Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essays in Political Economy

Autor:

Christos Mavridis

Director:

Ignacio Ortuño-Ortín

DEPARTAMENTO DE ECONOMÍA

Getafe, mayo 2016

Universidad Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essays in Political Economy

Autor: *Christos Mavridis*

Director: Ignacio Ortuño-Ortín

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Secretario:

Calificación:

Getafe, de de

Essays in Political Economy

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Economics
at Universidad Carlos III de Madrid

Christos Mavridis

Supervised by Ignacio Ortuño-Ortín

Spring 2016

To Laura, my Miss Private

Acknowledgments

This thesis is the result of more than five years of work. The moment these lines are written I am still not sure where the future will take me, being in a similar position like the one I was five years ago. But the similarities between my two selves more or less end here. During these years I grew academically and personally. I learned lots of new tools and theories, but more importantly I learned how to think: how to approach and solve problems, and how to not be satisfied with just any answer. I learned that whatever claim you make you have to be able to back it up, you have to be able to provide proof for it. This helped me come to terms with my own misconceptions and biases not only in terms of research but also in terms of attitude towards life. This academic and personal growth did not come easy; there were many days and nights of disappointment and for every up there was a bunch of downs. But in the end it was the ups that always brought me one step forward towards this final goal. During these years I met extraordinary colleagues and brilliant professors, and I consider myself lucky to have learned from all of them, and even to have befriended some of them. It was also during these years that I met and married my wife.

There are so many people that I wish to thank for, one way or another, being part of this journey of mine. First in that list is my advisor Ignacio Ortuño, to whom I am grateful for his guidance and support throughout the last four years. He was always more than happy to talk –often for hours– about research ideas or about problems that I had encountered. With his vast knowledge and understanding of the literature he was always able to suggest improvements in my work and to point me to the right research directions.

Second, I would like thank Antoine Loeper for giving me sometimes critical, but always extremely useful comments that improved my thesis significantly. Special thanks must also go to Luis Corchón and Emmanuel Petrakis as well as to Christos Agiakloglou from my undergrad years for inspiring me to go to graduate school and for all the help he provided in order for me to do so.

Agustin Casas deserves thanks for being a cool co-author but also for all the advice he never stopped throwing my way. I would also like to express my great appreciation to Francisco Marhuenda, Phillip Denter and Matilde Machado not only for being part of my internal defense thesis committee, but also for all their comments as well as for all the time they provided to help me prepare for the job market. For these same reasons I would also like to thank Marco Celentani, Jaime Millán, Angel Hernando, Jahn Stuhler, Irma Clots, Diego Moreno, Luigi Minale, Nicolas Motz and Carmelo Núñez.

I am thankful to Arancha Alonso, our Graduate Administrator, for always happily accommodating all my requests (and they were many!).

My friends and colleagues in Madrid both made my life there much more fun but also through numerous discussions helped me develop as a researcher. I would especially like to thank Mehdi for being there from the very start, Nikolas for his always optimistic attitude, Marco for teaching me how to dive head first into research questions, Pedro for teaching me how to grill meat (and whose work ethic has been an example to follow), Andrés for always be willing to get out of his way to provide help, and Matthias for designing mechanisms. The list of colleagues and friends that I would like to send a shout out to cannot be complete without Bernhard, Sebastian, Salvatore, Federico, F(r)ederico, Meng, José, Hasin, Emre, Tomás, Ursula and David. Thank you guys, hasta la próxima.

I wish to send my thanks and love to my mother who has always been my biggest fan and whose sacrifices throughout the years played a major role in allowing me to be here before one of the greatest milestones of my life.

Finally, for more reasons I would ever be able to express with words, I want to thank Laura, my partner, my pillar, my wife. I love you, you make my life perferct.

Abstract

This thesis is comprised of three chapters. In the first chapter, I examine a voting model where two political parties have fixed positions on a unidimensional policy space but where the implemented policy is the convex combination of the two positions and study the effects of opinion polls on election results and social welfare. Voters are completely agnostic about the distribution of preferences and gain sequential and partial information through series of opinion polls. Voters' behavior is driven in part by regret minimization. The mass of undecided voters decreases monotonically with the number of polls, but may not necessarily disappear. Voters who remain undecided have centrist ideologies. Finally, social welfare is not necessarily increasing in the number of polls: having more polls is not always better. Features of the model are confirmed by empirical evidence.

In the second chapter, which is a joint work with Agustin Casas and Guillermo Diaz, we evaluate the effect of an institutional provision designed to increase accountability of local officials, and we show that its implementation can lead to a distribution of power within the legislature which is not consistent with voters' true preferences. The cause of this inconsistency is the ballot design which asymmetrically affects the officials listed on it. We analyze the case of the Lima's 2013 city legislature recall referendum and show that the design of the referendum ballot had adverse and significant effects on the composition of the Lima's city legislature. We also show that the election results with more "neutral" ballot designs would have been significantly different, and the composition of the legislature would have been more representative of voters' true preferences. More specifically, we use our results to simulate the outcome of the election with a random order of candidates. Even though the voters' fatigue is still present, it affects all parties equally, obtaining a more faithful representation of the voters' preferences.

Finally, the third chapter is a joint work with Marco Serena. For small electorates, the probability of casting the pivotal vote drives one's willingness to vote, however the existence of costs of voting incentivizes ones abstention.

In two-alternative pivotal-voter models, this trade-off has been extensively studied under private information on the cost of voting. We complement the literature by providing an analysis under complete information, extending the analysis of Palfrey and Rosenthal [1983. A strategic calculus of voting. Public Choice. 41, 7-53]. If the cost of voting is sufficiently high at least for supporters of one of the two alternatives, the equilibrium is unique, and fully characterized. If instead the cost of voting is sufficiently low for everyone, we characterize three classes of equilibria and we find that all equilibria must belong to one of these three classes, regardless of the number of individuals. Furthermore we focus on equilibria which are continuous in the cost of voting. We show that this equilibrium refinement pins down a unique equilibrium. We conclude by discussing an application of our findings to redistribution of wealth.

Contents

1 Polling in a Proportional Representation System **8**
- 1.1 Related Literature . . . 11
- 1.2 Empirical Evidence . . . 14
- 1.3 Model Description . . . 17
  - 1.3.1 Regret Minimization . . . 18
  - 1.3.2 Complete information equilibrium . . . 19
  - 1.3.3 Polls . . . 21
- 1.4 Implemented Policy and Citizen Welfare . . . 28
- 1.5 Conclusions . . . 32
- 1.6 Appendix A: Position Uncertainty . . . 32
- 1.7 Appendix B: Further Empirical Analysis . . . 35

2 The Last Shall Be the First: Failed Accountability Due to Voters Fatigue and Ballot Design **53**
- 2.1 Literature . . . 55
- 2.2 Institutional Background and Data . . . 56
  - 2.2.1 Institutional Background . . . 56
  - 2.2.2 Data . . . 58
- 2.3 Empirical strategy . . . 59
  - 2.3.1 Estimation strategy . . . 59
  - 2.3.2 Identification strategy . . . 60
- 2.4 Results . . . 62
  - 2.4.1 Determinants of voting blank . . . 62
  - 2.4.2 Analysis of the probability of being removed from office 64
  - 2.4.3 Counterfactual Analysis and Discussion . . . 66
- 2.5 Conclusion . . . 67
- 2.6 Appendix . . . 68
  - 2.6.1 Institutional background . . . 68
  - 2.6.2 Multinomial Probit . . . 69

2.7 Tables and Figures 72
2.7.1 Figures 72
2.7.2 Tables 76

3 Costly voting under complete information 88
3.1 Model 91
3.2 Computing the equilibria 92
3.2.1 “Pure-Pure” and ”Pure-Mix” equilibria 92
3.2.2 “Mix-Mix” equilibria 96
3.3 Continuous Refinement and Uniqueness 100
3.4 Application - voting over redistribution of wealth 104
3.5 Appendix A 108
3.6 Appendix B 113
3.7 Appendix C 116

Chapter 1

Polling in a Proportional Representation System

The effects that opinion polls have on the electorate have been studied extensively. The conclusion of the literature is that the publication of polls influences how the electorate votes (Myerson and Weber, 1993; Morton et al, 2015). In many countries this is reflected on past or present restrictions on how, when and if a poll should be published. The restrictions can range from a ban on publishing polls for a specific time before the elections, like in Italy (fifteen days[1]) or in Greece (fifteen days until 2014, one day since[2]), to an outright ban during the entire campaign period like in Singapore.[3] In the last decade various organizations and watchdogs have been publishing reports on these restrictions (Article19 (2003) and Spangenberg (2003)) and at the same time in many countries there have been public discussions on whether restrictions should be posed or lifted. In India in 2004 the main parties argued that announcing exit poll results while the elections are still going should be banned as they believed this favored certain parties[4] and in 2012 India's Chief Election Commissioner stated that opinion polls and exit polls should both be banned.[5] A 2013 survey in the United Kingdom,

[1]Legge 22 Febbraio 2000, n. 28, "Disposizioni per la parità di accesso ai mezzi di informazione durante le campagne elettorali e referendarie e per la comunicazione politica", article 8.

[2]Law 4315/2014, article 32.

[3]Parliamentary Elections Act, Chapter 218, article 78C.

[4]Concern over India opinion polls (April 6, 2004). BBC News. Retrieved from: http://news.bbc.co.uk/2/hi/south_asia/3603741.stm

[5]Opinion, exit polls have no scientific basis: CEC (April 28, 2012). IBNLive. Retrieved from: http://ibnlive.in.com/news/

where there is not a ban on publishing polling results before elections in place, showed that three out of ten MPs supported the idea of such a ban.[6]

These restrictions and discussions show that the issue of the interaction between opinion polls and elections is important by both a theoretical and practical point of view. The motive behind such bans is usually to let the electorate vote "trully" without the influence of the results of opinion polls, an influence which is perceived as bad. The fear that polls may be biased, in which case the electorate might base its decision on false information, is one argument used to argue in favor of restrictions. Another argument is that basing the decision itself on the information generated by polling might distort the election results away from the "true" preferences of the electorate.[7] Opinion polls are useful inasmuch as they clear some uncertainty in the political environment leading to more efficient decision making from the side of the political actors. Examples of uncertainty include uncertainty about valence of the candidates as e.g. in Kendall et al (2015) or on the actual policy-ideology of candidates as in Baron (1994).

In modern parliamentary democracies where parties compete for parliament seats, these seats can be seen as a measure of political power; a ruling party that controls enough seats to just secure a majority will have to be much more moderate in its policies compared to a party that controls a larger fraction of the parliament. For this reason a moderate voter may choose to not vote for his favorite party if he feels that it may be too strong. Therefore a moderate voter who is happier with a more equal split of parliamentary seats would like to know how the rest of the voters will vote before he decides his own vote, implying that he would need to have some information about the distribution of the rest of the voters. This idea is analyzed by Ceci and Kain (1982) who using an experiment show that moderate voters, when instructed that the latest poll showed that a candidate "commanded a substantial lead over the other", tend to report that they would vote for the trailing one.

An interesting real life example that shows that parties understand and can try to exploit the motives of moderate voters comes from the Greek

---

opinion-exit-polls-have-no-scientific-basis-cec/252787-3.html

[6]Eaton, David. Should pre-election opinion polls be banned? A third of MPs think so (November 13, 2013). NewStatesman. Retrieved from: http://www.newstatesman.com/politics/2013/11/should-pre-election-opinion-polls-be-banned-third-mps-think-so

[7]Morton et al (2015) identification strategy exploits a 2005 French voting reform that came into place precisely to a avoid the situations where exit polls were public before some voters went to vote.

general elections of January 2015. The campaigning period of these elections was very short, and during this period it was made clear that the anti-austerity and left-wing Syriza party would win the elections, since it was systematically ahead in the polls. The real question was whether Syriza would be able to form a government on its own, meaning whether it would be able to secure at least 151 out of the 300 seats in the Greek parliament. From the parties expected to enter the parliament, the anti-austerity small right-wing party Anel was the one that Syriza would be most likely to cooperate with, on an anti-austerity basis. Anel themselves had made it clear that they would be willing to cooperate long before the elections were actually called. The previous government was a coalition of two pro-austerity parties the largest of which, the right-wing ND, was predicted to be in second place in the elections. Anel political campaigning focused on the fact that Syriza was going to win and the role Anel would play as a minority right-wing counterbalance in a government that was almost certain going to be dominated by the left. As such they targeted right-wing voters, who would otherwise vote for ND, and moderate ones who could see Syriza as being too radical, promising all their potential voters that they would control the government from within.

An important assumption of the paper is that voters behavior is driven by regret minimization. Regret can be defined as a painful emotional experience of feeling sorry of misfortunes or mistakes.[8] It is obviously a common human emotion, in fact, a study found it to be the second most frequently named emotion (Shimanoff, 1984). Regret minimization can be expressed concretely using the *minimax regret* decision rule (Savage, 1951) according to which agents try to minimize the maximum regret their actions can incur. This rule is particularly useful when an agent that has to make a decision is aware of the set of potential future states of the world, but cannot assign any probability to them.

The goal of this is paper to examine the effects in implemented policy and welfare that public opinion polling has in a proportional representation system, and under what conditions polls lead to complete information. To do so, I construct a simple voting model in which a continuum of voters

[8]In her 1993 book "Regret: Persistence of the possible", Janet Landman defines regret as follows: "*Regret is a more or less painful cognitive and emotional state of feeling sorry for misfortunes, limitations, losses, transgressions, shortcomings, or mistakes. It is an experience of felt-reason or reasoned- emotion. The regretted matters [...] may range from the voluntary to the uncontrollable and accidental; they may be actually executed deeds or entirely mental ones committed by oneself or by another person or group...*" (as cited in Li and Majumdar (2010)).

has to choose between two parties with fixed positions in a proportional representation system. Voters have imperfect information about each other's preferences and experience regret if they choose the wrong party. The voters care about the shares of the two parties because the implemented policy is a linear combination of the two parties' fixed policies, using their shares as weights. A voter's information about other voters' preferences can be improved through a series of polls. I find the following. In the cases where some voters have some sort of incomplete information the outcome may be quite different compared to the outcome under complete information. The voters most likely to report that they do not know what they will vote are the more moderate ones.

The main result of the paper is that under some condition on the distribution function of the citizens' preferences, a sequence of polls will be able to reveal the complete information case. With every poll the undecided mass weakly decreases, as more citizens are sure of what to vote. However, there can be instances that after a number of polls the undecided mass stops decreasing, and polls cannot provide the society with any more information. Finally, the implemented policy and citizen welfare are both not monotonic on the number of polls: the publication of one last additional poll before the elections may result in an implemented policy that is further away from the optimal policy or the policy that would be result under complete information.

To the best of my knowledge this is the only paper that examines polls and proportional representation together in a dynamic setting. It can also explain the following empirical facts at the same time, namely, the correlation between undecidedness and centrist ideologies, the non-monotonicity of election results with respects to the number of opinion polls and the decrease of the mass of undecided voters when more polls are published.

For simplicity the model of this paper only deals with two parties. However, the idea still holds to more than two parties.

1.1 Related Literature

Various works have examined how opinion polls affect voters and parties through the information they reveal. Whenever there is some uncertainty, an objective flow of information may help minimize it. Denter and Sisak (2015) show that poll results affect campaigning spending which in turn affect how voters vote. Bernhardt et al (2009), Morgan and Stocken (2008) and Meirowitz (2005) show that polls are used to clear uncertainty over voters preferences, but at the same time this can give rise to strategic poll answering on behalf of the voters. McKelvey and Ordeshook (1985) show that polls create a “bandwagon” effect in favor of the leading party, while Goeree and Großer (2006) and Taylor and Yildirim (2010) show that the opposite “underdog” effect can appear, through the mobilization of the voters of the trailing party. In both of these papers it is argued that through polls the “wrong” side may win which would result in a welfare loss. Klor and Winter (2006) in an experiment show that in both close and lopsided elections polls have a bandwagon effect, but the welfare effect of the polls is ambiguous, and in a later paper (Klor and Winter, 2014) they use US Gubernatorial elections data to show that the increase in participation in elections is greater for the supporters of the leading candidate. Großer and Schram (2010) find similar results. Herrmann (2014) examines polls under proportional representation and shows that voters act strategically, trying to anticipate the future coalitions after the elections take place.

The implemented policy as seen as a compromise among different ideological sides, or power sharing, is not a new idea (Alesina and Rosenthal, 1996; Grossman and Helpman, 1996; Llavador, 2006; De Sinopoli and Iannantuoni, 2007; Saporiti, 2014; Matakos et al, 2015; Herrera et al, 2015). Ortuño-Ortín (1997) discusses a two-party setting where implemented policy is a convex combination of the two parties’ platforms generated using a continuous function of the parties’ shares in the elections. Sahuguet and Persico (2006) point out that in a proportional system parties need to maximize their share (rather the probability of getting at least 50 per cent of the vote) and that the implemented policy is, is at least partly influenced by the minority party.

Various researchers have examined the idea of agents experiencing regret from their actions and how regret affects their actions themselves. Savage (1951) analyzes the minimax regret rule according to which agents try to minimize the maximum regret their actions can incur. Similarly, Loomes and Sugden (1982) argue that a theory of choice has to incorporate the fact that agents do feel regret and that they take this into account in their decision making. Zeelenberg (1999) provides a short overview of the issue of rationality of regret and argues that acting on anticipated regret can be rational, further pointing out that the minimax regret rule is useful when there is no knowledge at all about the probabilities of the possible outcomes. Most of the research on the application of regret in voting has focused on the context of the paradox of voting. Ferejohn and Fiorina (1974) point out the difference between “risk” (voters are able to assign probabilities to different states) and “uncertainty” (voters cannot assign probabilities to different states) and argue that rational behavior for a voter is not equivalent to assuming the voter is expected utility maximizer. They furthermore show that the minimax regret rule can give voter participation prediction results that are more in line with empirical observations in the presence of a cost of voting. In a later work (Ferejohn and Fiorina, 1975) they discuss the reactions to their first paper and provide empirical evidence in favor of the minimax regret rule over the expected utility maximization. Eager to test the minimax regret hypothesis directly Kenney and Rice (1989) conducted a survey which revealed that over one third of the participants could be identified as minimax regret voters. In another survey of university students Blais et al (1995) found that a third of the participants "strongly agreed" and about forty percent "agreed" with a statement that identified minimax regret voting, and that respondents that "strongly agreed" were significantly more likely to have voted in the previous elections. However, this strong relationship disappeared once they controlled for other variables such as civic duty. In a later work Blais (2000) provided a more analytical critique of Ferejohn and Fiorina. Pieters and Zeelenberg (2005) conducted a survey of Dutch voters which confirmed that regret is an emotion that voters can feel after the elections. Recently there has been some renewed interest on regret in voting. Li and Majumdar (2010) study voter turnout and incorporate the concept of regret but not in the minimax way: how much regret a voter experiences is inversely related to the margin of victory of his favorite party. (Also see Degan and Li (2015)). Since most of the papers incorporating regret in voting are concerned with the paradox of (not) voting, each voter's decision boils down to voting for his clearly preferred side, or abstaining. In these models the action that can generate regret is abstaining which happens when their failure to go and vote leads to a bad result for their favorite side. In my model however, the action that generates regret is voting for the wrong candidate.

Herrera et al (2015) in a proportional representation setting similar to the one of this paper show that the "marginal voter's curse" has the effect that voters with low quality information abstain and not vote for fear of casting a vote for the wrong side. Finally, voters may choose to support the party that they are least ideologically close an idea which is shared by the concept of the "protest vote", by which voters may want to punish, or control the power of their favorite party (see eg. Myatt (2015)).

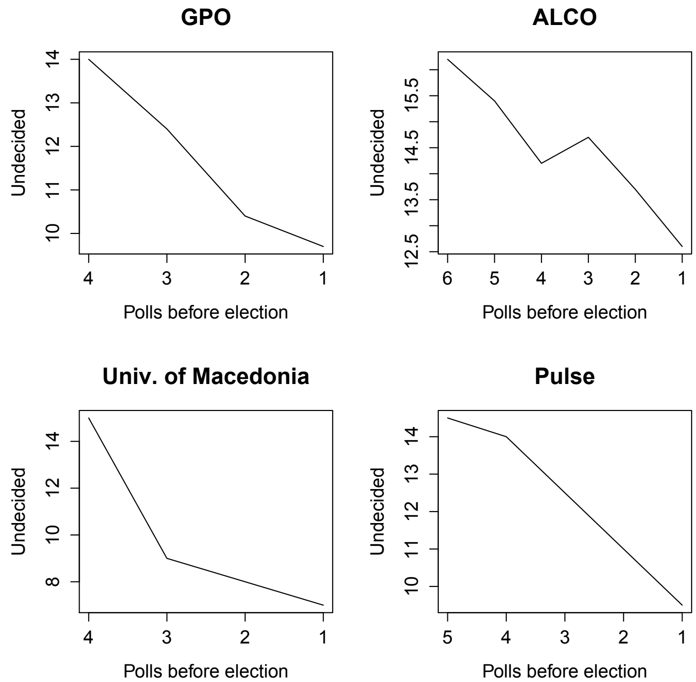

Figure 1.1: Polling for the 2015 Greek general elections

1.2 Empirical Evidence

I will use the Greek case to also briefly go over some empirical evidence regarding the influence of the number of polls on the undecided vote and winner's share. After the previous parliament failed to elect a new head of state constitutional provisions forced its dissolution. The announcement of the elections took place on December 29 with the date set on January 25, making the campaigning period a little more than three weeks long. There were plenty of polls conducted by a number of public opinion polling organizations and companies. Most organizations had time to conduct four to five polls before January 25. Figure 1.1 shows the percentages of undecided voters reported by four polling organizations:[9] GPO, ALCO, University of Macedonia Public Polling Unit and Pulse in chronological order. The x-axis in each of the graphs shows the number of polls conducted by each orga-

[9]The data was compiled from a number of news sources and the polling organizations own reports.

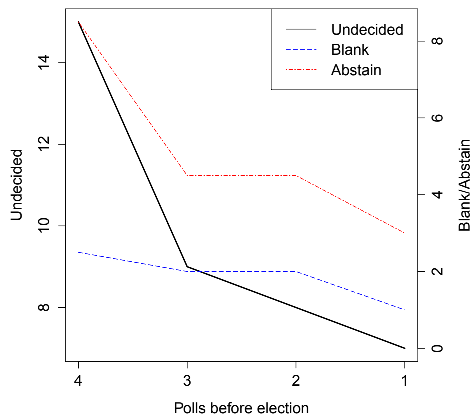

Figure 1.2: University of Macedonia polls

nization before the elections took place.[10] Only polls conducted after the elections were called are reported because at this moment the positions of the parties must be thought as given. Furthermore, under the assumption that respondents answered thoughtfully and given that the elections were imminent, the answers are taken to represent the respondents' true voting behavior.

From the results presented here, with the exception of the University of Macedonia results, "undecided" means any voter who when asked what he would vote he did not give a party as an answer, or put it differently he is classified as "undecided" if he answered any of the following: "I will vote blank", "I will abstain", or "I don't know/ I haven't decided yet." In reporting their results for the parties and "undecided" shares, the University of Macedonia first discarded those who indicated they would vote blank or that they would abstain, therefore their "undecided" share includes only those who specifically answered they had not decided yet. However they do report what was the share of original answers that were blanks and abstentions. Figure 1.2 shows the percentage of the undecided along with the percentage of blanks and abstentions. All graphs tell the same story: the percentage of the undecided voters, in any way we choose to define them, goes down with every single poll as the date of the elections comes closer.

In Figure 1.3 in the left panel I plot the Syriza expected results of the

[10]Meaning that 1 signifies the last poll before the elections, 2 the second to last poll before the elections and so on.

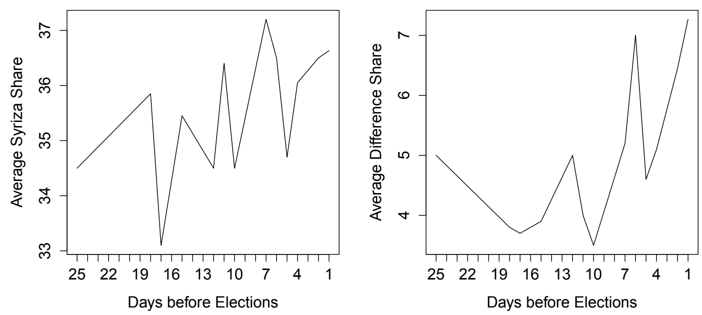

Figure 1.3: Syriza predicted share

four polling organizations together. When more than one organizations published results at the same day, I took the average of them. These results show, that abstracting from random noise, there is no monotonic effect of the number of polls on the predicted share of Syriza. This could imply the absence of a pure bandwagon or underdog effect in these elections. At the same time, in the right panel, I plot the difference between the same Syriza shares depicted on the left one, and the ND (the second party in the elections, leaving office) shares. The difference is positive, showing that Syriza was always the frontrunner, but it was not monotonic in the number of polls.

As mentioned in the introduction a finding of the model is that the citizens who are more likely to abstain are mostly the voters who are ideologically more moderate. In Appendix B, using the last wave of the World Value Survey[11] database I provide some evidence about that. The important finding is that there is a positive relationship between being ideologically in the center and being undecided. This positive relationship along with the findings that more polls imply less undecided voters, and the non-monotonic effect of polls on the share of the winning party and its difference from share of its contender are all features that the theoretical model that is presented next is able to accommodate.

[11]WORLD VALUES SURVEY Wave 6 2010-2014 OFFICIAL AGGREGATE v.20141107. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: Asep/JDS, Madrid, SPAIN

1.3 Model Description

There is a uni-dimensional policy instrument that can take values in $D = [-\eta, 1+\theta] \subset \Re$ with both η and θ positive. The voters have single-peaked preferences over this policy space and their optimal policies are distributed over D according to a distribution function $F(\cdot)$ which is assumed to be continuous, strictly increasing and differentiable with a corresponding density function $f(\cdot)$. The utility of the voter x_i from the implemented policy x is given by the function $-u(|x - x_i|)$, ie. it is a function of the absolute distance of the voter's optimal policy x_i from the implemented policy x. Function $-u(|x - x_i|)$ is decreasing and concave in $|x - x_i|$ and it has a unique maximum at $x = x_i$ such that $u(0) = 0$. These mean that the utility function of a voter is symmetric around his optimal policy point where the utility is maximum and normalized to zero. We will be identifying voters with their optimal policy, ie., voter x_i is the voter that has optimal policy x_i. There are two parties running in the elections, L and R each having a fixed position at 0 and 1 respectively. The actual implemented policy, x, will be the linear combination of the two policies using the parties' shares in the elections as weights:

$$x = p_L \times 0 + (1 - p_L) \times 1 = 1 - p_L,$$

where p_L is the share of votes party L gets. This share is defined as $p_L = \frac{m_L}{m_L+m_R}$, with m_L and m_R being the voter masses that voted for L and R respectively, meaning that the parties' shares are calculated over the mass of voters that actually voted. If u is the mass of the citizens that decide not to vote then $m_L + m_R + u = 1$. The setting is similar to Ortuño-Ortín (1999).

Notice that $p_L = \frac{m_L}{m_L+m_R}$ implies that the parties' shares of influence in the implemented policy are equal to their vote shares. Even though this assumption is made for simplicity and clarity, all the results follow if we instead assume the more flexible weight function

$$\hat{p}_L = \frac{(\frac{m_L}{m_L+m_R})^\gamma}{(\frac{m_L}{m_L+m_R})^\gamma + (\frac{m_R}{m_L+m_R})^\gamma}$$

with $\gamma \geq 1$[12] and finite (as in, for example, Saporiti (2014), Matakos et al (2015)). Obviously, for $\gamma = 1$, $p_L = \hat{p}_L$. As γ increases then the share of the bigger party is getting bigger, and that of the smaller party smaller.

[12] The $\gamma < 1$ case is unrealistic as it would give the smaller party a greater share of seats compared to its vote share.

Table 1.1: Decision of a voter at 0.4

L share	Implemented policy	Decision
0.8	0.2	R
0.3	0.7	L

This captures the fact that some parliamentary systems, while mostly proportional, give some bonus seats to the biggest party or demand a certain percentage threshold to be passed before they assign seats to smaller parties.

Since increasing the votes cast for a party increases its number of seats and its political power voters know that their votes will move the share of a party up or down, and as a result, the implemented policy left or right. If some voters expect the implemented policy to be, for example, too much to their left they will want to vote for R. We can think of the voters each deciding what vote to cast in order to bring the implemented policy as close to them as possible in a similar way as in Alesina and Rosenthal (1996). For this reason, each voter would like to know how the other voters are behaving before he chooses which party to vote for. For example, even though the voter might be ideologically closer to a party, if he expects too many people to vote this party, he might want to vote for the other one. Consider Table 1.1 where I analyze the decision of a voter located at 0.4. This voter is closer to 0 than to 1, and as such he identifies with L more than R. However, the decision of what to vote depends on what others will do.

1.3.1 Regret Minimization

The second feature of voters' behavior is regret minimization. In this setting regret minimization is modeled using the *minimax* regret decision rule. If the preference distribution is not known, some voters will need some additional information before they decide which party to vote for. This information will come in the form of polls. If however at the time of the election there are still some *undecided* voters not sure of what they should vote, they will vote according to the regret minimization criterion.

In this model, an undecided voter is a voter that cannot assign a probability to any of the following states of the world: i) the implemented policy being to his left, ii) the implemented policy being to his right and iii) the implemented policy being exactly on the voter's optimal position.

When the time of the election comes the undecided voter will use the minimax regret rule to decide whether he will vote for L or for R. As it will be made clear later, at any stage of the game, all voters know that the implemented policy can never be to the left of some minimum or to the right of some maximum implemented policy. Call them x_L and x_R respectively. Voters in (x_L, x_R) cannot predict where exactly the implemented policy will fall, because they do not know how the distribution of voters looks like. Then, when the elections are held, given (x_L, x_R), they will use the minimax regret rule.

Given the utility functions of the undecided voters, the greatest disutility for each voter comes from when the implemented policy is at one of the two corners of the interval (x_L, x_R). In particular, the greatest disutility for voters in $(x_L, \frac{x_L+x_R}{2})$ comes from when the implemented policy is near x_R, and for voters in $(\frac{x_L+x_R}{2}, x_R)$ the greatest disutility comes from when the implemented policy is near x_L.

For the undecided voters, the maximum regret of voting L, comes from when the implemented policy is too far to the left, near x_L. This is because, when the policy is too far to the left, all voters to the right of the policy should vote R. Similarly, the maximum regret of voting R comes from when the implemented policy is near x_R. Therefore, for voters in $(x_L, \frac{x_L+x_R}{2})$ the action that minimizes maximum regret is L and for the voters in $(\frac{x_L+x_R}{2}, x_R)$ it is R.[13]

1.3.2 Complete information equilibrium

If the distribution function of the citizens' optimal points $F(\cdot)$ were known to them then the regret minimization criterion would have no bite: citizens have all the available information they need in order to make the decision of what to vote. In this case the only driving force behind voting would be the optimal policies of the voters. In this complete information case the set of pure strategies for a voter $x_i \in D$ is $S = \{L, R\}$ with L and R representing him voting for the left and right party respectively. Since the implemented policy depends on the shares each party gets, voters strategies are not independent from what the other voters will do. If a voter thinks that the implemented policy will be too much on the left he may want to vote for the right, and if he thinks the implemented policy will be too much on the right he may want to vote for the left. A voter's strategy is a function $\sigma_{x_i} : S \to [0, 1]$ that assigns a probability of voting for L given the position of

[13] Let me note at this point that the results I will derive in the following are robust to using other decision criteria uncer uncertainty such as Wald's or Hurwicz's criterion.

the voter. I define as $x(\sigma)$ the implemented policy that results from strategy profile $\sigma = (\sigma_{x_i})_{x_i \in D}$.

We are now ready to define the complete information voting equilibrium.

Definition 1. *Strategy profile σ constitutes a Complete Information Voting Equilibrium if:*

- *if $x_i < x(\sigma)$ then $\sigma_{x_i} = 1$,*
- *if $x_i > x(\sigma)$ then $\sigma_{x_i} = 0$,*
- *and if $x_i = x(\sigma)$ then $\sigma_{x_i} = 0.5$.*[14][15]

The equilibrium defines a cut-off point $x(\sigma)$ that divides the electorate in two parts: one that vote for L that are to the left of it, and the one that vote for R. The theorem that follows establishes the existence and uniqueness of this cut-off.

Theorem 1. *There always exists a unique Complete Information Voting Equilibrium σ^* characterized by the cut-off point $x^* = 1 - F(x^*)$.*

Proof. By continuity of $F(\cdot)$ the point x^* exists, and by strict monotonicity this point is unique. Now, let x^* be the implemented policy. Then strategy profile σ^* implies that the voters who choose L are the ones to the left of x^* and the ones that choose R are those to the right of x^*, with masses $F(x^*)$ and $1 - F(x^*)$ respectively. In turn, these masses give rise to the the implemented policy $x^* = 1 \times F(x^*) + 0 \times (1 - F(x^*)) = 1 - F(x^*)$, which proves that an equilibrium always exists, and since x^* is unique there is only one equilibrium defined by $x^* = 1 - F(x^*)$.

The last step is to show that there is no other equilibrium. Suppose that there is an equilibrium implemented policy $x' \neq x^*$. The equilibrium strategy implies that everybody to the left of x' vote for L and the rest vote for R, giving rise to the implemented policy $x'' = 1 - F(x')$, but we have that $x'' \neq x'$, which leads to a contradiction.

□

Given their optimal policy points voters want to cast their vote in a way that will bring the implemented policy as close to their optimal point as possible, given the strategies of the other voters. There cannot be a

[14] In the the case of $x_i = x(\sigma)$, $\sigma_{x_i} = 0.5$ is taken without loss of generality. In fact, any $\sigma_{x_i} \in [0, 1]$ would work.

[15] This equilibrium can be easily shown to be the Strong Nash Equilibrium (Gerber and Ortuño-Ortín (1998) and Ortuño-Ortín (1999)).

situation in equilibrium where there exist some voters with optimal policy to the left of x^* who voted for R; if these voters would switch their vote, the implemented policy would be closer to their optimal one. Note that the equilibrium would be the same had we had expanded the strategy set to include the possibility of abstention. Another interesting thing to notice is that $x^* = 1 - F(x^*)$ does not necessarily correspond to the median voter x^M; the model does not in general converge to the median voter. The only case such that we have $x^* = x^M$ is when $x^M = 1/2$.[16]

1.3.3 Polls

When the distribution of voters is common knowledge, then the voters can calculate $x^* = 1 - F(x^*)$ and infer what they should vote. However, when the distribution is no longer known to the voters, we have incomplete information, and some voters do not have enough information to be certain which party they need to vote for.

A voter will be certain of what to vote if the interval where the future implemented policy will fall does not include his optimal policy. Therefore, since the implemented policy $x = 1 - p_L$ can never fall outside of $[0, 1]$, it is easy to see that the voters to the left of 0 would never want to vote for R. For all of them the state of the world for which $x_i < x$ happens with probability one: they know for sure that the implemented policy will fall to their right. By voting for L, p_L increases and x comes closer to 0. Voting R on the other hand, decreases p_L and x gets closer to 1. Obviously out of these two outcomes, it's the first one that the voters in $[-\eta, 0)$ prefer, so they will vote for L. Similarly, voters to the right of 1 would vote for R. I will call voters in $[-\eta, 0)$ as *left partizans* and voters in $(1, 1+\theta]$ *right partizans*. When referring to both left and right partizans I will simply use *partizans*. The rest of the voters, the ones in $[0, 1]$, are the initially undecided voters who would like some information from the polls to help them figure out what to do.

An initially undecided voter knows that the future implemented policy will be somewhere in [0,1] but he doesn't know where exactly. It can fall to the right of, to the left of or exactly on his optimal point, therefore it's not clear to him which party he should vote for. Consider again a voter with the optimal point at 0.4. If this voter votes for L but the elections result in an implemented policy of $x = 0.1$ then it would be clear to this voter that he should have chosen R, since voting for L actually moved the implemented

[16] Since in this case we would have $\frac{1}{2} = 1 - F(\frac{1}{2})$, so $F(\frac{1}{2}) = \frac{1}{2}$.

policy away from his optimal point.

The voters do not know how many polls there will before the election. The question asked in each poll is: "what will you vote in the elections?". A maintained assumption is that voters respond truthfully to polls which I believe is a mild assumption, as it is not clear what a voter would gain by lying. In this setting truthful responding means that the citizens are honest when they answer the question: if they are sure that the implemented policy will fall to their left they will respond R, if they are sure that the implemented policy will fall to their right they will respond L. Otherwise they will respond "I do not know." The only information the citizens need to answer in a poll or to vote in the elections is the information provided by by the previous poll.[17]

Each poll can provide the undecided voters with two pieces of information: the mass of the voters that will vote for party L, and the mass of the undecided voters at the time of the poll. Using this information they can calculate the two most extreme cases, the maximum possible implemented policy which is defined as x_R and the minimum possible one, defined as x_L. The way to find these two is to see what would be the implemented policy if all the currently undecided would vote for L and what would it be if all the undecided would vote for R.

Since the mass of the voters who have made up their mind consists of the partisan voters and the formerly undecided that are now absolutely sure of what to vote, this mass of voters forms a lower bound for the mass of total voters the party will get in the future. Therefore an undecided voter that reads the poll results and sees that the mass of voters that are going to vote for party L is m_L and the mass of the undecided is, say, u, can infer a few other things. First, that the mass of voters that are going to vote for party R is at least $1 - m_L - u$. Second, that the vote share that party L can get at this point is at most $\frac{m_L+u}{m_L+u+(1-m_L-u)} = m_L + u$ and similarly that the share of votes that R can get is at most $1 - m_L$. Third, using $m_L + u$ and $1 - m_L$ he can infer the potential implemented policies after each of these two most extreme results: $x_L = 1 - m_L - u$ and $x_R = 1 - m_L$.

A voter that is to the left of x_L would never vote for party R. He knows that the possible implemented policy cannot ever be closer to him than x_L. But since x_L is already to his right he would never vote for R because this way he would regret his decision with certainty. Therefore he votes for L. Similarly every voter to the right of x_R votes for party R. The voters that

[17] With the exception of the case where no poll took place. In this case, as it will be shown later, the relevant information is given by the fixed positions of the parties.

are located in $[x_L, x_R]$ remain undecided. I will use $[x_L^k, x_R^k]$ to indicate the interval of the undecided voters after k polls, with $x_L^0 = 0$ and $x_R^0 = 1$. Note also that

$$x_L^k \geq x_L^{k-1} \text{ and } x_R^k \leq x_R^{k-1}, \forall k \in \mathbb{N}. \tag{1.1}$$

This means that once uncertainty has been cleared for some undecided voters it never comes back. Furthermore, p_L^k is the percentage party L would get if there were elections right after k polls.

After k polls, voters in $[x_L^k, x_R^k]$ are called undecided at time k, voters in $[-\eta, x_L^k)$ are called L voters at time k, and voters in $(x_R^k, 1+\theta]$ are called R voters at time k. The voting equilibrium depends on how many polls took place before the election. Therefore, for each number of polls, we have a voting equilibrium which expresses how voters would have voted if there were elections right after this number of polls.

Definition 2. *After k polls strategy profile $\sigma^k = (\sigma_{x_i}^k)_{x_i \in D}$ is a Voting Equilibrium if:*

- *if $x < \frac{x_L^k(\sigma) + x_R^k(\sigma)}{2}$ then $\sigma_x^k = L$,*
- *otherwise $\sigma_x^k = R$.*

The two cut-off points $x_L^k(\sigma)$ and $x_R^k(\sigma)$ define the following regions of voters. For $x \leq x_L^k(\sigma)$ we have the L voters who know for sure that the implemented policy will be to their right; for $x \geq x_R^k(\sigma)$ we have the R voters who know for sure that the implemented policy will be to their left; and for $x \in (x_L^k(\sigma), x_R^k(\sigma))$, we have the voters who vote according to the minimax regret criterion.

Theorem 2. *Under incomplete information and after k polls, there always exists a unique equilibrium σ^{k*} characterized by the cut-off points $x_L^0 = 0$ and $x_R^0 = 1$ for $k = 0$, and $x_L^k = 1 - F(x_R^{k-1})$ and $x_R^k = 1 - F(x_L^{k-1})$ for $k > 0$.*

Proof. Consider the $k > 0$ case first. By continuity and strict monotonicity of $F(\cdot)$ we have that x_L^k and x_R^k exist and are unique. Given x_L^{k-1} and x_R^{k-1}, citizens calculate $x_L^k = 0 \times F(x_R^{k-1}) + 1 \times (1 - F(x_R^{k-1})) = 1 - F(x_R^{k-1})$, the lowest point the implemented policy can be in the future. Similarly they calculate x_R^k. Therefore, for $x \in [-\eta, x_L^k]$ we have $\sigma_x^k = L$, and for $x \in [x_R^k, 1+\theta]$ we have $\sigma_x^k = R$. Finally, $\forall x \in (x_L^k, \frac{x_R^k + x_R^k}{2})$, the minimax criterion implies that $\sigma_x^k = L$ and that $\forall x \in (\frac{x_R^k + x_R^k}{2}, x_R^k)$, the minimax criterion implies that $\sigma_x^k = R$. No voter has an incentive to deviate from this behavior. For $k = 0$ the argument is identical, with the difference that $x_L^0 = 0$ and $x_R^0 = 1$ since with no polls conducted, the citizens know that the implemented policy will have to fall into $[0, 1]$.

□

Note that if $x_L^k = x_R^k$ we are simply in the full information case. The preceding theorem defines a sequence of voting equilibria, one for every k. In the first poll, only the partizans declare their preference, the voters in $[0, 1]$ are undecided. When the first poll is published the masses of the partizans are revealed, and with them an updated pair of minimum and maximum possible implemented policies. In the second poll, the partizans declare their preference but now also some of the undecided have made up their mind. When the second poll is published the new voting masses are declared and so on. This mechanism means that the x_L^k and x_R^k are governed by the following system of equations:

$$x_L^k = 1 - F(x_R^{k-1}) \tag{1.2}$$

$$x_R^k = 1 - F(x_L^{k-1}),$$

If after a number of polls s with each subsequent poll the interval of the undecided does not change, then the process has reached a steady state where $x_L^s = x_L^{s+t} = x_L^*$ and $x_R^s = x_R^{s+t} = x_R^*$, $t \geq 1$. Plugging the expression for x_R^{k-1} into the expression for x_L^k we get:

$$x_L^k = 1 - F(1 - F(x_L^{k-2})). \tag{1.3}$$

This is a second order difference equation with initial values $x_L^0 = 0$ and $x_L^1 = 1 - F(x_R^0) = 1 - F(1)$.

Next, we will establish sufficient conditions for the sequence of polls to result to the complete information outcome.

Theorem 3. *If $F'(1 - F(x))F'(x) < 1$ $\forall x \in [0, 1]$ then a sequence of polls will lead to the complete information outcome.*

The theorem will be proved through the following three propositions.

Proposition 1. *The complete information outcome is a fixed point of function $h(x) = 1 - F(1 - F(x))$, $x \in [0, 1]$.*

Proof. This function is continuous as it is the composition of two continuous functions, it is strictly increasing and its range is a proper subset of $[0, 1]$ therefore it has a fixed point in that subset. Suppose that time s is the first time we observe $x_L^s = x_R^s$. Then the mass of undecided voters becomes 0, and we have: $x_L^{s+1} = \frac{F(x_L^s)+0}{F(x_L^s)+0+1-F(x_R^s)} \times 0 + (\frac{1-F(x_r^s)}{F(x_L^s)+0+1-F(x_R^s)}) \times 1 = 1 - F(x_R^s)$. Similarly $x_R^{s+1} = 1 - F(x_R^s)$, showing that $x_L^{s+1} = x_R^{s+1}$. We also know that $x_L^{s+1} \geq x_L^s$ and $x_R^{s+1} \leq x_R^s$. Combining the two last inequalities together we get that: $x_L^{s+1} \geq x_L^s = x_R^s \geq x_R^{s+1}$. But we have already shown that $x_L^{s+1} = x_R^{s+1}$, therefore we have to have: $x_L^{s+1} = x_L^s$ and $x_R^{s+1} = x_R^s$. $\square$

We have established existence of a fixed point of $h(\cdot)$, and we know that one of the fixed points will always be the *complete information* outcome: once it is reached all voters have made up their mind and are not going to change it. Therefore, if $h(\cdot)$ has a unique fixed point, it will be the complete information outcome. Uniqueness is not guaranteed however, and depends on the form of the distribution function $F(\cdot)$.

Proposition 2. *If $F'(1 - F(x))F'(x) < 1 \; \forall x \in [0, 1]$ then $h(\cdot)$ has a unique fixed point.*

Proof. Assume that there are two fixed points, x^* and x^{**} and that, without loss of generality, $x^* < x^{**}$, then by the Mean Value Theorem, there exists a $d \in (x^*, x^{**}) \subset (0, 1)$ such that

$$h'(d) = \frac{h(x^{**}) - h(x^*)}{x^{**} - x^*} = \frac{x^{**} - x^*}{x^{**} - x^*} = 1,$$

so

$$F'(1 - F(d))F'(d) = 1.$$

At the same time, since $h(0) > 0$ then $h(\cdot)$, crosses x^* from above then it has to be the case that it crosses x^{**} from below implying that there are points close to x^{**} for which $F'(1 - F(x))F'(x) > 1$, which concludes the contradiction. $\square$

The last proposition establishes when the sequence of polls will converge to a fixed point.

Proposition 3. *Let $\underline{x}^*$ be the smallest fixed point of $h(\cdot)$. If $F'(1-F(x))F'(x) < 1$ holds in $[0, \underline{x}^*]$ then it is a sufficient condition for x_L^k to converge to $\underline{x}^*$.*

Proof. Define the sequence $\{a_n\}_{n=0}^{\infty}$ with $a_n \equiv \underline{x}^* - x^n$. Now, let $x^0 \in [0, \underline{x}^*)$ and $x^1 \in (x^0, \underline{x}^*)$ and note that $\underline{x}^* - x^n = h(\underline{x}^*) - h(x^{n-2})$. Then, by the Mean Value Theorem, there exists $d^{n-2} \in (x^{n-2}, \underline{x}^*) \subset (0,1)$ such that:

$$h'(d^{n-2}) = \frac{h(\underline{x}^*) - h(x^{n-2})}{\underline{x}^* - x^{n-2}},$$

and therefore

$$h(\underline{x}^*) - h(x^{n-2}) = h'(d^{n-2})(\underline{x}^* - x^{n-2}).$$

or

$$\underline{x}^* - x^n = h'(d^{n-2})(\underline{x}^* - x^{n-2}),$$

so since $F'(1 - F(x))F'(x) < 1$ there exists a $K \in (0,1)$ such that:

$$\underline{x}^* - x^n \leq K(\underline{x}^* - x^{n-2}) \leq K(K(\underline{x}^* - x^{n-4}))$$

which by repeated substitution becomes:

$$\underline{x}^* - x^n \leq \begin{cases} K^{\frac{n}{2}}(\underline{x}^* - x^0), & n \ \ even \\ K^{\frac{n-1}{2}}(\underline{x}^* - x^1), & n \ \ odd \end{cases}$$

Splitting $\{a_n\}_{n=0}^{\infty}$ into two subsequences, one that has the odd-numbered elements and another that has the even-numbered ones, we see that each subsequence goes to zero, therefore since the two covering subsequences go to zero $\{a_n\}_{n=0}^{\infty}$ also goes to zero, which implies that

$$\lim_{n \to \infty} x^n = \underline{x}^*.$$

□

Therefore, if $F'(1 - F(x))F'(x) < 1$ holds in $[0,1]$ then the complete information outcome is a unique fixed point of $h(\cdot)$, and a sequence of possibly infinite polls will converge to it. In absence of fixed point uniqueness Proposition 3 also tells us that if there is a fixed point $\underline{x}^*$ such that $\underline{x}^* < x^*$ with x^* being the fixed point corresponding to the complete information outcome, and if $F'(1 - F(x))F'(x) < 1$ is satisfied in the interval $[0, \underline{x}^*]$, then we will have a mass of undecided voters located in $[\underline{x}^*, \bar{x}^*]$, with $\bar{x}^* = 1 - F(\underline{x}^*)$. In this case, the polls cannot lead to *complete information*, and there will always be some positive mass of undecided voters in $[\underline{x}^*, \bar{x}^*]$, no matter how many polls will be conducted afterwards.

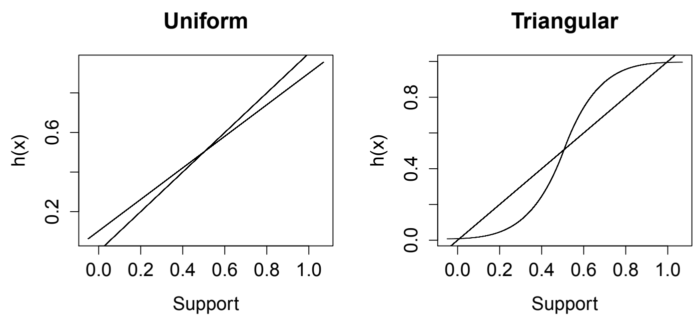

Figure 1.4: Examples of h(x)

Inequality $F'(1 - F(x))F'(x) < 1$ may seem a little strange; however it can be intuitively explained in the case of a symmetric distribution of the undecided at period 0. It provides a sort of upper bound of the the maximum of the probability density function, or putting it differently, an implicit lower limit on the variance of the distribution. When most of the mass of the undecided is located close to the mean of the distribution, their influence in the elections is potentially so large that there can be no improvement in information in the sense of having less undecided citizens.

To illustrate the insight of Theorem 3, I provide a couple of examples with different citizen preference distributions. First, assuming that the distribution is uniform, we have that:

$$F'(1 - F(x))F'(x) = \frac{1}{(1 + \theta + \eta)^2}. \tag{1.4}$$

Given $\theta + \eta > 0$, then $\frac{1}{(1+\theta+\eta)^2} < 1$, therefore the sufficient condition for uniqueness and convergence hold for all $x \in [0, 1]$. On the other hand, an extreme case when the condition fails is provided by a symmetric triangle distribution with small masses of partizan voters. In fact, a simple numerical application shows that for $\eta = 0.05$ and $\theta = 0.07$ and the triangle distribution, polls cannot improve the information since at least 98.5 per cent of the voters in $[0, 1]$ will remain undecided. In the next section I will provide further examples of distributions for which the minimum undecided mass can be between these two extremes examined here.

Figure 1.4 shows the plots of the $h(\cdot)$ function for the two cases mentioned above. In the uniform case, there exists only the fixed point corresponding to the complete information outcome, therefore, with enough polls, it can be achieved. On the contrary, in the triangular distribution case, we can see that there are three fixed points: the complete information one in the middle, the left one where x_L^k converges to, and the third one, which corresponds to the point x_R^k would converge to. It is easy to see that the condition for x_L^k to converge to the complete information outcome does not in fact hold. The two non-middle fixed points define the interval where the minimum possible mass of undecided citizens is located on.

In Appendix A, I examine an extension of the model analyzed here that allows for some day-to-day political noise that can shift the ideological position of the citizens to the left or the right. The only significant difference is that in this case the polls can never help reach the complete information outcome, so there will always be some undecided voters.

At this point it should be mentioned that the mechanism will work the same way if instead of only two parties we had three or more. The algebra is more technically demanding, but the intuition is the same.

1.4 Implemented Policy and Citizen Welfare

In this section I will discuss the influence of polls on implemented policy and on citizen welfare. After each poll we have a potential implemented policy defined as follows:

$$x^k = 1 - p_L^k = 1 - F(\frac{x_L^k + x_R^k}{2})$$

If the conditions for the complete information outcome hold then x^k will converge to x^*.

Social welfare takes the form of a simple utilitarian function:

$$W(x) = \int_{-\eta}^{1+\theta} -u(|x-y|)dF(y) \tag{1.5}$$

with x being the implemented policy and and y being the optimal policies of each citizen. Recall that x^k is the implemented policy we would have if there were elections right after the k-th poll. Using x^k in $W(\cdot)$ we find the social welfare that would be obtained from implemented policy x^k. In the case of linear utility $u(|x-y|) = |x-y|$, the welfare function is strictly concave in x as long as $f(x) > 0$ for all $x \in D$, and it has a maximum at the point x^M that satisfies:

$$\frac{1}{2} = F(x^M)$$

ie., at the median citizen. For $-u(|x-y|)$ strictly concave in x it is easy to see that this welfare function is also strictly concave in x but its

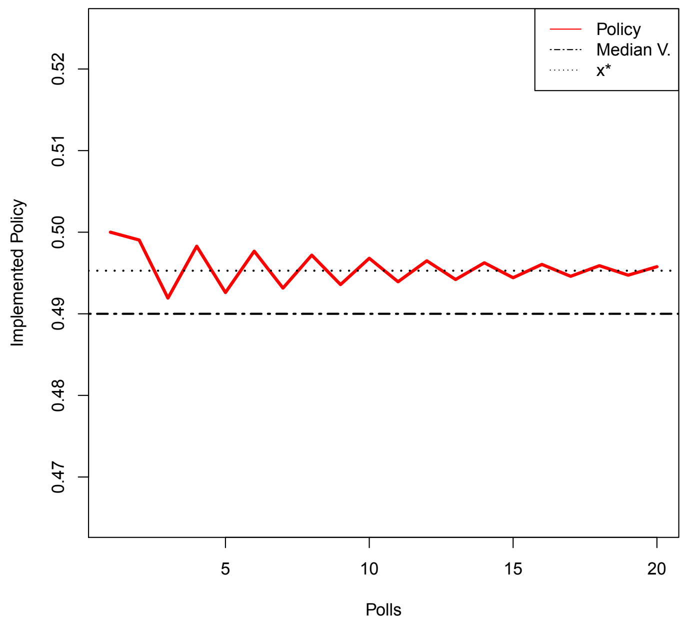

Figure 1.5: Uniform distribution example

maximum will generally not correspond to the median citizen. For instance, in the quadratic case, the maximum occurs at the mean of the distribution. Focusing on the linear utility case, the welfare is increasing as the implemented policy is getting closer to the median citizen and decreasing when the implemented policy moves away from him. A first, trivial thing to notice is that in case of complete symmetry, that is $\eta = \theta$ and symmetric distribution, the implemented policy will stay at 0.5, resulting in the maximum welfare possible, no matter the number of polls and no matter the mass of the undecided because with each poll each party gets the exact same mass of voters and they start with the same mass of partizans. The fact that the welfare is strictly concave in the implemented policy can help us analyze the social welfare using the levels of the implemented policy rather than the levels of the welfare itself. In particular, if the density function is symmetric around the median, then welfare is a strictly increasing function of the distance of the implemented policy from the median voter.[18] Take the

[18] The first and second derivatives of $W(\cdot)$ are respectively: $W'(x) = -2F(x) + 1$ and

case of a uniform distribution of the voters and without loss of generality assume that $\eta > \theta$ (since we have already discussed what happens in case of symmetric distribution and $\eta = \theta$ in the end of the previous section), an example of which, with $\eta = 0.07$ and $\theta = 0.05$ is depicted in Figure 1.5.

This figure shows how the implemented policy moves with every poll, and how far the implemented policy is from the median voter and the complete information outcome x^*. It's easy to see that neither the implemented policy nor the distance to the median are monotonic in the number of polls. In fact there are polls that are welfare reducing: the second poll is particularly bad, taking the implemented policy away from the median voter but also away from the complete information outcome. We have already seen that with the uniform distribution the complete information outcome can be reached. However this example shows that more polls are not always better from a social point of view.

For the next example I will focus on a left-skewed distribution, in particular the Beta distribution with parameters $\alpha = 1.4$ and $\beta = 1.1$. In Figure 1.6 we have $\eta = \theta = 0.05$. In this case we do not have convergence to the full information outcome: there will always be undecided voters and their mass cannot be lower than 67 per cent.

The implemented policy is decreasing on average, getting closer to the full information outcome, but it will never converge to it. The implemented policy converges to $x = 0.56$, while the complete information outcome is only $x^* = 0.54$: the distribution is skewed to the left so there are relatively more undecided citizens located in ideologies close to R. The polls cannot reveal information to the more centrist undecided citizens. Welfare is on average decreasing with more polls, as the implemented policy is getting farther from the median citizen. The polls cannot lead to complete information and on average more polls make the citizens worse-off. Furthermore we can see that if there are not many polls conducted then the welfare has high variability from one poll to the other, and there are polls that are particularly bad in terms of welfare.

If the conditions for the convergence to the complete information hold, then we know that a sequence of polls can lead us to the complete information outcome. However, even in this case, both the implemented policy and the citizen welfare are not necessarily monotonic in the number of polls.

An interesting question is if it is beneficial for the society to have plenty

$W''(x) = -2F'(x) = -2f(x)$. Therefore for symmetric densities around the median x^M the curvature of the welfare is the same for any two points that are the same distance from the median, so for any two such points the social welfare is the same.

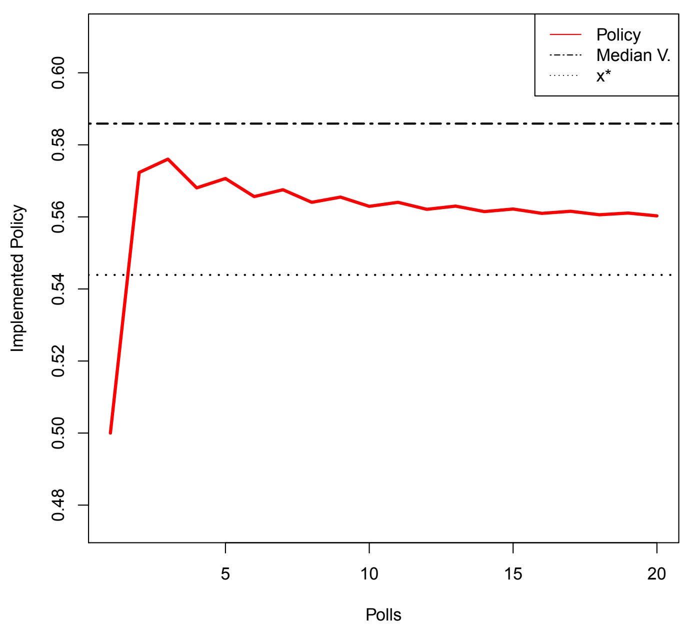

Figure 1.6: Beta distribution example, $\eta = \theta = 0.05$

of polls before the elections or whether there should be restrictions on their publication. The fact is that there is no way to know a priori if the polls will be welfare increasing or decreasing. Given a support, there is an infinite number of possible distributions, and in some of them the welfare will be on average increasing and in some others on average decreasing. What we can say however, is that with more polls, and as the implemented policy is converging, the variability of the welfare will decrease.

1.5 Conclusions

In this paper I examine a proportional representation voting model with polls and in which voters suffer from regret that is able to capture actual elections qualitative features, that to the best of my knowledge cannot be replicated by other models. First, the lack of monotonicity of the share of the leading party in the number of polls, second that the fact that a significant mass of voters declares to be undecided in polls when asked about what they are going to vote and that the mass of undecided is never increasing, and third that there is a positive correlation between having a centrist ideology and being undecided. The most important finding of the paper is that although each poll weakly decreases the mass of undecided voters, a sequence of polls may not always be able decrease this mass to zero and reveal the complete information outcome to the citizens. Polls have a non-monotonic effect on the implemented policy and social welfare. Finally, the average effect the effect of polls on welfare is ambiguous: more polls do not necessarily lead to higher social welfare.

1.6 Appendix A: Position Uncertainty

Even though the citizens have preferences over the implemented policy, it can often be the case that their opinions can change from day to day, due to random political noise. The assumption that this section maintains is that, while a citizen's ideological position can change due to random shocks, this position will not stray to far away from his initial ideological position. I will model this by assuming that the optimal policy for each voter can move randomly around a small interval $[-\epsilon, \epsilon]$, with $0 < \epsilon < \eta, \theta$. This is the only difference with the previous model. That is, using the distribution of voter optimal policies we have previously defined as a reference, each voter's optimal policy can change from day to day, moving randomly in this small interval.

The timing of the new model is as follows. At time t the citizens calculate $[x_L^t, x_R^t]$, a poll takes place to which the voters respond truthfully, and then a shock takes place. A shock for citizen with initial position y is a random draw from the set $[y-\epsilon, y+\epsilon]$. For some extreme L partizans we will have that $y - \epsilon < -\eta$ and similarly for some extreme right partisans we have that $y + \epsilon > 1 + \theta$. This means that there can be draws that can make these partisans "fall over" the support D. If such draw takes place for a partizan then for simplicity we will assume that the shock will just move this citizen to the corresponding extremum of the support. For example, take an L partizan with initial ideological position $y = -\eta + \epsilon/4$. For this citizen the set that his shocks belong to is: $[-\eta - 3\epsilon/4, -\eta + 5\epsilon/4]$. Now assume that this citizens draw is equal to $-\eta - \epsilon/2$ then this would lead the citizen to outside the support D, so his draw will be corrected to be equal to $-\eta$, the minimum point he can move to. The shock is the new ideological position of the citizen. Some partizan voters, the most extreme ones, will stay partizans even if they receive a very big shock in the opposite direction. Afterwards we can have the elections, or we start over with a new poll. The shock is not necessarily common to all voters. Now at the beginning of the game, each voter y is aware of their initial optimal policy and the interval $[y-\epsilon, y+\epsilon]$ that their optimal policy can fall into.

These random shocks from day to day, obviously affect the distribution as they are moving masses of voters around. There are two interesting "perturbed" distributions: the one that results if *all* voters receive the most extreme negative shock meaning that all voters move the most they can towards L, which we call $F_{-\epsilon}(\cdot)$ and one that results if if *all* voters receive the most extreme positive shock (the one that moves all the most towards R): $F_\epsilon(\cdot)$. To see how this difference in the model affects the results, we will examine how a partizan voter would vote.

Take the left partizan voters, the ones whose initial optimal policies are distributed in $[-\eta, 0)$. We know already that $x_L^0 = 0$ so the implemented policy cannot be lower than 0. With no possible shock all of them would want to vote for L. However with the possibility of a preference shock, the partizan voters that are close to 0 know that there is the possibility of finding themselves in the other side of 0. The partizan voters that are absolutely sure of what to vote are the ones that are to the left of $y_L^0 = 0 - \epsilon$. Similarly the right partizan voters that are sure are the ones to the right of $y_R^0 = 1 + \epsilon$. We will call the y_L^k and y_R^k the highest possible cutoff for left voters after k polls and the lowest possible cutoff for right voters after k polls respectively and we will define them as follows: $y_L^k \equiv x_L^k - \epsilon$ and $y_R^k \equiv x_R^k + \epsilon$.

Then the mass of left voters that will be reported by a first poll will be:

$F(-\epsilon) = F_\epsilon(0)$. The mass of the undecided will be: $F(1+\epsilon) - F(-\epsilon) = F_{-\epsilon}(1) - F_\epsilon(0)$. The minimum and maximum possible implemented policies are: $x_L^1 = 1 - F(y_R^0)$ and $x_L^1 = 1 - F(y_L^0)$. Now voters with initial optimal policies to the left of $y_L^1 \equiv x_L^1 - \epsilon$ will know that they will not regret their decision if they vote L, and so do the voters to the right of $y_R^1 \equiv x_R^1 + \epsilon$. If a second poll happens, then the we will have $x_L^2 = 1 - F(y_R^1)$ and $x_R^2 = 1 - F(y_L^1)$ and generally:

$$x_L^k = 1 - F(y_R^{k-1}) \tag{1.6}$$

$$x_R^k = 1 - F(y_L^{k-1}).$$

The proof of existence and uniqueness of the voting equilibrium follows the same lines as before and is therefore skipped. Rewriting the previous two equations and plugging the second one into the other we get:

$$x_L^k = 1 - F(1 - F(x_L^{k-2} - \epsilon) + \epsilon). \tag{1.7}$$

If similarly to the previous subsection we define the function:

$$h_\epsilon(x) = 1 - F(1 - F(x - \epsilon) + \epsilon), \tag{1.8}$$

we see that all the conditions for the existence of a fixed point still hold. Now however, there is no fixed point that corresponds to full information in the sense we have defined it previously. In other words, there does not exist a fixed point where all the voters have made up their minds.

Proposition 4. *The full information outcome cannot be achieved with uncertainty in preferences.*

Proof. Suppose not. Then there exists an s such that $y_L^s = y_R^s = y_L^{s+1} = y_R^{s+1} = y_L^* = y_R^*$. Then we have to have:

$$x_L^* - \epsilon = x_R^* + \epsilon$$

or,

$$x_L^* = x_R^* + 2\epsilon$$

which is a contradiction as x_L^* cannot be greater than x_R^*. □

We can also see that the results of Propositions 2 and 3 follow immediately using:

$$F'(1 - F(x - \epsilon) + \epsilon)F'(x - \epsilon) < 1 \tag{1.9}$$

The natural next question is to find the fixed point that corresponds to the "second best", where the mass of the undecided voters is as low it can get. An obvious candidate fixed point is a point where $x_L^* = x_R^* = x^*$, that is, the point that there is no uncertainty about the position of the implemented policy anymore. Then, we will have $y_L^* = x^* - \epsilon$ and $y_R^* = x^* + \epsilon$.

Proposition 5. *There is no fixed point of $h_\epsilon(\cdot)$ of the form $x_L^* = x_R^* = x^*$.*

Proof. Suppose that time s is the first time we observe $x_L^s = x_R^s$. Then the mass of undecided voters becomes $F(x_R^s + \epsilon) - F(x_L^s - \epsilon)$, and we have: $x_L^{s+1} = (F(x_L^s - \epsilon) + F(x_R^s + \epsilon) - F(x_L^s - \epsilon)) \times 0 + (1 - F(x_R^s + \epsilon)) \times 1 = 1 - F(x_R^s + \epsilon)$. Similarly $x_R^{s+1} = 1 - F(x_L^s - \epsilon)$, showing that $x_L^{s+1} \neq x_R^{s+1}$. □

Proposition 5 tells us that there cannot be a situation where there is no uncertainty about the implemented policy, which further implies that the ideological distance between the last voter to vote for L and the first voter to vote for R cannot be less than 2ϵ. The exact distance cannot be pinned down without first knowing the distribution function.

What we take from the preceding analysis is that the added uncertainty of preferences makes not only the full information outcome completely unattainable, but also puts a lower bound on the measure of undecided voters. As ϵ is getting smaller, more and more people are voting and we are getting closer and closer to full information. On the other hand, with larger ϵ, more voters stay away from the polls.

1.7 Appendix B: Further Empirical Analysis

The World Values Survey aims to provide insight on the political, social, economic, and in general life attitudes of citizens from 59 countries. To this end, it utilizes an extensive questionnaire. Using a measure of Undecidedness the interest lies in showing that there is positive correlation between that and being a moderate voter. In particular, the specification to be estimated is the following:

$$\begin{aligned} Undecided_i = \beta_0 + \beta_1 Center_i + \beta_2 NeverNational_i \\ + \beta_3 NeverLocal_i + \beta_4 NoTrust_i + \epsilon_i \end{aligned} \tag{1.10}$$

Where i refers to a respondent and $Undecided$ is a variable that takes the value 1 if the respondent indicated that, in case there were elections the next day, he either would not know what to vote, he would vote null or blank, or not vote at all, and 0 otherwise. In the survey, Question 95 asks the interviewees to state where they position themselves in the Left-Right political spectrum, giving a number from 1 to 10, with 1 being Left and 10 being Right. The answers to this question constitute the variable $Ideology$. Using the stated $Ideology$ I constructed a series of variables called $Center$ as follows: $Center56$ assigns the value 1 if the respondent declared he places himself on 5 or 6 ($Ideology = 5$ or $Ideology = 6$) and 0 otherwise, and similarly for $Center4567$ and $Center345678$. The variables $NeverNational$ and $NeverLocal$ take the value 1 if the respondent indicated that he never votes in national or local elections respectively and 0 otherwise. $NoTrust$ is a variable that takes the value 1 if the respondent answered that he does not trust political parties in the country and 0 otherwise. $NeverNational$ and $NeverLocal$ are included to control for voters that (according to their own answer) would never vote in elections, no matter who was running. Similarly, $NoTrust$ controls for voters that are in general mistrustful towards parties which can make them vote less. Using the full country data and after dropping observations for which there were no valid answers for the variables examined, there are 53734 observations.

Table 1.2 depicts the results of the least squares estimations of (1.10) using the three different $Center$ definitions and with and without Country fixed effects. Country fixed effects are implemented by simply adding a dummy variable for each country, the coefficients of which are not reported in the table for the economy of space. Since these estimations are of linear probability models, the results of columns (1) to (3), are found by employing weighted least squares, using the fitted values of a first step ordinary least squares estimation to construct appropriate weights.[19] The last three columns use a heteroskedasticity-robust variance-covariance matrix for calculating the standard errors because the WLS method was not appropriate.[20] The results of each column tell the same story: voters that never vote in elections are naturally more likely to be Undecided.

More interestingly, the results show that belonging ideologically to the center implies a higher probability of being Undecided. Columns (1) and (4) define the center in the strictest way. There we see that being in the

[19]More precisely, each model was estimated by OLS first. Then I calculated the estimated standard deviation as follows: $\hat{\sigma}_i = (\hat{y}_i(1-\hat{y}_i))^{1/2}$, provided that $\forall i \; 0 < \hat{y}_i < 1$. Then the WLS estimation was conducted by simply using $1/\hat{\sigma}_i$ as weights.

[20]Since not all fitted values of the initial OLS where such that $0 < \hat{y}_i < 1$

Table 1.2: LS Estimation Results of Specification (1.10)

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.078***	0.071***	0.068***	0.136***	0.118***	0.108***
	(0.003)	(0.003)	(0.003)	(0.023)	(0.023)	(0.023)
Center56	0.097***			0.080***		
	(0.004)			(0.003)		
Center4567		0.077***			0.075***	
		(0.003)			(0.003)	
Center345678			0.059***			0.071***
			(0.004)			(0.004)
NeverNational	0.018*	0.017*	0.017*	0.067***	0.066***	0.066***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
NeverLocal	0.177***	0.180***	0.183***	0.148***	0.150***	0.153***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
NoTrust	0.108***	0.110***	0.114***	0.057***	0.058***	0.059***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	53,734	53,734	53,734	53,734	53,734	53,734
Fixed Effects	No	No	No	Country	Country	Country
R^2	0.235	0.234	0.233	0.241	0.240	0.237
Adjusted R^2	0.235	0.234	0.233	0.240	0.239	0.236

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ (4)-(6) use heteroskedasticity-robust s.e.

Table 1.3: Logit Estimation Results of Specification (1.10)

	(1)	(2)	(3)
Constant	-2.293***	-2.350***	-2.355***
	(0.027)	(0.029)	(0.033)
Center56	0.588***		
	(0.021)		
Center4567		0.503***	
		(0.023)	
Center345678			0.399***
			(0.027)
NeverNational	0.160***	0.149***	0.149***
	(0.042)	(0.042)	(0.042)
NeverLocal	0.859***	0.878***	0.892***
	(0.042)	(0.042)	(0.042)
NoTrust	0.785***	0.793***	0.804***
	(0.027)	(0.027)	(0.027)
Observations	53,734	53,734	53,734
Log Likelihood	-26,871.970	-26,997.450	-27,137.050
Akaike Inf. Crit.	53,753.930	54,004.890	54,284.100
Note:	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$		

center, ceteris paribus, increases the probability of being Undecided by 9.7 and 8 per cent respectively. As the definition of the center is expanded, the effect of being in the center on the probability of being undecided is still positive and significant but falls monotonically. This shows that it is indeed the more ideologically central voters that are more likely to be undecided.

As a robustness check Table 1.3 provides the coefficients of logit estimations of (1.10). The results are qualitatively similar to the least squares results. Furthermore, the least squares estimation was conducted after dropping all the countries that are not parliamentary democracies. These results are not reported here, but they are also qualitatively similar to the original least squares ones.

As a further robustness check, I examined the effect of the ideological position on the probability of being undecided using a measure of distance from the center. For this I calculated the variable $Distance_1$ as follows:

$$Distance_1 = \begin{cases} 5 - Ideology & \text{if } Ideology \leq 5 \\ Ideology - 6 & \text{if } Ideology \geq 6 \end{cases}$$

Table 1.4: Results of Specification (1.11)

	(1)	(2)
Constant	0.175***	-1.705***
	(0.004)	(0.027)
Distance	-0.088***	-0.579***
	(0.004)	(0.026)
Distance^2	0.016***	0.112***
	(0.001)	(0.007)
NeverNational	0.017**	0.160***
	(0.008)	(0.042)
NeverLocal	0.178***	0.859***
	(0.008)	(0.043)
NoTrust	0.106***	0.784***
	(0.003)	(0.027)
Method	WLS	Logit
Observations	53,734	53,734
R^2	0.234	
Adjusted R^2	0.234	
Log Likelihood		-26,845.590
Akaike Inf. Crit.		53,703.180
Residual Std. Error	1.010 (df = 53728)	
Note:	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$	

And then I estimated the following specification:

$$\begin{aligned} Undecided_i = \gamma_0 + \gamma_1 Distance_i + \gamma_2 Distance_i^2 + \gamma_3 NeverNational_i \\ + \gamma_4 NeverLocal_i + \gamma_5 NoTrust_i + \epsilon_i. \end{aligned} \tag{1.11}$$

In Specification 1.11 the square of $Distance_1$ is included so as to capture the effects of the ideological distance that are potentially non-linear. The results of the estimation of this specification are reported on Table 1.4. The results show that an increase in the ideological distance from the center (positions 5 and 6), results in a decrease in the probability of being undecided, albeit in a decreasing way.

The preceding analysis shows that, controlling for other variables, there is a positive relationship between having a centrist ideology and being undecided. However, since it has been based on aggregate data, I will analyze a few cases to see how it works in the country level. For this illustration purposes I chose Sweden, Spain, Estonia and Greece.[21] In addition, for the

[21] Other countries that were analyzed but whose results are not reported to save space

individual country cases I also defined $Distance_2$ to measure the distance from the country median:

$$Distance_2 = |Ideology - median(Ideology)|.$$

The first country to analyze is Sweden. A simple correlation coefficient between $Undecided$ and the two measures of the distance is: -0.22 and -0.24 respectively. The regression results for Sweden are presented in Tables 1.5 and 1.6.

Since the median Ideology in Sweden is equal to 5, I have included alternative specifications of the $Center$ variables: $Center5$, $center456$ and $Center34567$. The Swedish results are consistent with the aggregate findings, with the only difference that $NeverNational$ does not seem to play a significant role in the probability of someone being Undecided and the same holds for $NoTrust$. Dropping $NoTrust$ does not result in different qualitative results. Both distance specifications have the correct negative sign, meaning that moving away from the center decreases the probability of being undecided.

As for the Spanish results, the simple correlation between $Undecided$ and the two distance measures is: -0.23 and -0.30. Tables 1.7 and 1.8 are also consistent with the aggregate: being in the center implies a higher probability of being undecided.

The two countries analyzed so far are consistent with the aggregate results. For comparison, I include the results of Estonia (Tables 1.9 and 1.10), where being in the center has apparently no predictive power over the probability of being Undecided. The simple correlation between $Undecided$ and the two distance measures is: -0.03 and -0.04.

The last country to analyze is Greece, which unfortunately was not included in the last wave of the World Values Survey so for this reason I had to use the European Values Study of 2008.[22] Although the two surveys examine similar issues, the datasets are a little different. For example, in the EVS there are no questions regarding the frequency of participation in elections. Therefore, the variables $NeverNational$ and $NeverLocal$ cannot be used. The variables that have to do with center and $Trust$ are defined in the same way. The variable $NoInterest$ is a dummy variable that takes the value 1 if the respondent indicated that he is not very interested or not

are Australia, Germany, New Zealand, Poland, Slovenia and South Africa. The first four of which give consistent results with the aggregate.

[22] EVS (2011): European Values Study 2008: Integrated Dataset (EVS 2008). GESIS Data Archive, Cologne. ZA4800 Data file Version 3.0.0, doi:10.4232/1.11004

at all interested in politics. A final important, for our purposes, difference between the two is how the variable *Undecided* is constructed. In the WVS the relevant question asked what the interviewees would vote, if there were elections the next day. On the other hand, in the EVS there are three questions examining this. First, question 263 asks whether the interviewee would vote, had there been elections the next day. If the interviewee would answer "yes" then question 264 asks what they would vote. If the interviewee would asnwer "no" to question 263 then question 264 would be skipped and the interviewer would ask question 265: "which party appeals to you the most". In this case, I coded as Undecided those who in question 263 stated that they would not vote at all, those who stated they did not know whether they would vote and those who gave no answer or answered "I don't know" to question 264. Table 1.11 is consistent to the previous findings.

From the table we see that *Center*5 itself is not statistically significant but all the other *Center* variables are. Dropping *NoTrust* does not change the results. The results of Greece for the distance specifications are in Table 1.12.

Distance as defined with the variable $Distance_1$ is significant but not $Distance_2$. Moving away from ideological positions 5 and 6, results in a decrease in the probability of being undecided by 4.9 per cent. The simple correlation between *Undecided* and the measures of distance is: -0.13 -0.12. The results for Greece, although not as clear-cut as for Sweden or Spain, show that there is a positive relationship between being ideologically in the Center and being undecided.

Table 1.5: WLS Sweden

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.063***	0.026***	0.024***	0.038***	0.030***	0.034***
	(0.012)	(0.010)	(0.007)	(0.010)	(0.011)	(0.008)
Center5	0.164***					
	(0.035)					
Center456		0.155***				
		(0.022)				
Center34567			0.114***			
			(0.019)			
Center56				0.182***		
				(0.027)		
Center4567					0.112***	
					(0.019)	
Center345678						0.080***
						(0.020)
NeverNational	-0.051	-0.032	-0.016	-0.044	-0.022	-0.009
	(0.041)	(0.034)	(0.042)	(0.034)	(0.041)	(0.062)
NeverLocal	0.356***	0.335***	0.353***	0.341***	0.349***	0.357***
	(0.085)	(0.083)	(0.102)	(0.081)	(0.086)	(0.125)
NoTrust	0.027	0.036**	0.013	0.035**	0.030*	0.011
	(0.018)	(0.015)	(0.016)	(0.016)	(0.016)	(0.021)
Observations	989	989	989	989	989	989
R^2	0.131	0.133	0.095	0.137	0.126	0.079
Adjusted R^2	0.126	0.129	0.091	0.132	0.121	0.074
Residual Std. Error (df = 984)	0.997	0.994	1.227	0.983	1.014	1.460
Note:				$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$		

Table 1.6: Distance Sweden

	(1)	(2)	(3)	(4)
Constant	0.206***	-1.467***	0.232***	-1.312***
	(0.027)	(0.197)	(0.031)	(0.220)
Distance_1	-0.128***	-1.116***		
	(0.026)	(0.257)		
Distance_1^2	0.020***	0.146**		
	(0.006)	(0.074)		
Distance_2			-0.109***	-0.756***
			(0.024)	(0.230)
Distance_2^2			0.013***	0.038
			(0.005)	(0.059)
NeverNational	-0.013	-0.106	-0.017	-0.122
	(0.046)	(0.407)	(0.045)	(0.408)
NeverLocal	0.326***	1.985***	0.327***	2.006***
	(0.089)	(0.475)	(0.087)	(0.475)
NoTrust	0.029	0.353	0.027	0.331
	(0.019)	(0.220)	(0.019)	(0.221)
Observations	989	989	989	989
Method	LS	Logit	LS	Logit
R^2	0.107		0.110	
Adjusted R^2	0.103		0.106	
Log Likelihood		-308.166		-306.817
Akaike Inf. Crit.		628.331		625.633

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$
(1) and (3) use heteroskedasticity-robust s.e.

Table 1.7: LS Spain

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.168***	0.115***	-0.030	0.156***	0.123***	-0.010
	(0.033)	(0.031)	(0.037)	(0.033)	(0.033)	(0.042)
Center5	0.219***					
	(0.036)					
Center456		0.233***				
		(0.030)				
Center34567			0.283***			
			(0.028)			
Center56				0.181***		
				(0.032)		
Center4567					0.168***	
					(0.030)	
Center345678						0.239***
						(0.035)
NeverNational	0.258**	0.278**	0.312***	0.261**	0.295***	0.331***
	(0.114)	(0.112)	(0.113)	(0.114)	(0.114)	(0.114)
NeverLocal	0.038	0.040	0.047	0.063	0.051	0.038
	(0.116)	(0.115)	(0.115)	(0.117)	(0.117)	(0.115)
NoTrust	0.110***	0.097***	0.144***	0.113***	0.108***	0.139***
	(0.037)	(0.034)	(0.038)	(0.037)	(0.036)	(0.039)
Observations	937	937	937	937	937	937
R^2	0.425	0.424	0.128	0.419	0.413	0.099
Adjusted R^2	0.422	0.421	0.124	0.416	0.410	0.095

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$, (3) and (6) use heteroskedasticity-robust s.e.

Table 1.8: Distance Spain

	(1)	(2)	(3)	(4)
Constant	0.328***	-0.851***	0.382***	-0.649***
	(0.037)	(0.227)	(0.045)	(0.236)
Distance_1	-0.113***	-0.462***		
	(0.031)	(0.175)		
Distance_1^2	0.008	0.009		
	(0.007)	(0.054)		
Distance_2			-0.129***	-0.452***
			(0.030)	(0.174)
Distance_2^2			0.006	-0.030
			(0.007)	(0.050)
NeverNational	0.295***	1.381***	0.299***	1.557***
	(0.111)	(0.530)	(0.115)	(0.571)
NeverLocal	0.034	0.136	0.012	-0.085
	(0.113)	(0.547)	(0.116)	(0.586)
NoTrust	0.121***	0.654***	0.125***	0.682***
	(0.031)	(0.221)	(0.038)	(0.223)
Observations	937	937		937
Method	WLS	Logit	LS	Logit
R^2	0.425		0.149	
Adjusted R^2	0.421		0.144	
Log Likelihood		-555.126		-537.549
Akaike Inf. Crit.		1,122.251		1,087.098
Residual Std. Error	1.002 (df = 931)			

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$, (3) uses heteroskedasticity-robust s.e.

Table 1.9: LS Estonia

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.006	-0.011	-0.010	-0.010	-0.008	-0.023
	(0.008)	(0.009)	(0.011)	(0.008)	(0.011)	(0.015)
Center5	0.003					
	(0.013)					
Center456		0.010				
		(0.011)				
Center34567			0.006			
			(0.012)			
Center56				0.011		
				(0.011)		
Center4567					0.003	
					(0.012)	
Center345678						0.020
						(0.015)
NeverNational	0.129***	0.129***	0.129***	0.128***	0.129***	0.128***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
NeverLocal	0.078***	0.078***	0.078***	0.077***	0.078***	0.080***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
NoTrust	0.020**	0.019*	0.020*	0.019*	0.020**	0.020*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Observations	1076	1076	1076	1076	1076	1076
R^2	0.120	0.120	0.120	0.120	0.120	0.121
Adjusted R^2	0.116	0.117	0.117	0.117	0.116	0.118
Note:	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ all specifications use heteroskedasticity-robust s.e.					

Table 1.10: Distance Estonia

	(1)	(2)	(3)	(4)
Constant	-0.0004	-5.138***	-0.0002	-5.122***
	(0.010)	(0.550)	(0.012)	(0.565)
Distance_1	-0.007	-0.310		
	(0.013)	(0.442)		
Distance_1^2	0.001	0.055		
	(0.004)	(0.126)		
Distance_2			-0.007	-0.253
			(0.012)	(0.382)
Distance_2^2			0.001	0.045
			(0.003)	(0.100)
NeverNational	0.128***	1.180***	0.129***	1.190***
	(0.043)	(0.422)	(0.043)	(0.422)
NeverLocal	0.078***	2.211***	0.078***	2.201***
	(0.025)	(0.501)	(0.025)	(0.498)
NoTrust	0.020*	0.801*	0.020*	0.802*
	(0.010)	(0.466)	(0.010)	(0.466)
Observations		1,076	1,076	1,076
Method	LS	Logit	LS	Logit
R^2	0.120		0.120	
Adjusted R^2	0.116		0.119	
Log Likelihood		-130.552		-130.750
Akaike Inf. Crit.		273.105		273.499

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$ LS specifications use heteroskedasticity-robust s.e.

Table 1.11: LS Greece

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.159***	0.135***	0.091***	0.144***	0.096***	0.079**
	(0.028)	(0.029)	(0.029)	(0.028)	(0.028)	(0.032)
Center5	0.027					
	(0.029)					
Center456		0.073***				
		(0.026)				
Center34567			0.119***			
			(0.026)			
Center56				0.076***		
				(0.027)		
Center4567					0.115***	
					(0.025)	
Center345678						0.104***
						(0.029)
NoInterest	0.129***	0.130***	0.122***	0.123***	0.121***	0.126***
	(0.027)	(0.027)	(0.026)	(0.027)	(0.026)	(0.027)
NoTrust	0.062**	0.055*	0.057*	0.056*	0.068**	0.068**
	(0.031)	(0.031)	(0.030)	(0.031)	(0.030)	(0.030)
Observations	1,165	1,165	1,165	1,165	1,165	1,165
R^2	0.281	0.282	0.284	0.283	0.285	0.283
Adjusted R^2	0.278	0.280	0.282	0.280	0.283	0.281
Note:				$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$		

Table 1.12: Distance Greece

	(1)	(2)	(3)	(4)
Constant	0.224***	-1.198***	0.200***	-1.307***
	(0.035)	(0.201)	(0.035)	(0.208)
Distance_1	-0.049*	-0.230		
	(0.029)	(0.157)		
Distance_1^2	0.003	0.004		
	(0.007)	(0.043)		
Distance_2			0.006	0.112
			(0.024)	(0.135)
Distance_2^2			-0.008*	-0.069**
			(0.005)	(0.032)
NoInterest	0.121***	0.622***	0.127***	0.656***
	(0.027)	(0.135)	(0.026)	(0.135)
NoTrust	0.063**	0.232	0.064**	0.206
	(0.030)	(0.187)	(0.029)	(0.188)
Observations	1,165	1,165	1,165	1,165
Method	LS	Logit	LS	Logit
R^2	0.285		0.283	
Adjusted R^2	0.282		0.280	
Log Likelihood		-659.747		-659.291
Akaike Inf. Crit.		1,329.494		1,328.583
Note:		$^{*}p < 0.1$; $^{**} < 0.05$; $^{***}p < 0.01$		

Bibliography

Alesina A, Rosenthal H (1996) A theory of divided government. Econometrica 64(6):1311–1341

Article19 (2003) Comparative Study of Laws and Regulations Restricting the Publication of Electoral Opinion Polls

Baron DP (1994) Electoral competition with informed and uniformed voters. American Political Science Review 88(1):33–47

Bernhardt D, Duggan J, Squintani F (2009) Private polling in elections and voter welfare. Journal of Economic Theory 144(5):2021–2056

Blais A (2000) To vote or not to vote?: The merits and limits of rational choice theory. University of Pittsburgh Press, Pittsburgh

Blais A, Young R, Fleury C, Lapp M (1995) Do People Vote on the Basis of Minimax Regret? Political Research Quarterly 48(4):827–836

Ceci SJ, Kain EL (1982) Jumping on the Bandwagon with the Underdog: The Impact of Attitude Polls on Polling Behavior. Public Opinion Quarterly 46(2):228–242

De Sinopoli F, Iannantuoni G (2007) A spatial voting model where proportional rule leads to two-party equilibria. International Journal of Game Theory 35(2):267–286

Degan A, Li M (2015) Psychologically-based voting with uncertainty. European Journal of Political Economy 40:242–259

Denter P, Sisak D (2015) Do Polls Create Momentum in Political Competition? Journal of Public Economics 130:1–14, DOI 10.1016/j.jpubeco.2015.07.003, URL `http://papers.ssrn.com/sol3/papers.cfm?abstract{\_}id=2224731`

Duch RM, May J, Armstrong Da (2010) Coalition-directed Voting in Multiparty Democracies. American Political Science Review 104(4):698–719

Ferejohn JA, Fiorina MP (1974) The paradox of not voting: A decision theoretic analysis. The American Political Science Review 68(2):525–536

Ferejohn JA, Fiorina MP (1975) Closeness counts only in horseshoes and dancing. The American Political Science Review 69(3):920–925

Gerber A, Ortuño-Ortín I (1998) Political compromise and endogenous formation of coalitions. Social Choice and Welfare 15(3):445–454

Goeree JK, Großer J (2006) Welfare Reducing Polls. Economic Theory 31(1):51–68

Großer J, Schram A (2010) Public Opinion Polls, Voter Turnout, and Welfare: An Experimental Study. American Journal of Political Science 54(3):700–717

Grossman GM, Helpman E (1996) Electoral Special Competition and Special Interest Politics. The Review of Economic Studies 63(2):265–286

Herrera H, Llorente-Saguer A, McMurray J (2015) The Marginal Voter's Curse. Mimeo

Herrmann M (2014) Polls, coalitions and strategic voting under proportional representation. Journal of Theoretical Politics 26(3):442–467

Kendall C, Nannicini T, Trebbi F (2015) How Do Voters Respond to Information? Evidence from a Randomized Campaign. American Economic Review 105(1):322–353

Kenney PJ, Rice TW (1989) An Empirical Examination of the Minimax Hypothesis. American Politics Research 17(2)

Klor EF, Winter E (2006) The welfare effects of public opinion polls. International Journal of Game Theory 35(3):379–394

Klor EF, Winter E (2014) On Public Opinion Polls and Voters' Turnout. Mimeo

Li M, Majumdar D (2010) A psychologically based model of voter turnout. Journal of Public Economic Theory 12(5):979–1002

Llavador H (2006) Electoral Platforms, Implemented Policies, and Abstention. Social Choice and Welfare 27(1):55–81

Loomes G, Sugden R (1982) Regret Theory: An Alternative of Rational Choice Under Uncertainty. The Economic Journal 92(368):805–824

Matakos K, Troumpounis O, Xefteris D (2015) Turnout and Polarization under Alternative Electoral Systems. In: Schofield N, Caballero G (eds) The Political Economy of Governance: Institutions, Political Performance and Elections, Springer International Publishing, Cham, Switzerland, pp 335–362

McKelvey RD, Ordeshook PC (1985) Elections with limited information: A fulfilled expectations model using contemporaneous poll and endorsement data as information sources. Journal of Economic Theory 36(1):55–85

Meirowitz A (2005) Polling games and information revelation in the Downsian framework. Games and Economic Behavior 51:464–489

Morgan J, Stocken PC (2008) Information aggregation in polls. The American Economic Review 98(3):864–896

Morton RB, Muller D, Page L, Torgler B (2015) Exit polls, turnout, and bandwagon voting: Evidence from a natural experiment. European Economic Review 77:65–81

Myatt DP (2015) A Theory of Protest Voting. The Economic Journal

Myerson RB, Weber RJ (1993) A theory of voting equilibria. American Political Science Review 87(1):102–114

Ortuño-Ortín I (1997) A spatial model of political competition and proportional representation. Social Choice and Welfare 14:427–438

Ortuño-Ortín I (1999) A Model of Voting with Incomplete Information and Opinion Polls. IVIE WP-AD 99-25

Pieters R, Zeelenberg M (2005) On bad decisions and deciding badly: When intention-behavior inconsistency is regrettable. Organizational Behavior and Human Decision Processes 97:18–30

Sahuguet N, Persico N (2006) Campaign spending regulation in a model of redistributive politics. Economic Theory 28:95–124

Saporiti A (2014) Power sharing and electoral equilibrium. Economic Theory 55(3):705–729

Savage LJ (1951) The theory of statistical decision. Journal of the American Statistical Association 46(253)

Shimanoff SB (1984) Comonly Named Emotions in Everyday Conversations. Perceptual and Motor Skills 58(2):514–514

Spangenberg F (2003) The freedom to publish opinion poll results: Report on a worldwide update. Tech. rep., ESOMAR/WAPOR

Taylor CR, Yildirim H (2010) Public information and electoral bias. Games and Economic Behavior 68(1):353–375

Zeelenberg M (1999) The use of crying over spilled milk: A note on the rationality and functionality of regret. Philosophical Psychology 12:325–340

Chapter 2

The Last Shall Be the First: Failed Accountability Due to Voters Fatigue and Ballot Design

In 2000, Al Gore lost a significant amount of votes to Patrick Buchanan in Palm Beach Country, Florida. Due to a confusing ballot design, around two thousand voters may have voted for the latter by mistake (Wand et al, 2001). Later on, George W. Bush won the election in the state of Florida by a thousand votes and became president of the United States. If a voter's behavior may be affected by the way in which the alternatives are presented to him in strong democracies, we should pay attention to less developed ones as well. It may be the case that the layout of ballots have a more pervasive influence in younger democracies, because of either unintentional or deliberate manipulation. Either way, the voters' preferences would not be reflected in the electoral results, which undermines the principle of accountability in democracies.

In this paper, we use polling-booth level data to show that the ballot design explains around 3% of the variation of the votes in a (removal) referendum in Lima in 2013. Moreover, we show that different layouts would have led to completely dissimilar electoral results, changing the composition of Lima's local legislature in a significant way. Furthermore, we show that the final political outcome may not have been representative of the preferences of the electorate due to ballot-induced "choice fatigue", i.e., when voters are faced with simultaneous elections they are more likely to cast a vote in the ones listed first than in the ones listed last (Augenblick and Nicholson, 2015).

In 2013, in an unprecedented measure to increase political accountability, Lima's citizens had the opportunity to remove each of the forty local legislators, voting them out one by one. However, due to a combination of poor ballot design and voters' fatigue, rather than reinforcing accountability, this removal referendum penalized disproportionately the majority coalition, whose legislators were placed first in the ballot.

This removal referendum allows us to investigate the effect of choice fatigue on aggregate outcomes cleanly because voters' preferences are uncorrelated with the candidates' positioning in the ballot. Our identification follows from the (quasi) random allocation of voters to polling booths, and the location of the legislators in the ballot, which did not depend on their popularity or other observable characteristics. Thus, controlling for precinct unobservables and candidates observables, we can estimate the causal effect of voters' fatigue (due to the legislator's position in the ballot) on the removal referendum.

Our main finding shows the presence of choice fatigue: the share of blank votes increases rapidly and continuously between the first and fortieth candidate. Moreover, since candidates where listed in two columns, we also identify a large jump of blank votes between the last candidate in the first column and the first one in the second column.

Secondly, we estimate the effect of the existing fatigue on the probability that any given candidate is removed from the local council, taking into account that voters face three alternatives which do not need be independent of each other – `YES`, `NO` and `BLANK`, where the first option is a favorable vote toward recalling the candidate. Using the yes, no and blank shares by polling booth (that is, 170 voters on average) we show that, not only the ballot design had an effect on the vote shares, but also, more reasonable designs would have had completely different outcomes. For instance, in the spirit of King et al (2000), we simulate the electoral results under different ballot designs and we show that with a ballot with a random order of legislators, everybody but the mayor would have been removed from office.

As a result, a constitutional provision meant to increase accountability, the removal referendum, combined with a poor implementation that did not take into account (predictable) choice fatigue might have had counterproductive consequences. Since the twenty legislators listed first in the ballot, representing almost the whole majoritarian coalition, had to be replaced with a by-election, the removal referendum may have caused Lima's council to be less representative of voters preferences: after the new council was put together, around 46% of politicians came from the main opposition party (PPC - *Partido Popular Cristiano*), whose seats increased from 9 to 16, even though the maximum share of votes it had obtained in this period was less than 30%[1].

The paper is organized as follows: in Section 2.2 we discuss the institutional background, in Section 2.3 we explain our econometric strategy, in Section 2.4 we show evidence of voters' fatigue and we quantify the effect of the ballot design on the electoral results by exploring the effects of using alternative ballots. Finally, in Section 2.5, we discuss the implication of our results and we conclude.

2.1 Literature

Our findings indicate the presence of a political "Peltzman effect" (Peltzman, 1975): the introduction of a provision meant to increase accountability may have the opposite consequences if combined with poor implementation. In our case, the unintended consequences arise due to a poorly designed ballot, which did not take into account documented evidence about choice fatigue (Augenblick and Nicholson, 2015) and other behavioral anomalies related to the order in which the options are presented to consumers (Rubinstein and Salant, 2006).

More generally, the puzzle of voting has caught the economists' attention since Downs (1957): if voting is costly, why do people vote in large elections when the probability of being pivotal is arguably zero? The early theories of the calculus of voting Riker and Ordeshook (1968); Downs (1957); Enelow and Hinich (1981) draw on voters' heterogeneity to explain abstention decisions. Feddersen and Pesendorfer (1996) highlights that a comprehensive theory of turnout and abstention should also explain the phenomenon of *roll-off*[2]. Closer to the point in our paper, ballot design is not innocuous as it may favor the candidates ranked first - priming effect (Esteve-Volart and Bagues, 2012; Meredith and Salant, 2013)), or ranked last - anti-priming effect (Alvarez et al, 2006). Augenblick and Nicholson (2015) suggests the presence of voting fatigue using a natural experiment in California, and they show that when an election appears sooner in a ballot, there is a significant lower number of abstentions.

[1] And *Peru Posible*, another opposition party, moved from 1 to 7 legislators, even though it obatined only 11% of the votes.

[2] As explained in Feddersen and Pesendorfer (1996): "Roll-off occurs when voters who are already at the polls decide not to vote on a race or issue".

There is much evidence about the effect of the set and order of alternatives on consumers' choice outside political economy. Among others, Feenberg et al (2015) show that the order in which papers appear at a NBER mailing list affects the short run number of citations, Liu and Simonson (2005) show that pairwise comparison of products leads to more consumption than a sequential one, and Glejser and Heyndels (2001) show that participants who perform later in a music contest are more likely to get more positive evaluations from the jury.[3]

2.2 Institutional Background and Data

Perú is a federal representative democracy subdivided into 25 regions and the capital city, Lima. Each region is composed by provinces and districts. Institutionally and politically, the Municipality of Lima has the status of a region rather than a city, despite its name. It is also the largest city of Peru with 8.5 million inhabitants and hosts the executive, judiciary and legislative branches of the national government of Peru. In 2010, Susana Villaran runs for mayor of Lima as the leader of a center-left coalition of parties. In October of that same year, Villaran is the first elected woman to become mayor of Lima.

2.2.1 Institutional Background

The city is run by a mayor (the maximum administrative authority of the executive branch) and thirty nine city legislators - *regidores*.[4] These forty politicians are chosen in a municipal election every four years by popular vote with a closed-list proportional rule that gives an automatic majority (twenty out of forty regidores) to the party with most votes. All remaining seats are assigned proportionally.[5]

Recall referendum: The Constitutional reform of 1993 added the possibility of calling for a recall referendum, known as *Consulta Popular de*

[3] From a theoretical perspective, Rubinstein and Salant (2006) point out that an increasing number of choices are made out of list sets, and study the conditions under which the independence of irrelevant alternatives axiom (IIA) can be extended to a list-IIA. Kamenica (2014) shows that choice overload can be exploited by a profit-maximizing monopolist by introducing a premium loss leader product.

[4] As stipulated in *Ley Organica de Municipalidades.*

[5] Closed-list proportional rule implies that each party proposes a pre-defined list of 40 candidates and his substitutes, and citizens choose a party, without any interference on the list composition.

Revocatoria (CPR). This provision, meant to keep politicians accountable, implies that all subnational politicians holding office can be exposed to a non-confidence vote at (almost) any time of their mandate, except the first and last year.

A referendum takes place only after a formal request signed by 25% of the citizens that live in the jurisdiction of the politician under scrutiny. Up until 2015, there was a cap on the signatures needed, which implies that in the case of Lima only 400 thousand signatures are needed to proceed with the recall referendum.

The electoral rule used is simple majority rule: if more than 50% of the registered voters participate in the recall referendum and if the non-confidence votes (`YES` votes) to a politician are more than the confidence ones (`NO` votes), then he is recalled and he must be replaced. If more than thirteen legislators are recalled, a new election takes place to replace them. Otherwise, they are substituted by a party member who was on the closed list presented by the party in the original election. If needed, in the by-election, the new legislators are elected by a closed-list proportional rule (without a bonus to the winner).

On the one hand, the potentiality of a recall should provide the incumbents with incentives to listen to their constituency and perform according to their needs. On the other hand, this provision has given the elections' losers an instrument to prevent the winner from governing: since its introduction there were more than five thousand recall referenda, and 20% of the recall initiatives were promoted by candidates who did not win the election.[6]

Electoral Management

The ONPE ("Oficina Nacional de Procesos Electorales") is an independent body in charge of organizing and administering the elections and referenda in Peru. It was created in 1993, and among other tasks, is in charge of designing the ballot papers. For the removal referendum of March 17th 2013, when forty candidates were up for removal, the final ballot design sticked to the following rules: parties are ordered downward according to the number of legislators and, within the party order, legislators are ordered according to their order in the 2010 closed list, which was determined by the parties. Hence, the ballot used in 2013 had the forty candidates listed in two columns without including any partisan identification or picture, as shown in Figure 2.1.

[6]According to ONPE (2010), in 2009, 36% of the initiatives were promoted by loosing candidates (to mayor or legislators).

Voters: Voting is mandatory in Peru. When citizens turn 18 years old, they are issued a national I.D. (D.N.I.) and they are assigned an "electoral number" that determines the polling booth where they have to vote. This number is assigned according to the order in which citizens got their I.D. As we discuss later, a direct implication of this mechanism is that within a polling center, any two polling booths, have the same ex-ante distribution of votes.

2.2.2 Data

In 2010, there was a city-wide election to choose the municipal authorities. Six parties or coalitions had managed to elect at least a legislator (see Table 2.7.2). In 2013, the forty legislators were put up for a confidence vote, in which citizens had to decide, for each legislator, whether to remove him/her (e.g. vote YES), or not (e.g. vote NO). In November 2013, a by-election to elect the substitutes of the recalled candidates was held. We use the electoral data of these elections and of each of the candidates subjected to the removal referendum. Except for the electoral data by polling booth – which was provided by ONPE after our request – all the data is available at ONPEs website (www.onpe.gob.pe), unless stated otherwise.

Electoral data: The mayor and twenty one legislators belonged to coalition *Fuerza Social* (FS), thirteen to coalition *PPC - Unidad Nacional*, two from party *Restauracion Nacional*, one from party *Cambio Radical*, one from party *Somos Peru*, and one from party*Siempre Unidos*.

There are 36,740 booths divided in 888 centers throughout Lima's 43 districts. Polling booths do not all have the same number of registered voters, but they cannot be larger than 300 voters, by constitutional design. In 2013, the total number of the eligible voters was 6,357,243, with a turnout rate of 83,7%. Figure 2.3 shows that the blank votes for each candidate display a large variance, which increases with the position of the candidates in the ballot. Figure 2.4, the actual votes for Yes and No for each candidate, summarizes the results from the referendum: the mayor was kept in her place, while candidates in positions 2 to 21, from *Fuerza Social* were recalled. Also candidates in position 26 and 31, both from *PPC - Unidad Nacional* were recalled (discussed below).

The support for all the parties is relatively stable across districts (see Table 2.7.2), with the noticeable exception of *PPC - Unidad Nacional*, with the largest standard deviation among the parties that have won a seat. This party obtained large support in less populated areas, with an average of 40%, even though the total actual share was 37.5%.

Candidates characteristics: Table 2.7.2 shows the candidates observable characteristics, ordered by their position in the ballot. The median legislator is a 49 years old male politician with college degree, who had won a local election once, but with no experience in national politics. As it can be seen by media exposure, the median candidate is almost unknown to the voters. This variable is constructed with the mentions of the forty legislators or *regidores* in the period between 2010 and March 2013, in the five most important newspapers from Perú (El Comercio, La República, Perú 21, Gestión and Correo). Although it is the variable with the most disparity, the median candidate has been mentioned only three times in the three years previous to the removal referendum. The candidates with the most mentions are the mayor (Villarán, with 1068 mentions), and Luis Castañeda Jr, the son of a previous mayor[7].

2.3 Empirical strategy

To examine the relationship between the results of the removal referendum for a legislator and her position in the ballot, we estimate two models, described below.

2.3.1 Estimation strategy

To analyze the effect of candidates positioning in the ballot on his electoral performance, we begin by analyzing the following linear model:

$$Y_{idl} = \alpha\theta_d + \beta X_l + \gamma Z_l + \varepsilon_{idl} \tag{2.1}$$

In Equation 2.1 we explain the share of blank votes of legislator in position l in polling booth i in precinct d, Y_{idl}, as a result of the politicians individual characteristics, X_l, and the position on the ballot, Z_l. Last, θ_d captures the precincts characteristics, either as a fixed effect or as other variables such as poverty, depending on the specification. More importantly, precincts refer to the geographical location of the polling booth, either the polling center or the district, depending on the specification. To show the

[7]Castañeda Jr., who was allegedly behind the call for a removal referendum, may also be confused with Luis Castañeda Sr., so we may be overstating his media exposure). Incidentally, Castañeda Sr. was elected mayor when Villaran finished her mandate

robustness of our results, we also explore the effect of the ballot design on the share of blank votes per candidate with a non-linear probability model with aggregate data. In all specifications we cluster the standard errors by polling center, and in all estimations of Equation 2.1, the precinct fixed effects are at the polling center level. [8]

We also run a multinomial probit to explore non-linear effects on the share of `YES` and `NO` votes in our specification. Our approach to the multinomial estimation is based on a model of voting with additive noise, a la Banks and Duggan (2005), explained in detail in the Appendix. Moreover, these estimates are the main input for building the counterfactuals in the next section.

Even though precincts fixed effects control for unobserved heterogeneity, there might be other confounding voters' or candidates' characteristics. Both can be ruled out by design, as we explain in the next section. Regarding the former, within a polling center, the allocation of voters to polling booths is (quasi) random, hence in any two polling booths within a polling center we should expect similar outcomes. Regarding the latter, the position of the politicians in the ballot does not depend on individual characteristics, but the parties'.

2.3.2 Identification strategy

Our identification relies on two sources of exogenous variation: the predetermined order of candidates in the ballot during the 2013 removal referendum, and the (quasi) random allocation of voters to polling centers and polling booths.

In the first place, the candidates order in the ballot is determined jointly by the performance of their party/coalition in the 2010 elections, and their position in the closed list for that same election. Hence, the main determinant of the legislators position of the ballot, specially the column, is not determined by his personal characteristics but by the party's popularity (nonetheless, we will also control for the candidates' characteristics). Moreover, since *Fuerza Social* and *PPC-Unidad Nacional* were coalitions of parties, a politician's position within the coalition would depend more on his party bargaining power rather on his own skills or electoral potential.[9]

[8] Even though it should not be concern in our case due to the lack of information about the legislators, we also use multiway clustering a la Cameron and Miller (2015). As we will show in Table 2.11, the standard errors increase – as expected – but all our main coefficients are statistically significant at the 1% level.

[9] In the next section we will provide some empirical evidence regarding the absence of

To reinforce the idea that the votes for a politician in 2013 and his position in the ballot cannot be jointly determined, we argue that the voters' information on the candidates is very poor. The ballot design does not include any other information than the candidates' names, and except for notable exceptions (the mayor and the son of the former mayor), all these local legislators are relatively unknown to the population (except for Villaran – in the first position – and Castañeda Jr. – in position thirty one, as commented above).[10]

The second source of variation is also potentially important. Unobserved heterogeneous preferences for parties or candidates across geographic areas could potentially affect the estimates of the effect of the ballot design. In areas where right-wing candidates are preferred, votes of confidence (no recall) could be more likely for candidates located at the end of the ballot, and the opposite would be the case for the candidates located above. A similar phenomenon, with the opposite direction could occur on areas with more left-wing preferences. All in all, a direct comparison between the candidates results without accounting for unobserved heterogeneity could bias the estimated effect of the ballot design.

In order to avoid this, we include school fixed effects, given that the citizens' political preferences are orthogonal to the voters' allocation to polling booths within precincts, by design: as citizens turn 18 years old and get their I.D.s, they are automatically registered as voters in a given polling booth and center. Thus, this allocation only depends on the order in which citizens get their I.D. Concerns about cohort effects - if voters ideology depends on age - are taken care of with the inclusion of polling-center fixed-effects, which controls for heterogeneity across centers.[11]

Validity of the identification strategy

As shown in Figures 2.1 (the actual ballot) and 2.2 (the ballot with party identification), at the time of voting, citizens have no information about the candidates except for their last names. Concerns with our identification strategy may arise if voters knowledge about the seating legislators was correlated with their position in the ballot. Nonetheless, these local politicians are not only unknown, but also they lack significant previous political experience (see Table 2.4).

Moreover, if voters had used other informational cues, e.g. being first

correlation between *Order* and individual characteristics.

[10] In particular, there are not even party identifiers in the ballot.

[11] Cantú (2013) and Casas et al (2014) use a similar identification strategy.

within the closed-party list, we should observe jumps in the number of blank votes. In Figure 2.5, that shows the change in blank vote shares between any two consecutive candidates in the ballot, we can see that the only jumps in that variable occur for the mayor, with the change of column and with Castañeda. In all other cases, including the *jump* from one party to the other, the increase in blank shares is steady.

Last, despite Lima's large population density, it could be argued that within a polling center, across polling booths, there could be selection of voters, undermining our identification strategy. Fortunately, we can even add polling booth fixed effects to account for this possible source of bias due to omitted variables. While in the appendix we show that our results would not change, confirming that there is no such problem of endogeneity – the coefficients are statistically identical to our main regressions – we prefer the polling center fixed effects.

2.4 Results

2.4.1 Determinants of voting blank

The effect of order, column and the candidates' characteristics on the share of blank votes per candidate are reported in Table 2.7. Remarkably, in all specifications the signs of almost all of our estimates are the same, even when we omit the polling center fixed-effects, suggesting that our estimates do not depend of the distribution of preferences within a polling center.

In all models, the position of the legislator on the ballot has a very strong effect on his share of blank votes: the coefficients of $Order$, $Column$ and their interaction are all statistically significant at the 1% level. While $Order$ and $Column$ are positive, indicating that being further down the ballot leads to more blank votes, the interaction coefficient is negative. Across all specifications, being ten positions further away implies an increase in blank votes by 1.3 percentage points. Furthermore, being in the second column increases the blank votes in 4.4 percentage points. This evidence indicates that citizens follow the order in which candidates appear in the ballot, instead of jumping across columns.[12]

Some remarks are in order: although the effect of the order in which the legislator is placed is an important predictor of the share of blank votes, this effect decreases in the second column. The blank share is increasing

[12] One could think that voters begin by looking at the first two candidates in the first row – candidate 1 and 21 – instead of the first two candidates in the first column – candidates 1 and 2. This is ruled out by the data.

at a lower rate compared to the first column. One possible interpretation of this result is that voters who keep voting in the second column may have a lower cost of voting, i.e., there is a selection of voters. Hence, the rate at which they stop voting is smaller. This interpretation may have implications beyond the decision to abstain: if voters "self-select" into the second column, their preferences and their voting behavior may differ in comparison to the set of voters in the first column. It could be worrying that the cost of voting is correlated with preferences. A first attempt to correct this issue is in column (3) of the same table, where we control for the voters preferences in the previous election (the variable "Party share in 2010"). We observe no change in our estimates, which suggests that the above mentioned correlation does not take place here.[13]

Both the coefficients of for Villaran (position 1 - *Mayor*) and *Castañeda* are significant with a negative sign, showing that controlling for ballot effects, party affiliation and other individual characteristics voters were more likely to express their opinion about these two compared to the remaining ones. These findings are consistent with what we see in Figure 2.3.

Focusing in column (3), where we have the extended set of covariates of individual characteristics we see that none of them has a great effect on the blank vote behavior, in comparison to *Order* and *Column*. This is also in line with our identifying assumption: with the exception of the two famous faces of the campaign (Villaran and Castañeda) the public cannot distinguish among different officials in the ballot. Nonetheless, if anything, exposure has a negative effect: having been a candidate in a previous national election, won or not, and having won a local election, decreases the blank shares. We find the same effect with the number of mentions in national newspapers (the variable "media exposure").

Finally, belonging to the right-wing coalition (*PPC - Unidad Nacional*) further increases the share of blank votes by 4.4 percentage points. Table 2.8 shows the squared semi-partial correlation coefficients for the linear estimation of column 3 of Table 2.7. *Order* and *Column* account for most of the variance in the model. Columns (4) to (6) indicate that our results do not depend on the linearity imposed in the first three columns. [14]

[13]We also address this issue in our multinomial specification.

[14]Even though our identifying assumptions are not too demanding, it could be argued that within a polling center, there may be significant differences across polling booths that may be biasing our results. In Column (2) of Table (2.11), we show the results with polling booth fixed effects, and the coefficients are identical. Moreover, in the first column of that table we cluster the standard errors in a different way, following Cameron and Miller (2015) multi-way clustering. This clustering is more demanding and increases the

As a further robustness check we compare what happens when we drop all covariates and in their place we substitute forty dummies, one for each position in the ballot. Pairwise comparisons of the coefficients show that all are statistically significant from each other. Figure 2.6 depicts the coefficients of the dummies.

Before we continue to the next subsection, we need to examine if there are any empirical evidence regarding the correlation of the *Order* variable with any personal characteristics of the officials. To examine this, we run a simple regression, *Order* against all valid personal characteristics, for all forty officials, but also by restricting our attention to the left-wing and to the right-wing coalition. The estimation results are in Table 2.6. The sample is very small (only forty observations) but using the full sample we see that the only two significant variables are the *College* dummy and the *Elected in national elections* one. However, the elected variable is essentially a dummy since there is only one official that has ever been elected at the national level, so it would naturally be highly correlated with the *Order* variable. The *College* dummy is very significant and positive which is driven by the fact that for all right-wing officials the dummy is equal to 1 (this is the reason why, the *College* dummy is dropped when we restrict the sample only on the right-wing officials). When we examine the two other samples separately we see that the *Elected in national elections* variable is dropped, since within coalitions there is no variation. In the fourth column of the table, we use as a dependent variable the order of the officials in their own party's closed list.[15] The results are not qualitatively different, with the exception of the significance of media exposure. In any case, the sample size is even smaller in these cases, so naturally we expect large standard errors. Nevertheless, we do not have any evidence that would suggest that the order is correlated with any observable characteristics.

2.4.2 Analysis of the probability of being removed from office

Table 2.9 depicts the results of the multinomial probit in which the probability of voting blank is the baseline. The upper table shows the results for the `NO` and the lower part the `YES` ones, always with respect to blank. Our preferred specification is in the third column, where we control both for *District* fixed effects and the extended set of covariates, but the three

standard errors, but still all our coefficients of interest are significant at the 1% level.

[15] This means, for example, that the officials in positions 23 to 35 are re coded from 1 to 13.

estimated models are qualitatively similar.[16] The effect of *Order*, *Column* and their interaction are in line with our previous results: the share of `YES` and `NO` decreases with respect to the blank votes as the candidates are further down in the ballot. Interestingly, while both coefficients for *Mayor* are positive, indicating a relatively divided opinion about keeping her or not, the coefficients on *Castañeda* are negative for `NO` vote and positive for `YES`. This result shows a large level of agreement on removing him from the local council in Lima, in line with Figure 2.4. The effects of *Party Share in 2010 elections* have the expected signs: the higher the party share of an official, the more likely that this official receives a no vote compared to a blank one, which is also more likely compared to a no one.

Since the coefficients of `YES` and `NO` vote of each covariate are often quite similar, pairwise statistical comparisons were in order. In all columns, all coefficients are statistically different between the `YES` and `NO` equations with the exception of the minor party dummy coefficient, who is only statistically different between the two equations only in the third column. The difference in model fitness between the first column and the other two may indicate the presence of heterogeneity among the districts which we address below.

Analysis by district. Table 2.10 shows the results of the multinomial probit in two districts: Villa El Salvador and San Isidro. Villa El Salvador is the district where "Fuerza Social" (the incumbent party) obtained their highest share in the 2010 election, whereas San Isidro is the district where "PPC-Unidad Nacional" got their highest share, which is the same district where Fuerza Social got their lowest share. Although all the coefficients are consistent with our previous results, some very interesting patterns emerge. The effects of *Order* and *Column* are larger in column (1), a left-wing district, than in column (2), a right-wing one. Hence, the design of the ballot seems to have more pervasive effects in the district where the mayor

[16] To estimate the probit model with district fixed effects, we create a dummy variable for each district. Being a non-linear model, this could be problematic, given that these models are known to have small sample bias when the group sizes are small (Greene, 2004). In our case, however, the average number of tables per district is around 840, so this problem would be less of a concern. Additionally, the results shown that controlling for the unobserved district characteristics does not seem to greatly affect the estimated coefficients, so that the panel model could be consistent with a model with random, rather than fixed, effects. Heckman (1981) showed that the small sample bias in this case would be "surprisingly small". Last, we are not concerned about consistency of the estimators, as the models with and without fixed effects result in estimates of the same direction and magnitude. Nonetheless, in our simulations we use by-district specifications, as we explain below.

was elected with the largest share. Although this results does not threaten our identification strategy, in order to make policy recommendations we have to take into account the districts' heterogeneity of preferences, as we do in the next section.[17]

2.4.3 Counterfactual Analysis and Discussion

Based on our previous results, we analize what would have happened in Lima if the design of the ballots were different. Since the legislators' location is the main explanatory variable, we explore designs that change their position in the ballot: in particular, we randomize the order of the legislators and/or we eliminate the columns.

Since we have to incorporate the districts' hetoregeneity to provide the most accurate counterfactuals, we do the analysis by district (for the 43 districts, as with El Salvador and San Isidro above), we collect the estimated coefficients and, taking into account the districts size, we perform our exercises. Figure 2.7 shows the fitted values of the multinomial probit model (run by district). Comparing with the actual results in Figure 2.4 we see that the model does a good job reproducing the referendum results.

In order to see how the ballot design affected the referendum result we force the *Column* and *Order* coefficient to be zero for all officials. Hence we obtain the predicted values as if the ballot design had no effect: Figure 2.8 shows this result.[18] For all officials except the mayor, the `YES` votes are a majority, meaning that everybody except the mayor would had been recalled. In practical terms, this result shows what the outcome would had been if every voter received a different single-column ballot where the order of the officials was completely randomized.

In order to be more precise, we simulate the elections varying the order and the columns. First, we assigned each district a different random order in the ballot, calculated the referendum result predictions, and repeated this process one thousand times. Then we collected the results of the repetitions and calculated the share of recalls for each official. Figure 2.9 shows the result of this process when assuming a two-column ballot and Figure 2.10 shows the result assuming a single column. Starting with the latter, we see

[17]Related, but less surprising is the fact that the share of `NO` votes for *Villaran*, the mayor, is larger with respect to the blank votes in the left-wing district than in the righwing one (San Isidro). Similarly, *Castañeda* has larger support in San Isidro, while the voters in Villa El Salvador do not show any differential treatment of this candidate.

[18]The figure where we shut down only the column differs from Figure 2.8 only in the sense that both `YES` and `NO` vote lines have a negative trend without intersecting.

that in almost all repetitions, it is only the mayor who is not getting recalled. On the other hand, with two columns in almost all repetitions it is only the mayor who survives the recall from her coalition. As for the opposition, Castañeda gets recalled almost always, while the rest of the officials of the right-wing coalition get recalled some of the time.

The predicted values and the simulations point into the same direction: the ballot design had given an "unfair" advantage to all the legislators located in the second column. A proper randomization, even keeping two columns, would have ended up in completely different electoral outcomes. For simplicity of interpretation, let us focus on the case with a single column. If all legislators had been recalled, the final composition of the city council after the by-election for the new members would have ended up having only less than a third of legislators from the *Partido Popular Cristiano* (PPC). Instead, because the ballot design gave a safer position to the legislators located in the second column, as it was the case for the PPC ones, these legislators were not removed. As a result, the final composition of the council included the new seven legislators elected in the by-election plus the other nine elected in the 2010 election. The sixteen legislators from PPC represent 41% of the council, even though this party never obtained more than 30% of the votes. Hence, the possibility of the recall referendum, intended to keep politicians accountable during their term, ends up having adverse effects on the representativity of the political institutions.

2.5 Conclusion

On one hand, Peru's Constitution allows for an ad-hoc recall of politicians during their tenure. This provision, together with regular elections, makes politicians more accountable to their citizenry. In particular, it is meant to address the issue of moral hazard by increasing the instances in which citizens can remove politicians from office. Moreover, as instance of direct democracy, it also allows for a more direct supervision of the politicians activities and decisions.

On the other hand, we showed that the implementation of this provision and the "removal referenda" have to be taken very seriously. A poorly designed ballot may have consequences that go against the spirit of the rule of law, as it may have been the case in Lima's 2013 referendum. Rather than their performance, the order in which legislators were listed in the ballot and the column in which they were placed were the main determinants of being removed from office: the legislators position and column are better predictors of the electoral results than education, partisan affiliation, political experience, and presence in media – jointly, they are two thousand times better predictors than the other variables (see Tables 2.7 and 2.8). As a result, the politicians listed in the second column were more likely to be confirmed in office, only due to their position in the ballot.

Unfortunately, this case shows yet another instance in which institutional design has unintended consequences. The old adagio – institutions matter – may still be true, but institutional reform has to be accompanied with a thoroughly thought implementation. Furthermore, our evidence indicates that electoral transparency scholars and international electoral observers may also need to devote attention to ballot design. If these "details", like ballot design, may tilt elections in developed and strong democracies, younger democracies' accountability mechanisms may be threaten by a greater exposure to manipulation.

Last, we also proposed an under-exploited identification strategy: the combination of polling-booth level data with a deep understanding of the assignment of voters to polling booths – in our case according to the I.D. number. Yet, beyond the purposes of this paper, this strategy can be used in different countries and contexts in which this disaggregation of data is available.

2.6 Appendix

2.6.1 Institutional background

According to the *Ley organica de Elecciones*, Law number 26859, the following articles define how voters are allocated to polling booths. Article 52 states that there must be between 200 and 300 voters per polling booth[19], and Article 53 states that they are allocated according to their order of registration in the district[20].

According the *Ley de elecciones Municipales* 26864, the apportionment of the regidores is as follows: Artículo 25^o Elección de Regidores del Concejo Municipal

[19] Artículo 52^o En cada distrito político de la Repblica se conforman tantas mesas de sufragio como grupos de 200 (doscientos) ciudadanos hábiles para votar como mínimo y 300 (trescientos) como máximo existan. El número de ciudadanos por mesa de sufragio es determinado por la Oficina Nacional de Procesos Electorales.

[20] Artículo 53^o Las mesas tienen un nmero que las identifica y las listas de electores por mesa se hacen sobre la base de los ciudadanos registrados en la circunscripción, en orden numérico.

Los Regidores de cada Concejo Municipal son elegidos por sufragio directo para un período de cuatro (4) años, en forma conjunta con la elección del Alcalde. La elección se sujeta a las siguientes reglas:
1. La votación es por lista.

2. A la lista ganadora se le asigna la cifra repartidora o la mitad más uno de los cargos de Regidores del Concejo Municipal lo que más le favorezca, según el orden de candidatos propuestos por las agrupaciones políticas. La asignación de cargos de Regidores se efectúa redondeando el número entero superior.

3. La cifra repartidora se aplica entre todas las demás listas participantes para establecer el número de Regidores que les corresponde.

4. El Jurado Nacional de Elecciones dentro de los quince (15) días siguientes a la vigencia de la presente Ley, aprobará las directivas que fuesen necesarias para la adecuada aplicación de lo dispuesto en el presente artículo.

Texto modificado por el Artículo 1^o de la Ley N^o 27734, publicada el 28-05-2002.
CONCORDANCIA:
Ley N^o 26859, Art. 29^o
R. N^o 1248-2006-JNE

2.6.2 Multinomial Probit

Having found an influence from the location in the voting ballot on the decision to not make a decision (blank vote), we also try to study whether the location can also affect the decision to finally mark a yes or no in the ballot. We do not expect an influence of location on preferences, but since the latter can be correlated with unobserved characteristics of the voters (different motivation to participate in this particular electoral process by political leaning, for example), there could be an influence into the actual voting. On the other hand, while this model (under its own assumptions) can also help to generate counterfactual scenarios of voting behavior under different configurations of the ballot, it has to be taken as a guidance of the empirical work rather than a literal explanation of the voting behavior.

We follow the discrete choice framework. In this formulation we consider a voter that has to choose between three alternatives for each candidate: vote in favor of the recall (yes), vote against (no), or express no preference (voting blank). We denote the voter as i, who is located in the district d (in the estimations we use district, polling center or polling booth, but here for simplicity we call it district generically). His or her choice for candidate c is denoted by j, where:

$$j = \begin{cases} 2 & \text{if the vote is } \texttt{YES} \\ 1 & \text{if the vote is } \texttt{NO} \\ 0 & \text{if the vote is } \texttt{BLANK} \end{cases}$$

We assume that the drivers behind this decision are the utilities that the voter enjoys from taking each decision, which we call U_{ijcd}. The indexes indicate that this is the utility that voter i, who votes in district d, enjoys for choosing alternative j for candidate c. This utility can be affected by several characteristics of the candidates, as well as the preferences of the voter. Some of the candidates characteristics are observed, like their location in the ballot (the order and/or the column) so we include them in vector x_c (we also add here a dummy variable for the major Villarán and candidate 31, Castañeda Jr., who were particularly salient during the whole process). Another set of utility drivers could be unobserved, for example, electoral preferences: left wingers (the majors political leaning), or voters that were distrustful of the main sponsor of the whole recall process (Castañeda, the father of the candidate 31 and former major), would probably obtain higher satisfaction from alternative `NO`, and conversely. To capture the influence of additional unobservables we also include two random drivers of the utility. The first is the random variable θ_d, which captures the average preferences in district d. The final term, ϵ_{ijc} is a zero mean shock that captures any idiosyncratic preference for option j for candidate c for voter i. Finally, we assume a linear shape, so that the utility takes the form:

$$U_{ijcd} = x_c\beta_j + \theta_{jd} + \epsilon_{ij} \tag{2.2}$$

for $j = \{0, 1, 2\}$. The voter i will pick option j if this option gives greater utility than any other. That is, the probability that voter i picks option j for candidate t is:

$$P_i(y = j) \quad = \quad Prob(U_{ijcd} \geq U_{ij'cd}, \forall j' \neq j).$$

Let $\epsilon_{ij} = (\epsilon_{i2}, \epsilon_{i1}, \epsilon_{i0})$ be the vector of idiosyncratic utility terms for the alternative of voting yes (2), no (1) or blank (0) of voter i. Let $\phi(\epsilon_2, \epsilon_1, \epsilon_0)$ be its multivariate normal distribution, so we can estimate a multinomial probit. Finally, let $V_{jcd} = x_c\beta_j + \theta_{jd}$ be the mean utility obtained from choice j, given that the ϵ terms has a zero mean. Hence

$$P_i(y = j) \quad = \quad \int I(V_{jcd} - V_{j'cd} \geq \epsilon_{ijc} - \epsilon_{ij'c}) dF(\epsilon_{ij}), \forall j' \neq j,$$

For simplicity, we write V_{jcd} for yes (2), no (1) and blank (0) as follows: v_2, v_1 and v_0, and similarly with the random variables ϵ_2, ϵ_1 and ϵ_0. Hence the probability of voting `YES` is

$$P_i(y=2) = \int_{\epsilon_2=-\infty}^{\infty} \int_{\epsilon_1=-\infty}^{v_2-v_1+\epsilon_2} \int_{\epsilon_0=-\infty}^{v_2-v_0+\epsilon_2} \phi(\epsilon_2, \epsilon_1, \epsilon_0) d\epsilon_2 d\epsilon_1 d\epsilon_0$$

Minimizing the observed `YES` and `NO` shares observed for each candidate in each table with the ones predicted according to the previous equations we identify estimates for the sets of parameters β and θ for `YES` and `NO`, with respect to blank. That is, we identify the terms in parenthesis in the equations below.

$$V_{2cd} - V_{0cd} = x_c(\beta_2 - \beta_0) + (\theta_{2d} - \theta_{0d}) \quad (2.3)$$

$$V_{1cd} - V_{0cd} = x_c(\beta_1 - \beta_0) + (\theta_{1d} - \theta_{0d}) \quad (2.4)$$

2.7 Tables and Figures

2.7.1 Figures

Figure 2.1: Ballot as seen by the voters

Figure 2.2: Ballot as published in a newspaper

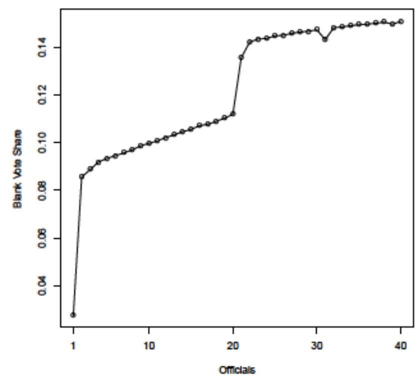

Figure 2.3: Share of blank votes

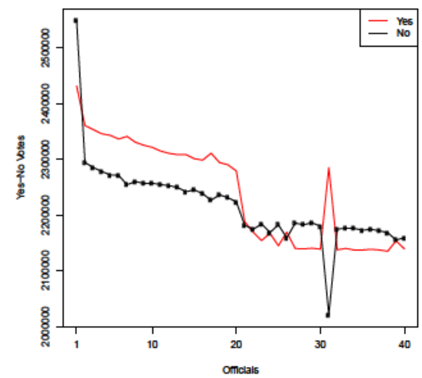

Figure 2.4: Yes (red) and No votes

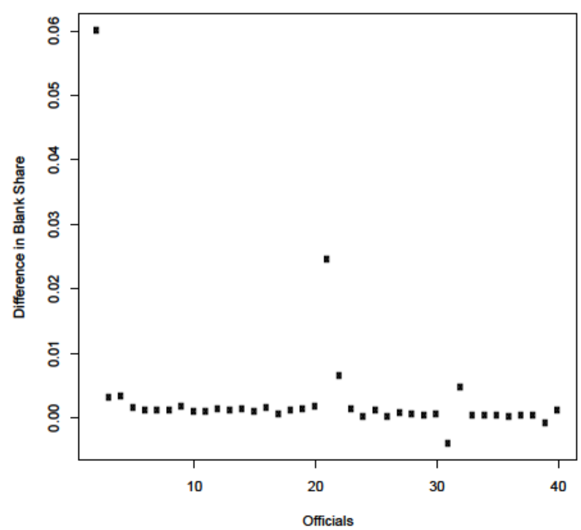

Figure 2.5: Difference in blank shares obtained by official t and t-1.

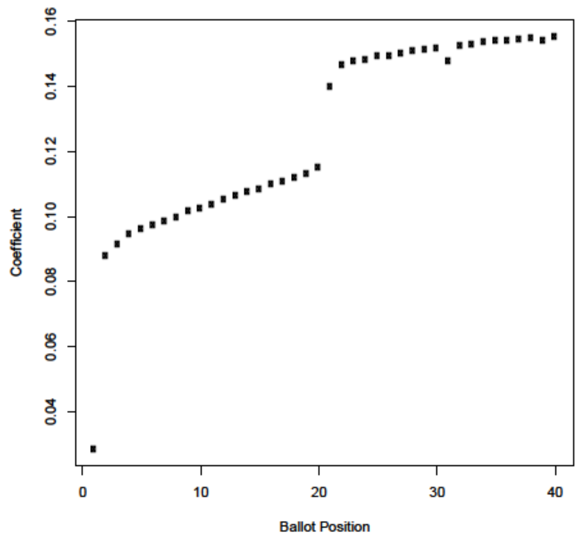

Figure 2.6: Coefficients of linear model using only dummies for positions in the ballot

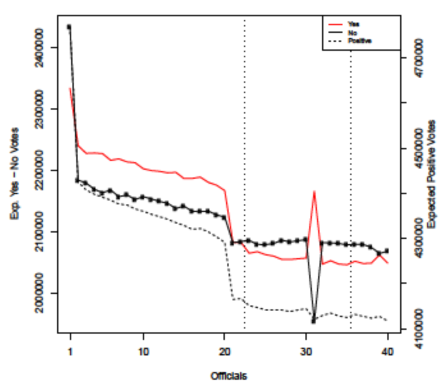

Figure 2.7: By-District Multinomial Probit Expected Votes

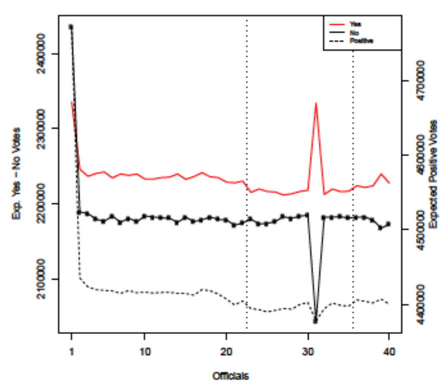

Figure 2.8: By-District Multinomial Probit, no column, no order effect

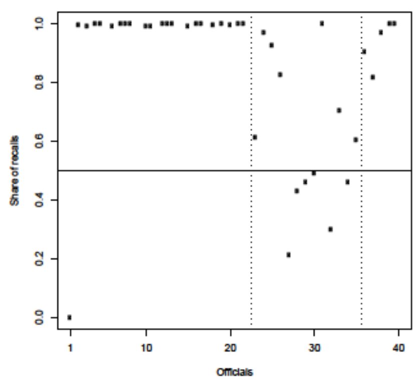

Figure 2.9: Simulation Recall Shares: two columns, random order across districts

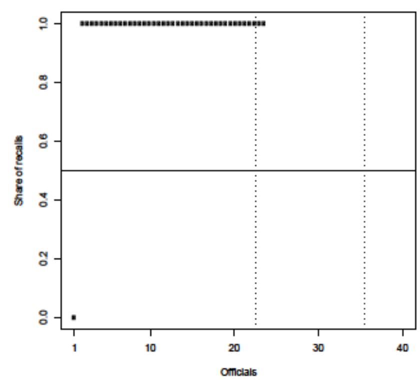

Figure 2.10: Simulation Recall Shares: one column, random order across districts

2.7.2 Tables

Eligible Voters	6295952	Booths	36386	
Turnout	5273790	Centers	888	
Turnout %	83.8	Districts	43	
Relative Sizes	Min	Max	Average	Sd
Voters per Booth	112	240	173	28
Voters per Center	1088	52460	7090	4467
Booths per Center	7	242	41	25
Voters per District	1298	631100	146417	134506
Booths per District	7	3529	846	770
Centers per District	1	79	21	18

All numbers are calculated after we dropped all observations for which all votes cast for an official at a specific booth where coded as "null."

Table 2.1: Description of 2013 Lima Referendum Data

Table 2.2: Individual Characteristics of Elected Officials

Official	Party	Age	Gender	National Level		Local Level		Education	Media
				Times Run	Elected?	Times Run	Times Elected	Level	Exposure
SUSANA VILLARAN	FS	65	Female	1	No	1	1	Bachelor (not completed)	1068
EDUARDO ARIEL ZEGARRA MENDEZ	FS	51	Male	1	No	1	1	PhD (Economics)	111
MARISA GLAVE REMY	FS	33	Female	0	N/A	2	2	Secondary	41
RAFAEL EDUARDO GARCIA MELGAR	FS	59	Male	0	N/A	2	2	Bachelor (not completed)	4
PERFECTO VICTOR RAMIREZ CIFUENTES	FS	65	Male	2	No	2	1	Secondary	5
ZOILA ELENA REATEGUI BARQUERO	FS	34	Female	0	N/A	1	1	Bachelor (not completed)	4
LUIS VALER CORONADO	FS	61	Male	3	0	4	3	Technical School	16
MARCO ANTONIO ZEVALLOS BUENO	FS	40	Male	0	N/A	1	1	Master (not completed)	21
SIGIFREDO MARCIAL VELASQUEZ RAMOS	FS	54	Male	1	No	2	1	Post-graduate specialization	1
VICTORIA DE SOTOMAYOR COTRADO	FS	26	Female	0	N/A	1	1	Bachelor (in progress)	3
LUISA MERCEDES MARTINEZ CORNEJO	FS	33	Female	0	N/A	1	1	Bachelor	1
DORA BEATRIZ HERNANDO SANCHEZ	FS	53	Female	0	N/A	1	1	Bachelor	2
INES CECILIA RODRIGUEZ VELASQUEZ	FS	56	Female	0	N/A	1	1	Bachelor	1
JOSE LIBORIO ESTEVES ROBLES	FS	60	Male	0	N/A	2	1	Bachelor	2
MONICA GISSELLA ERAZO TRUJILLO	FS	33	Female	0	N/A	1	1		2
MANUEL ABELARDO CARDENAS MUOZ	FS	52	Male	0	N/A	1	1		1
CAYO TITO QUILLAS	FS	54	Male	0	N/A	2	1	Post-graduate specialization	14
OLGA CELINDA MORAN ARAUJO	FS	50	Female	0	N/A	1	1	Bachelor	0
RONALD GONZALES PINEDA	FS	29	Male	0	N/A	1	1	Bachelor (not completed)	3
MAIA LIBERTAD ROJAS BRUCKMANN	FS	37	Female	0	N/A	1	1	Master	2
PEDRO JAVIER LOPEZ TORRES TUBBS	FS	36	Male	0	N/A	1	1	Master (in progress)	1
HERNAN NUEZ GONZALES	FS	29	Male	0	N/A	1	1	Bachelor (not completed)	1
WALTER ARCESIO GUILLEN CASTILLO	PPC	51	Male	0	N/A	3	3	Master	1
JAIME EDUARDO SALINAS LOPEZ TORRES	PPC	51	Male	3	No	4	1	Master	4
EDGARDO RENAN DE POMAR VIZCARRA	PPC	52	Male	1	No	3	1	Master	6
JOSE ALBERTO DANOS ORDOEZ	PPC	67	Male	0	N/A	1	1	Master	2
MONICA EMPERATRIZ SARAVIA SORIANO	PPC	45	Female	1	NO	1	1	Master	0
JORGE RAFAEL VILLENA LARREA	PPC	35	Male	0	N/A	1	1	Bachelor	4
LUZ MARIA DEL PILAR FREITAS ALVARADO	PPC	63	Female	1	NO	1	1	Master	0
PABLO ALBERTO SECADA ELGUERA	PPC	43	Male	0	N/A	1	1	Master	55
LUIS MANUEL CASTAEDA PARDO	PPC	29	Male	0	N/A	2	2	Bachelor	237
TERESA DE JESUS CANOVA SARANGO	PPC	54	Female	0	N/A	1	1	Post-graduate specialization	8
ALBERTO VALENZUELA SOTO	PPC	43	Male	0	N/A	2	1	Master	26
OSCAR JAVIER IBAEZ YAGUI	PPC	37	Male	0	N/A	1	1	Master	3
LUIS FELIPE CALVIMONTES BARRON	PPC	48	Male	0	N/A	1	1	Bachelor	7
RUBEN SANTIAGO GAVINO SANCHEZ	RN	54	Male	1	No	1	1	Bachelor	2
IVAN BECERRA HURTADO	RN	42	Male	0	N/A	1	1	Master	3
GERMAN RICARDO APARICIO LEMBCKE	SP	70	Male	0	N/A	5	5	Bachelor	3
FERNAN ROMANO ALTUVE-FEBRES LORES	CR	46	Male	2	Yes	2	1	Master	21
LUIS FELIPE CASTILLO OLIVA	SU	38	Male	0	N/A	1	1	Master	8

The “Run at” columns display how many times this official has run in local or national elections before the 2010 elections. The “Elected at National Level” column shows if an official had been elected at the national level before the 2010 elections, and the “Elected at Local Level” show how many times the official has been elected at the local level before the 2010 elections with N/A implying “Not Applicable”. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.3: Results of 2010 Municipal Election of Lima

Party	Min. Share	Max. Share	Av. Share	Sd. of Shares	Seats
Fuerza Social	0.202	0.463	0.353	0.07	22
Partido Popular Cristiano - Unidad Nacional	0.239	0.665	0.404	0.113	13
Restauración Nacional	0.039	0.183	0.081	0.026	2
Somos Perú	0.017	0.115	0.044	0.021	1
Cambio Radical	0.008	0.192	0.043	0.032	1
Siempre Unidos	0.005	0.138	0.027	0.033	1
Acción Popular	0.012	0.118	0.024	0.016	0
Alianza Para el Progreso	0.001	0.111	0.016	0.021	0
Partido Fonavisto del Perú	0.002	0.019	0.009	0.004	0

The second and third column refer to the maximum and minimum share a party got in a district. “Average Share” and “Sd. of share” give the average and standard deviation of the district shares. Seats gives us the number of seats the party got.

The first two “parties” are coalitions. Fuerza Social coalition includes: Fuerza Social, Movimiento Tierra y Libertad, Movimiento Nueva Izquierda, Movimiento Lima para Todos and independent candidates. PPC-Unidad Nacional coalition includes: PPC-Unidad Nacional, Solidaridad Nacional, Peru Posible, and independent candidates.

	Minimum	Maximum	Average	Median	Sd
Age	26	70	46.95	49	11.97
Times Candidate National	0	3	0.42	0	0.81
Times Candidate Local	1	5	1.57	1	0.98
Times Elected Local	1	5	1.27	1	0.78
Imputed Years of Education	11	20	15.47	15.5	1.94
Media Exposure	0	1068	42.35	3	171.3

The years of education have been imputed using the highest known educational level of the official, taking half of the duration if this educational level was not completed. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.4: Summary Statistics of Elected Officials

Party	Min. Share	Max. Share	Av. Share	Sd. of Share	Seats
Partido Popular Cristiano	0.155	0.555	0.316	0.097	7
Somos Perú	0.168	0.515	0.268	0.06	6
Peru Posible	0.04	0.19	0.099	0.033	2
Siempre Unidos	0.021	0.335	0.092	0.08	2
Acción Popular	0.068	0.298	0.102	0.037	2
Tierra y Dignidad	0.033	0.11	0.075	0.015	2
Partido Humanista Peruano	0.04	0.19	0.099	0.033	1

The second and third column refer to the maximum and minimum share a party got in a district. Average Share and Sd of share give the average and standard deviation of the district shares. Seats gives us the number of seats the party got.

Table 2.5: Results of 2013 Municipal By-Election of Lima

	Full Sample	Only Left	Only Right	Full Sample (by coalition)
College	13.9***	1.46		-2.70
	(3.24)	(3.37)		(2.35)
Age	-0.15	-0.14	-0.10	-0.050
	(0.12)	(0.14)	(0.13)	(0.083)
Gender	5.85	2.08	-1.31	-0.81
	(3.71)	(3.52)	(3.02)	(2.26)
Candidate in national elections	-0.21	-2.25	-2.56	-3.18*
	(3.26)	(2.82)	(1.93)	(1.88)
Candidate in local elections	-2.22	2.30	0.14	1.36
	(4.04)	(4.78)	(2.31)	(1.97)
Elected in national elections	15.3***			-3.58
	(5.09)			(2.60)
Elected in local elections	3.91	-4.68	-3.64	-3.51
	(3.64)	(5.17)	(1.91)	(2.11)
Media Exposure	-0.0045	-0.0055	0.0091	-0.0078***
	(0.0043)	(0.0045)	(0.014)	(0.0023)
Constant	12.1**	19.7***	40.1***	17.7***
	(5.11)	(6.36)	(7.36)	(4.24)
N	40	22	13	40
R^2	0.4727	0.4142	0.6192	0.3527
$\bar{R}^2$	0.3366	0.1213	0.2383	0.1856

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robust standard errors. College is a dummy variable that takes value 1 if the official had at least finished college. Gender is a dummy taking the value 1 if the official is a man. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.6: Estimation of *Order* against personal characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Order	0.0013***	0.0013***	0.0013***	0.0076***	0.0076***	0.0072***
	(0.000013)	(0.000013)	(0.000013)	(0.000057)	(0.000056)	(0.000059)
Column	0.044***	0.044***	0.042***	0.23***	0.24***	0.22***
	(0.000329)	(0.000329)	(0.000349)	(0.0015)	(0.0014)	(0.0015)
Column × Order	-0.00084***	-0.00084***	-0.00077***	-0.0054***	-0.0054***	-0.0048***
	(0.000012)	(0.000012)	(0.000013)	(0.000062)	(0.000062)	(0.000066)
Mayor	-0.061***	-0.061***	-0.031***	-0.57***	-0.58***	-0.38***
	(0.00037)	(0.00037)	(0.00099)	(0.0036)	(0.0035)	(0.0057)
Castañeda	-0.0043***	-0.0043***	0.0026***	-0.018***	-0.018***	0.025***
	(0.000126)	(0.000125)	(0.00022)	(0.00057)	(0.00058)	(0.0011)
Right Party	0.0041***	0.0041***	0.0044***	0.018***	0.018***	0.019***
	(0.000098)	(0.000097)	(0.00020)	(0.00040)	(0.00040)	(0.00046)
Minor Parties	0.0030***	0.0030***	-0.0051***	0.014***	0.014***	0.0032**
	(0.000127)	(0.000126)	(0.00039)	0.00055)	(0.00054)	(0.0014)
Party Share in 2010 elections			-0.031***			-0.036***
			(0.00133)			(0.0047)
College			-0.00071***			-0.0034***
			(0.000039)			(0.00020)
Age			0.0000076***			0.000034***
			(0.0000013)			(0.0000062)
Gender			0.00014***			0.00065**
			(0.000029)			(0.00014)
Candidate in national elections			-0.000023			-0.00049**
			(0.000021)			(0.00011)
Candidate in local elections			-0.000020			0.00053**
			(0.000029)			(0.00013)
Elected in national elections			-0.0015***			-0.0032***
			(0.000095)			(0.00041)
Elected in local elections			-0.00021***			-0.0012***
			(000032)			(0.00014)
Media Exposure			-0.000029***			-0.00019***
			(0.0000009)			(0.0000050)
Constant	0.089***	0.089***	0.10***	-1.35***	-1.33***	-1.31***
	(0.00069)	(0.00026)	(0.00048)	(0.0042)	(0.00093)	(0.0019)
FE	No	Center	Center	No	Center	Center
Method	LS	LS	LS	Probit	Probit	Probit
N	1451215	1451215	1451215	2902430	2902430	2902430
R^2	0.2354	0.3868	0.3882			
$\bar{R}^2$	0.2354	0.3864	0.3879			
Log-Likelihood	2237147.2	2397277.3	2398980.3	-76153626.0	-75668239.8	-75667520.4
ΔR^2	0.0282	0.0282	0.017			

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All regressions are clustered at the Center level. Minor Parties is a dummy that takes value 1 if the official belonged neither to the party of the Mayor neither to the right-wing coalition. College is a dummy variable that takes value 1 if the official had at least finished college. Gender is a dummy taking the value 1 if the official is a man. Party share in 2010 elections is the share that the party's official obtained in the 2010 elections in a district. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election. ΔR^2 shows the difference in R^2 between the full model and a restricted model without $Order$, $Column$ and $Column \times Order$.

Table 2.7: Estimations of blank share votes by table

Order	0.0040	College	0.00001
Column	0.0031	Age	0.00001
Column × Order	0.0004	Gender	0.00001
Mayor	0.0001	Candidate in national elections	0.00000
Castañeda	0.00001	Candidate in local elections	0.00000
Right Party	0.0001	Elected in national elections	0.00001
Minor Parties	0.0001	Elected in local elections	0.00000
Party Share in 2010 elections	0.0013	Media Exposure	0.00001

Table 2.8: Squares of Semi-partial correlation Coefficients

	No Share	No Share	No Share
Order	-0.0096***	-0.0096***	-0.0092***
	(0.000073)	(0.000072)	(0.000076)
Column	-0.29***	-0.29***	-0.28***
	(0.0019)	(0.0019)	(0.0020)
Column × Order	0.0071***	0.0071***	0.0063***
	(0.000082)	(0.000082)	(0.000087)
Mayor	0.76***	0.77***	0.51***
	(0.0047)	(0.0047)	(0.0077)
Castañeda	-0.036***	-0.035***	-0.094***
	(0.0013)	(0.0012)	(0.0022)
Right Party	-0.019**	-0.019***	-0.019***
	(0.00076)	(0.00074)	(0.00065)
Minor Parties	-0.017	-0.017***	0.016***
	(0.00083)	(0.00081)	(0.0024)
Party Share in 2010 elections			0.11***
			(0.0077)
College			0.0044***
			(0.00026)
Age			-0.000079***
			(0.0000083)
Gender			-0.0023***
			(0.00020)
Candidate in national elections			0.00051***
			(0.00014)
Candidate in local elections			-0.0013***
			(0.00017)
Elected in national elections			0.00035
			(0.00060)
Elected in local elections			0.0025***
			(0.00019)
Media Exposure			0.00025***
			(0.0000069)
Constant	1.16***	1.25***	1.20***
	(0.0074)	(0.011)	(0.011)
	Yes Share	**Yes Share**	**Yes Share**
Order	-0.0098***	-0.0098***	-0.0093***
	(0.000073)	(0.000072)	(0.000077)
Column	-0.31***	-0.31***	-0.29***
	(0.0019)	(0.0019)	(0.0020)
Column × Order	0.0067***	0.0067***	0.0060***
	(0.000083)	(0.000083)	(0.000090)
Mayor	0.69***	0.70***	0.46***
	(0.0043)	(0.0043)	(0.0073)
Castañeda	0.080***	0.082***	0.029***
	(0.0025)	(0.0026)	(0.0028)
Right Party	-0.027***	-0.027***	-0.030***
	(0.00077)	(0.00079)	(0.00080)
Minor Parties	-0.018***	-0.018***	-0.030***
	(0.00087)	(0.00088)	(0.0033)
Party Share in 2010 elections			-0.037***
			(0.010)
College			0.0041***
			(0.00027)
Age			-0.0000047
			(0.0000088)
Gender			0.00066***
			(0.00019)
Candidate in national elections			0.00075***
			(0.00015)
Candidate in local elections			-0.000068
			(0.00018)
Elected in national elections			0.0075***
			(0.00065)
Elected in local elections			0.00041**
			(0.00020)
Media Exposure			0.00023***
			(0.0000064)
Constant	1.19***	1.21***	1.22***
	(0.0054)	(0.010)	(0.011)
FE	No	District	District
N	4353645	4353645	4353645
Log-Likelihood	-199726345.0	-198379308.7	-198375258.8

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All regressions are clustered at the Center level. Minor Parties is a dummy that takes value 1 if the official belonged neither to the party of the Mayor neither to the right-wing coalition. College is a dummy variable that takes value 1 if the official had at least finished college. Gender is a dummy taking the value 1 if the official is a man. Party share in 2010 elections is the share that the party's official obtained in the 2010 elections in a district. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.9: Multinomial Probit estimations of structural model at the City Level

	Villa El Salvador	San Isidro
	No Share	**No Share**
Order	-0.011***	-0.0080***
	(0.00087)	(0.00039)
Column	-0.30***	-0.27***
	(0.015)	(0.013)
Column × Order	0.0072***	0.0071***
	(0.00090)	(0.00057)
Mayor	0.54***	0.31***
	(0.050)	(0.051)
Castañeda	-0.040***	-0.23***
	(0.0061)	(0.0096)
Right Party	-0.020***	-0.11**
	(0.0051)	(0.054)
Minor Parties	-0.0073	0.083 ***
	(0.0097)	(0.020)
Party Share in 2010 elections	0.021	0.35***
	(0.032)	(0.12)
College	0.0077***	-0.0015
	(0.0017)	(0.0013)
Age	-0.00014***	-0.000066**
	(0.000045)	(0.000028)
Gender	-0.00011	-0.0094***
	(0.00091)	(0.00066)
Candidate in national elections	0.0011*	-0.0010
	(0.00068)	(0.0011)
Candidate in local elections	-0.0013*	0.0010
	(0.00077)	(0.0014)
Elected in national elections	-0.0024	0.033***
	(0.0025)	(0.0040)
Elected in local elections	0.0019***	-0.00065
	(0.00068)	(0.0014)
Media Exposure	0.00013***	0.00062***
	(0.000032)	(0.000046)
Constant	1.05***	1.37***
	(0.022)	(0.022)
	Yes Share	**Yes Share**
Order	-0.011***	-0.0090***
	(0.00092)	(0.00039)
Column	-0.32***	-0.26***
	(0.016)	(0.012)
Column × Order	0.0074***	0.0043***
	(0.00097)	(0.00055)
Mayor	0.50***	0.31***
	(0.045)	(0.048)
Castañeda	0.0027	0.15***
	(0.0069)	(0.013)
Right Party	-0.028***	0.22***
	(0.0046)	(0.079)
Minor Parties	-0.025***	-0.25***
	(0.0088)	(0.031)
Party Share in 2010 elections	-0.023	-0.84 ***
	(0.031)	(0.18)
College	0.0072***	0.0014
	(0.0019)	(0.0018)
Age	-0.00016***	0.000046
	(0.000049)	(0.000040)
Gender	0.0016*	-0.0046***
	(0.00084)	(0.0012)
Candidate in national elections	0.00090	-0.000054
	(0.00071)	(0.00100)
Candidate in local elections	-0.00045	0.0053***
	(0.00091)	(0.0016)
Elected in national elections	0.0035*	0.0052
	(0.0020)	(0.0036)
Elected in local elections	0.0013	0.000083
	(0.00083)	(0.0017)
Media Exposure	0.00013***	0.00044***
	(0.000030)	(0.000038)
Constant	1.16***	1.43***
	(0.026)	(0.039)
N	181416	47757
Log-Likelihood	-9138692.5	-1879681.3

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
All regressions are clustered at the Center level. Minor Parties is a dummy that takes value 1 if the official belonged neither to the party of the Mayor neither to the right-wing coalition. College is a dummy variable that takes value 1 if the official had at least finished college. Gender is a dummy taking the value 1 if the official is a man. Party share in 2010 elections is the share that the party's official obtained in the 2010 elections in a district. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.10: Multinomial Probit estimations of structural model for Villa El Salvador and San Isidro

	(1)	(2)
Order	0.0013***	0.0013***
	(0.000047)	(0.0000083)
Column	0.042***	0.042***
	(0.0022)	(0.00028)
Column × Order	-0.00077***	-0.00077***
	(0.000059)	(0.000012)
Mayor	-0.031***	-0.031***
	(0.0095)	(0.00095)
Castaneda	0.0026	0.0027***
	(0.0021)	(0.00021)
Right Party	0.0044**	0.0044***
	(0.0020)	(0.000095)
Minor Parties	-0.0051**	-0.0051***
	(0.0023)	(0.00025)
Party Share in 2010 elections	-0.031***	-0.031***
	(0.0033)	(0.00083)
College	-0.00071	-0.00070***
	(0.00065)	(0.000037)
Age	0.0000076	0.0000071***
	(0.000013)	(0.0000013)
Gender	0.00014	0.00014***
	(0.00030)	(0.000026)
Candidate in national elections	-0.000023	-0.000019
	(0.00019)	(0.000021)
Candidate in local elections	-0.000020	-0.000029
	(0.00023)	(0.000026)
Elected in national elections	-0.0015**	-0.0016***
	(0.00064)	(0.000078)
Elected in local elections	-0.00021	-0.00020***
	(0.00025)	(0.000028)
Media Exposure	-0.000029***	-0.000030***
	(0.0000090)	(0.00000092)
Constant	0.12***	0.10***
	(0.0016)	(0.00030)
FE	Center	Booth
Clustering	Center× Order (multiway)	Booth
N	1451215	1451215
R^2	0.39	0.88

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
College is a dummy variable that takes value 1 if the official had at least finished college. Minor Parties is a dummy that takes value 1 if the official belonged neither to the party of the Mayor neither to the right-wing coalition. Gender is a dummy taking the value 1 if the official is a man. Party share in 2010 elections is the share that the party's official obtained in the 2010 elections in a district. Candidate variables count how many times the official was a candidate in the past and Elected count how many times an official was elected in the past. Media Exposure is the total number the official was mentioned in the leading Peruvian newspapers from 2010 until the day of the election.

Table 2.11: Blank Share of Votes: Additional Robustness Checks.

Bibliography

Alvarez RM, Sinclair B, Hasen RL (2006) How Much Is Enough? The" Ballot Order Effect" and the Use of Social Science Research in Election Law Disputes. Election Law Journal 5(1):40–56, URL http://online.liebertpub.com/doi/abs/10.1089/elj.2006.5.40

Augenblick N, Nicholson S (2015) Ballot Position , Choice Fatigue , and Voter Behavior. Review of Economic Studies pp 1–27, URL http://faculty.haas.berkeley.edu/ned/Choice{_}Fatigue.pdf

Banks JS, Duggan J (2005) Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-motivated Candidates. Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S Banks pp 15–56

Cameron A, Miller D (2015) A practitioner's guide to cluster-robust inference. Journal of Human Resources 50(2):317–372, URL http://jhr.uwpress.org/content/50/2/317.short

Cantú F (2013) Identifying Irregularities in Mexican Local Elections. American Journal of Political Science 58(4):1–31, DOI 10.1111/ajps.12097, URL http://doi.wiley.com/10.1111/ajps.12097

Casas A, Díaz G, Trindade A (2014) Who monitors the monitor?: effect of party observers on electoral outcomes URL http://e-archivo.uc3m.es/handle/10016/19455

Downs A (1957) An economic theory of political action in a democracy. The Journal of Political Economy 65(2):135–150

Enelow J, Hinich MJ (1981) A New Approach to Voter Uncertainty in the Downsian Spatial Model. American Journal of Political Science 25(3):483–493, DOI 10.2307/2110815, URL http://www.jstor.org/stable/10.2307/2110815

Esteve-Volart B, Bagues M (2012) Are women pawns in the political game? Evidence from elections to the Spanish Senate. Journal of Public Economics 96(3-4):387–399, DOI 10.1016/j.jpubeco.2011.12.004, URL http://www.sciencedirect.com/science/article/pii/S0047272711001848http://dx.doi.org/10.1016/j.jpubeco.2011.12.004

Feddersen TJ, Pesendorfer W (1996) The swing voter's curse. The American economic review 86(3):408–424, URL http://www.jstor.org/stable/10.2307/2118204

Feenberg D, Ganguli I, Gaule P, Gruber J (2015) It's Good to be First: Order Bias in Reading and Citing NBER Working Papers URL http://www.nber.org/papers/w21141

Glejser H, Heyndels B (2001) Inefficiency in the ranking in competitions: The case of the Queen Elisabeth Music Contest. Journal of Cultural Economics 25(2):109–129, DOI 10.1023/A:1007659804416, URL http://www.springerlink.com/index/K2W324444MJJ2J07.pdf

Greene W (2004) The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. The Econometrics Journal 7(1):98–119, DOI 10.1111/j.1368-423X.2004.00123.x, URL http://www.blackwell-synergy.com/links/doi/10.1111/j.1368-423X.2004.00123.x

Heckman J (1981) Statistical Models for Discrete Panel Data. In: McFadden D, Manski CF (eds) Structural Analysis of Discrete Data with Econometric Applications, vol 999, MIT Press, Cambridge, MA, p 477

Kamenica E (2014) Contextual Inference in Markets : On the Informational Content of Product Lines. American Economic Review 98(5):2127–2149, URL http://ideas.repec.org/a/aea/aecrev/v98y2008i5p2127-49.html

King G, Tomz M, Wittenberg J (2000) Making the most of statistical analyses: Improving interpretation and presentation. American journal of political science URL http://www.jstor.org/stable/2669316

Liu W, Simonson I (2005) Take It or Leave it?'-The Effect of Explicit Comparisons on Commitment to Purchase. Advances in Consumer Research 32(1):453, URL https://scholar.google.com/scholar?oi=gsb40{\&}q=Takeitorleaveit{\%}3F{\%}E2{\%}80{\%}99{\%}E2{\%}80{\%}94Theeffectofexplicitcomparisonsoncommitmenttopurchase{\&}lookup=0{\&}hl=en{\#}0

Meredith M, Salant Y (2013) On the Causes and Consequences of Ballot Order Effects. Political Behavior 35(1):175–197, DOI 10.1007/s11109-011-9189-2, URL http://link.springer.com/10.1007/s11109-011-9189-2

ONPE (2010) Consultas Populares de Revocatoria 1997-2009 y Nuevas Elecciones Municipales 2005-2010. Tech. rep., Oficina Nacional de Procesos Electorales

Peltzman S (1975) The Effects of Autombile Safety Regulation. The Journal of Political Economy 83(4):677–726, URL http://www.jstor.org/stable/1830396

Riker WH, Ordeshook PC (1968) A Theory of the Calculus of Voting. American Political Science Review 62(01):25–42, DOI 10.2307/1953324, URL http://www.jstor.org/stable/1953324

Rubinstein A, Salant Y (2006) A model of choice from lists. Theoretical Economics 1:3–17, URL http://tspace.library.utoronto.ca/handle/1807/4780

Wand JN, Shotts KW, Sekhon JS, Mebane WR, Herron MC, Brady HE (2001) The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. The American Political Science Review 95(4):793–810, URL http://www.jstor.org/stable/3117714$\backslash$nhttp://www.jstor.org/page/

Chapter 3

Costly voting under complete information

Pivotal voter models with costly voting started under complete information with the seminal contribution of Palfrey and Rosenthal (1983) (henceforth, PR), where two groups of individuals each preferring one of two alternatives simultaneously decide between abstaining or voting for their preferred alternative. The winner is decided by majority rule. Technical difficulties and multiplicity issues allowed them to analyze only special cases.[1] Two years later, the same authors proposed in Palfrey and Rosenthal (1985) to drop the assumption of complete information, and from then onwards the literature has almost exclusively focused on private information on the cost of voting. Some of the most relevant contributions which analyze private-information costly-voting pivotal voter models are Campbel (1999), Börgers (2004), Feddersen and Sandroni (2006), and Taylor and Yildirim (2010).[2]

[1]In particular, PR's setting is one of identical individual benefits from winning across groups, and they analyze: i) identical group sizes and symmetry of strategies across groups, and ii) aggregate probabilities of voting across individuals of different groups summing to 1. Besides these special cases of our analysis, they also analyze two different settings: one where there is a status quo (ties are broken in favor of one group, instead of randomly), and one, namely "k equilibria", where individuals of one group mix with identical probability, whereas among the individuals of the other group, k vote with probability 1 and the remaining with probability 0.

[2]Another way to exploit private information to simplify the analysis is provided in Myerson (1998a) (and also in Myerson (1998b) and Myerson (2000)). He suggests an alternative model in which the size of the electorate is a Poisson random variable. As Krishna and Morgan (2012) claim, Myerson's approach "has the important advantage of considerably simplifying the analysis of pivotal events". In Myerson's games, citizens' preferences are determined via a *privately* observed stochastic draw.

We believe the complete information setting also deserves attention, hence the goal of this paper is to complement the existing private information literature, and to move the PR analysis a step forward.

The literature has shown that assuming private information on the cost of voting typically has technical advantages. In particular the equilibrium is unique and the strategies are completely characterized by a single cut-off value: supporters of A (B) vote if and only if the cost of voting is lower than a threshold c_A (c_B).[3] As PR showed, uniqueness and cut-off strategies are not generally present in models of complete information. Nevertheless, we first show that, if the cost of voting is sufficiently high at least for the supporters of one of the two alternatives, the equilibrium is unique. We fully characterize it. If instead the cost of voting is sufficiently low for all individuals, we characterize three classes of equilibria, and show that any equilibrium must belong to one of these three classes, regardless of the number of individuals. Furthermore, we propose a novel equilibrium refinement that always singles out a unique equilibrium. This refinement says that the equilibrium probability of voting is continuous in the cost of voting. In fact, it would be hard to claim that negligible changes in the cost of voting could bring about drastic changes in the probabilities of voting. For example, if the voting center station moves slightly away from the home of an individual, her probability of voting also changes negligibly. The unique equilibrium pinned down by the continuity refinement is proved to belong always to the same class of equilibria out of the three classes previously characterized. The features of the equilibrium that we analyze are as follows. First of all, we find a turnout upper-bound: the sum of the equilibrium probabilities of voting of an m-individual and an n-individual is less than 1. Moreover, members of the majority group with higher cost-to-benefit ratio have higher probability of being pivotal in equilibrium, and thus –intuitively– vote with a higher probability in equilibrium. This carries over even if the two groups are asymmetric only in the cost-to-benefit ratio and symmetric in size. If instead the two groups are symmetric in cost-to-benefit ratio and asymmetric in size, members of the minority group vote with a strictly higher probability than those in the majority do. This latter result already exists in models of private information on the cost of voting, see Taylor and Yildirim (2010). The fact that members of the minority vote more than members of the majority is called the "underdog effect".[4] However, we find that if the minority

[3]Note that all papers cited in the main text of the previous paragraph assume private information and have only equilibria in cut-off strategies, but PR (83).

[4]Laboratory experiments confirm the underdog effect, see Levine, D., Palfrey, T., 2007. The paradox of voter participation? A laboratory study. Amer. Polit. Sci. Rev. 101 (1),

has a sufficiently higher cost-to-benefit ratio than the majority, the underdog effect disappears (i.e., the members of the minority group votes with lower probability than members of the majority group). This result is trivial per se (infinite cost-to-benefit ratio yields necessary 0 probability of voting) unless something is said about the asymmetry in cost-to-benefit needed to break the underdog effect: we show that if the ratio of cost-to-benefit ratios is greater than the number of members of the minority then the underdog effect is contradicted and the members of the minority group vote with 0 probability regardless of their (strictly positive) cost-to-benefit ratio. We furthermore perform comparative statics in the number of individuals, and give the intuition behind it, along with the construction of the equilibrium itself.

Our complete information setting springs from PR, however we depart from them in that an individual's benefit of having the favorite alternative win can be asymmetric according to whether the individual supports one or the other alternative, meaning that the personal benefit can differ between individuals supporting one alternative or the other. Our generalization of PR to asymmetric benefits is natural especially when the group sizes are asymmetric. For instance, think of an economics department consisting of several microeconomists and a few macroeconomists, all called to vote over who to hire between two job market candidates: a microeconomist and a macroeconomist.[5] Both types of economists are better-off if the newly hired candidate is of their same type. Furthermore, the benefit for a macroeconomist from hiring another macroeconomist is greater than the one for a microeconomist from hiring a microeconomist because of the asymmetric size of the two groups; that is, since macroeconomists are fewer, having another macroeconomist in the department sharply increases each macroeconomist's coauthoring possibilities, whereas the benefit for a microeconomist from having a new microeconomist in the department is lower because they are already plenty. In other words, the benefit is asymmetric across individuals of different groups. This asymmetry gives rise to asymmetric willingness to vote.[6]

1431 58.

[5]The cost of voting in this case is the opportunity cost of showing up to vote that day instead of, for example, being on vacation.

[6]Another example of asymmetric benefit across groups is the following. Residents of two neighborhoods are called to vote over the location of a new school in one of the two neighborhoods. In neighborhood 1 there is already a school, in neighborhood 2 there is none: thus, despite the fact that each resident strictly prefers the school to be located in her neighborhood, residents in neighborhood 1 care less than residents in neighborhood 2 about the location of the school since there is already a school in neighborhood 1.

3.1 Model

Consider a complete information setting where there are two groups of individuals of size m and n, with $m, n \in \mathbb{N}^+$. Throughout the paper we assume $m > n > 1$: the analysis of $n = 1$ is ruled out to avoid dealing with trivial cases, and the analysis of $m = n$ produces peculiar results and is left to Appendix C. We use sub-index $i \in \{m, n\}$ to identify the group with a slight abuse of notation. The individuals are called to cast a vote between two alternatives, M and N. An individual of group m prefers alternative M, and an individual of group n prefers alternative N. That is, if M wins, the payoff of an individual m increases by $\Delta\pi_m \in \mathbb{R}_{++}$, and similarly by $\Delta\pi_n \in \mathbb{R}_{++}$ for individuals n if N wins. Individuals choose whether to vote for their preferred alternative or to abstain, since voting for the non-preferred alternative is strictly dominated. If an individual casts a vote, she faces a group-specific cost of voting, $c_i \in \mathbb{R}_{++}$. Thus, the increase in payoff –net of cost of voting– for an individual i when her preferred policy wins is $\Delta\pi_i - c_i$ if she voted, and $\Delta\pi_i$ if she did not vote. Individuals vote simultaneously and the winning alternative is decided by majority rule. Ties are broken by a fair coin toss.

Each individual i chooses her probability of voting, denoted by p_i, that maximizes her expected payoff, given the choices of all other individuals. We consider Quasi-Symmetric Nash Equilibria (QSNE), that is, individuals of group i follow the same equilibrium strategy p_i^*. Besides being used in PR, the assumption of QSNE has been used in private-information pivotal-voter models to obtain that individuals adopt cut-off strategies. See for instance Börgers (2004) and Taylor and Yildirim (2010).

In a QSNE a pair (p_i^*, p_j^*) is an equilibrium if an individual of group $i \in \{m, n\}$ would not want to deviate from p_i^* if she expects every other individual of group i to also play p_i^* and all individuals of group j with $j \neq i$ to play p_j^*. A QSNE can be of one of the following three types:

1. "Pure-Pure": $(p_m^*, p_n^*) \in \{0, 1\}^2$
2. "Pure-Mix": $p_m^* \in \{0, 1\}, p_n^* \in (0, 1)$ or $p_m^* \in (0, 1), p_n^* \in \{0, 1\}$
3. "Mix-Mix": $(p_m^*, p_n^*) \in (0, 1)^2$.

Define A_i to be the probability that the vote of an individual of group i is pivotal. An individual of group i will cast a vote if:

$$A_i \Delta\pi_i \geq c_i$$

or

$$A_i \geq \frac{c_i}{\Delta\pi_i} \equiv B_i \tag{3.1}$$

for $i \in \{m, n\}$. The probabilities of being pivotal are defined as follows:[7]

$$A_i = \sum_{s=0}^{n} \binom{i-1}{s} \binom{j}{s} p_i^s (1-p_i)^{i-s-1} p_j^s (1-p_j)^{j-s} + \sum_{s=0}^{n-1} \binom{i-1}{s} \binom{j}{s+1} p_i^s (1-p_i)^{i-s-1} p_j^{s+1} (1-p_j)^{j-s-1} \tag{3.2}$$

for $i, j \in \{m, n\}$, $i \neq j$.

We now explain how the expression (3.2) is constructed. A single individual m, who computes her probability of being pivotal, takes as given the probabilities of voting (p_m, p_n) of all other individuals. The individual m is pivotal when her vote either breaks a tie, or when it creates one. In (3.2) the first summation is her probability of breaking a tie, and the second of creating a tie. She can break a tie with her vote, if the number of individuals that vote for m equals the number of individuals that vote for n. Call this number s. Out of $m-1$ other m-individuals, exactly s vote with probability $\binom{m-1}{s} p_m^s (1-p_m)^{m-s-1}$. On the other hand, out of n n-individuals, exactly s vote with probability $\binom{n}{s} p_n^s (1-p_n)^{n-s}$. The second summation of (3.2) is similarly constructed: individual m can create a tie with her vote, if the number of individuals that vote for m, (which is again called s), is one less than the number of individuals that vote for n.

3.2 Computing the equilibria

In this section we compute the QSNE of the voting game, given the relative costs of voting B_i and B_j, and classify them according to their type: "Pure-Pure", "Pure-Mix" or "Mix-Mix". We start with the first two types, ie. equilibria in which some individuals have pure strategies.

3.2.1 "Pure-Pure" and "Pure-Mix" equilibria

If $A_i < B_i$ or $A_i > B_i$, then individual i's dominant strategy is to abstain or to vote, respectively (i.e. pure strategy). Whereas when $A_i = B_i$, i is indifferent between voting and not (i.e. mixed strategy). B_i is therefore the minimum probability of being pivotal such that an individual i will vote.

[7] Note that they already take into account that within groups individuals vote with the same probability.

For that reason, an individual i whose B_i is greater than 1 does not vote in equilibrium.[8] We formalize this result in the following lemma.

Lemma 1. *If* $B_i \geq 1$ *then* $p_i^* = 0$.

Proof. Let $p_i^* > 0$. First, if $B_i > 1$, by (3.1) we have $A_i > 1$, which is a contradiction, since A_i is a probability. Second, if $B_i = 1$, by (3.1) we have $A_i = 1$. Then, we cannot have $p_i^* \in (0,1)$ because of the following: if all individuals are randomizing between voting and not, then they cannot be pivotal with certainty, so $A_i \neq 1$ leading to a contradiction. Therefore we only need to rule out $B_i = p_i^* = 1$. For this we need to distinguish the following cases.

Case 1. If $p_m^* = 1$ all m individuals vote, which means that they win regardless of p_n^* because they are the largest group. But then, no n individual would want to face the cost of voting, $p_n^* = 0$. Thus, for a single m-individual deviation to $p_m = 0$ is profitable, leading to a contradiction.

Case 2. If $p_n^* = 1$ then in order to sustain $A_n = 1$ the n individuals must be certain that either n or $n-1$ of the m individuals vote. However, this can happen neither if $p_m^* \in (0,1)$ nor if $p_m^* = \{0,1\}$ (because this would imply m or 0 votes cast by group m). Thus, we reach a contradiction. □

The previous lemma shows that if relative costs of voting are high enough for individuals of both groups, the only equilibrium that exists is the "Pure-Pure" one in which nobody votes.

Obviously, the situation described in Lemma 1 is not very interesting. Therefore, next we allow one of the two relative costs of voting to be low enough such that individuals from one of the two groups might consider voting, in other words $B_i \geq 1$ only for individuals i. Then Lemma 2 yields a simple and unique characterization of j's equilibrium strategy:

Lemma 2. *("Pure-Mix") For* $B_i \geq 1$ *and* $B_j \in (0,1)$ *the unique QSNE is that* $p_i^* = 0$ *and* $p_j^* = 1 - B_j^{\frac{1}{j-1}}$*, for all* $i, j \in \{m, n\}$*,* $i \neq j$*.*[9]

Proof. By Lemma 1, $p_i^* = 0$. Suppose $p_j^* = 0$. Then any single j individual would have an incentive to deviate and vote for sure in order to single-handedly decide the election in favor of the j-group. Thus $p_j = 0$ is not an

[8] If the cost of voting for i is greater than the benefit of winning with certainty of being pivotal, in the unique equilibrium no member of i will vote (that is, $c_i > \Delta\pi_i$ contradicts $A_i \in [0,1]$).

[9] As a reminder, in the beginning of this section we have assumed $m > n > 1$. It is easy to see that for $j = 1$ the unique equilibrium is: $p_i^* = 0$ and $p_j^* = 1$

equilibrium. On the other hand, suppose $p_j^* = 1$. This means that j group wins for sure with a margin of j votes. Then any single j-individual would have an incentive not to pay the cost without affecting the outcome. Thus $p_j = 1$ is not an equilibrium. Therefore $p_j^* \in (0, 1)$. Plugging $p_i^* = 0$ in A_j we have $A_j = (1 - p_j)^{j-1}$, and since (3.1) must hold with equality for individuals j to mix, we have:

$$(1 - p_j)^{j-1} = B_j$$

or equivalently:

$$p_j = 1 - B_j^{\frac{1}{j-1}}.$$

□

For $x \in (0, 1)$ the expression $1 - x^{\frac{1}{j-1}}$ is strictly decreasing in x. Therefore $p_j^* = 1 - B_j^{\frac{1}{j-1}}$ is strictly decreasing in c_j and strictly increasing in $\Delta\pi_j$. Higher individual cost-payoff ratio results in j-individuals voting with lower probability.

The two previous Lemmas examine cases in which individuals from at least one group find it too costly to vote, not matter what individuals from the other group are doing. These two cases gave rise to two types of equilibria "Pure-Pure" in which everybody's strictly dominant strategy is to not vote (Lemma 1), and "Pure-Mix" in which individuals from one group have a strictly dominant strategy to not vote and individuals from the other group mix (Lemma 2).

We examine next what happens when individuals from neither group have a strictly dominant strategy to abstain, ie. what happens when $B_m < 1$ and $B_n < 1$. Under these conditions individuals of both groups may vote with positive probability. This causes strategic interactions that may generate multiple equilibria.

It is easy to see that when $B_m < 1$ and $B_n < 1$ no "Pure-Pure" equilibria exist.

Lemma 3. *For $B_i < 1$ and $B_j < 1$, no "Pure-Pure" QSNE exist, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Assume $p_i^* = 0$. Then $p_j^* = 0$ cannot be a QSNE because $A_j = 1$ so deviation to voting for a j individual would be profitable. Also $p_j = 1$ cannot be a QSNE because a j-individual deviating to $p_j = 0$ would not affect the outcome of the election and save her cost of voting.

The only case left to analyze is $p_m^* = p_n^* = 1$. The n individuals lose for sure, and thus they would be better-off not to vote. $\square$

After proving that for $B_m < 1$ and $B_n < 1$ no "Pure-Pure" equilibria exist, the next lemma establishes that for $B_m < 1$ and $B_n < 1$ a "Pure-Mix" equilibrium does exist. First, we need to define $\underline{B_i} = jB_j - (j-1)B_j^{\frac{j}{j-1}}$

Lemma 4. *For $B_i < 1$ and $B_j < 1$, there exists a "Pure-Mix" QSNE with $p_i^* = 0$ and $p_j^* = 1 - B_j^{\frac{1}{j-1}}$, for all $i, j \in \{m, n\}$ and $i \neq j$ if and only if $B_i \geq \underline{B_i}$.*

Proof. An equilibrium where $p_i^* = 0$ implies: $A_j = (1-p_j)^{j-1}$ and $A_i = (1-p_j)^j + jp_j(1-p_j)^{j-1}$. The former means that an individual j is pivotal only if none of her groupmates vote (her vote breaks the tie in which nobody votes). The latter means that an individual i is pivotal if none of j individuals vote or if only one of them votes. In order for the i-individuals to not want to vote we must have:

$$A_i \leq B_i,$$

or equivalently

$$(1-p_j)^j + jp_j(1-p_j)^{j-1} \leq B_i,$$

and similarly, for the j-group individual to mix we must have:

$$(1-p_j)^{j-1} = B_j, \tag{3.3}$$

dividing the two conditions and rearranging we get:

$$\begin{aligned} 1 - p_j + jp_j &\leq \frac{B_i}{B_j} \\ (j-1)p_j &\leq \frac{B_i}{B_j} - 1 \end{aligned} \tag{3.4}$$

Isolate p_j in (3.3) and plug it in (3.4) to get

$$(j-1)\left(1 - B_j^{\frac{1}{j-1}}\right) \leq \frac{B_i}{B_j} - 1 \tag{3.5}$$

Or equivalently,

$$\begin{aligned} -B_j^{\frac{j}{j-1}} &\leq \frac{B_i - jB_j}{j-1} \\ B_i &\geq jB_j - (j-1)B_j^{\frac{j}{j-1}} \equiv \underline{B_i} \end{aligned} \tag{3.6}$$

$\square$

$\underline{B_i}$ is an increasing bijection from $[0,1]$ to $[0,1]$, such that if $B_j = 0$, $\underline{B_i} = 0$, and if $B_j = 1$, $\underline{B_i} = 1$.

Note that the equilibria pinned down by Lemma 2 and Lemma 4 are essentially the same, the difference being that Lemma 2 provides the range of B_i's for which the equilibrium is unique, and Lemma 4 provides the range of B_i's for which that equilibrium continues to exist although not necessarily uniquely. This is an important finding that we will discuss further in the next section where we analyze our continuity refinement.

Lemma 4 is silent with respect to which of the two groups will be mixing and which will not be voting. What it says is that if $B_m < 1$ and $B_n < 1$ it can be either that the m individuals do not vote and the n individuals mix, or that the n individuals do not vote and the m individuals mix. The next lemma shows that for a given pair (B_i, B_j) these two "Pure-Mix" equilibria of Lemma 4 cannot co-exist.[10]

Lemma 5. *For $B_i < 1$ and $B_j < 1$, $B_i \geq \underline{B_i}$ and $B_j \geq \underline{B_j}$ are mutually exclusive, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Suppose not and consider the $(B_i, B_j)-$space. We first show that $\underline{B_i} > B_j$, or equivalently:

$$\begin{aligned} (j-1)B_j &> (j-1)B_j^{\frac{j}{j-1}} \\ 1 &> B_j^{\frac{1}{j-1}}. \end{aligned}$$

For the same reason we also have $\underline{B_j} > B_i$. Then, $B_i \geq \underline{B_i}$ and $\underline{B_i} > B_j$ imply $B_i > B_j$, while $B_j \geq \underline{B_j}$ and $\underline{B_j} > B_i$ imply $B_j > B_i$ leading to a contradiction. □

3.2.2 "Mix-Mix" equilibria

Lemmas 1 to 5 completely characterized the "Pure-Pure" and "Pure-Mix" equilibria of our voting game. We are left to analyze the "Mix-Mix" equilibria. Obviously, in any "Mix-Mix" equilibrium the voting conditions (3.1) for the two groups hold with equality. That is, the best reply under mixing for the m individuals is defined by:

$$A_m = B_m \tag{3.7}$$

[10] Meaning that for a given pair (B_i, B_j) we either have m individuals not voting and n mixing *or* n individuals not voting and m mixing (but not both).

and the best reply under mixing for the n individuals is defined by:

$$A_n = B_n \tag{3.8}$$

Since we have imposed the condition that all individuals within a group employ the same strategy, it suffices to focus on the mixing condition of a single m and on that of a a single n individual and analyze the intersections between these two in the (p_m, p_n)-space. These intersections are "Mix-Mix" equilibrium pairs (p^*_m, p^*_n), which are what we are after in this Subsection.

In contrast with the "Pure-Pure" and "Pure-Mix" cases, the "Mix-Mix" case entails solving a system of two polynomial equations of arbitrary power –expressions (3.7) and (3.8)–, and thus there is no general algebraic solution for equilibrium strategies (by the Abel-Ruffini theorem which states that there is no general algebraic solution to polynomial equations of degree five or higher with arbitrary coefficients). Instead, in order to analyze them we will use a number of indirect results about the space these equilibria lie on. For this it is useful to distinguish among the three cases: $B_m = B_n$, $B_m > B_n$, and $B_m < B_n$. Thus, by (3.7) and (3.8), these translate into $A_m = A_n$, $A_m > A_n$, and $A_m < A_n$. Analyzing $A_m = A_n$ will greatly help the analysis of the other two cases.

The set of points for which $A_m = A_n$ is depicted by the two black lines of Figure 3.1; the increasing and the decreasing one. These two black lines divide the (p_m, p_n)-space in four regions depending on the ranking of A_m and A_n. Keep in mind that the two mixing conditions are defined in the same space; dividing the space in these four regions will help us analyze how the intersections of the two mixing conditions behave.[11]

In Appendix A we characterize the set $A_m = A_n$, through a series of lemmas. First, we find the four points where these two lines touch the edges of the (p_m, p_n)-space (Lemma 6). Second, we characterize the decreasing line connecting the top-left corner with the bottom-right corner (Lemma 7). Finally we characterize the increasing line (Lemmas 8 and 9).[12] We summarize this series of lemmas in Proposition 6.

[11] At this stage it is interesting to compare our analysis with the one of PR. In particular in Section 6 of PR, they discuss "totally quasi-symmetric equilibria", which are what we call "Mix-Mix" equilibria. However they analyze two special cases, which in our notation are: i) $p_m = p_n$ and $m = n$ and ii) $p_m + p_n = 1$. In terms of our Figure 3.1 it means that they analyze equilibria that might arise along the two diagonals (in the case of the 45-degree line, they also assume $m = n$).

[12] Note that, as it will be explained in more detail later, the fact that the increasing line is in fact increasing is not needed. What is only needed is that it crosses the decreasing line once.

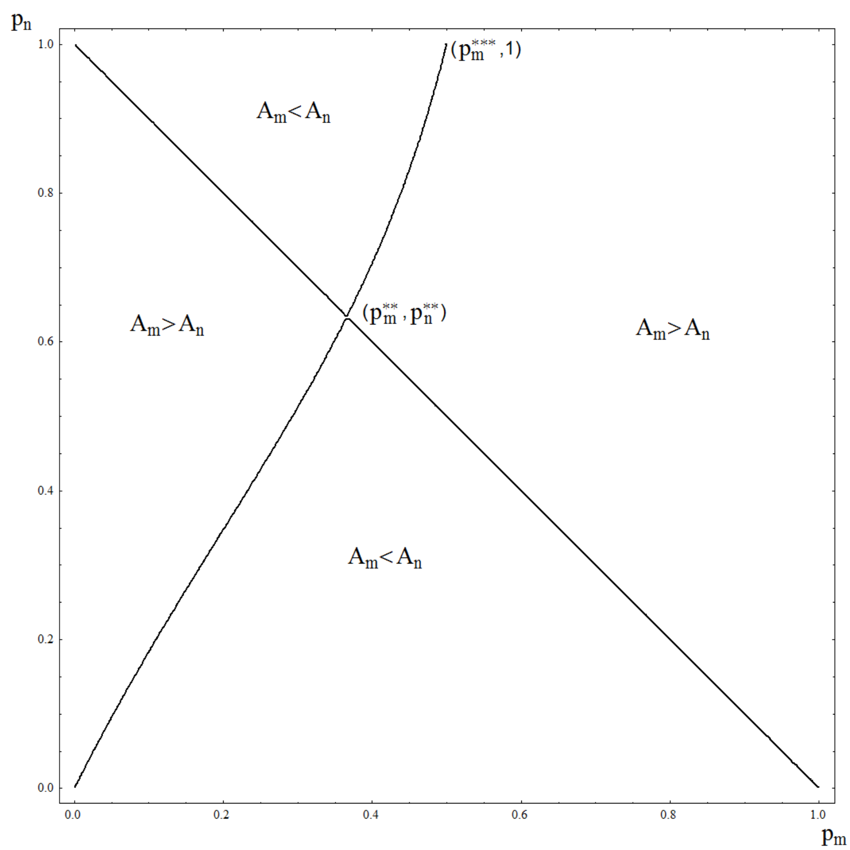

Figure 3.1: Set of points in the (p_m, p_n)-space according to whether $A_m \gtreqless A_n$ when $m = 3$ and $n = 2$.

Proposition 6. *The only points* $(p_m, p_n) \in (0,1)^2$ *satisfying condition*

$$A_m = A_n \tag{3.9}$$

are the points along the line $p_m + p_n = 1$ *and along a continuous line that goes from* $(0,0)$ *to* $(p_m^{***}, 1)$ *where* $p_m^{***} = \frac{n(n-1)}{n(n-1)+\sqrt{n(n-1)(m-n)(m-n+1)}}$.

Proof. See Appendix A. □

The two lines that we have characterized divide the (p_m, p_n)-space in four regions: two where $A_m > A_n$ and two where $A_m < A_n$. Since we are examining "Mix-Mix" equilibria, these two inequalities directly translate into conditions on B_m and B_n that must be satisfied in "Mix-Mix" equilibria (see (3.7) and (3.8)). Thus when $B_m > B_n$ all equilibria lie in one of the two regions where $A_m > A_n$, and when $B_m < B_n$ all equilibria lie in one of the two regions where $A_m < A_n$.

What we prove next (see Appendix B) is that the two mixing conditions cross at most once in each of the four regions delimited by the set of points such that $A_m = A_n$ (see Figure 3.1). This implies that we always have at most two "Mix-Mix" equilibria, one where $p_m^* + p_n^* < 1$ which we name "Mix-Mix 1", and one where $p_m^* + p_n^* > 1$ which we name "Mix-Mix 2". Including the "Pure-Mix" equilibrium (Lemma 4 characterizes it and Lemma 5 proves it is unique), this shows that we have at most three equilibria. Hence, we prove the following:

Theorem 4. *If* $B_i > \underline{B_i}$ *for individuals of both groups* $i \in \{m, n\}$*, there exists a unique QSNE such that* $p_m^* = p_n^* = 0$.

If $B_i > \underline{B_i}$ *and* $B_j \leq \underline{B_j}$ *with* $i \in \{m, n\}$ *and* $i \neq j$ *there exists a unique QSNE such that* $p_i^* = 0$ *and* $p_j^* = 1 - B_j^{\frac{1}{j-1}}$.

If $B_i \leq \underline{B_i}$ *for all* $i \in \{m, n\}$*, there are at most three QSNE, one "Pure-Mix" and two "Mix-Mix" ones.*

Proof. The first statement follows from Lemma 1. The second statement follows from Lemmas 2 to 5. The third statement is proved in Appendix B. □

Theorem 4 concludes the first part of the paper, establishing that there are at most three equilibria. Returning to PR, the analysis that they carry out for two special cases makes them "[...] conjecture that the class of all totally quasi-symmetric equilibria is much larger than those we have been able to investigate."[13] However, we showed that this class of equilibria – which we call "Mix-Mix"– admits, at most, two equilibria.

What we do in the next Section is propose a refinement and show that this refinement always pins down a unique equilibrium. We first deliver the intuition on how the continuity refinement works by means of a numerical example.

3.3 Continuous Refinement and Uniqueness

In this section we consider the simplest non-trivial example: $m = 3$ and $n = 2$. We compute and depict the three equilibria and give the intuition how the continuity refinement pins down a unique equilibrium. All qualitative features of this numerical example hold for any m and n.

For the sake of the numerical example, we fix $\Delta\pi_m = \frac{n}{m}\Delta\pi_n$. This parametrization complies with the application we discuss in Section 5 but it also has a simple interpretation: the individual gain of winning of an individual belonging to the big group m, is smaller than the individual gain of an individual of group n. Moreover, how smaller is governed by the ratio of the two group sizes. This implies that we can write $B_n = B$ and $B_m = \frac{m}{n}B$, and thus we have only one parameter B simplifying the exposition greatly.

For any B we examine all the types of equilibria: "Pure-Pure", "Mix-Mix", and "Pure-Mix". We find that for low B there are three different types of equilibria, which are depicted in the first row of Figure 3.2, while for larger B we only have the equilibrium depicted in the second row. In fact the unique equilibrium for B sufficiently large corresponds to the characterization in Lemmas 1 to 5.

Now, take for example $B = 1/3$ and consider Figure 3.2. There are three different equilibria we can have: two equilibria where both types of individuals are mixing ("Mix-Mix 1" and "Mix-Mix 2") and one where the n individuals vote for sure and the m individuals mix. This last equilibrium is of the form "Pure-Mix". However, it involves individuals from one group voting with certainty, and thus it is in sharp contrast with the equilibrium characterized in Lemmas 1 to 5, where individuals from one group do not vote. Since this equilibrium will be ruled out by our continuity refinement, throughout the rest of the paper we refer to "Pure-Mix" as the one characterized in Lemmas 1 to 5 (with $p_i^* = 0$). These three equilibria correspond to Theorem 4. On the other hand for larger (but not too large) B, say $B \in (2/3, 1)$, there is only one type of equilibrium, the one where the n

[13] PR, page 10.

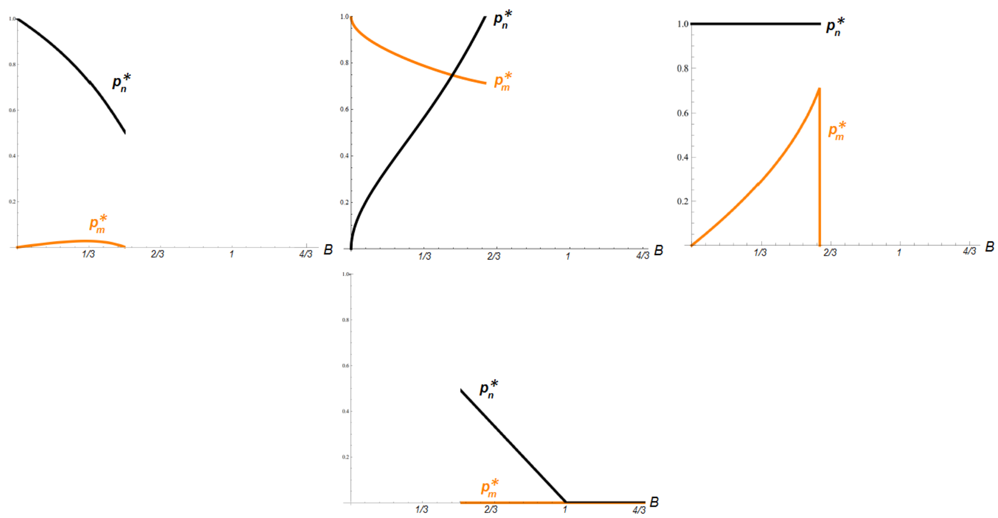

Figure 3.2: First row, first panel: "Mix-Mix 1". First row, second panel: "Mix-Mix 2". Second row: "High B" equilibrium.

individuals are mixing, and the m individuals are abstaining for sure (note that in this case $B_m < 1 < B_n$). And naturally, for $B \geq 1$, no-one votes. These two last cases are the "High B" equilibrium in Figure 3.2 which is composed of the "Pure-Pure" equilibrium of Lemma 1 and the "Pure-Mix" equilibrium of Lemmas 1 to 5.

Starting from sufficiently high B we have to be in the "High B" equilibrium. As we keep decreasing B, at some point we need to switch to one (or more) equilibria from the first row of Figure 3.2. However the only equilibrium out of the three that involves no jumps in the probabilities of voting, is the "Mix-Mix 1" equilibrium. We show that this continuous equilibrium (depicted in Figure 3.2) exists and is unique for all (m, n) and for all (B_i, B_j).

The "Mix-Mix 1" and "Mix-Mix 2" equilibria can also be seen in the mixing condition graphs depicted in Figure 3.4 in the (p_m, p_n)-space. The red (blue) lines represent the mixing condition for a n (m) individual for four values of B, $\{0.166, 0.333, 0.5, 0.61\}$. In particular, we chose the third value to be the minimum B such that the "Mix-Mix 1" equilibrium disappears (bottom-left panel), and the fourth value to be the minimum B such that even the "Mix-Mix 2" equilibrium disappears. The black lines in Figure 3.4

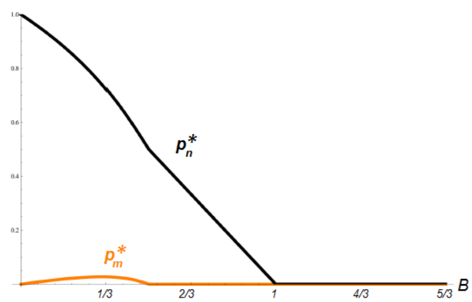

Figure 3.3: Unique continuous equilibrium

are the set of (p_m, p_n) satisfying $A_m = A_n$ as in Figure 3.1.

In each panel, the equilibrium on the left side of the panel is the "Mix-Mix 1" equilibrium. It converges to the "Pure-Mix" equilibrium $(p_m^*, p_n^*) = (0, 0.5)$ as $B \to 0.5$. The equilibrium that lays on the right side of the panel is the "Mix-Mix 2" equilibrium.

We use a standard definition of continuity (see for example page 943 of Mas-Colell et al (1995)). This definition means that, for all $i, j \in \{m, n\}$, $i \neq j$, there is a single continuous selection of the equilibrium correspondences $p_i^*(B_i, B_j)$ mapping (B_i, B_j) to equilibrium probabilities of voting.

Which of the three equilibria together with the "High B" unique equilibrium satisfies continuity? First, notice that the three equilibria need that $B_m, B_n < 1$. This means that we can focus on the part of the "High B" equilibrium that is defined for $B_m, B_n < 1$ as well. Hence, we need to find among the three equilibria the one for which p_i^* goes to 0 and at the same time, p_j^* is interior.[14] For this to happen we need that $p_m^* + p_n^* < 1$, which contradicts the third equilibrium in Figure 3.2 because this equilibrium requires that one of the two p_i^*'s is equal to one. Out of the two "Mix-Mix" equilibria, "Mix-Mix 2" contradicts $p_m^* + p_n^* < 1$.

Therefore we are left with only one candidate, namely "Mix-Mix 1", that together with "High B" equilibrium, might satisfy continuity. First notice that Lemma 4 gives us for every $B_j \in (0, 1)$ the lowest B_i for existence of the "Pure-Mix" equilibrium $p_i^* = 0$ and $p_j^* = 1 - B_j^{\frac{1}{j-1}}$. This satisfies the system $A_i \leq B_i$ and $A_j = B_j$, which implies that the system $A_i = B_i$ and $A_j = B_j$ is also satisfied (A_i's are continuous in p_i's).

[14]In terms of Figure 3.1 and Figure 3.4, equilibria must converge to the horizontal or vertical axis. This happens in the third panel of Figure 3.4.

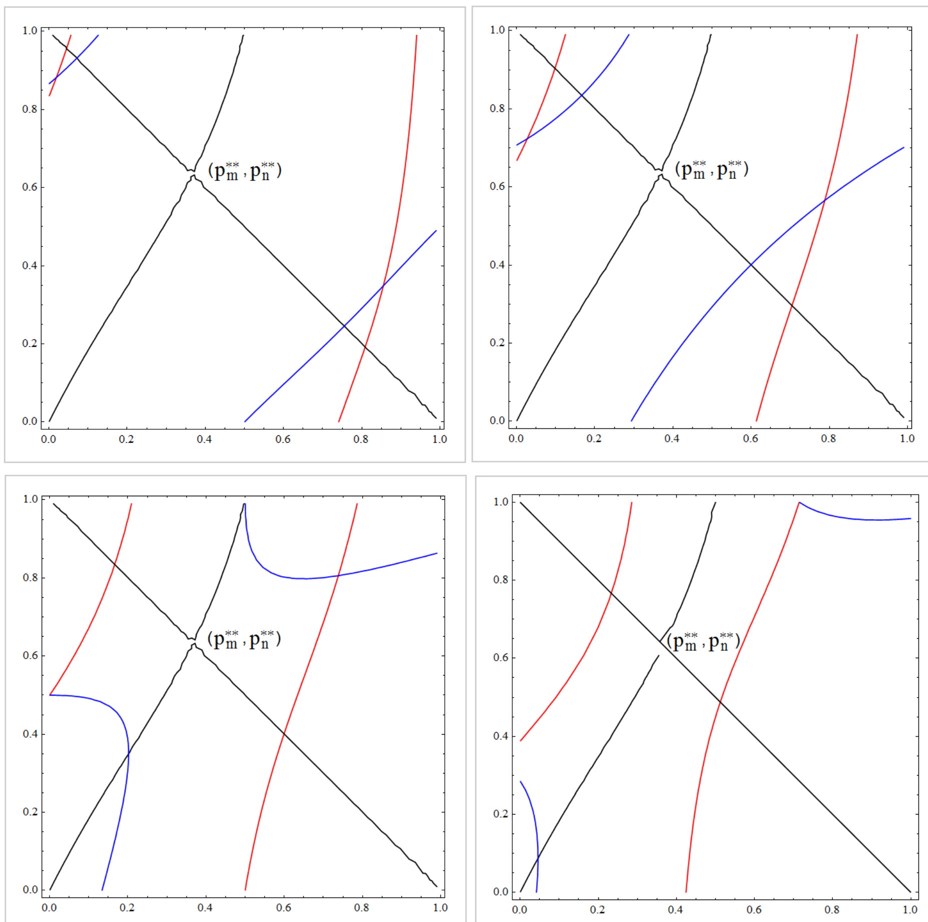

Figure 3.4: Mixing conditions in the (p_m, p_n)-space respectively for $B = \{0.1666, 0.3333, 0.5, 0.61\}$.

Thus we proved the following:

Theorem 5. *There exists a unique pair* (p_m^*, p_n^*) *such that continuity holds at all* $(B_m, B_n) \in \mathbb{R}^2_{++}$*, and it is composed of "Mix-Mix 1" and "High B".*

3.4 Application - voting over redistribution of wealth

Voting over wealth redistribution is a neat setting where our model applies: there are more poor individuals than rich, thus redistribution of wealth yields greater harm to a single rich individual than benefit to a single poor individual, thus complying with our setting of asymmetric individual benefits.

There are m poor individuals whose wealth we normalize to 1 and n rich individuals whose wealth we normalize to 2.[15] Individuals are called to vote between two extreme redistribution policies: no redistribution (alternative N) and full redistribution (alternative M). Since the wealth of the poor is less than the average wealth, a poor m individual would always prefer the higher level of redistribution (alternative M) and, at the same time, a rich n individual would always prefer the lower level (alternative N). As a reminder, we assume that $m > n > 1$.

Under full redistribution each individual ends up with the average wealth in the economy, which is:

$$\frac{2n+m}{n+m}$$

and under no redistribution everyone keeps her original wealth. Thus,

$$\begin{aligned} \Delta\pi_m &= \frac{2n+m}{n+m} - 1 = \frac{n}{n+m} \\ \Delta\pi_n &= 2 - \frac{2n+m}{n+m} = \frac{m}{n+m} \end{aligned}$$

Notice that $\Delta\pi_m = \frac{n}{m}\Delta\pi_n$, as in our parametrization in Section 4.

The policy is decided by majority rule with ties being broken evenly, and for simplicity the cost of voting is assumed to be identical for each individual - that is, $c_m = c_n = c > 0$. This means that $B_m = \frac{c(n+m)}{n}$ and $B_n = \frac{c(n+m)}{m}$. The decision of the individuals is whether to cast a vote for their preferred alternative and pay c, or to not vote at all. Thus, an individual of group i may cast a vote if (3.1) is met: that is, if $A_m \geq \frac{c(n+m)}{n}$ for a poor individual,

[15] It will be clear that these wealth level assumptions are qualitatively without loss of generality, since different wealth levels would just re-scale the results in the cost parameter c.

and $A_n \geq \frac{c(n+m)}{m}$ for a rich individual. The individuals' payoff is equal to their final wealth minus, possibly, the cost of voting.

From Lemma 4 we know that the "Pure-Mix" $p_m^* = 0$ and $p_n^* = 1 - B_n^{\frac{1}{n-1}}$ exist if and only if $B_m > nB_n - (n-1)B_n^{\frac{n}{n-1}}$. Therefore, if the "Pure-Mix" still exists as c goes to 0 (equivalently, B goes to 0), then "High B" is the unique equilibrium without any need for continuity, and the poor individuals will never vote for any $c > 0$. Notice that the second term of the right-hand side of $B_m \geq nB_n - (n-1)B_n^{\frac{n}{n-1}}$ goes to 0 faster than the other terms in the inequality as B's go to zero, and thus in the limit it is negligible. Then sufficiently close to 0 we are left with only $B_m > nB_n$. By plugging the expressions for B_m and B_n we get:

$$m \geq n^2. \tag{3.10}$$

This is a necessary and sufficient condition for $p_m^* = 0$ to hold in the unique equilibrium for any $c > 0$. Also, it has a nice interpretation. If the society is sufficiently polarized ($m < n^2$), the poor might vote and redistribution has a chance of winning. However, in a sufficiently non-polarized society ($m \geq n^2$), poors are doomed to lose the election.

How does a change of m affect (p_m^*, p_n^*)? We answer with the support of Figure 3.5. We fix $n = 3$, and set m so as to initially have a polarized society ($m = 4$, top-left), and gradually decrease the polarization ($m = 5$, top-right, and $m = 6$, bottom-left), until we hit polarization $m = n^2$ ($m = 9$, bottom-right). When we hit this polarization threshold condition (3.10) is satisfied and the "Mix-Mix 1" equilibrium (which survives continuity) disappears, and we have only "High B" ("Pure-Mix" and "Pure-Pure").

Consider $n = 3$ and $m = 4$. Since the society has (slightly) more poor than rich individuals, the average wealth is closer to the wealth of a poor individual than to the wealth of a rich individual, thus if full redistribution wins, the individual loss of a single rich individual is greater than the individual gain of a single poor individual. For this reason, an n rich individual is willing to vote for greater B's than an m poor individual. In other words, a rich has more at stake than a poor, and thus is willing to face a greater cost of voting. Therefore, p_n^* turns positive for greater B's than p_m^*, as we can see in Figure 3.5.

Consider an increase of m ($n = 3$ and $m = 5$, or 6). This has the effect of sharpening the asymmetry in willingness to face the cost of voting between rich and poor: in fact, now, p_n^* turns positive for even greater B's (the rich has even more at stake to lose in case of full redistribution), while p_m^* turns positive for even lower B's (the poor has even less at stake to win in case of full redistribution). This widens the "Pure-Mix" region (see Figure 3.5).

If the increase in m reaches the polarization threshold when $m = n^2$ ($n = 3$ and $m = 9$), the poor has so little at stake that she is nowhere willing to face the cost of voting with positive probability in equilibrium.[16] A further increase in m would still imply $p_m^* = 0$ everywhere, and further increases the willingness to vote of the rich (i.e., p_n^* increases for any given B, and p_n^* turns positive for greater B's).

This could be read as a "poverty trap": the greater is the share of poor in a society, the less likely is redistribution of resources to be the outcome of a democratic process (and if $m \geq n^2$ this probability is zero). Thus, the poor might have an incentive to attempt non-democratic channels to exit the policy trap.

[16]Remember that if $m \geq n^2$ the equilibrium is unique without the need for continuity selection.

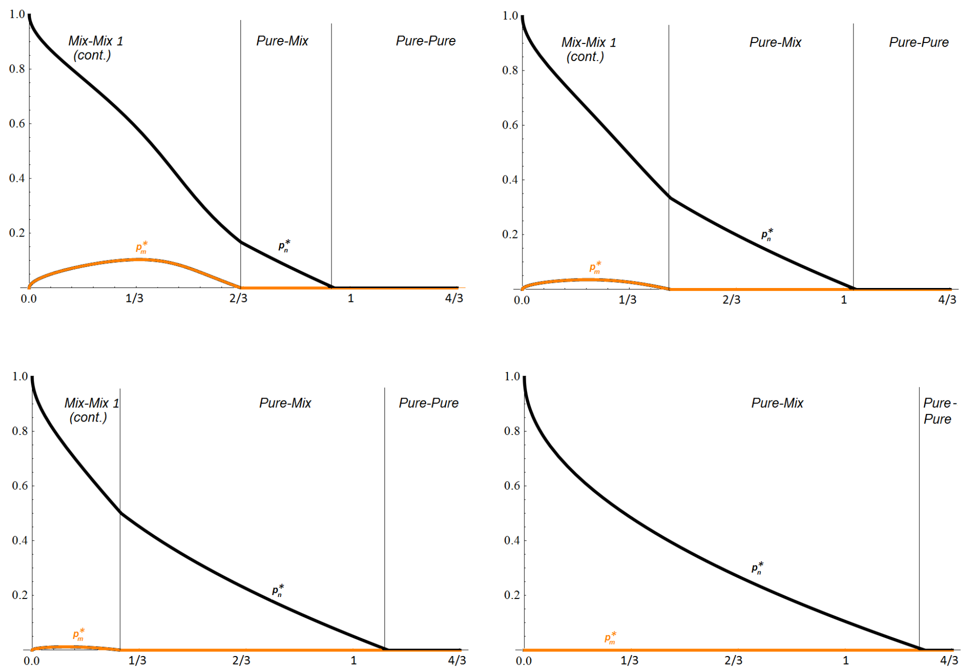

Figure 3.5: Effects on voting probabilities of increasing m, keeping $n = 3$. First row left panel $m = 4$, right panel $m = 5$. Second row, left panel $m = 6$ right panel $m = 9$.

3.5 Appendix A

We prove Proposition 6 by way of the following lemmata. See Figure 3.1.

Lemma 6. *The only points satisfying (3.9) and* $(p_m, p_n) \in \{0,1\}^2$ *are:* $(0,0)$, $(0,1)$, $(1,0)$, *and* $(p_m^{***}, 1)$, *with* $p_m^{***} = \frac{n(n-1)}{n(n-1)+\sqrt{n(n-1)(m-n)(m-n+1)}}$. *Also,* $p_m^{***} = 1$ *iff* $m = n$.

Proof. By continuity of A_m and A_n in p_m and p_n, in order to analyze the behavior of A_m and A_n in $(p_m, p_n) \in \{0,1\}^2$ we can compute the following limits for A_m

$$\begin{aligned}
\lim_{p_m \to 0} A_m &= np_n(1-p_n)^{n-1} + (1-p_n)^n \\
\lim_{p_m \to 1} A_m &= \begin{cases} p_n^n + np_n^{n-1}(1-p_n) & \text{if } m = n \\ \binom{m-1}{n} p_n^n & \text{if } m = n+1 \\ 0 & \text{if } m > n+1 \end{cases} \\
\lim_{p_n \to 0} A_m &= (1-p_m)^{m-1} \\
\lim_{p_n \to 1} A_m &= \binom{m-1}{n-1} p_m^{n-1}(1-p_m)^{m-n} + \binom{m-1}{n} p_m^n (1-p_m)^{m-n-1}
\end{aligned}$$

and for A_n

$$\begin{aligned}
\lim_{p_m \to 0} A_n &= (1-p_n)^{n-1} \\
\lim_{p_m \to 1} A_n &= \begin{cases} p_n^{n-1} & \text{if } m = n \\ 0 & \text{if } m > n \end{cases} \\
\lim_{p_n \to 0} A_n &= m(1-p_m)^{m-1} + (1-p_m)^m \\
\lim_{p_n \to 1} A_n &= \binom{m}{n} p_m^n (1-p_m)^{m-n} + \binom{m}{n-1} p_m^{n-1}(1-p_m)^{m-n+1}
\end{aligned}$$

From the above,

- if $p_m = 0$, (3.9) holds iff $p_n = 0$ or $p_n = 1$
- if $p_m = 1$, (3.9) holds iff $p_n = 0$ or (see Appendix C) $p_n = 1$ and $m = n$
- if $p_n = 0$, (3.9) holds iff $p_m = 0$ or $p_m = 1$
- if $p_n = 1$, (3.9) is equivalent to

$$\begin{aligned}
\binom{m-1}{n-1} p_m^{n-1}(1-p_m)^{m-n} + \binom{m-1}{n} p_m^n (1-p_m)^{m-n-1} &= \binom{m}{n} p_m^n (1-p_m)^{m-n} + \binom{m}{n-1} p_m^{n-1}(1-p_m)^{m-n+1} \\
\binom{m-1}{n-1}(1-p_m) + \binom{m-1}{n} p_m &= \binom{m}{n} p_m(1-p_m) + \binom{m}{n-1}(1-p_m)^2 \qquad (3.11)
\end{aligned}$$

If $m = n$, (3.11) boils down to

$$1 - p_m = p_m(1 - p_m) + m\,(1 - p_m)^2$$

whose unique solution is $p_m = 1$.

If $m > n$, (3.11) boils down to

$$\frac{(1 - p_m)}{m - n} + \frac{p_m}{n} = \frac{m p_m(1 - p_m)}{n(m - n)} + \frac{m(1 - p_m)^2}{(m - n)(m - n + 1)}$$

Solving the simple polynomial in the last expression we see that p_m^{***} is indeed one of its two roots (the second root has to be discarded since it is greater than 1). □

Lemma 7. *Equation $p_m + p_n = 1$ solves $A_m = A_n$ $\forall (p_m, p_n) \in [0, 1]^2$.*

Proof. From (3.2) plug A_m and A_n into (3.9), simplify for $(1-p_m)^m(1-p_n)^n$, and use $p_n = 1 - p_m$ to obtain

$$\sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} p_m^{s-s}(1-p_m)^{-s-1+s} + \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1} p_m^{s-s-1}(1-p_m)^{-s-1+s+1}$$
$$= \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} p_m^{s-s-1}(1-p_m)^{-s+s} + \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s} p_m^{s+1-s-1}(1-p_m)^{-s-1+s}$$

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{s} + (1 - p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{s+1}$$
$$= (1 - p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{s} + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{s}$$

$$p_m \sum_{s=0}^{n} \binom{m-1}{s}\binom{n}{n-s} + (1 - p_m) \sum_{s=0}^{n-1} \binom{m-1}{s}\binom{n}{n-s-1}$$
$$= (1 - p_m) \sum_{s=0}^{n-1} \binom{m}{s}\binom{n-1}{n-s-1} + p_m \sum_{s=0}^{n-1} \binom{m}{s+1}\binom{n-1}{n-s-1}$$
$$p_m\binom{m+n-1}{n} + (1-p_m)\binom{m+n-1}{n-1} = (1-p_m)\binom{m+n-1}{n-1} + p_m\binom{m+n-1}{n}$$
$$0 = 0$$

where in the second-to-last step we used the symmetry rule for binomial coefficients, and in the last step we used Vandermonde's identity.[17] □

[17]Vandermonde's identity states that $\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$ for $m, n, r \in \mathbb{N}_0$.

Next we characterize the set of points $A_m = A_n$ that are depicted by an increasing line in the (p_m, p_n)-space by means of two lemmas. In Lemma 8 We show that there exists a point (p_m^{**}, p_n^{**}) along the decreasing line which divides the neighborhoods of the decreasing line into two parts:

1. The first part is the one connecting (p_m^{**}, p_n^{**}) and $(1, 0)$, where we prove that increasing p_m (i.e., moving to the right of the line), increases A_m *more* than A_n. Since exactly along the line $A_m = A_n$, this result implies that to the right of the segment connecting (p_m^{**}, p_n^{**}) and $(1, 0)$ we have $A_m > A_n$ and to its left we have $A_m < A_n$.

2. The second part is the one connecting $(0, 1)$ and (p_m^{**}, p_n^{**}), where we prove that increasing p_m (i.e., moving to the right of the line), increases A_m *less* than A_n. Since exactly along the line, $A_m = A_n$, this result implies that to the right of the segment connecting $(0, 1)$ and (p_m^{**}, p_n^{**}) we have that $A_m < A_n$ and to its left $A_m > A_n$.

Lemma 8. *$\exists!(p_m^{**}, p_n^{**}) \in (0, 1)^2$ with $p_m^{**} + p_n^{**} = 1$ such that*

$$\left.\frac{\partial A_m}{\partial p_m}\right|_{p_m+p_n=1} > \left.\frac{\partial A_n}{\partial p_m}\right|_{p_m+p_n=1} \quad \textit{iff } p_m > p_m^{**} \textit{ (or equivalently } p_n < p_n^{**}\textit{)}$$

Also, if $m = n$, then $p_m^{**} = p_n^{**} = \frac{1}{2}$, and if $m > n$, then $p_m^{**} \in (0, \frac{1}{2})$ and $p_n^{**} \in (\frac{1}{2}, 1)$.

In particular,

$$p_m^{**} = \frac{n(n-1)}{n(n-1) + \sqrt{m(m-1)n(n-1)}} \quad \textit{and } p_n^{**} = 1 - p_m^{**}$$

Proof. For notation simplicity and for the sake of space we define the following

$$\begin{aligned} \tilde{p}_{s,m} &= \left(\frac{p_m}{1-p_m}\right)^s \\ \tilde{p}_{s,n} &= \left(\frac{p_n}{1-p_n}\right)^s \end{aligned}$$

Then,

$$\left.\frac{\partial A_m}{\partial p_m}\right|_{p_m+p_n=1} > \left.\frac{\partial A_n}{\partial p_m}\right|_{p_m+p_n=1}$$

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+p_m}{p_m(1-p_m)^2}+\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+p_m}{p_m(1-p_m)^2}\frac{p_n}{1-p_n}\Bigg|_{p_m+p_n=1}>$$
$$\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s}{p_m(1-p_m)(1-p_n)}+\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\frac{s+1}{(1-p_m)^2(1-p_n)}\Bigg|_{p_m+p_n=1}$$

by noticing that $p_m + p_n = 1$ implies $\tilde{p}_{s,m}\tilde{p}_{s,n} = 1$ the above inequality simplifies to

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\frac{s+p_m}{p_m(1-p_m)^2}+\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\frac{s+p_m}{p_m^2(1-p_m)}>$$
$$\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\frac{s}{p_m^2(1-p_m)}+\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\frac{s+1}{p_m(1-p_m)^2}$$

$$\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}p_m(s+p_m)+\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}(1-p_m)(s+p_m)>$$
$$\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}(1-p_m)s+\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}p_m(s+1)$$

Note that some summands in the above inequality contain s only in the binomial coefficients. By applying to these terms the same procedure at the end of Lemma 7 (i.e. symmetry rule for binomial coefficients and Vandermonde's identity), we get

$$p_m\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}s+p_m^2\binom{m+n-1}{n}+(1-p_m)\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}s+p_m(1-p_m)\binom{m+n-1}{n-1}>$$
$$(1-p_m)\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}s+p_m\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}s+p_m\binom{m+n-1}{n}$$

$$p_m\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}s+(1-p_m)\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}s>$$
$$(1-p_m)\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}s+p_m\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}s+p_m(1-p_m)\left[\binom{m+n-1}{n}-\binom{m+n-1}{n-1}\right]$$

we now analyze the summations left (containing s not only in the binomial coefficient), and use the fact that $\sum_{s=0}^{n}\binom{m}{s}\binom{n}{s}s=n\binom{m+n-1}{n}$ and that

$\sum_{s=0}^{n}\binom{m}{s}\binom{n}{s+1}s = m\binom{m+n-1}{n-2}$,[18] and get

$$np_m\binom{m+n-2}{n} + (m-1)(1-p_m)\binom{m+n-2}{n-2} >$$
$$(n-1)(1-p_m)\binom{m+n-2}{n-1} + (n-1)p_m\binom{m+n-2}{n} + p_m(1-p_m)\left[\binom{m-n-1}{n} - \binom{m-n-1}{n-1}\right]$$

$$(m-1)(1-p_m)\binom{m+n-2}{n-2} >$$
$$(n-1)(1-p_m)\binom{m+n-2}{n-1} - p_m\binom{m+n-2}{n} + p_m(1-p_m)\left[\binom{m-n-1}{n} - \binom{m-n-1}{n-1}\right]$$

and simplifying by $\frac{(m+n-2)!}{(m-2)!(n-2)!}$ we get

$$\begin{aligned}
\frac{1-p_m}{m} &> \frac{1-p_m}{m-1} - \frac{p_m}{n(n-1)} + p_m(1-p_m)\frac{(m-n)(m+n-1)}{m(m-1)n(n-1)} \\
-n(n-1)(1-p_m) &> -m(m-1)p_m + p_m(1-p_m)(m-n)(m+n-1) \\
(m-n)(m+n-1)p_m^2 + 2n(n-1)p_m - n(n-1) &> 0 \\
p_m &> \frac{n(n-1)}{n(n-1)+\sqrt{m(m-1)n(n-1)}} = p_m^{**}
\end{aligned}$$

If $m = n$ it is trivial to see that $p_m^{**} = \frac{1}{2}$. But notice also that p_m^{**} decreases in m, and hence by $m > n$, $p_m^{**} \in (0, \frac{1}{2})$ and $p_n^{**} \in (\frac{1}{2}, 1)$. □

Lemma 9 concludes the characterization of the increasing line.

Lemma 9. *There exists a unique and continuous line in the (p_m, p_n)-space which satisfies $A_m = A_n$ and connects $(0,0)$ and $(p_m^{***}, 1)$ Furthermore, this line crosses the $p_m + p_n = 1$ line once at (p_m^{**}, p_n^{**}).*

---

[18]

$$\begin{aligned}
\sum_{s=0}^{n}\binom{m}{s}\binom{n}{s}s &= \sum_{s=0}^{n}\binom{m}{s}\frac{n!}{s!(n-s)!}s \\
&= \sum_{s=0}^{n}\binom{m}{s}\frac{n!}{(s-1)!(n-s)!} = \sum_{s=0}^{n}\binom{m}{s}\frac{n!}{(s-1)!(n-1-s+1)!} \\
&= \sum_{s=0}^{n}\binom{m}{s}\frac{(n-1)!}{(s-1)!((n-1)-(s-1))!}n = n\binom{m+n-1}{n}
\end{aligned}$$

Where the last equality follows from Valdemore's identity. The calculations for the other summation are similar.

Proof. Lemma 7 establishes that the decreasing line connects two out of the four points satisfying $A_m = A_n$ along the edges. The line connecting the remaining two points is continuous and by Lemma 8 crosses the decreasing line once, at (p_m^{**}, p_n^{**}). □

3.6 Appendix B

The goal of this Appendix is to show that the two mixing conditions cross at most once in each of the four regions delimited by the set of points such that $A_m = A_n$ (see Figure 3.1). This implies that there are at most two "Mix-Mix" equilibria.

In Figure 3.4 we depict these mixing conditions of both types of individuals for different B's, and for $m = 3$ and $n = 2$. Note that i's mixing condition might cross the $p_m + p_n = 1$ line either twice, once or zero times. If they cross twice, one crossing will be above and to the left of the point (p_m^{**}, p_n^{**}), and the other will be below and to the right of it. If they only cross once the crossing coincides with point (p_m^{**}, p_n^{**}). This is what we prove in Proposition 7. Moreover, in Proposition 8 we show that the mixing conditions of the n-individual are steeper (flatter) in all points satisfying $A_m \geq A_n$ ($A_m \leq A_n$). Thus the mixing conditions cross once in each of the two regions of Figure 3.1 where $A_m > A_n$ ($A_m < A_n$) for sufficiently low B's. This, together with the fact that continuity imposes $p_m + p_n < 1$ yields the result.

Proposition 7. *Define* $\hat{B}_i = \max_{p_m+p_n=1} A_i$*. The number of intersections between* $p_m + p_n = 1$ *and i's mixing condition* $A_i = B_i$ *are:*

1. *two, if* $B_i < \hat{B}_i$*. In particular, one with* $p_m < p_m^{**}$ *and one with* $p_m > p_m^{**}$
2. *one, if* $B_i = \hat{B}_i$*. In particular, the one solving* $p_m = p_m^{**}$ *and* $p_n = 1 - p_m^{**}$
3. *zero, if* $B_i > \hat{B}_i$

   $\forall i \in \{m, n\}$.

Proof. Consider Figure 3.6. On the vertical axis there is A_i *conditional* on being along the $p_m + p_n = 1$ line, and on the horizontal axis there is p_m. We are going to show that A_i conditional on being along $p_m + p_n = 1$ is increasing in p_m if and only if $p_m < p_m^{**}$. Therefore there exists a $\hat{B}_i = \max_{p_m+p_n=1} A_i$ such that the "Mix-Mix" condition $A_i = B_i$ under $p_m + p_n = 1$ has two, one or zero solutions according to whether $B_i < \hat{B}_i$, $B_i = \hat{B}_i$ or $B_i > \hat{B}_i$.

We use the same manipulations of A_m and A_n used in the Proof of Lemma 7, and hence (3.7) with $p_n = 1 - p_m$ reads

$$(1-p_m)^{m-1}p_m^{n-1}\left[p_m\binom{m+n-1}{n} + (1-p_m)\binom{m+n-1}{n-1}\right] = B_m \quad (3.12)$$

and (3.8) with $p_n = 1 - p_m$ reads

$$(1-p_m)^{m-1}p_m^{n-1}\left[p_m\binom{m+n-1}{n} + (1-p_m)\binom{m+n-1}{n-1}\right] = B_n \quad (3.13)$$

and notice that the left-hand sides are clearly identical by construction because we know that along $p_m + p_n = 1$ line we have that $A_m = A_n$ (see Lemma 7), whereas the right-hand sides could be unequal. We now analyze the left-hand side, which can be rewritten as:

$$\binom{m+n-1}{n}(1-p_m)^{m-1}p_m^{n-1}\left[p_m + (1-p_m)\frac{n}{m}\right]$$

or, equivalently,

$$\binom{m+n-1}{n}(1-p_m)^{m}p_m^{n}\left[\frac{1}{1-p_m} + \frac{n}{mp_m}\right]$$

which takes value 0 if $p_m \in \{0,1\}$. Also it increases in p_m if and only if $p_m < p_m^{**}$, because:

$$\begin{aligned}
\frac{\partial}{\partial p_m}\left[\binom{m+n-1}{n}(1-p_m)^{m}p_m^{n}\left[\frac{1}{1-p_m} + \frac{n}{mp_m}\right]\right] &> 0\\
\frac{\partial}{\partial p_m}\left[m(1-p_m)^{m-1}p_m^{n} + n(1-p_m)^{m}p_m^{n-1}\right] &> 0\\
m\left[-(m-1)(1-p_m)^{m-2}p_m^{n} + n(1-p_m)^{m-1}p_m^{n-1}\right] + n\left[-m(1-p_m)^{m-1}p_m^{n-1} + (n-1)(1-p_m)^{m}p_m^{n-2}\right] &> 0\\
n(n-1)(1-p_m)^2 - m(m-1)p_m^2 &> 0
\end{aligned}$$

and the right-hand side of the last inequality has a unique root in $p_m \in (0,1)$ which coincides with p_m^{**}.

Thus, consider different values of B_i's which could satisfy (3.12) and (3.13): if $B_i < \hat{B}_i$ then the mixing condition of i crosses the $p_m + p_n = 1$ line twice; if $B_i > \hat{B}_i$ it does not cross the $p_m + p_n = 1$ line; and if $B_i = \hat{B}_i$ it crosses the $p_m + p_n = 1$ line exactly once in p_m^{**}. □

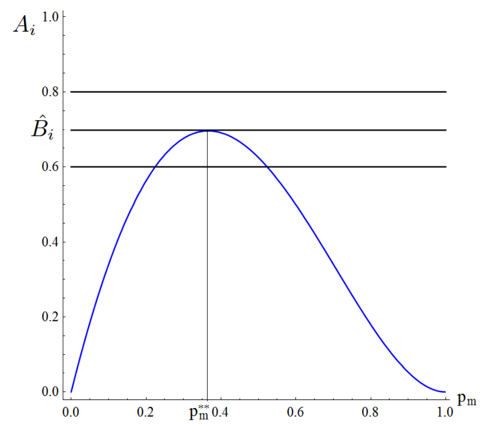

Figure 3.6: A_i conditional on $p_m + p_n = 1$ as a function of p_m, and its crossings with B_i

Definition 3. *We can express the mixing conditions as a function of* p_m *as follows:* $BR_m : p_m \to p_n$ *and* $BR_n : p_m \to p_n$ *respectively.*

Note that they are both continuous by continuity of the "Mix-Mix" conditions (3.7) and (3.8) in both p_m and p_n. Note furthermore that we defined both BR's as functions of p_m into p_n. We do this so we can compare their slopes in the following proposition.

Proposition 8. *For a given* (p_m, p_n)*,* $A_m \geq A_n$ *if and only if* $\frac{\partial BR_n(p_m)}{\partial p_m} \geq \frac{\partial BR_m(p_m)}{\partial p_m}$.

Proof. We write condition (3.7) as $A_m(p_m, BR_m(p_m)) = B_m$. Note that we have substituted into p_n $BR_m(p_m)$ and by the implicit function theorem we get

$$\frac{\partial BR_m(p_m)}{\partial p_m} = -\frac{\frac{\partial A_m(p_m, BR_m(p_m))}{\partial p_m}}{\frac{\partial A_m(p_m, BR_m(p_m))}{\partial BR_m(p_m)}}$$

and similarly for individual n

$$\frac{\partial BR_n(p_m)}{\partial p_m} = -\frac{\frac{\partial A_n(p_m, BR_n(p_m))}{\partial p_m}}{\frac{\partial A_n(p_m, BR_n(p_m))}{\partial BR_n(p_m)}}$$

Therefore $\frac{\partial BR_n(p_m)}{\partial p_m} \geq \frac{\partial BR_m(p_m)}{\partial p_m}$ is equivalent to

$$\frac{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial p_m}}{\frac{\partial A_m(p_m,BR_m(p_m))}{\partial BR_m(p_m)}} \geq \frac{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial p_m}}{\frac{\partial A_n(p_m,BR_n(p_m))}{\partial BR_n(p_m)}} \tag{3.14}$$

and using a similar notation as in the Proof of Lemma 8 for $\tilde{p}_{s+1,m}$ and $\tilde{p}_{s+1,n}$:[19]

$$\begin{aligned}\tilde{p}_{s+1,m} &= \left(\frac{p_m}{1-p_m}\right)^{s+1} \\ \tilde{p}_{s+1,n} &= \left(\frac{p_n}{1-p_n}\right)^{s+1}\end{aligned}$$

we get that (3.14) is equivalent to

$$\frac{\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s+1,m}\tilde{p}_{s,n}\left(m-\frac{s+1}{p_m}\right)+\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(m-\frac{s}{p_m}\right)}{\sum_{s=0}^{n-1}\binom{m}{s+1}\binom{n-1}{s}\tilde{p}_{s+1,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-(n-1)\right)+\sum_{s=0}^{n-1}\binom{m}{s}\binom{n-1}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-(n-1)\right)} \geq$$
$$\frac{\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s+1,n}\left(m-1-\frac{s}{p_m}\right)+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(m-1-\frac{s}{p_m}\right)}{\sum_{s=0}^{n-1}\binom{m-1}{s}\binom{n}{s+1}\tilde{p}_{s,m}\tilde{p}_{s+1,n}\left(\frac{s+1}{p_n}-n\right)+\sum_{s=0}^{n}\binom{m-1}{s}\binom{n}{s}\tilde{p}_{s,m}\tilde{p}_{s,n}\left(\frac{s}{p_n}-n\right)} \tag{3.15}$$

It remains to be mathematically proven the fact that condition (3.15) is equivalent to $A_m \geq A_n$, but is confirmed by *Mathematica*. The code is available upon request. □

3.7 Appendix C

This appendix will deal with the case of $m = n$. By the logic of QSNE it is natural to assume that in equilibrium $p_m^* = p_n^* = p^*$. Therefore we will have that $A_m = A_n$.

First note that for $B_i > 1$, $B_j > 1$ Lemma 1 holds regardless of m and n. The rest of the cases will be analyzed using a series of lemmas. Table 3.1 provides a summary of all the cases and corresponding lemmas (notice that it is symmetric).

Lemma 10. *For $B_i = 1$ and $B_j = 1$, there are only two equilibria: the "Pure-Pure" $(0,0)$ and $(1,1)$ for all $i, j \in \{m, n\}$ and $i \neq j$.*

[19] For the sake of space we use p_n instead of $BR_n(p_m)$ and $BR_m(p_m)$.

	$B_i < 1$	$B_i = 1$	$B_i > 1$
$B_j < 1$	Lemma 12		
$B_j = 1$	Lemma 11	Lemma 9	
$B_j > 1$	Lemma 10		Lemma 1

Table 3.1: Summary of Cases and Lemmas when $m = n$

Proof. When $B_i = 1$ then the only way for an individual i to vote with positive probability is when $A_i = 1$ as well, meaning that individual i is pivotal for sure.Thus he must know with certainty the number of individuals i and j that will vote. But this can only happen if i and j individuals are using pure strategies. Let $p_i = 1$ and $p_j = 0$. In this case, an individual i has an incentive to deviate and not vote, so $(1, 0)$ or $(0, 1)$ cannot be equilibria. However, if $p_i = 1$ and $p_j = 1$ then no-one has an incentive to deviate because since $A_i = B_i(= 1)$ an i individual would be indifferent between voting or not, making $(1, 1)$ an equilibrium. For the same reason (voters being indifferent) we can also sustain $p_i = 0$ and $p_j = 0$. Since no-one is voting, everybody is pivotal and $A_i = B_i(= 1)$ still holds. □

Lemma 11. *For $B_i \leq 1$ and $B_j > 1$, the unique equilibrium is $p_i^* = 0$ and $p_j^* = 0$ if the equality ($B_i = 1$) holds and $p_i^* = 1 - B_i^{\frac{1}{i-1}}$, $p_j^* = 0$ when the equality does not hold, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. By Lemma 1 $p_j^* = 0$. If the equality ($B_i = 1$) holds, then the only case for i to vote with positive probability is if $A_i = B_i = 1$. However since no-one from j is voting, the only way to have $A_i = 1$ is by not voting at all. Therefore the unique equilibrium is $p_i^* = 0$ and $p_j^* = 0$. If the equality does not hold, the only case for i to vote with positive probability is if $A_i \geq B_i$. Then $p_i = 0$ cannot be sustained in equilibrium because in this case $A_i = 1$, and any i voter has an incentive to go and vote. Furthermore, $p_i = 1$ cannot support an equilibrium because any i individual would have an incentive to deviate and not vote. Therefore the i individual must be mixing. Plugging $p_j = 0$ into the expression for A_i, we get: $A_i = (1 - p_i)^{i-1}$, so the mixing condition becomes: $B_i = (1 - p_i)^{i-1}$. Solving for p_i we get the equilibrium probability of voting for i: $p_i^* = 1 - B_i^{\frac{1}{i-1}}$.

□

Lemma 12. *For $B_i < 1$ and $B_j = 1$, $p_i^* = 1$ and $p_j^* = 1$, for all $i, j \in \{m, n\}$ and $i \neq j$.*

Proof. Since $B_j = 1$, an individual j votes only if $A_j = 1$, which cannot happen under mixed strategies. Then $p_i^* = 0$, $p_j^* = 0$ cannot be an equilibrium because the i voters have incentive to vote since in this case $A_i = 1 > B_i$ due to the fact that all i voters are pivotal. However, $p_i^* = 1$ and $p_j^* = 0$ is not an equilibrium either because then an i individual would prefer to not vote, and by the same logic, $p_i^* = 0$ and $p_j^* = 1$ is not an equilibrium either. Then the only remaining "Pure-Pure" case is: $p_i^* = 1$ and $p_j^* = 1$. This is an equilibrium because under this we have: $A_i > B_i$ and $A_j = B_j = 1$. □

Lemma 13. *Without loss of generality let $B_i \leq B_j < 1$ for all $i, j \in \{m, n\}$ and $i \neq j$. Then there exists only one "Pure-Pure" QSNE $p_i^* = 1$, $p_j^* = 1$, and at most two "Mix-Mix" equilibria.*

Proof. For the first part, notice that neither $p_i^* = 0$, $p_j^* = 1$ nor $p_i^* = 1$, $p_j^* = 0$ can be equilibria, since members of the voting group have an incentive to not vote. For both of the other two candidate "Pure-Pure" equilibria ($p_i^* = 0$, $p_j^* = 0$ and $p_i^* = 1$, $p_j^* = 1$) we have $A_j = A_i = 1 > B_j \geq B_i$. This means that $p_i^* = 0$, $p_j^* = 0$ cannot be an equilibrium because any individual would have an incentive to deviate and vote, and that $p_i^* = 1$, $p_j^* = 1$ is in fact an equilibrium because nobody has an incentive to deviate. For the second part, since we need $A_i = B_i$ and $A_j = B_j$, the result follows from the analysis in Appendix B.

□

Bibliography

Börgers T (2004) Costly Voting. American Economic Review 94(1):979–1002

Campbel CM (1999) Large Electorates and Decisive Minorities. Journal of Political Economy 107(6):1199–1217

Feddersen T, Sandroni A (2006) A Theory of Political Participation in Elections. American Economic Review 96(4):1271–1282

Krishna V, Morgan J (2012) Voluntary voting: Costs and Benefits. Journal of Economic Theory 147(6):2083–2123

Mas-Colell A, Whinston MD, Green J (1995) Microeconomic Theory. New York: Oxford University Press

Myerson R (1998a) Extended Poisson games and the Condorcet Jury Theorem. Games and Economic Behavior 25:111–131

Myerson R (1998b) Population Uncertainty and Poisson Games. International Journal of Game Theory 27:373–392

Myerson R (2000) Large Poisson Games. Journal of Economic Theory 94:7–45

Palfrey TR, Rosenthal H (1983) A Strategic Calculus of Voting. Public Choice 41(1):7–53

Palfrey TR, Rosenthal H (1985) Voter Participation and Strategic Uncertainty. American Political Science Review 79(1):32–78

Taylor CR, Yildirim H (2010) A Unified Analysis of Rational Voting with Private Values and Group-specific costs. Games and Economic Behavior 70(2):457–471