Budapest, Hungary

2011
Abstract

This study aims at contributing to the literature on the effect of political competition on turnout. I test the Downsian Closeness Hypothesis (DCH) on data from runoffs in General Elections in Hungary. The expected closeness of the runoffs is proxied with first round margins. Also I consider estimation procedures that take it into account that the dependent variable is a fraction. The findings are consistent with the DCH: increases in margins between parties in the first round significantly decrease turnout in the second, even when turnout in the first round is controlled for. This is in line with the theoretical considerations of the DCH but contrary to a large part of the existing empirical literature.
Acknowledgements

I would hereby like to express my gratitude to the people who helped me in writing this thesis. First and foremost I would like to thank Gabor Kezdi, my supervisor who provided with patient but demanding guidance throughout the research. Insightful comments from Gabor Toka and other members of the Political Behavior Research Group also proved extremely useful. Robin Bellers gave important comments concerning the structure and style of the paper. Also, I am heavily indebted to David Kovacs and Marton Kovesdy for their indispensable help in assembling the data I worked with.
# Table of contents

Abstract.................................................................................................................................. i
Acknowledgements ................................................................................................................ ii

1. Introduction ........................................................................................................................1

2. Rational choice theories of voting .......................................................................................3
   2.1. Decision theoretic approaches ..................................................................................3
   2.2. Game theoretic approaches .......................................................................................4
   2.3. Alternative theories of voting ...................................................................................5
   2.4. What do rational choice models say about closeness and turnout?.........................6

3. A summary of existing empirical studies.............................................................................7
   3.1. The determinants of voter turnout.............................................................................7
   3.2. The effect of closeness on turnout.............................................................................8
   3.3. Methodological issues ..............................................................................................9

4. The Hungarian general elections in 2002 and 2006...........................................................11
   4.1. The Hungarian Electoral System ............................................................................11
   4.2. The General Elections of 2002 and 2006 ...............................................................12
   4.3. Closeness and turnout in Hungary – A look at the data ...........................................13

5. Estimation strategy ...........................................................................................................16
   5.1. A statistical model of runoff elections.....................................................................16
   5.2. Model specification and estimation strategy ...........................................................17

6. Estimation results .............................................................................................................20

7. Conclusion .......................................................................................................................23

References............................................................................................................................25
“Saying that closeness increases the probability of being pivotal is like saying that tall men are more likely than short men to bump their heads on the moon” (Schwartz, 1987, p.118)

1. Introduction

Understanding the determinants of voting turnout has been one of the most important enterprises of political science. On one hand, electoral participation is generally thought of as key element in the working of political system. On the other hand, the literature on turnout has been controversial both in terms of theoretical and empirical results.

This study aims at contributing to the empirical literature on the effect of expected political competition on turnout. The insight that voting turnout is linked to the expected closeness of an election (the so called Downsian Closeness Hypothesis) goes back as far as the seminal work of Anthony Downs, *An economic theory of democracy* (1957). He argued that the voting behavior is based on an expected cost-benefit analysis with the benefit of voting being the chance of influencing election results. This claim has been transformed to the *Paradox of Voting*, a sort of skeleton in the cupboard for positive political economy: why do people actually vote if the chance for them to make any difference is so little?

There are two quite different contexts in which the answer to this question is particularly important. First, from a normative point of view, it would be rather attractive to see whether two desirably characteristics of a political system – party competition and political participation – are related to one another. If that was the case, than institutional arrangements or direct policies targeted at one of the two could in principle indirectly affect the other. Second, the extent to which people are rational (or, at least their behavior can be described by rational choice models) has long been an important issue in the research of electoral behavior. The hypothesis that voters “react rationally” to changes in the political environment, such as the degree of party competition is a sort of stronghold for scholars advocating rational choice modeling.

The reason why the Paradox of voting is still subject to research is probably the difficulty of establishing a causal relationship between election closeness and turnout. Not surprisingly, the closeness of competition and the turnout at a given election is determined at the same time.

---

1 See Grofman (1993) who asks the question explicitly in the title: is turnout the paradox that ate rational choice theory?
Thus it can be argued that each of the previous studies that used actual election results to “proxy” the expected closeness of the same elections “encounter[ed] a serious problem of endogeneity” (Fauvelle-Aymar and Francois, 2006 p.473.). Finally, it is also argued that the effect of closeness should be estimated using “within variation” of the same electorate. The reason for that is that the ceteris paribus effect is more plausible when the same groups of voters are exposed to different political environment (Grofman, 1993).

In this study I am going to test the link between expected closeness and turnout in a way that mitigates the methodological flaws present in existing literature. The approach taken here will be similar to that of Fauvelle-Aymar and Francois (2006) and Indridason (2008) in that I use data on two-round election (Hungary, 2002 and 2006) to find causal links between expected closeness and actual turnout. As I will argue later on, the Hungarian elections in 2002 and 2006 provide an “ideal setting” to test the effect of closeness on turnout for three reasons. First, two-round elections allow for proxying expected closeness of the runoff by actual closeness in the first round, similarly to the case of France (as in Fauvelle-Aymar and Francois, 2006). Second, defining the closeness of an election is more straightforward in this case since in 2002 and 2006 basically only two parties competed for single-seat constituencies a rare situation in two-round elections. Third, making use of the similar structure of the two elections I can test hypotheses in a panel context (i.e. to use two waves of elections for the same constituencies) which is likely to ensure more precise estimates and the identification of the effects from variation within constituencies.

The results presented in this paper support the 50 years old theoretical insight: voters who live in constituencies where the political competition is expected to be close are in fact more likely to vote. Results are not sensible to the specification of the model: the effect of closeness stays roughly the same in linear and non-linear models. Also, the case of Hungary offers a setting where indirect effects of closeness through mobilization can be logically excluded.

The remainder of the paper is organized as follows. Section 2 and 3 summarize the previous theoretical and empirical literature respectively. In section 4 I provide some arguments for the rationale of using the Hungarian General Elections of 2002 and 2006 for the study and also present some basic facts about the two elections. Section 5 gives details about the empirical strategy. Section 6 provides the estimation results and Section 7 concludes.

---

2 Fauvelle-Aymar and Francois (2006) used data on French general elections, which also have two rounds. The significance of two-round elections will be shown later.
2. Rational choice theories of voting

In this section I give a brief summary of the major contributions to the rational choice theories of electoral participation. It is worth to note that even though economists and political scientist have been engaged with explaining turnout for more 40 years, “there is not a canonical rational choice model of voting in elections” (Feddersen, 2004 p. 100). That is why I feel that it is appropriate to present a brief discussion of the most influential theories of turnout.

2.1. Decision theoretic approaches

The theoretical foundation of the causal link between election closeness and turnout goes back as far as Downs (1957). He was the first to suggest that voting decision reflects a rational calculation of expected utilities depending on the voting decision. In the classic framework developed further by Riker and Ordeshook (1968) the underlying cost and benefits associated with voting are captured in the following equation:

\[ R = PB - C + D \]  

(1)

where \( R \) is the expected utility of voting, \( P \) is the probability of casting a decisive vote, \( B \) is the benefit of the preferred candidate being the winner, \( C \) is the costs of voting and \( D \) is the utility of voting regardless of the outcome, often labeled as “civic duty”. The model predicts that only those people will show up in the Election Day for whom

\[ PB + D > C \]  

(2)

A notable implication of this theory is that voters should be more likely to go to the ballots when they think they have more chance to influence the results (i.e. they have more chance to cast the decisive vote). This conjecture about the effect of the \( P \) term of the equations above become a famous and controversial one, usually referred to as the Downsian Closeness Hypothesis (DCH)(Matsusaka and Palda, 1993).

The reason the DCH has induced so much debate is that the probability of being pivotal is tiny, both theoretically and empirically. Under some distributional assumptions on the distribution of turnout one can determine the probability of election results as a function of the size of the electorate and the distribution of preferences for the parties (Mueller, 2003). Then, the probability of a tie in a large scale election can be simulated for various population
parameters (as is done by Beck, 1975 and Margolis, 1977). As it turns out, even for an election which is extremely close \textit{ex ante}, the probability of an exact tie quickly converges to zero as the number of voters increase (Beck, 1975). Also, the empirical frequency of tied elections has been extremely low (Mulligan and Hunter, 2003).

The fact that the probability of being a pivotal voter on a large scale election was proved to be negligible called the Downsian Closeness hypothesis into question. It was pointed out that with such minuscule probability of tied election; even a small net cost of voting should imply zero turnout.

\textbf{2.2. Game theoretic approaches}

In the 80’s a number of scholars questioned the validity of thinking about election in terms of models in which the probability of a pivotal vote is taken as an exogenous parameter. Their typical argument was as follows:

\begin{quote}
"If each rational voter were to decide not to vote because her vote has to small of a chance of affecting the outcome, and all voters were rational, no one would vote. But then, any one voter could determine the outcome of the election by voting... The greater the number of voters I expect will rationally abstain, the more rational it is for me to vote" (Mueller, 1989, 351-52)
\end{quote}

Treating the probability of being pivotal as an endogenous variable is the major contribution of the “game theoretic” approach to turnout. In this study I will not consider strategic interaction between voters; however I sketch the most important results of this strand. A good review of strategic models of turnout can be in Feddersen (2004).

The first game theoretic accounts of turnout were suggested by Ledyard (1981, 1984). He set up a model in which the equilibrium turnout is positive if candidate platforms are different; even if each of the voters faces strictly positive net costs. However, if the model is extended in a way that the platforms of candidates are also determined endogenously, as platforms converge, the participation decreases to zero.

Palfrey and Rosenthal (1983) consider elections in which candidates take different platforms that are fixed and voters face an identical positive net cost of voting. They show that if the two competing candidates have nearly identical number of supporters positive turnout can be achieved as an equilibrium outcome for even a large electorate. On the contrary, in a follow-
up article (Palfrey and Rosenthal, 1985) the same authors present a modified model in which the party preference and costs of voters are treated as private information. Then, the Perfect Bayesian Equilibria of the game are analyzed. Palfrey and Rosenthal find that in such a setup the equilibrium probability of being pivotal converges to zero in large electorates. Consequently, the model predicts negligible turnout regardless of the true distribution of preferences for candidates.

Yet another approach emphasizes the role of asymmetric information across voters. The idea that people may not show up on the Election Day simply because the find it to costly to get information about candidates is as old as Downs (1957). The phenomenon of rational ignorance was formalized by Matsusaka (1995) who set up a two stage model of voting in which the level of information about candidate positions acquired by individuals is endogenous. The insight that information about candidates is a key issue in voting is embedded in the game theoretic approach taken by Feddersen and Pesendorfer (1996, 1999). They show that in large scale elections even if voting is not costly, some uninformed voters may abstain specifically because they do not want to be pivotal. In such a setup the extent to which the electorate is informed about candidate positions is a crucial element in the determination of turnout.

2.3. Alternative theories of voting

In this section I briefly discuss some other theories of turnout that do not fit into the categories specified above. The two approaches to be discussed here are models of voter mobilization and models with “ethical voters”.

Theories of mobilization move away from the individual level and analyze elections as games between political leaders. Studies taking this approach emphasize the role of elites who make costly efforts in order to get out their supporters to vote. This view on elections appears implicitly in the study of Cox and Munger (1989) who argue that “perhaps closeness [of elections] causes elite actors to increase their activity in the electoral process, thus stimulating turnout” (p. 218). Shachar and Nalebuff (1999) set up a full fledged general equilibrium model of US presidential elections mobilization efforts of leaders are determined endogenously as a function of the expected closeness of state and national races. Their model predict an indirect effect of expected closeness: political leader have greater incentives to mobilize voters in close elections.
Finally, a recent strand of research emphasizes the role of civic duty in the voting decision. Models of ethical voters specify *sociotropic preferences* to voters, that is, they get payoffs if they act according to some rule which is optimal for a subset of the electorate. Feddersen and Sandroni (2002) develop a model in which the “ethical rule” means a cut-off point in terms of voting costs: those facing smaller costs then the threshold value are “supposed to” vote, while does who do not are supposed to abstain. Analyzing Bayesian equilibria of such a game, the authors find that the having more similar groups in terms of size increases equilibrium turnout and decreases the margin of victory.

The theories described above seem to dominate current research on electoral participation. As compared to earlier models of turnout, they seem to be more appropriate in explaining empirical regularities. Nevertheless, as Feddersen (2004) points out, these model have failed to explain “how people join or identify with their groups in the first place” (p. 110)

### 2.4. *What do rational choice models say about closeness and turnout?*

After this survey of the theoretical literature on electoral participation, it could prove useful to summarize the findings of each of the approaches concerning closeness and turnout. This can serve as a motivation of why the exact goal of this paper – that is, to test the relationship between closeness and turnout empirically – is of particular importance.

As Blais (2000) puts it “the trademark of rational choice [models of turnout] is in the P term” (p. 14). That is, each of the accounts of turnout presented above features the probability of being pivotal as a key explanatory variable. This is equivalent to the statement that in any rational choice model of electoral participation, one of the expected benefits that make people show up in the Election Day should be the hope of influencing the outcome of the voting.

At the same time, one should recognize the differences in the ways in which P works in the models above. In the decision theoretic framework, the pivotal probability is an exogenous parameter of the model and thus determines turnout directly. In the game theoretic models of voting P evolves as an endogenous variable as a function of the distribution of individual characteristics, namely that of voting costs and the proportion of various voter types in the electorate. Finally, in the case of mobilization models, expected closeness of the race affects turnout indirectly, through the actions of the elites (or political leaders).
3. A summary of existing empirical studies

In this section I summarize the empirical evidence on electoral participation. First, I summarize the results on turnout in general and then the effect of closeness in particular. Finally, I point out some of the methodological concerns about existing evidence. For a more extensive review of the literature see Blais (2000), Mueller (2003) and Geys (2006).

3.1. The determinants of voter turnout

Even if one does not convinced by rational choice theories of turnout, one can look at the evidence on the determinants of voting without believing to any of the competing models of electoral participation. There are three particular sets of variables to look at: institutional arrangements and political variables. I will consider the first two here and then elaborate on the effect of political variables in section 3.2.

There is convincing evidence that institutional arrangements of elections in different countries or other polities have a large effect on turnout. First, there are institutional arrangements that affect turnout through the cost of voting. It probably comes as no surprise that turnout is generally higher in countries where there are institutionalized incentives to vote. In particular Blais (2000) reports that compulsory voting increases turnout by as much as 11% (p. 27). Along the same lines, turnout is reported to be lower in countries where institutional settings are such that people who intend to vote have to engage in registration procedures (Kelley et al., 1967). Finally, it is also argued that turnout is higher in concurrent elections i.e. when multiple elections are held at the same day (Hill and Leighly, 1993).

Institutional characteristics may affect turnout through channels other than the cost of voting. There is evidence of systematic differences in terms of turnout across countries with different electoral and party systems. As for electoral systems, they, by definition, determine the way in which votes are mapped to actual political outcomes, that is the distribution of seats (Farell, 2001). One can think about more than one channel through which electoral system could affect turnout. On one hand Ladner and Miller (1999) argue that people should be more inclined to participate in elections characterized by proportional representation (PR) as it is generally believed to go with a stronger association between the distribution of votes and seats. On the other hand, the same authors also argue that the fact that PR is harder to
understand and more prone to party coalitions may bring about smaller turnout. Thus the way different electoral systems affect turnout across remains an empirical question. Most of the empirical studies come to the conclusion that turnout is actually higher in PR systems (Geys, 2006).

Another stream of studies analyzed the differences in voting behavior across different socio-economic groups. The effects of various socio-economic variables are usually studied in micro-level analyses, that is, using survey data on political participation. The theoretical bases of such effects are at best heterogeneous but the patterns in the results of existing empirical studies are unequivocal.

The first study in which the role of socio-economic characteristics was systematically investigated was authored by Wolfinger and Rosenstone (1980) who analyzed US survey data. They found education to be the single most important characteristic explaining turnout. Specifically, their estimates point out a 38% gap between people with college degree and the others. Also, they found that the propensity to vote also increases with age and income (Wolfinger and Rosenstone, 1980).

Other studies, analyzing voters of European countries find similar empirical patterns. Blais (2000) analyzes turnout in a pooled sample of nine different countries. His results support the findings of Wolfinger and Rosenstone, (1980): older, richer and more educated people are more likely to vote as well as religious and married members of the electorate. Employment status is not found to be a significant determinant of electoral participation Blais (2000).

While the empirical patterns presented above explain a large proportion of the variance in turnout they are not suited to test the rational choice of model. For example the effects of socio-economic characteristics might be used to underpin an alternative, sociological interpretation of voting Blais (2000). Thus, I now turn my attention to empirical relationship which is central to this paper, the effect of closeness on turnout.

3.2. The effect of closeness on turnout

Arguably, closeness is „by far the most analyzed element in the turnout literature” (Geys, 2006, p. 647). One reason for that is the key role of closeness (which is used as a proxy for the probability of being pivotal) I the rational choice theories of turnout. The other reason is that existing studies have come up with rather different estimates concerning the effect of closeness on turnout.
Mueller (2003) lists more than 50 papers and shows parameter estimates of proxies of closeness. These estimates vary a lot both in terms of magnitude and significance, but all significant coefficients have the right sign. In a meta-analysis of the existing aggregate level empirical literature on turnout Geys (2006) lists 362 estimates of the effect of closeness on turnout in 52 studies. Out of the 362 the (expected) significant, positive effect is measured in 206 which makes the 57% of the tests presented by Geys (2006, Table 3, p 646). Of course one should be cautious with the interpretation of meta-analyses of this sort but it seems clear that the evidence is far from convincing. All in all Geys (2006) reports that “the estimated size of the effect is such that an increase in closeness with one standard deviation unit increases turnout rates by approximately 0.58-0.69 standard deviation units.” (p. 647)

Even more puzzling is the pattern of empirical findings in the studies that related closeness and turnout using micro-data. This strand of the research typically analyzed surveys which contain data about individuals’ (self reported) turnout and the perceived closeness of the election they participated in (Riker and Ordeshook, 1968, Ashenfelter and Kelley, 1975 and Blais, 2000). These studies typically found a positive relation between closeness and turnout. The problem with this approach is that one can suspect that those people who actually voted are more likely to have overestimated the closeness of the given election.

Matsusaka and Palda (1993, 1999) tackle this problem by relating survey data on turnout to aggregate measures of election closeness using data on Canadian elections. They find no systematic correlation between participation and closeness and conclude that aggregate evidence is likely to suffer from aggregation bias. This stream of research is still inconclusive as it generally hard to assemble datasets in which individual and aggregate characteristics could be linked. Throughout reviews of the literature can be found in Matsusaka and Palda (1998), Blais (2000) as well as in Endersby et al. (2002).

### 3.3. Methodological issues

In this section I review the main points of criticism concerning aggregate studies of the effect of closeness on turnout.

On one hand, there is a growing concern with the use of *ex post* election results to proxy *ex ante* expectations about the closeness of the race. The theoretical and empirical problems with this approach are pointed out among others by Kirchgassner and Zu Himmern (1997), Shachar and Nalebuff (1999) and Fauvelle-Aymar and Francois (2006). First, from a purely
theoretical point of view it’s hard to justify the assumption that voters foresee election outcomes perfectly. As Kirchgassner and Zu Himmern (1997) put it „in the previous literature about the closeness effect it has hardly ever been mentioned that a precondition for such an effect is that voters and/or candidates are able to form expectations about the electoral outcome“. Second, the same problem has raised severe concerns about the endogeneity of closeness in the usual regression equations. That is, the left hand side variable (the closeness of the election) is jointly determined with the right hand side variable (turnout) which in turns causes bias in the estimates.

On the other hand, the effect of expected closeness might be reinforced by campaign effects. That is, if the race is expected to be close politicians and parties may find it reasonable to put more effort into mobilizing voters (Key, 1957; Denver and Hands, 1974). These two effects can be separated both analytically and empirically and indeed there is considerable evidence on the indirect effect of closeness through mobilization of the voters (see Cox and Munger, 1988). As mentioned above Shachar and Nalebuff (1999) provide a general equilibrium framework in which mobilization efforts are determined endogenously as a function of the expected closeness of Congress election. They structurally estimate their model and find that much of the closeness-turnout relationship can be explained in terms of their proposed mobilization model.

Finally, existing estimates of the closeness-turnout relationship have not taken the limited nature of the dependent variable into consideration. If one acknowledges that turnout is actually a proportion then the empirical model should reflect this, setting up the problem in a non-linear framework. This issue is particularly important if we hypothesize that the marginal effects of closeness is not constant.
4. The Hungarian general elections in 2002 and 2006

In the following section I provide some information of the Hungarian electoral system and the elections of 2002 and 2006. In doing so I will focus on the aspects of them that render these elections appropriate for the purposes of this study. For a detailed description of the Hungarian electoral system see Körösényi (1998).

4.1. The Hungarian Electoral System

In this section I rely heavily on Körösényi, Török and Tóth (2007). The electoral system of Hungary is usually referred to as “mixed” as it contains elements of both majoritarian and proportional. Votes are mapped to mandates according to these two principles and three branches. Voters have two votes: one may be cast for candidates in the single seat constituencies, and the other on a regional party list. As for the first vote, the mandates are distributed in a majoritarian system, i.e the “first past the post” gets the seat. As for the party lists, seats are distributed proportionally to the votes. An important element of the system is the national, or “compensation” seats: these are distributed according to the number of votes cast on parties in constituencies where they lost.

I confine myself here to the detailed description of the single-seat constituencies only, since only those are relevant for the purpose of this paper purposes. The election of the MP-s from single seat constituencies comprises of two rounds except when one of the candidates gains the absolute majority of the votes in the first round already. In the case of single seat constituencies coalitions and mutual withdrawal of candidates between allied parties often proves decisive. In the second round, the race is often decided by the third strongest candidates who may step back in favor of one of the other two. Of course for that a group of voters that can be persuaded is very much needed.

One peculiarity of the Hungarian system is that voters face very different incentives to show up in the two rounds. In the first round, each of the votes counts: even if one votes for a candidate that loses her constituency, her vote will be channeled to the so called “compensation list”, which in turn, earns seats to the preferred party of the give voter in the
However, in the runoff, a vote “counts” only if it helps a given party to win the single seat constituency. Thus, in the first round people may find it rational to vote even in constituencies, where they expect a large margin of victory; whereas the same logic does not apply to the runoff.

4.2. The General Elections of 2002 and 2006

Now, I sketch some of the tendencies and communalities appearing in the two elections analyzed in the paper. I focus on two key phenomena: the concentration of the party system and growing awareness in voting strategies.

The single most important similarity of the General Elections of 2002 and 2006 was that by that type the Hungarian party structure essentially transformed into a two-party competition between Fidesz (Alliance of Young Democrats) and MSZP (Hungarian Socialist Party) (Szoboszlai, 2003). The concentration of the party structure had been growing since the fall of the socialist system: this is underpinned by the decline and then stagnation in the number of effective parties and growth of the share of votes cast on the two major parties. Of course this does not mean that we can speak of a classic two-party system: both Fidesz and MSZP were in a great need of finding potential allies to form coalitional governments. However we can speak of a quasi-two party system following Tóth (2001) as the two strongest parties together gained the 83 and 85% of the votes, respectively.

In particular, MSZP entered a close alliance with SZDSZ (Alliance of Free Democrats) a small liberal party with which MSZP governed the country in a coalition between 1994 and 1998. This resulted in mutual withdrawal of candidates in the favor of each other in both 2002 ad 2006. On the other hand, in 2002 FIDESZ established a similar agreement with MDF (Hungarian Democratic Forum), their coalitional partner between 1998 ad 2002 and the proposed joint candidates in the single seat constituencies. This alliance was broken in 2006 when MDF decided to have her own independent candidates.

The close competition between the two major parties brought about an unforeseen mobilization of the voters: the 2002 election produced an average turnout of 71%, the highest since ‘89 with a still greater participation in the second round. In 2006 turnout declined to 68% in the first round and 64% in the second, still being the second largest historically. These

---

3 An important caveat should be made here: the votes cast on parties who gain less than 5% of the multi-member list are actually lost.
figures show that by the 2000’s the Hungarian electorate “learnt” that the second round often proves decisive in the final results of the elections.

In short, in the second round only the candidates of the two major parties (or party alliances) had any chance to win. That is, the expected closeness of the election can well be thought of as the difference in the votes candidates of the two large parties got in the first round.

4.3. Closeness and turnout in Hungary – A look at the data

In this section I provide a preliminary look at the data on the Hungarian general elections of 2002 and 2006. Also, I give some details of measurement issues and the definitions of the main variables of interest.

The data on election results was downloaded from a website [www.vokscentrum.hu](http://www.vokscentrum.hu) which collects data on Hungarian elections. The dataset contains constituency level information of elections taking part after 1989. I used information on the number of eligible voters, the number of valid votes and the number of votes cast for the major parties in the two rounds of the two elections.

Turnout, the dependent variable in the empirical analysis is defined as the ratio of the number of valid votes and the number of eligible voters in a constituency. This measure was preferred to the number of votes as it is easier to interpret the estimates in such a framework. As the number of eligible voters intuitively will not change between the two rounds I use the figures reported for the first round.

Closeness is more problematic. First, it is ambiguous which parties to compare in terms of the number of votes. As I have mentioned above, in the second round in both 2002 and 2006, a vast majority of the constituencies was won by either Fidesz or MSZP. Exceptions are: one independent victory in 2006 and a couple of withdrawals in favor of SZDSZ by MSZP. Thus, I define closeness as the percentage difference between votes cast for the two parties who gained the most votes in the first round. As an alternative specification, percentage difference between “party blocs” is considered, i.e. the percentage difference between the vote share of the right (FIDESZ and MDF) and left (MSZP and SZDSZ). I publish the descriptive statistics of the variables I used in the analysis for the two elections in Table 1 and 2.

---

4 The data was downloaded in December, 2010.
5 The majority of the studies on turnout use this measure, and some papers (such as Indridason, 2008) confirms that results are robust in terms of the definition of the dependent variable.
In Table 1 I report descriptive statistics of constituency level turnout. One can immediately see that around two-thirds of the electorate got out to vote in both 2002 and 2006. Turnout increased between the two rounds in 2002 (presumably due to the closer race in the national contest) and declined in 2006. Also, one can observe that the variation of turnout in different constituencies was considerable in each of the election rounds.

<table>
<thead>
<tr>
<th>Year</th>
<th>1st round</th>
<th>2nd round</th>
<th>Change in turnout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Standard deviation</td>
<td>Average</td>
</tr>
<tr>
<td>2002</td>
<td>69.7%</td>
<td>5.7%</td>
<td>72.9%</td>
</tr>
<tr>
<td>2006</td>
<td>67.0%</td>
<td>5.3%</td>
<td>63.9%</td>
</tr>
<tr>
<td>Average</td>
<td>68.3%</td>
<td>5.5%</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

Table 2 shows vote shares of the two major parties, Fidesz and MSZP in the single-seat constituencies as well as the margin of victory. The figures are striking: the average share of votes gained by two major parties was extremely close in both years. Looking at margins in the constituency levels reveals that in no less than 33 contests the margin of victory was less than 1%. Figure 1 shows the distribution of margins in a histogram.

<table>
<thead>
<tr>
<th>Year</th>
<th>FIDESZ</th>
<th>MSZP</th>
<th>Margin of victory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Standard deviation</td>
<td>Average</td>
</tr>
<tr>
<td>2002</td>
<td>39.5%</td>
<td>8.5%</td>
<td>41.7%</td>
</tr>
<tr>
<td>2006</td>
<td>42.3%</td>
<td>8.3%</td>
<td>42.3%</td>
</tr>
<tr>
<td>Average</td>
<td>40.9%</td>
<td>8.4%</td>
<td>42.0%</td>
</tr>
</tbody>
</table>

A natural starting point for the empirical analysis of the effect of closeness on turnout is to check the correlation between 1st round closeness and change in turnout between the two rounds (as in Indridason, 2008). One would expect a negative relation between expected margin and change in turnout in the data.

Figure 2 provides a clear confirmation of this hypothesis. There is a clear negative relationship between the margin of victory in the first round and the change in turnout in both elections. One can immediately observe that there is a shift in the distribution of turnout change with participation increased in 2002 and declined in 2006 but the slope of the relation seems to be rather similar. This suggests that the effect of expected closeness was roughly the same in the two elections and consequently it makes sense to pool the two years.
Figure 1: The distribution of margins

Figure 2: Margin and change in turnout
5. Estimation strategy

In this section I set up the framework of the empirical analysis. First, I build a simple statistical model of electoral participation á la Riker and Ordershook (1968). Then, I discuss the specification of the empirical model and estimation issues.

5.1. A statistical model of runoff elections

In this section I lay out a simple statistical model of voter turnout of which Downsian predictions can be derived. My goal is not to come up with a new model of electoral participation, instead to formalize the decision-theoretical approach in a way that can take the nature of uncertainty of voting costs into account. Also, one should be careful in the interpretation of such a model: I do not claim here that the model presented here offers a full explanation of turnout. Nevertheless the structure presented here may be useful in explaining differential effects of political competition on turnout.

Let us consider a two round election in with each member of the electorate makes her decision about participation according to the Riker and Ordershook (1968) framework. That is, person \( i \) votes if and only if \( P^e B_i + D_i > C_i \), where \( D_i \) and \( C_i \) denote benefits and costs of voting irrespective of the outcome, \( B_i \) the differential utility associated with the favored outcome of voter \( i \) and \( P^e \) is the expected probability of being pivotal in the constituency in question. To simplify notation let \( c_i = C_i - D_i \) denote the net costs of voting. Also, without loss of generality let us normalize \( B_i = 1 \). Then, we can rewrite the condition of turnout as \( P^e > c_i \). That is, individual \( i \) will vote if and only if the expected utility from influencing the election is greater than the net cost of voting.

I first focus on turnout in the second round. In this case we can assume that individuals form their expectation of the closeness of the second round on the basis of the first round results, that is \( P^e \) is the closeness of the first round. Also, I assume that the net cost of voting that individuals face is distributed normally with mean \( \bar{c} \) and variance \( \sigma^2 \).

Now, observe that the probability that a randomly chosen voter shows up in a in given constituency in the second round is

\[
P(c_i < P^e) = \Phi \left( \frac{c_i - \bar{c}}{\sigma} < \frac{P^e - \bar{c}}{\sigma} \right) = \Phi \left( \frac{P^e - \bar{c}}{\sigma} \right).
\]
If we take a frequentist interpretation of probability we can interpret this probability as the actual ratio of people going to the ballots. But then we already have a formula for turnout in the second round given by Equation 3.

As for the first round, I assume that the expected closeness of the election in the constituency level does not vary across districts. This assumption sounds quite unrealistic but there are two rather strong arguments in favor of it. First, it can be argued that first round closeness is actually very hard to predict. In Hungary there exist no pre-election polls in the constituency level and the presence of small parties is likely to cause strategic voting, which in turn makes results even less predictable. Second, it can be argued that as a result of compensation lists, the motive to “influence the results” may not be applicable in the classical sense. In the first round, even the votes cast for the losing parties “count” to the national level race, so the standard formulation of the expected benefit of voting is not likely to influence turnout at this point. In contrast, in the second round, a vote only counts if it actually decides the race for the single-seat district.

This means that we can conclude that in a given election, differences between turnout levels across constituencies are only due to differences in the net cost that voters face in each of the districts. Using this insight, we can proxy the costs of voting with the turnout in the first round.

5.2. Model specification and estimation strategy

There are two ways of spelling the empirical model. First, following the existing literature, one can ignore the nonlinearity implied by the statistical model and specify some linear equation relating closeness and turnout including some control variables. This is the approach suggested by Barzel and Silberberg (1973). In a panel data setup this implies the following equation:

\[
\text{Turnout}_{\text{second},i,t} = \alpha_i + \mu_{t,2} + \beta \text{Closeness}_{\text{first},i,t} + \delta \text{Turnout}_{\text{first},i,t} + \lambda X_{i,t} + \varepsilon_{i,t}
\]  

(4)

Index \(i\) denotes constituencies and \(t = 1,2\) denotes which year the election took place. The term \(\mu_{i,t}\) is a dummy variable that is meant to capture different intercepts (i.e. average change in turnout) for the two election years. On the other hand \(\alpha_i\) are constituency effects– whether they should be treated as fixed or random is an open question at this point and is to be decided using specification tests.
Estimation in such a setup is straightforward. One can either estimate the coefficients of the model using random effects GLS (assuming the strict exogeneity of individual effects), or use the within estimator to get rid of unobserved heterogeneity upfront. The viability of the exogeneity assumption can be tested by a Hausman specification test (see Woolridge, 2002).

An alternative specification is also considered to account for the nonlinearity of the model of turnout. In such a setup, the dependent variable is turnout in the 2nd round and I explicitly model it as a random variable constrained between 0 and 1. Therefore, the conditional expectation of turnout is formulated as a Probit function and the regression equation is written up in terms of a Generalized Linear Model (GLM) (see for example Cameron and Trivedi, 2005).

\[ \text{Turnout}_{\text{second}}^{\text{second}} = \Phi\left( \alpha_i + \mu_i + \delta \text{Turnout}_{\text{first}}^{\text{first}} + \beta \text{Closeness}_{\text{first}}^{\text{first}} + \gamma X_{i,t} \right) + \varepsilon_{i,t} \]  

(5)

This model was introduced to the applied econometrics literature by Papke and Woolridge (1996, 2008) under the name of fractional probit. To my knowledge this paper is the first to use this framework in the analysis of voter turnout, acknowledging that turnout is actually a limited dependent variable.6

In a panel setting, fractional probit model can be estimated by pooled Bernoulli Quasi Maximum Likelihood (QMLE), or the Generalized Estimating Equation \( \text{GEE} \) (Papke and Woolridge 2008). These models differ in the way they specify the nature of unobserved heterogeneity. Notably, the QMLE procedure makes use of the assumption that the unobserved heterogeneity is exogenous, while the GEE framework allows for correlated random effects (individual heterogeneity that might be correlated with the time-means of variables).

In both setups, the effect of closeness can be estimated and it will correspond to the P-term in the empirical model if two additional assumptions are satisfied. First, we have to rule out indirect effects of closeness through mobilization ad campaigning. To be sure, there is a theoretical possibility that parties try to mobilize voters in constituencies where the race is expected to be close. However, similarly to the case of France, in Hungary the time between the two rounds is only one week so I agree with Indridason (2008) who points out that in such a setting the importance of mobilization should be at least limited. Also, to my knowledge

6 To my knowledge the only paper that applies fractional probit in the political science literature has been Gardeazabal, J. (2010).
7 GEE is a generalized population averaged panel model in which the conditional mean is modeled in some non-linear fashion. See details in Cameron and Trivedi (2005)
there is no existing constituency level data which could be used to proxy campaign effects so I would not be able to separate mobilization effects from the total effect of expected closeness anyway.

Second, the assumption that voting costs are invariant across the two rounds should also be satisfied. This assumption is less straightforward, as in principle, the two rounds of the same election could be fundamentally different. For example, a large margin of victory in the national level is likely to deter turnout as it reduces the stake of the second round. Nevertheless, we can still maintain the assumption that the relative costs of voting across constituencies will stay the same as the time between the two rounds is so short. Then, introducing a year fixed effect will control for changes in the costs of voting that affect each constituencies in the same way.

I also include two additional variables to control for factors that could affect changes in turnout between rounds. First, I include the vote share of parties other than MSZP, FIDESZ as I suspect that perceived closeness of the race would affect the behavior of their supporters differently. Typically, small parties were not present in the runoffs (at least with a significant chance to win) so one could expect that a large part of the supporters of small parties are likely to stay at home in the second round. Second, I include the size of the electorate, that is the number of eligible voters. A majority of studies report that turnout is generally lower in larger constituencies (Geys, 2006). This could be interpreted as evidence supporting the Riker-Ordeshook model, as the probability of being pivotal is greater when the number of eligible voters is small.

As a test of robustness I estimated some of the regressions with an alternative proxy for closeness. In that specification I took the difference between “party blocks” and used it to capture the margin of victory in a more robust way. With this approach I allow for the possibility that people who voted for a small party in the first round showed up in the runoff and gave their votes for the larger ally of their preferred party. However, the estimates do not differ from those in the basic specification so I do not report them here.
6. Estimation results

I present the first set of results in Table 3. The two columns show the estimated coefficients of equation 4 by random effects GLS (RE) and fixed effect between estimator. Apparently, the point estimates are rather similar in the two cases: using the Hausman specification test I cannot reject the null that there is no systematic difference between the estimated coefficients. Thus, we can maintain the exogeneity of the individual effects and thus claim that the random effect estimator is consistent and efficient.

<table>
<thead>
<tr>
<th>Dependent variable: Runoff turnout</th>
<th>RE</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnout 1st round</td>
<td>0.797***</td>
<td>0.820***</td>
</tr>
<tr>
<td></td>
<td>[0.029]</td>
<td>[0.105]</td>
</tr>
<tr>
<td>Eligible voter (log)</td>
<td>-0.016**</td>
<td>-0.139**</td>
</tr>
<tr>
<td></td>
<td>[0.008]</td>
<td>[0.058]</td>
</tr>
<tr>
<td>Margin of victory</td>
<td>-0.204***</td>
<td>-0.232***</td>
</tr>
<tr>
<td></td>
<td>[0.040]</td>
<td>[0.056]</td>
</tr>
<tr>
<td>Votes for small parties</td>
<td>-0.045***</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>[0.017]</td>
<td>[0.022]</td>
</tr>
<tr>
<td>2006</td>
<td>-0.081***</td>
<td>-0.081***</td>
</tr>
<tr>
<td></td>
<td>[0.002]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>Constant</td>
<td>0.371***</td>
<td>1.678***</td>
</tr>
<tr>
<td></td>
<td>[0.082]</td>
<td>[0.615]</td>
</tr>
</tbody>
</table>

Observations: 241 241
R-squared: 0.967 0.969
Number of constituencies: 138 138

Robust standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

The results confirm the Downsian Closeness Hypothesis: closer results in the first rounds are associated with higher turnout in the runoff. In particular, a 10% point decrease in the relative margin of the victory of the party who got the most of the votes in the first round is expected to increase turnout in the runoff by 2%. This estimate is substantially larger than that reported by Indridason (2008).

In line with theory (but contrary to most of the earlier studies) the effect of the size of the electorate is found to be negative and significant as well. That is, holding other factors fixed, in smaller constituencies turnout was higher. This result is particularly convincing, as in
contrast with the margin of victory, the variation of the size of the electorate between the two elections is definitely exogenous.

Also, turnout is found to be lower in the runoff in districts where more people voted for small parties in the first round. A reason for this is that the benefit of voting was lower for people who knew that their preferred party was out of the race.

I report the fractional probit estimates in Table 4. Apparently, the point estimates of the marginal effects are quite similar to the coefficients of the linear models. Each of the coefficients are significant at the 5% level and the size of the estimates is roughly the same as the ones in Table 3. This suggests that the linear specification is a good approximation around the mean of the explanatory variables. Summarizing the results of Table 3 and 4 it should be pointed out that the relation between closeness and turnout is remarkably stable, both with regards to functional form and to the specification of unobserved heterogeneity.

<table>
<thead>
<tr>
<th>Table 4: The effect of closeness on turnout (fractional probit models)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: Runoff turnout GLM GEE</td>
</tr>
<tr>
<td>Turnout 1st round</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Eligible voter (log)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Margin of victory</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Votes for small parties</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2006</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>Number of constituencies</td>
</tr>
</tbody>
</table>

Reported coefficients are marginal effects at the mean of each LHS variables

Robust standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1

However, based on the statistical model of turnout, I suspect that the partial effect of closeness should vary across constituencies with different baseline characteristics (i.e. constituencies which behaved differently in the first round). A way to investigate these differences is to look at the marginal effect of closeness at some points of interest and check whether they in fact differ from each other. In particular I calculated the partial effect of the margin of victory at twelve different points, for the combination of three “interesting values”

---

8 An alternative explanation is that variation in the number of eligible voters is due to people who changed their place of residence. Then, in constituencies with large inflows turnout could decrease because new arrivers are less likely to vote.
of first round closeness and for values of the margin of victory. The values of first round turnout used are 60%, 70% and 80%; and the values of margin are 1%, 0.1% and 0.01%. The other RHS variables are set to their mean, and the election year is set to 2006.

Table 5 reports the marginal effect of closeness at various combinations of closeness and turnout. Apparently, closeness has a roughly linear effect, that is the marginal effect is the same across various level of the margin of victory. On the other hand, one can observe that the marginal effects decrease in first round turnout. The intuition behind this result is that in district in which participation was large even in the fist round, the perceived closeness of the election played a minor part as compared to constituencies where few voters showed up in the first round of the election.

Table 5: The marginal effect of closeness at "typical points"

<table>
<thead>
<tr>
<th>Turnout first round</th>
<th>Margin of victory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01%</td>
</tr>
<tr>
<td>60%</td>
<td>-0.236</td>
</tr>
<tr>
<td></td>
<td>[0.042]</td>
</tr>
<tr>
<td>70%</td>
<td>-0.218</td>
</tr>
<tr>
<td></td>
<td>[0.038]</td>
</tr>
<tr>
<td>80%</td>
<td>-0.191</td>
</tr>
<tr>
<td></td>
<td>[0.033]</td>
</tr>
</tbody>
</table>

Top entries are marginal effects of margin from the GEE model
Robust standard errors in brackets

Table 5 reports the marginal effect of closeness at various combinations of closeness and turnout. Apparently, closeness has a roughly linear effect, that is the marginal effect is the same across various level of the margin of victory. On the other hand, one can observe that the marginal effects decrease in first round turnout. The intuition behind this result is that in district in which participation was large even in the fist round, the perceived closeness of the election played a minor part as compared to constituencies where few voters showed up in the first round of the election. Of course, the differences between the estimates in Table 5 are modest so it would be a mistake to consider them as firm evidence for the decreasing impact of closeness.
7. Conclusion

In this study I presented an empirical analysis of the effect of expected closeness on turnout. In a summary of the existing theoretical approaches to electoral participation I pointed out why testing this relationship is of crucial importance to rational choice theories of voting behavior. Also, in a critical survey of existing empirical evidence, I showed the main points of controversies regarding aggregate level studies of turnout and introduced an empirical framework in which much of the flaws in previous research could be mitigated.

The paper contributes to the empirical literature on electoral participation in three aspects. First, to my knowledge this is the first paper that models district level heterogeneity explicitly, making use of panel data. As a consequence of that, the effect of expected closeness can be identified from within variation in constituencies in the case of linear models, solving a problem pointed out by Grofman (1993). Second, for the first time in the literature in turnout, I explicitly model turnout as a limited dependent variable, and use an estimation method (fractional probit) that is well suited for such a problem. Third, in a substantive sense this paper provides empirical evidence about the Downsian Closeness Hypothesis in an \textit{ex ante} framework. To my knowledge, apart of French elections analyzed by Fauvelle-Aymar and Francois (2006) and Indridason (2008) Hungarian election are the only to be analyzed in such a way.

My hypotheses were that expected closeness of runoff elections in single seat constituencies proxied by the closeness of the first round race increases turnout. Making use of similarities of the Hungarian General Elections of 2002 and 2006 I used panel data to test hypotheses. The results presented in the study are consistent with the Downsian Closeness Hypothesis. In particular, constituencies experiencing a close contest in the first round saw greater participation in the runoff. Quantitatively, holding everything constant, a 5% percent increase in the margin of victory in the first round is expected to decrease turnout in the runoff by 1%. This estimate is large in the context of previous results: Indridason (2008) estimated a 0.5% decrease in turnout as an effect of such a change. The study also contains some evidence that the effect of expected closeness might be contingent on the costs of voting. Examining marginal effect estimated in the non-linear model suggests that the effect of closeness was smaller in those districts where turnout was already high in the first round.
A problem not addressed in this paper is that any aggregate level study of turnout is potentially subject to the problem of ecological fallacy (Matsusaka and Palda, 1993). It should be emphasized that providing bulletproof evidence on the determinants of voting behavior would require a solid micro-foundation both theoretically and in terms of the empirical model used. Consequently, making use of runoff elections to rule out the endogeneity of closeness as an explanatory variable is just a first step towards a fully convincing model of electoral participation. The final step would be to assess the effect of closeness in a two rounds setup in a way in which the micro-foundation of the perception of the expected benefits of voting are modeled explicitly. However, the robustness of the results presented here might convince students of voting behavior that the quest for the understanding the relationship between closeness and turnout is worth to be continued.
References


