Université catholique de Louvain
Faculté des Sciences Économiques, Sociales et Politiques
Département des Sciences Économiques
Center for Operations Research and Econometrics

Essays in Political Economics

Margherita Negri

Thèse présentée en vue de l'obtention du grade de docteur
en sciences économiques et de gestion

Composition du Jury:

Promoteur:	Prof. François Maniquet	Université catholique de Louvain
Membres:	Prof. Massimo Morelli	Columbia University
	Prof. Gani Aldashev	Université de Namur
	Prof. Micael Castanheira	Université Libre de Bruxelles
Président:	Prof. Julio Davila	Université catholique de Louvain

Louvain-la-Neuve, Belgique, Juin 2014

Acknowledgments

During the years that led to the completion of this thesis, I had the incredible luck of being guided by two great supervisors, François and Massimo. I met François when I was still an inexperienced master student, who had no idea of what being a researcher means - and probably did not even think about becoming one. During these five years, François transformed me into what I am now. He taught me concepts, tools and methods. He challenged me with questions and remarks. He showed me how to be a great teacher. But, most importantly, he supported and encouraged me during all the difficult moments of my PhD. One day, I hope to be for my students the supervisor that François has been for me.

Massimo was not mentioned among my official supervisors for purely bureaucratic reasons. For me, he really is one. Much of the material contained in this thesis would not exist if I had not met him. His help and support were fundamental in the most crucial moments of my PhD. When I was about to give up with some ideas because I thought they could not lead me anywhere, Massimo always suggested me the right approach to the problem. And those ideas became papers. He introduced me to the highest and most advanced research in Political Economics and I will never thank him enough for this.

I also want to express my deep gratitude to the other two members of my thesis committee, Gani and Micael. The discussions we had throughout all my PhD and their comments during my private defense have strongly contributed to increase the quality of this thesis. A sincere thank to Julio as well, who very kindly accepted to be the president of my committee.

CORE has been for me an inspiring research environment and a friendly work place. I am proud of having had the possibility to spend five years there and my gratitude goes to all its members. I particularly want to thank the participants to the Welfare Economics Seminar, which I had the pleasure to co-organize for two years. All PhD students and Post-docs at CORE and IRES have been wonderful colleagues during these years and I hope to remain in touch with all of them after my departure. Finally, practical aspects of my research at CORE have been made much easier by the presence of a fantastic administrative staff. Catherine, in particular, has always been of great help.

Many other academics have been very important for the development of this project. Above all, my coauthor Yves Sprumont. His contribution to the third chapter of the thesis was simply invaluable and I really learned a lot by working with him. I am also indebted to Henirqueta Aragones, Miguel Angel Ballester, Douglas Bernheim, Alessandra Casella, Benoit Crutzen, Torun Dewan, Alexander Hirsch, John Huber, Matias Iaryczower, Alessandro Lizzeri, Alan Miller, Salvatore Nunnari, Pietro Ortoleva, Francesco Squintani and John Weymark.

Ai miei genitori, per il loro incondizionato amore e supporto, va il mio ringraziamento più grande. Questo e tutti gli altri traguardi che ho raggiunto nella mia vita sono anche i loro, perché non ci sarei mai arrivata se non li avessi sempre avuti al mio fianco. I miei zii Laura, Livio, Paola e Paolo e i miei cugini Enrico e Marta sono gli altri membri di una famiglia davvero importante per me. Un enorme grazie a tutti loro. Un pensiero va anche ai miei nonni e a mia zia Renza, perché non penso di averli mai ringraziati come avrei dovuto.

Let me also thank my "American" family, Annamaria, Alberto and their kids Adrian, Erik and Angelica. They welcomed me during my stay in New York and made me feel at home from the very first day.

And now, it is time for friends. Let me start with the "Belgians". Some of them were with me since the very beginning of my experience in this country. Others arrived later or left before its end. Some were initially colleagues and then became very good friends, others entered in my life through different ways. Independently of where they are now or how I met them, they have all contributed to make these six years really special ones. They are Mara S., Andrea, Claudia, Marion, Salome, Joaquin, Mara P., Paolo P., Mikel, Abdel, Stephane, Marco M., Daria, Simone, Ester, Alessandro, Corinna, Roberta Z., Antonio, Jacques, Ariane, Elena, Tommaso A., Tommaso T., Roberta T., Tim, Maia, Chiara, Giacomo and Paolo B., among many others. Six of these people deserve a particular mention: Mara S., with whom I shared the most since the very first day of my Belgian experience; Andrea, my squashmate and job market companion - and a great cook; Claudia, an irreplaceable colleague and friend; Marion, pour le français qu'elle m'a appris (mais pas seulement); Salome, for her energy and happiness; Joaquin, an (almost) ideal flatmate and a great friend. Among the many "non-Belgian" friends, I particularly want to thank Federica, Giovanna, Agnese, Audinga and Sara.

Finally, I am deeply indebted to the Fonds de la Recherche Scientifique - F.N.R.S, for having financed my research during these years. Their financial support was fundamental for me to develop my research and participate to numerous important conferences around the world.

Contents

Acknowledgments **i**

Introduction **v**

1 Minority Representation in Proportional Representation Systems **1**
1.1 Introduction 1
1.2 Literature 5
1.3 The Model 6
1.4 Results 9
1.5 Discussion 15
1.5.1 Party Leaders 15
1.5.2 Voters' Behavior under Indifference 18
1.5.3 Higher number of Parties 18
1.6 Ideology and FPTP 21
1.7 Conclusions 23
1.A Appendix 25

2 Why do Good Politicians Take Sub-Optimal Decisions? **35**
2.1 Introduction 35
2.2 Literature 38
2.3 The Model 39
2.4 Results 41
2.5 Extensions and Robustness 44
2.5.1 Addressing both Issues 45
2.5.2 Unique State of the World 47
2.6 Term Limits: a Welfare Analysis 47
2.7 Conclusions 49
2.A Appendix 49

3 An Axiomatic Approach to the Measurement of Ethnic Voting **53**
3.1 Introduction 53
3.2 Notation 56
3.3 Axioms and Results 57
3.4 Conclusions 67
3.A Appendix A: Existing Measures 67

3.B Appendix B: Proofs 71

Conclusions **77**

Bibliography **79**

Introduction

In all representative democracies, citizens agree to delegate their decision making power to Parliaments, whose composition is determined through elections. Elections in these systems serve a twofold goal. The first one consists in facilitating policy making process through the aggregation of the diverse preferences in the population. A key element for this is the type of electoral system that is chosen to select representatives. Different rules governing the aggregation process can lead to very different compositions of Parliaments. Furthermore, they can induce different strategic behaviors by voters, resulting in even more distorted outcomes. The second goal of elections is to improve the quality of representatives and their performance. This is obtained by making politicians accountable for their actions and allowing voters to replace those whose performance was not satisfying.

This thesis focuses on both goals. In the first chapter we analyze the effect of different forms of proportional representation system on the final composition of the Parliament. As the name says, proportional representation is considered to be the system delivering a Parliament composition that most closely mirrors population preferences. Previous analyses of this system have however only focused on how seats are allocated to parties. No existing theoretical work ever considered the rules that govern seat allocation *within them*. As a consequence, candidates' behavior under proportional representation system has mostly been ignored. Empirical observation, however, shows a large cross-country variation in the way seats are assigned to candidates within party-lists. Two big classes can be distinguished. Under the so called *closed list* systems, candidates' election depends on a ranking that is decided by party leaders. Citizens are only allowed to vote for party-lists and cannot modify this ranking. Such a system is used, for example, to elect the Spanish House of Representatives and it was used in Italy to elect both Chambers until 2013. In countries like Brazil, Greece or Belgium, instead, voters can vote for a list and cast a *preferential vote* for some candidates within it. Under this *open list* system, candidates' election depends on the amount of preferential votes that they received. The number of preferential votes that voters are allowed to cast can differ from country to country: it is equal to one in Brazil, it varies from one to five (depending on the size of the electoral district) in Greece, it coincides with the number of candidates to be elected in Belgium.

In Chapter 1, we study the effect of adopting a closed or open list system on the representation of minorities in Parliaments. We show that the these groups are proportionally represented only under open list, when voters can cast a limited number of preferential votes. Under this system, a candidate is sure to be elected if he or she receives enough individual votes. This creates incentives for some candidates to appeal to the minority in order to obtain their votes and be elected for sure. When voters can express many preferential votes, instead, the majority can always assign a large number of votes to a large number of politicians, preventing the election of any minority candidate. Thus, minorities will never be represented under this system. The same outcome is reached under closed list proportional representation. If voters can only vote for parties, candidates are forced to think in terms of lists. In other words, their incentives tend to coincide with the maximization of the number of seats obtained by their party. The best way to do so is to appeal to majority voters: whenever a party looks too much like a minority party, it will lose the votes of the majority, reducing the number of seats won. An important conclusion to be drawn from these results is that minorities can never be over-represented in Parliament.

In Chapter 2, we focus on the second goal of elections. As already discussed before, the threat of being replaced by another candidate should create incentives for incumbents to act in the interest of voters. Our analysis in this chapter springs from the observation that, instead, politicians do not always choose to address issues that are most relevant for the electorate. Political economy literature has already addressed this point. Such a behavior has been related, for example, to incumbents' *pandering* to public opinion, to the to need to focus on divisive issues as a signaling device for the majority of voters or to the higher level of risk associated to important issues. The model we analyze in Chapter 2 contributes to the existing literature by proposing a new explanation for sub-optimal decision making by politicians. We claim this might arise because of a negative correlation between importance of an issue and the likelihood of being able to solve it. Incumbents care about remaining in power and know that re-election is conditional on their ability to successfully address issues in a country. If the most relevant issues are difficult to solve, politicians might prefer to focus on other problems, which, despite being less relevant for the electorate, guarantee higher probabilities of success. The most interesting aspect of our result is that it holds even for incumbents who care about issue resolution *per se*. In absence of re-electoral concerns (for example, if they have reached their term limit) these politicians would always try to solve the most important issue. However, when electoral pressure is introduced, the desire of being in power in the second period induces them to behave sub-optimally. Our model therefore highlights a possible drawback of politicians accountability.

In Chapter 3, a joint work with Yves Sprumont, we return on the role of elections as a mechanism to aggregate population preferences. The focus of this chapter is on the driving forces behind citizens' voting decisions. In particular, we are interested in assessing the extent of *ethinic voting* in different countries. As a broad definition, ethnic voting occurs when individuals of the same ethnicity vote for the same party and when there exist parties whose electoral support primarily comes from specific ethnic groups. Our goal in Chapter 3 is the axiomatic definition of a measure for this phenomenon. To do so, we let each country be represented by a matrix whose cells contain the number of people belonging to an ethnic group and voting for a party. We then define axioms characterizing the ranking of these matrices according to their degree of ethnic voting.

The central axiom in our analysis imposes invariance with respect to groups or party sizes: ethnic voting should not depend on the total number of people belonging to an ethnic group, as long as the distribution of this group's support across different parties remains constant; likewise, ethnic voting should not depend on the total number of individuals voting for a party, if the share of this party's support coming from each group does not vary. The *Size*

Neutrality axiom we propose is completely new in the literature. Existing measures of ethnic voting do not satisfy the axiom. The same holds for all the indices that have been proposed to capture similar phenomena.

Our third chapter identifies necessary and sufficient conditions for a measure of ethnic voting to be size neutral. It also proposes different axioms of monotonicity, specifying the changes in the cells of the matrix that are associated to higher degrees of ethnic voting. We propose a concept of monotonicity that is compatible with size neutrality. The axiom is based on the concept of *universally favorable or unfavorable bias* between a group and a party. A universally favorable bias exists between a party and a group if the ratio between the number of supporters of that party and the number of supporters of any other party is largest for that specific ethnic group. The axiom requires that any increase in the support that the party receives from that group raises the degree of ethnic voting in the country. Whether strongest monotonicity axioms compatible with Size Neutrality can be defined remains an unsolved question, which will be addressed in future research.

Chapter 1

Minority Representation in Proportional Representation Systems

Abstract: The paper examines the effect of different forms of proportional representation (PR) on minority representation within Parliaments. We focus on the way candidates are ranked within party lists. Under closed list PR, rankings are decided by party leaders; under open list PR, the electorate determines the ranking by casting votes for individual candidates. The paper goes beyond the standard distinction between open and closed list PR by also considering variations in the number of candidates voters can select under the first system. We show that minority representation is lowest under closed list PR and under open list PR when voters can select many candidates within a list. It is highest when voters can only approve a limited number of politicians. This suggests a non-monotonic relationship between voters' control on the selection of elected candidates and minority representation.

Keywords: proportional representation, open list, closed list, preferential vote, minority representation.
JEL Classification: D02, D72.

1.1 Introduction

Among all electoral systems, *proportional representation* is deemed to be the one that most closely translates people's preferences into Parliament compositions. By assigning seats to parties in proportion to the number of votes won, it guarantees that the number of seats obtained by each of them mirrors the share of the electorate that ranks the party as first. As a direct consequence, the representation of minority groups within a country should be better protected under this system than under any other one.

The premise behind this widespread belief is that the most relevant characteristic of an electoral system is the rule governing seat allocation to parties. This has shaped the theoretical study of proportional representation systems in two important ways: first, the assignment of seats to *candidates within party-lists* has never been incorporated in the analysis; and, as a consequence, attention has never been drawn on the behavior of individual politicians appearing on a party-list.

Nevertheless, empirical observation shows a great degree of variation in the rules specifying which candidates are elected within party-lists. In some countries (among many others, Argentina, Israel, Italy, Spain and South Africa), seat allocation follows a ranking of candidates that is determined by party leaders before election. Voters cannot modify the ranking and can only cast a vote for a party-list as a whole. In other countries (e.g. Belgium, The Netherlands or Sweden), voters can modify the rankings decided by party leaders by expressing a specific vote (a *preferential* or *preferencel vote*) for some candidates within the list. If some candidates obtain enough individual votes, they are elected independently of their position on the original ranking. Finally, there are countries (e.g. Finland and Norway) where rankings are solely determined by the preferential votes given by electorate. The seats won by a party are distributed to the candidates obtaining the highest number of individual votes. The three specifications are known as *closed*, *flexible* (or *semi-open*) and *open list* proportional representation systems, respectively.

Preferential votes under open and flexible list are equivalent to approval voting within party lists: voters decide to approve (and not rank) some of the politicians appearing on the ballot. The number of approval votes that can be expressed constitutes another source of variation among countries. Just to mention some examples, this is equal to one in Sweden, Finland, Denmark and Brazil, it varies from one to five - depending on the electoral district - in Greece and it coincides with the total number of candidates on the list in Belgium.

Research conducted in Political Science clearly highlights how the different forms of proportional representation affect candidates' behavior. This literature is mostly concerned with the importance for candidates to cultivate personal reputation. Its main conclusion is that open list proportional representation increases the value of personal reputation with respect to party reputation. This happens both because open list enhances intra-party competition and electoral uncertainty (Carey and Shugart (1995), Chang (2005)) and because it induces voters to focus more on candidates' characteristics and less on parties' positions (Shugart et al. (2005)). Ames' study of Brazilian electoral and political context (Ames (1995a) and Ames (1995b)) supports these conclusions and clearly highlights the very weak role played by national parties in the country[1].

Drawing on these findings, this paper constitutes the first attempt to construct a theoretical analysis of proportional representation systems that i) incorporates the different rules for the allocation of seats within party-lists and ii) investigates how candidates and voters react to these rules. In particular, we are interested in understanding how the standard conclusions about minority representation in proportional representation systems are affected by the introduction of these two new components. Our focus will be on whether voters are or are not allowed to vote for individual candidates within party-lists (i.e. on the distinction between closed and open list proportional representation) and, when they are, on the number of preferential votes that they can express.

[1]Furthermore, cultivating personal reputation requires a consistent amount of resources and, if very valuable, might induce politicians to resort to illegal sources of financing. This explains why open list proportional representation systems are found to be positively associated to corruption (Chang and Golden (2007)).

In the attempt to keep the analysis as simple as possible, we consider a population that is divided in only two groups, the *majority* and the *minority*, each of which supports the implementation of a different policy. The role of parties in our model is reduced to the minimum and the players of the game are the candidates within them. During the electoral campaign preceding a Parliamentary election, each candidate commits to support the position preferred by the majority or the one preferred by the minority, if elected. We will often refer to a candidate choosing the majority position as a *majority candidate* or as a candidates supporting the majority. The corresponding terminology will be used for candidates supporting the minority position. Candidates' primary goal is to be elected and they only secondarily care about the number of seats won by their party. On their part, voters want to maximize the number of elected politicians supporting their favorite policy. The assumption captures the idea that having more candidates supporting one particular policy increases the chances for it to be adopted[2]. We define *minority representation* as the number of elected candidates that committed to support the position of the minority.

Our first result shows that only majority candidates will be elected under closed list proportional representation. It actually says even more than that: in equilibrium, all candidates, elected and not, will support the policy favored by the majority. As voters cannot vote for individual politicians, candidates are forced to "think in terms of list" under this system. If some candidates within a party supported the minority, they would make their party appear as a minority one. This would induce the majority to vote for a more "majority-committed" party, decreasing the total amount of seats that can be won and therefore the amount of candidates that can be elected.

The number of parties in our model is fixed to two. However, in an extension of the paper we show that the result holds even in presence of a larger number of parties, provided that the minority is small enough. Furthermore, our analysis suggests an interesting conjecture about endogenous party creation, which we plan to investigate in future research. Under closed list proportional representation, minorities are represented if and only if they manage to form their own party in a country where many others already exist. This might never be the case if party creation is costly.

Our second result focuses on open list proportional representation. We show that minority representation is considerably increased with respect to closed list whenever voters can cast a limited number of preferential votes. If sufficiently many individuals express a preferential vote for one candidate, this candidate is elected for sure, independently of the behavior of the rest of the electorate. This clearly creates incentives for some candidates to support the policy favored by the minority. In equilibrium, minority representation will be perfectly proportional, in the sense that it will coincide with the relative size of the group in the population.

One might think that the possibility of casting a large number of preferential votes could never harm minority voters, as it would only increase their ability to vote for their favorite candidates. Our results contradict this intuition. Allowing voters to express many preferential votes will *at best* guarantee a perfectly proportional representation of the minority. Most surprisingly, if too many preferential votes can be expressed, the minority will be as poorly

[2]For example, when legislative bargaining is done on multiple topics, some representatives of the majority might accept to vote in favor of the minority issue in exchange for minority's support on another dimension.

represented under open list proportional representation as it is under closed list. This happens because majority voters will always give many preferential votes to many candidates, outnumbering the votes that the other group can assign to their favorite ones.

The possibility of casting preferential votes for specific candidates within a party-list can be interpreted as a form of control by voters on the selection of their representatives. Such control is lowest under closed list proportional representation while it increases with the number of preferential votes that can be expressed under open list. Our results suggest a non-monotonic effect of such control on minority representation. When voters cannot express any preferential vote, no candidate representing the minority will be elected. As soon as the control on politicians' selection is increased, the minority will be proportionally represented in Parliament. When too much control is given to voters, however, minority representation goes back to the closed list levels.

The thresholds for the number of preferential votes determining whether the minority is perfectly represented or not represented at all depend on the size of this group. More precisely, the minority will be perfectly represented if the number of preferential votes that can be expressed is at most equal to its relative size in the population. Our analysis therefore produces clear policy indications: if the representation of a minority is to be protected, then a country should adopt an open list system, with a number of preferential votes not larger than the size of the group. However, there might as well be minorities whose representation in Parliament should be avoided. For instance, consider an interest group that politicians can target with resources subtracted from public good provision (Myerson (1993b), Persson and Tabellini (1999), Lizzeri and Persico (2001), Milesi-Ferretti et al. (2002)). In this case, the electoral system to be adopted is either closed list or open list with many preferential votes. These systems would guarantee that no resources are diverted from public good provision as no candidate targeting the interest group would be elected in Parliament. In short, our results allow to answer the following question: if a country's priority is to protect or avoid the representation of a specific group, which type of proportional representation system should it adopt?

The paper clearly challenges the common belief that proportional representation systems are associated to higher minority representation in Parliaments. Such belief is based on the idea of parties as unique and cohesive players of the electoral game. As our results show, however, the introduction of candidates as autonomous players significantly changes these conclusions. In Section 1.6, we adapt our benchmark model to the analysis of first-past-the-post systems in presence or absence of primaries. Primaries can be interpreted as a way to increase voters' control on candidate selection under first-past-the-post. We show that equilibria where seats are assigned to minority candidates can only exist if primaries are not held. Furthermore, minority representation in these equilibria can be much higher than its perfectly proportional level. When comparing these results with those found for proportional representation, two important conclusions emerge. First of all, minority representation under first-past-the-post can be strictly higher than under any form of proportional representation, a result that strongly contradicts standard conclusions about the two systems. Second, increasing voters' control on candidate selection produces very different effects in the two systems: under proportional representation, as long as this control is limited, it improves representation of minorities; under the first-past-the-post, the effect is always negative.

In the discussion of the model, we also address the effect of stronger party discipline on minority representation. More precisely, we assume that party leaders have a sufficiently strong control on their candidates to dictate to each of them the policy position they should support. We show that minority representation can only decrease under these circumstances. More precisely, equilibria where the minority is not perfectly represented appear under open list. While candidates primarily care about election, party leaders care about maximizing the number of seats won. The incentives to support the majority are therefore stronger for them. In other words, when supporting the minority is beneficial for one candidate, but reduces the number of seats won by the party, the leader will always impose to the candidate to support the majority. As a consequence, there exist fewer situations in which minority candidates are elected.

The remainder of the paper is organized as follows. In Section 1.2, we summarize the literature that is relevant for our paper. In Section 1.3 and 1.4 we introduce the formal model and present the results. In Section 1.5, we comment on the assumptions of the paper and show the extent to which our conclusions are robust to alternative specifications. In Section 1.6, we introduce the comparison between proportional representation and first-past-the-post systems. Section 1.7 concludes.

1.2 Literature

Consistent research in the literature has proven how proportional representation of preferences in Parliament may fail to be achieved in presence of *strategic voting* by the electorate. In Cox and Shugart (1996), strategic voting arises as a direct consequence of assigning seats through the Largest Remainder Method[3]. Consider a voter that ranks party A first, party B second and party C third. Assume she expects the remainder for party A to be small, and those of B and C to be equal. If the voter is strategic, she will give her vote to B, as voting for A would only increase the chances for C to obtain an additional seat. In Austen-Smith and Banks (1988), Baron and Diermeier (2001) and Maniquet (2011), the source of strategic voting is voters' anticipation of coalition formation and bargaining process within the Parliament. Quoting Baron and Diermeier (2001),

> In parliamentary systems governments with the support of a parliamentary majority control the legislative process, and elections determine representation but not which government forms. Voters thus must anticipate which policies the possible governments would choose and may not cast their ballot for their most preferred candidate or party if voting for another party leads to a coalition government that would choose a more preferred policy[4].

As already discussed in the introduction, Political Science literature has largely investigated the effect of closed and open list proportional representation on the incentives for

[3] The method comprises two stages: in the first, the number of votes won by each party is divided by a quota. The integer part of the division determines the number of seats that each party is allocated in the first stage. In the second stage, the remaining seats are allocated to the parties with the largest remainders. A formal description of this method is contained in Section 1.3.

[4] Baron and Diermeier (2001), p. 933.

candidates to cultivate personal vote. An important variable in this analysis is district magnitude, defined as the number of seats that are allocated within each district. Increases in district magnitude are associated to higher importance of personal vote under open list proportional representation and with lower under closed list (Shugart et al. (2005), Chang and Golden (2007)). Under open list, higher district magnitude coincides with stronger intra-party competition, which in turn increases the importance for candidates to obtain many individual votes. As preferential votes cannot be expressed under closed list, any effort exerted by candidates during the electoral campaign only increases the total amount of votes obtained by the party. The larger the number of politicians on the list, therefore, the higher the incentives to free-ride on the effort provided by other candidates. Moderate values of district magnitude seem to combine at best the need for stability of governments and representation of different preferences in the population (Carey and Hix (2011)). The trade-off between stability and representation is well known. Single-member districts are supposed to produce more stable governments at the expenses of worse representation in Parliaments. On the contrary, multimember districts favor representation, but are typically associated to coalition governments. Intermediate values of district magnitude capture the positive aspects of both extremes, allowing to achieve representation without giving up too much stability. A study of the behavior of Swiss legislators by Portmann et al. (2012) suggests, however, that they are more likely to support majority positions if elected in districts of low district magnitude (but see also Carey and Hix (2013)).

Political Economic literature has typically limited the analysis of electoral systems to the distinction between first-past-the-post and proportional representation systems. The two electoral rules have been contrasted under a large variety of aspects, among which the probably most important ones are party creation (Duverger (1959), Taagepera and Shugart (1989), Palfrey (1989), Feddersen (1992), Lijphart and Aitkin (1994), Fey (1997), Morelli (2004), etc.), public good provision and transfers (Myerson (1993b), Persson and Tabellini (1999), Lizzeri and Persico (2001), Milesi-Ferretti et al. (2002)) and corruption (Myerson (1993a), Persson et al. (2003)). The first element of comparison is probably the most relevant to our purposes. Classical Duvergerian conclusions about the number of parties under the two systems seem to imply that representation of minorities can be better achieved under proportional representation.

Little attention has so far been paid to finer classifications of electoral systems. A notable exception is Crutzen (2013), who examines the effect of different forms of candidates selection on the effort exerted by politicians. Under closed list proportional representation, politicians compete by exerting effort both to be included in the list and to be placed in high positions. Under open list proportional representation, the position on the list does not matter any longer, but an additional source of competition is introduced by the need to attract a sufficiently high number of individual votes. Which of the two systems induces higher effort provision by politicians depends on voters' responsiveness to effort.

1.3 The Model

We consider the election of a Parliament composed of S representatives and let all the seats be allocated in a unique national district. For simplicity, assume that S is an even number.

The set of voters within the country is denoted by $\mathcal{V} = \{1, \ldots, V\}$. Each individual $i \in \mathcal{V}$ supports the implementation of policy $x_i \in \{x_M, x_m\}$. More precisely, we assume that the *majority* of the electorate favors policy x_M, while only a *minority* prefers x_m. We define

$$Maj = \{i \in \mathcal{V} : x_i = x_M\}$$

and

$$Min = \{i \in \mathcal{V} : x_i = x_m\}$$

with $Maj \cup Min = \mathcal{V}$ and $Maj \cap Min = \emptyset$. There are $2S$ candidates, belonging to two different parties, A and B. Denote a candidate occupying position s on party A's and B's list by a_s and b_s, respectively. Then,

$$A = \{a_1, \ldots, a_S\} \quad B = \{b_1, \ldots, b_S\}$$

More generally, we let $\mathcal{C} = A \cup B$ be the set of all candidates and denote by $c \in \mathcal{C}$ a generic candidate within. Before election, candidates commit to support the policy preferred by the majority or the one preferred by the minority. Formally, each candidate c chooses $x_c \in \{x_M, x_m\}$. Candidates' preferences are lexicographic over the two following goals:

1. Win a seat

2. Maximize the number of seats won by the party

In other words, whenever two strategies are equivalent in terms of personal outcome (i.e., either they both guarantee election or none of them does), candidates will prefer the one that is best for their party. Voters care about the final composition of the Parliament and want to maximize the number of elected politicians supporting their favorite position. Formally, letting W be the set of candidates obtaining a seat in the Parliament, voter i's utility can be written as

$$U_i(W) = u_i\left(\sum_{c \in W} \mathbb{1}_{[x_c = x_i]}\right)$$

with $u_i(\cdot)$ strictly increasing.

Seat allocation across parties follows the Largest Remainder Method. We denote the quota of votes that allows to receive a seat by $\alpha = V/S$. Thus, a party obtains a seat every α votes received. For example, 2α votes guarantee two seats, 3α votes guarantee three seats and so on. If any seat is left unallocated after this step, it is assigned to the party with the largest remainder. Formally, let v^P denote the amount of votes received by party $P \in \{A, B\}$, and let δ^P be defined as

$$\delta^P = \frac{v^P}{\alpha} - \left\lfloor \frac{v^P}{\alpha} \right\rfloor < 1$$

where for all $y \in \mathbb{R}$, $\lfloor y \rfloor = \max\{x \in \mathbb{Z} | x \leq y\}$. Then, the number of seats assigned to party P is

$$S^P = \begin{cases} \left\lfloor \frac{v^P}{\alpha} \right\rfloor + 1 & \text{if } \delta^P > \delta^{P'} \\ \left\lfloor \frac{v^P}{\alpha} \right\rfloor & \text{if } \delta^P < \delta^{P'} \end{cases}$$

Whenever $\delta^P = \delta^{P'}$, each party receives the additional seat with probability one-half. To keep the analysis simple, we make the following assumption about the size of the two groups of voters.

Assumption 1. *There exists $k_m \in \mathbb{N}_{++}$ such that $|Min| = \alpha k_m$ and $|Maj| = \alpha(S - k_m)$.*

In words, Assumption 1 states that the size of the two groups is an exact multiple of α[5]. For convenience, assume k_m is even. Notice that Assumption 1 does not necessarily imply that the number of votes received by each party will always be a multiple of α. On the contrary, as our analysis will show, there are cases where optimal voting behavior by the two groups will be such that $\delta^P > 0$ for some $P \in \{A, B\}$.

Our focus in this paper is on the way the S^P seats won by a party are distributed across candidates within its list. The two systems we consider are

Closed List PR. *Seats are assigned to the first S^P candidates appearing on the party-list. Voters cannot modify this ranking and can only vote for a party.*

Open List PR. *Voters vote for a party and, within its list, they can approve up to $\pi \in \{1, \dots, S\}$ candidates. Seats are assigned to the S^P candidates that were approved by the highest number of voters. Ties are broken by assigning the seat to each entitled candidate with equal probability.*

Thus, π denotes the number of *preferential votes* for individual candidates that the electorate can express within a party-list. For example, $\pi = 1$ in Finland, Sweden or Brazil, $\pi \in \{1, \dots, 5\}$ in Greece and $\pi = S$ in Belgium. The timing of the game is the following:

$t = 1$ Candidates select their policy position $x_c \in \{x_M, x_m\}$. Under closed list, they do so knowing their ranking on the list;

$t = 2$ Elections take place and seats are distributed first across and then within parties;

The implicit assumption behind our timing choice is that party-list formation occurs before candidates make their electoral promises. This is equivalent to requiring that party discipline is relatively mild, and candidates can deviate from party lines if this increases their chances of being elected. The alternative case of strong party discipline is examined in Section 1.5.1, where we assume that party leaders can dictate to their candidates the policy position they must support.

We require the equilibria of the game to be subgame perfect and we additionally impose Strong Nash (Aumann (1959)) in the voting subgame. A profile of strategies is a Strong Nash equilibrium of a game if there exists no profitable deviation for any *coalition* of players. This clearly imposes a high degree of coordination on voters and we are aware of the contrast with empirical results about *positional misvoting*[6]. The choice of considering only Strong Nash

[5] Our results would not change if we assumed that $|Min| = \alpha k_m + \delta_m$ and $|Maj| = \alpha(S - k_m) + \delta_M$, with $\delta_m + \delta_M < \alpha$.

[6] Positional misvoting occurs when some candidates obtain more votes simply because they are listed at a particular position on the ballot. In particular, candidates on top of the list might enjoy an electoral advantage. Evidence of this phenomenon was found, for example, in the Spanish Senate (Lijphart and Pintor (1988) and Esteve-Volart and Bagues (2012)) and in Australian elections (Kelley and McAllister (1984)).

equilibria allows us to rule out coordination failure by voters as a possible explanation for the presence or absence of minority representation in Parliaments. Furthermore, the refinement consistently reduces the number of equilibria of the game, increasing the tractability of our analysis. As common in voting games, the set of Nash equilibria in our model is too wide to produce any informative conclusion. When faced with the trade-off between sharper predictions and less demanding assumptions, we gave priority to the first. Moreover, the set of Strong Nash equilibria in our model coincides with that of Coalition-Proof Nash equilibria (Bernheim et al. (1987) and Bernheim and Whinston (1987)).

In addition to the Strong Nash requirement, we adopt the following tie-breaking rule.

Assumption 2. *If a coalition of voters in a group is indifferent between two parties,*

i) They vote for the party with the highest number of preferred candidates

ii) If this is equal among the two parties, they equally divide their votes across the two parties.

The first part of Assumption 2 is in line with the idea of *partial honesty* proposed by Dutta and Sen (2012). With regard to the second part, we believe this is the most intuitive way of modeling voters' behavior in this situation. Our results partially depend on this assumption, in a way that will be discussed at the end of the paper.

Finally, to strengthen the idea that candidates belong to a party, we assume some form of coordination among them. In particular, we will select equilibria that are robust to deviations by any coalition of candidates *within the same party.*

1.4 Results

All the results we are going to present are expressed in terms of *minority representation* (MR), which we define as the number of elected candidates supporting policy x_m, that is

$$MR = |\{c \in W : x_c = x_m\}|$$

Perfectly proportional representation of the minority is achieved when the share of members of Parliament supporting x_m, MR/S, coincides with the relative size of the minority in the population, $\alpha k_m/V$. This immediately implies that

$$\frac{MR}{S} = \frac{\alpha k_m}{V} = \frac{\alpha k_m}{\alpha S}$$

so that the minority is proportionally represented when $MR = k_m$. We begin the exposition of our results with closed list. Lemma 1 characterizes equilibrium voting behavior in subgames where only one party contains candidates supporting the minority. Denote by n_m^P the number of candidates in party P supporting policy x_m,

$$n_m^P = |\{c \in P : x_c = x_m\}|$$

Likewise, $n_M^P = S - n_m^P$ will represent the number of candidates in P that chose to support x_M.

Lemma 1. *Assume the electoral system is closed list PR and consider a voting subgame where, for some $P \in \{A, B\}$, $n_m^P > 0$ while $n_m^{P'} = 0$ for $P' \neq P$. Equilibrium voting behavior in this subgame is such that all voters in the minority vote for P and all voters in the majority vote for P'.*

Proof. Voting for P is a weakly dominant strategy for the coalition of voters in the minority, as it is the only one that might allow to elect some candidates supporting x_m. By Assumption 2, all minority voters will vote for P. Remember that $|Min| = \alpha k_m$ and a party receives a seat every α votes. This implies that party P will for sure obtain k_m seats and candidates $\{c_1, \dots, c_{k_m}\} \in P$ will be elected. Consider voters in the majority now. If αx of them vote for P, they will elect candidates $\{c_{k_m+1}, \dots, c_{k_m+x}\} \in P$. If $x_c = x_m$ for some $c \in \{c_{k_m+1}, \dots, c_{k_m+x}\}$, then deviating to P' would be profitable, as no candidate in P' supports x_m. If instead $x_c = x_M$ for all $c \in \{c_{k_m+1}, \dots, c_{k_m+x}\}$, then they are indifferent and, by Assumption 2, they should vote for P'. This holds for all coalitions of αx voters in Maj, for all $x \in \{1, \dots, S - k_m\}$. □

Lemma 1 implies that in subgames where only party P contains candidates supporting the minority, $S^P = k_m$ and $S^{P'} = S - k_m$. We will make large use of this result to prove our main result for closed list.

Proposition 1. *Under closed list proportional representation, no minority candidate is elected in equilibrium, i.e. $MR = 0$.*

Proof. To show that an equilibrium where $MR = 0$ exists, assume no candidate decides to support the minority, i.e. consider the voting subgame where $n_m^A = n_m^B = 0$. By Assumption 2, each party obtains $S^A = S^B = S/2$ seats and the first $S/2$ candidates in each list will be elected. It is obvious that no coalition of voters can profitably deviate from this strategy. Proceeding backward, let us now consider possible deviations by coalitions of candidates within the same party. Without loss of generality, assume that a coalition of $\tilde{n}_m^A > 0$ candidates in party A deviates to x_m and denote by $\tilde{S}^P$ the number of seats that party P wins in this subgame. As $\tilde{n}_m^A > 0$ and $n_m^B = 0$, Lemma 1 immediately implies that

$$\tilde{S}^A = k_m < S/2 \qquad \tilde{S}^B = S - k_m > S/2$$

so that only the first k_m candidates in A are elected. The deviating coalition, therefore, cannot contain candidates occupying positions $\{a_{k_m+1}, \dots, a_{S/2}\}$. These are candidates that were elected in the original subgame (as $S^A = S/2$) and would lose their seat by deviating. For all other candidates, instead, the deviation does not affect whether they obtain or not the seat: candidates in $\{a_1, \dots, a_{k_m}\}$ are elected before and after the deviation, while those in $\{a_{S/2+1}, \dots, a_S\}$ are never elected anyway. Given that these candidates care about the total number of seats obtained by their party, they will never be part of the deviating coalition. Thus, it must be that $\tilde{n}_m^A = 0$.

Let us now assume by contradiction that an equilibrium with $MR > 0$ exists. This happens only if either i) both parties contain minority candidates, i.e. $n_m^P > 0$ for both $P \in \{A, B\}$ or ii) only one of the two parties contains minority candidates, i.e. $n_m^P > 0$ for some $P \in \{A, B\}$ and $n_m^{P'} = 0$ for $P' \neq P$. Consider case i) first. In equilibrium, either $S^A = S^B = S/2$ or one of the two parties obtains more seats than the other. Let $S^B \geq S^A$.

Now assume that all candidates supporting x_m in A deviate to x_M, so that, in the new subgame, $\tilde{n}_m^A = 0$.

By Lemma 1, this deviation guarantees to the party $\tilde{S}^A > S/2 \geq S^A$ seats. Thus, no candidate loses her seat by deviating. Furthermore, the candidates in positions $\{a_{S^A}, \dots, a_{\tilde{S}^A}\}$, who were not elected in the original subgame, are now assigned a seat and therefore strictly gain from the deviation. For all other candidates, the deviation is profitable as it increases the number of seats won by the party. Independently of the position occupied by the deviating candidates on the list, therefore, the deviation will be profitable for all of them.

Consider the second case now and assume $n_m^A > 0$ while $n_m^B = 0$. By Lemma 1 party A obtains $S^A = k_m$ seats. As before, a deviation to x_M by all candidates supporting the minority, would allow the party to obtain $S/2 > k_m$ seats and would be profitable for all candidates within the party. □

We now turn to open list proportional representation. Under this system, if party $P \in \{A, B\}$ obtains S^P seats, a candidate within it is elected if and only if she is among the S^P candidates obtaining the highest number of preferential votes. Let us start with the analysis of the voting subgames and consider optimal behavior by the minority. These voters will always vote cohesively for one party. Such strategy maximizes the number of preferential votes that minority candidates can receive and therefore the chances that at least one of them is elected. Furthermore, whenever the minority is indifferent between voting for one party or the other, optimality of this strategy is a direct consequence of the tie-breaking rule specified in Assumption 2. The only situation where minority voters vote for different parties is the one where they equally split their support between A and B. By Assumption 2, this happens if $n_m^A = n_m^B$ and no other strategy can guarantee a better outcome.

Consider the behavior of the majority now: their goal is to elect the highest number of candidates supporting x_M. Thanks to the properties of the zero sum game under consideration, this is equivalent to minimizing the number of elected candidates supporting x_m. The optimal strategy by the majority should therefore aim at maximizing the number of majority candidates receiving more preferential votes than minority ones. Assume that some minority voters vote for party P. If the majority wants to prevent the election of at least one minority candidate in this party, it needs to assign to some of its favorite candidates an amount of preferential votes that is strictly greater than those obtained by minority ones. To do so, however, they must vote for P, thus increasing the total number of seats won by the party.

Remark 1. If the number of majority candidates ranked above minority ones is weakly lower than the additional seats assigned to P by the votes of the majority, it will be impossible for this group to prevent the election of candidates supporting x_m.

To understand the statement, consider the following simple example. Let $\alpha = 1$, $k_m = 2$ and $S = 8$. In other words, a seat is assigned to a party if one voter votes for it; the minority contains two voters ($\alpha k_m = 2$), and there are $\alpha(S - k_m) = 6$ voters in the majority. Assume all minority voters vote for party A, so that the party obtains at least two seats. To prevent the election of some minority candidates, at least 3 voters in the majority must vote for A as well: if they all cast their preferential votes for the same candidates in A, these will receive a strictly higher number of votes than those obtained by minority ones. Notice that, under this strategy profile, the majority assigns 3 additional seats to A, so that $S^A = 5$. The number of elected minority candidates will be strictly lower than 2 if and only if there are strictly more than 3 majority candidates receiving enough preferential votes.

Our discussion shows that the ability for the majority to prevent the election of some minority candidates strictly depends on the number of preferential votes that can be cast by voters, π. Assume $\pi \leq k_m$. In this case, the following Lemma holds.

Lemma 2. *Under open list proportional representation with $\pi \leq k_m$, if a candidate receives $\alpha\pi$ preferential votes, she is elected for sure.*

Assume a minority candidate in party P receives $\alpha\pi$ preferential votes. If majority voters want to prevent her election, they need to assign at least $\alpha\pi + 1$ preferential votes to a majority candidate in P. As no voter can express more than one preferential vote for the same candidate, at least $\alpha\pi + 1$ majority voters must vote for P. Thus, the party obtains at least π additional seats. The maximum number of majority candidates that can receive $\alpha\pi + 1$ votes, however, is π. By Remark 1, the candidate supporting the minority will therefore be elected anyway. Clearly, a symmetric reasoning holds for candidates supporting the majority.

Now notice that if all minority voters vote for the same party, they can cast a total amount of preferential votes equal to $\pi\alpha k_m$. This implies that they can assign $\alpha\pi$ votes to

$$\left\lfloor \frac{\pi\alpha k_m}{\alpha\pi} \right\rfloor = k_m$$

minority candidates in the party. By Lemma 2, these candidates will be elected for sure. By the same reasoning, if all majority voters vote for the same party, they can assign $\alpha\pi$ preferential votes to

$$\left\lfloor \frac{\pi\alpha(S - k_m)}{\alpha\pi} \right\rfloor = S - k_m$$

candidates supporting x_M. This is a central step in the proof of the following Proposition.

Proposition 2. *Under open list proportional representation with $\pi \leq k_m$, the minority is proportionally represented in Parliament, i.e. $MR = k_m$.*

A complete proof of Proposition 2 can be found in the Appendix. We there show that, in equilibrium, there will always be enough minority and majority candidates to vote for.

Combining the results found in Propositions 1 and 2, we know that the minority can be at best proportionally represented in Parliament. Can there be situations where the group is over-represented, meaning that $MR > k_m$? It turns out that the answer to this question is no. Indeed, larger values of π only help the majority to elect more candidates supporting their favorite position. Assume $\pi > k_m$. The maximum number of preferential votes that a minority candidate in party P can receive is αk_m. This happens when all minority voters vote for the candidate. If $\alpha k_m + 1$ majority voters vote for P as well, they can assign strictly more than αk_m preferential votes to $\pi > k_m$ majority candidates while increasing the number of seats obtained by the party by only k_m. They might therefore be able prevent the election of some minority candidates.

Proposition 3. *Under open list proportional representation with $\pi > k_m$, the minority is never over-represented, i.e. $MR \leq k_m$.*

Proof. Assume that strictly more than k_m minority candidates are elected in equilibrium. If they all belong to the same party P it must be that $S^P > k_m$ and at least α majority voters must be voting for P. Let αx be the number of such voters, $x \geq 1$, so that $S^P = k_m + x$. The number of elected minority candidates is greater than k_m if there are strictly less than x majority candidates receiving $\alpha k_m + 1$ preferential votes. This means that the majority is assigning seats to P knowing that they will not be won by majority candidates. Such a strategy is optimal only if there are no more majority candidates to elect in party P': a deviation to P' by some of the αx majority voters would indeed only lead to the election of a different minority candidate. Clearly, if this is the case, a profitable deviation exists by at least one minority candidate in P'. By deviating to x_M, this candidate will be elected for sure.

Now assume that the elected minority candidates belong to different parties. Let αy^A of them vote for A and αy^B vote for B, with $y^P \geq 1$ for both P and $y^A + y^B = k_m$. Assume $n_M^A \geq y^A + 1$ and $n_M^B \geq S - k_m - y^A - 1$. If $\alpha y^A + 1$ majority voters vote for A, these voters will be able to elect at least $y^A + 1$ majority candidates. Moreover, the $\alpha(S - k_m - y^A) - 1$ voters voting for B will be able to assign $\alpha y^B + 1$ preferential votes to

$$\left\lfloor \frac{[\alpha(S - k_m - y^A) - 1] \min\{\pi, n_M^B\}}{\alpha y^B + 1} \right\rfloor \geq S - k_m - y^A - 1$$

candidates in this party. Thus, at least $S - k_m$ majority candidates will be elected. Now notice that, in equilibrium, it must indeed be that $n_M^A \geq y^A + 1$ and $n_M^B \geq S - k_m - y^A - 1$. If one of the two conditions did not hold, a profitable deviation would exists for at least one minority candidate. □

The existence of pure strategy Strong Nash equilibria in the voting subgames where $MR < k_m$ depends on the size of the majority. To understand why, consider the following example: let $k_m = 4$, $S = 10$ and $\pi > 4$. If all minority voters vote for party A, the best response by the majority is to have $4\alpha + 1$ voters voting for the same party. Since $\pi > 4$, this will prevent the election of at least one minority candidate. Notice that the at least some of the remaining $2\alpha - 1$ majority voters must vote for party B, since otherwise $S^A = S$ and all minority candidates in A would be elected. Thus, at most $k_m - 1$ minority candidates will be elected. The maximum number of preferential votes that a majority candidate in B can obtain, however, is $2\alpha - 1$. This immediately implies that the best response by the minority to the strategy played by the majority is to vote for B and assign 4α preferential votes to as many minority candidates as possible. This will guarantee the election of a strictly higher number of minority candidates than those elected if voting for A. Notice that, however, this cannot be an equilibrium either: if all minority voters vote for B, the best response by the majority is to assign $4\alpha + 1$ votes to as many majority candidates as possible in the party. As a consequence, however, majority candidates in A will obtain at most $2\alpha - 1$ preferential votes, triggering a deviation by the minority towards party A.

To see how this cycling behavior does not occur when the majority is numerous enough, let $k_m = 2$ while keeping S fixed. Thus, there are $(S - k_m)\alpha = 8\alpha$ voters in the majority now. This group is always able to assign at least $2\alpha + 1$ votes to each party and to some of their favorite candidates within it. Independently of which party the minority votes for, therefore, it will be impossible to elect $k_m = 2$ minority candidates, as there will always be at least $\pi > k_m$ majority candidates receiving a strictly greater amount of preferential votes.

Larger values of π increase the number of majority candidates that can be assigned $\alpha k_m + 1$ votes without increasing the number of seats won by the party. As a consequence, when the majority is numerous enough, minority representation decreases with the number of preferential votes that can be expressed. In particular, if π is very large, no candidate supporting the minority will ever be elected in equilibrium. This result is formally stated in Proposition 4. Before introducing it, let us highlight some interesting features about candidates' behavior.

For the sake of exposition, fix $\pi = S$, so that each voter could potentially vote for all the candidates on a party-list. Furthermore, assume that the majority is large enough. By our discussion above, these two assumptions allow the majority to give many votes to many candidates, preventing the election of *any* minority one. A necessary condition for this to happen, clearly, is that there are enough majority candidates to vote for in the two parties. Assume all minority voters vote for party P and remember that the maximum number of preferential votes they can assign to a candidate is αk_m. Then, no minority candidate is elected in P if the number of majority candidates receiving at least $\alpha k_m + 1$ preferential votes is at least equal to the number of seats won by the party. This immediately implies that the number of majority candidates in party P must be

$$n_M^P \geq S^P = k_m + \ \# \text{ seats assigned to } P \text{ by majority} \tag{1.1}$$

As at least $\alpha k_m + 1$ majority voters vote for P, $S^P \geq 2k_m$. Furthermore, notice that condition (1.1) must hold for both $P \in \{A, B\}$, as otherwise the minority would be able to elect at least one candidate in one of the two parties. Thus, $MR = 0$ in equilibrium if and only if

$$\begin{cases} n_M^P \geq 2k_m \quad \text{for both } P \in \{A, B\} \\ n_M^A + n_M^B = 2k_m + (S - k_m) = S + k_m \end{cases} \tag{1.2}$$

where the first equality in the second condition follows from the fact that the total amount of seats that the majority can assign to the two parties is $S - k_m$.

A key step in the proof of Proposition 4 is to show that the conditions listed in (1.2) are always satisfied in equilibrium. At first, one might think that supporting x_M is a dominant strategy for a candidate. As we show in the proof, this intuition is flawed: supporting the minority can be optimal for a candidate belonging to a party where many other candidates support x_m, if the other party is mainly composed of majority candidates. Assume party B only contains candidates supporting x_M while A also contains some minority candidates. Clearly, the minority will vote for A. If majority voters vote for B, they will be able to elect $S - k_m$ candidates. Is there an alternative strategy that allows to elect strictly more majority candidates? The answer depends on n_M^A. First of all, notice that the maximum number of votes that the majority is willing to give to A is $\alpha k_m + 1$: as $\pi = S$, these voters are enough to assign to all candidates supporting the majority in A an amount of votes that is strictly higher than those received by minority ones. Our tie-breaking rule implies that all other voters must vote for B[7]. When $\alpha k_m + 1$ majority voters vote for A, $S^A = 2k_m$. The number

[7] Incidentally, notice that assigning to A any amount of votes in between zero and $\alpha k_m + 1$ would also be suboptimal, as majority candidates in the party would never receive enough preferential votes.

of elected majority candidates in A will therefore be the minimum between S^A and n_M^A. As the remaining majority voters vote for B, this party will obtain $S - 2k_m$ seats[8], all of which will be assigned to majority candidates. Giving some votes to A is therefore optimal for the majority if and only if

$$\min\{2k_m, n_M^A\} + S - 2k_m > S - k_m$$

which is satisfied if and only if $n_M^A > k_m$. When the condition does not hold, only the minority will vote for A. Thus, exactly k_m minority candidates and no majority one will be elected in the party. This proves that supporting the minority is not a dominated strategy.

Another interesting observation to make is that, when $k_m \leq n_M^A \leq 2k_m$, all n_M^A majority candidates are elected for sure. An ideal situation for a candidate is therefore to support the majority when enough, but not too many, other candidates do the same. Or, in other words, a candidate would like to be a majority candidate in a predominantly minority party. Such party attracts the votes of the minority and, together with them, the votes of a part of the majority that wants to prevent the election of minority candidates. A direct consequence of this is that, if $n_M^A < 2k_m$, a profitable deviation will always exist for some minority candidates in A that are elected with probability strictly lower than one. To derive this conclusion, we assumed that B only contained majority candidates. In the proof, we show that the same reasoning holds in more general cases, namely when $n_M^B > (S + k_m)/2$.

Proposition 4. *Under open list proportional representation with $\pi \geq 2(k_m + 1)$, if $k_m < S/3 - 1$, no minority candidate is elected in equilibrium, i.e. $MR = 0$.*

1.5 Discussion

In this section, we discuss the effects of relaxing some of the assumptions we made in the benchmark model. In Section 1.5.1, we assume politicians are fully subject to party control and examine the case where party leaders can instruct their candidates on which policy to support. Section 1.5.2 removes some of the restrictions on voters' behavior that were imposed in Assumption 2. Finally, in Section 1.5.3, we consider the existence of more than two parties.

1.5.1 Party Leaders

Assume the only players of the game are the leaders of the two parties. Before election, they must instruct their candidates on which policy to support. That is, party P's leader chooses $x_c \in \{x_M, x_m\}$ for all $c \in P$. Possibly, $x_c \neq x_{c'}$ for $c, c' \in P$.

Given our assumptions on candidates' utility in the benchmark model, results for closed list are not affected by this alternative setting. Indeed, party P's leader has a profitable

[8] The number of votes received by the two parties is

$$v^A = 2\alpha k_m + 1$$

$$v^B = \alpha(S - k_m) - \alpha k_m - 1 = \alpha(S - 2k_m - 1) + \alpha - 1$$

The remainder from the seat allocation process is $1/\alpha$ for A and $(\alpha - 1)/\alpha$ for B. Thus, $S^A = 2k_m$ and $S^B = S - 2k_m$.

deviation in this setting if and only if there exists a coalition of P's candidates that can profitably deviate in the benchmark model. To see why, consider a coalition of candidates in party P and assume that, for a given the strategy of the other players, a profitable deviation for them exists. This can happen in two cases: first, if the deviation guarantees them a seat which they could not obtain before; second, if the deviation has no impact on whether they win a seat or not, but it helps the party to win a higher number of seats. Clearly, politicians' incentives in the second case are completely aligned to party leaders' ones. With regard to the first case, notice that a deviation allows a candidate to obtain a seat that she was not getting before if and only if it increases the number of seats won by the party. Thus, if the coalition is willing to deviate, party P's leader must have a profitable deviation as well, and vice versa.

Under open list PR with $\pi \leq k_m$, this alternative setting increases the number and type of equilibria of the game. Providing a full characterization of the different classes of equilibria is beyond our goal and we will leave it for future research. We here limit ourselves to show an example of equilibrium that arises in this framework. In particular, we show that minority representation can be smaller than k_m. Key for the result is the partial misalignment of candidates' and party leaders' incentives arising under open list: while candidates primarily care about winning a seat, party leaders always want to maximize the number of votes obtained. Given this misalignment, a deviation allowing a candidate to be elected while reducing the total number of seats won by the party was profitable in the benchmark model, but it is not any more in this new framework. The intuition behind the example is then the following: starting from a subgame where less than k_m candidates supporting the minority are elected in the Parliament, it is always possible to find equilibria in the alternative (out-of-equilibrium) voting subgames such that any deviation by a party leader would reduce the number of seats obtained by the party.

Example 1 ($\pi = 1$*:* $MR < k_m$)**.** Assume the following values for the parameters of the model

k_m	S	π
4	12	1

Thus, there are 4α minority voters and 8α majority ones. Consider a subgame where $n_m^A = n_m^B = 1$. That is, party leaders A and B instructed exactly one candidate in their parties to support x_m. Let voters equally split among the two parties: 2α voters in the minority and 4α voters in the majority vote for each of them (see Figure 1.1). Under this voting strategy, $S^A = S^B = 6$. Now assume party B's leader deviates and instructs $\tilde{n}_m^B > 1$ candidates on the list to support the minority. Equilibrium voting behavior in this subgame is such that all voters in the majority vote for A. Thus, B obtains at most all the votes of the minority, i.e. $4 < S^B$. Thus, this is not a profitable deviation for party B. The same reasoning applies to party A, proving that there exist an equilibrium where less than k_m candidates supporting the minority are elected under open list.

The conclusions arising from this section are clear. Stronger party discipline can only decrease representation of minorities in Parliaments.

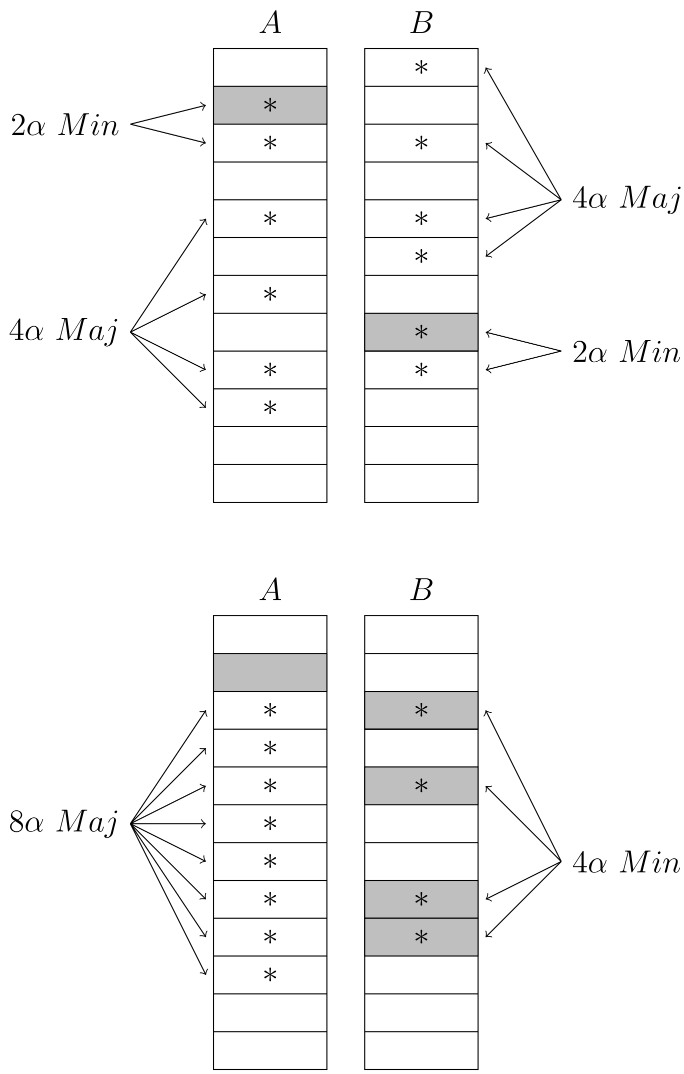

Figure 1.1: (Example 1) The top figure represents the voting subgame on the equilibrium path: one candidate in each list was instructed by party leaders to choose $x_c = x_m$ (gray rectangles). Voters distribute their votes as indicated by the arrows: the two groups equally split their votes between the two parties. Each party wins $S^A = S^B = 6$ seats (elected candidates are denoted by the symbol $*$). The bottom figure represents the voting equilibrium of a subgame where party leader B deviated and instructed 4 candidates to choose $x_c = x_m$. In this subgame, all majority voters vote for A and all minority voters vote for B. The seats won by B in this subgame are strictly less than those it was winning in the original one.

1.5.2 Voters' Behavior under Indifference

In this section, we examine the effects of relaxing the second part of Assumption 2. This assumption is essential only for the closed list case, so we focus our attention on this system[9]. The main consequence of relaxing the assumption is that multiple equilibria arise in the voting subgame where all candidates support position x_M. Indeed, since voters are completely indifferent, any strategy profile is an equilibrium. Depending on the voting behavior in each subgame, it is possible to construct equilibria in the complete game where $MR \in \{1, \ldots, k_m\}$. As before, we provide the intuition with an example (see Figure 1.2 for a graphical illustration).

Example 2 ($MR > 0$ *under closed list*)**.** Consider the subgame where $0 < n_m^A \leq k_m$ candidates in A decided to support the x_m and these candidates are the first ones appearing on party A's list. Formally, $x_c = x_m$ for all $c \in \{a_1, \ldots, a_{n_m^A}\}$, $x_c = x_M$ for all $c \in \{a_{n_m^A+1}, \ldots, a_S\}$. Furthermore, assume $n_m^B = 0$. Equilibrium voting behavior in this subgame is such that all majority voters vote for B and all minority voters vote for A. Thus, the n_m^A minority candidates in A are elected. Assume these candidates deviate to x_M, so that the new subgame is such that $x_c = x_M$ for all $c \in \mathcal{C}$. When Assumption 2.ii holds, voters will equally split among the two parties. The deviation therefore guarantees to A a strictly higher number of seats without endangering the election of the deviating candidates. If Assumption 2.ii does not hold, however, a voting profile where all voters vote for B is an equilibrium of the new subgame. If candidates expect this to be the equilibrium arising after their deviation, they will never deviate. Thus, an equilibrium with $n_m^A > 0$ candidates supporting x_m exists.

1.5.3 Higher number of Parties

Conclusions about minority representation in first-past-the-post and proportional representation systems rely on the idea that, for a given cost of creating a party, party formation is more rewarding under proportional representation than first-past-the-post. Higher minority representation in the former is therefore associated with the possibility of creating *minority parties* that support the interests of this group.

Given that the number of parties is fixed in our model, our results might not seem to be directly comparable with those ones. This section shows that, for sufficiently small sizes of the minority, our conclusions do not change when we allow for the creation of one (or more) additional parties. Our focus will be on the systems where $MR = 0$ in equilibrium (i.e. closed list and open list with large π), as these are the ones that challenge standard conclusions. Let us begin with an example for closed list and consider the possibility that a third party is created. Our question is, can there be an equilibrium where this party supports the minority? Remark 2 shows that, if the minority is sufficiently small, the answer is always no.

[9] Removing assumption 2.ii would change the equilibrium strategy profile under open list with $\pi \geq 2(k_m + 1)$, but would not alter the fact that only candidates supporting the majority would be elected in equilibrium under this system.

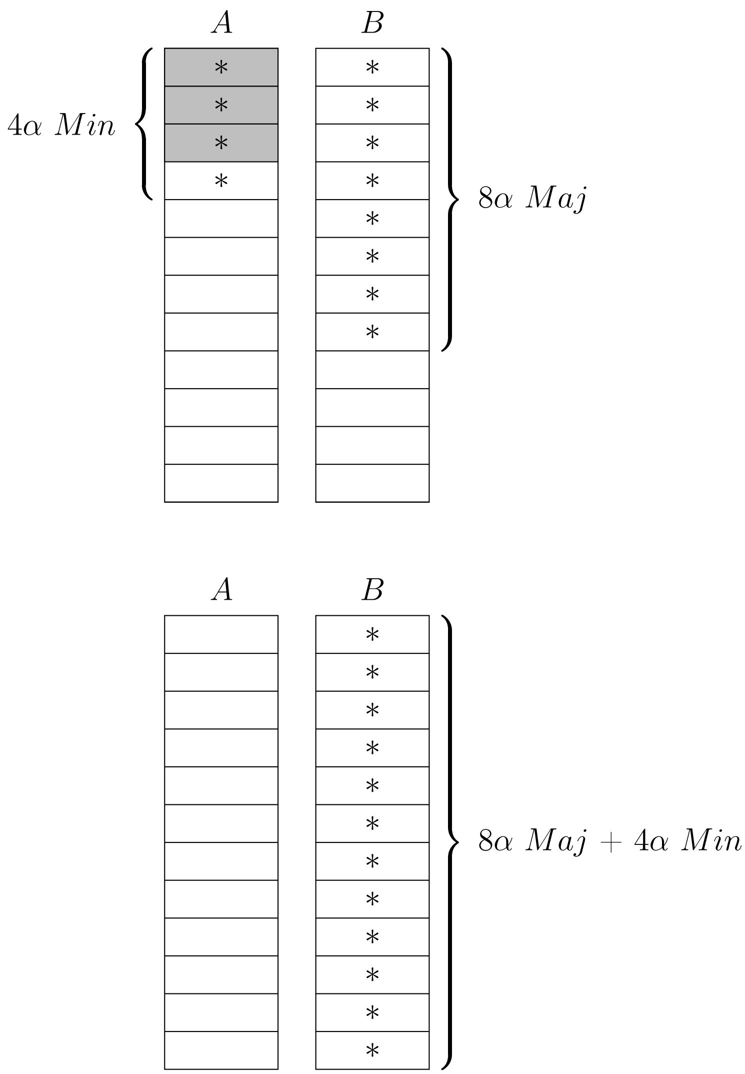

Figure 1.2: (Example 2) As for Figure 1.1, we set $k_m = 4$ and $S = 12$. The top figure represents the voting subgame on the equilibrium path: $x_c = x_m$ for all $c \in \{a_1, a_2, a_3\}$, while all other candidates support the majority. All minority voters all vote for party A and all majority voters all vote for B. All the symbols and colors should be interpreted as in Figure 1.1. The bottom figure represents an alternative subgame where all candidates supporting x_m in A deviated to x_M. All voters voting for party B is an equilibrium of this subgame.

Remark 2 (*Closed list: creation of a third party*)**.** Let $k_m < S/3$ and assume a third party C is created. Furthermore, assume it contains candidates supporting the minority. Then, three possible situations can occur:

i) only C contains candidates supporting the minority: $n_m^C > 0$, $n_m^A = n_m^B = 0$;

ii) C and another party, say A, contain candidates supporting the minority, B fully supports the majority: $n_m^C > 0$, $n_m^A > 0$, $n_m^B = 0$;

iii) all parties A, B and C contain candidates supporting the minority: $n_m^P > 0$ for all $P \in \{A, B, C\}$.

Consider case i) first. By Assumption 2, majority voters will equally divide their votes between A and B, while C will receive the votes of the minority. Thus, $S^C = k_m$. If all the n_m^C minority candidates in C deviated to x_M, then each party would get $S/3 > k_m$ seats. Thus, i) cannot happen in equilibrium. In case ii), party B is the only party that obtains the votes of the majority. The other two parties compete for the αk_m votes of the minority. This implies that $S^P \leq k_m/2$ for at least one $P \in \{A, C\}$. Deviating to x_M is again a profitable deviation for the n_m^P minority candidates in P, as it would imply that party P gets half of the votes of the majority and $(S - k_m)/2 > S^P$ seats. Finally, let us consider case iii). Notice that in this case there must exist a party that obtains $S^P \leq S/3$ seats. If minority candidates in this party deviate to x_M, the party would obtain all the votes of the majority and it would win $S - k_m > S/3$ seats. Thus, no equilibrium can exist where C contains candidates supporting the minority.

To conclude, assume all candidates support the majority. Then, $S^P = S/3$ for all P. Any deviation by a coalition of politician in party P would lead to a seat outcome equivalent to the one described in case i). That is, the party containing the deviating candidates would get $k_m < S/3$ seats. This proves that $n_m^P = 0$ for all $P \in \{A, B, C\}$ is the only equilibrium strategy for candidates.

A generalization of Remark 2 proves the following proposition.

Proposition 5. *If $k_m < S/\bar{Q}$, for some $\bar{Q} \in \mathbb{N}_{++}$, $\bar{Q} < S$, and the number of parties is $Q \leq \bar{Q}$, then $MR = 0$ in equilibrium under closed list proportional representation.*

The effects of introducing endogenous party creation will be investigated in future research. Notice that Proposition 5 already allows us to formulate conjectures about the possible outcomes: a party containing minority candidates could be created when many other parties already exist in the country. If party creation is costly, however, this might never be the case in equilibrium.

The analysis of open list proportional representation with large π in presence of many parties would also require a separate analysis. The general intuition, however, should not be different from the one proposed in the benchmark model: if the majority is sufficiently large, then these voters will always be able to give many votes to many candidates in all the parties, preventing the election of any minority candidate.

1.6 Ideology and FPTP

In this section, we show how our model can be extended to a more general one, where candidates are not *ex ante* identical. For example, they might belong to leftist or rightist parties and therefore be associated by voters to a specific ideological position. A straightforward way to extend our results to this case is to assume the existence of two parties for each ideology and apply our model to the interaction between them.

Formally, assume there now exist four parties, A_1, A_2, B_1, B_2. As before, let each of them be composed of S candidates. Furthermore, let the policy space be bi-dimensional. More precisely, assume there exists a second policy dimension $y \in \{y_A, y_B\}$, so that the policy space is now the the set $\{(x, y) \in \{x_M, x_m\} \times \{y_A, y_B\}\}$. Let candidates be *ex ante* differentiated over y. Denoting by y_c candidate c's position on y, we assume

$$y_c = \begin{cases} y_A & \text{if } c \in A_1 \cup A_2 \\ y_B & \text{if } c \in B_1 \cup B_2 \end{cases}$$

In other words, candidates are bound to follow their party's line on dimension y and compete as before by selecting their position on dimension x. We now distinguish four groups of voters in the population, which we denote by Min_A, Min_B, Maj_A and Maj_B, with $|Min_A| < |Maj_A|$, $|Min_B| < |Maj_B|$ and $|Min_A| + |Maj_A| = |Min_B| + |Maj_B|$. As before, they care about the composition of the Parliament. Their utility function now also depends on candidates' position over y.

$$U_i(W) = \sum_{c \in W} u_i(x_c, y_c)$$

Our benchmark model can be straightforwardly extended to this case if we assume that $u_i(x_c, y_c)$ represents the following preferences over candidate c's position on (x, y),

Min_A	Min_B	Maj_A	Maj_B
(x_m, y_A)	(x_m, y_B)	(x_M, y_A)	(x_M, y_B)
(x_M, y_A)	(x_M, y_B)	(x_m, y_A)	(x_m, y_B)
(x_m, y_B)	(x_m, y_A)	(x_M, y_B)	(x_M, y_A)
(x_M, y_B)	(x_M, y_A)	(x_m, y_B)	(x_m, y_A)

In words, voters value dimension y more than dimension x: a candidate proposing their favorite position on y will always be preferred to one proposing the alternative position on that dimension, independently of her choice on x. Clearly, this implies that voters preferring y_A will never vote for candidates in $B_1 \cup B_2$ and voters supporting y_B will never vote for candidates in $A_1 \cup A_2$. The model therefore reduces to two separate and identical sub-models where two parties (A_1, A_2 and B_1, B_2) compete on x. This is precisely the framework of our benchmark case and the results provided before can be replicated here.

Under this new framework, it is interesting to compare minority representation under proportional representation and under first-past-the-post systems (FPTP). We here consider FPTP with and without primaries. Assume the S seats are allocated by majority rule in S single member districts. There are two national parties A and B, represented by a candidate

a_s and b_s in each district s. As before, candidates in A and B support party position y_A and y_B, respectively, on dimension y and they can freely choose where to stand on x. Given that the election of one candidate is completely independent of the behavior of other candidates of the same party, we can simply assume that politicians care about winning the seat in their own district (i.e., their utility does not depend on the total number of seats obtained by their parties). The population of voters is identical to the one we just described for proportional representation, but is replicated S times across the different districts[10].

We here consider two different ways of selecting the candidates that will represent party P at the election. Let (a_s^1, a_s^2) and (b_s^1, b_s^2) be two couples of politicians competing to become the candidate representing party A and B, respectively, for district s.

FPTP with primaries. *Before elections, voters within each district select the candidate that will represent the party. We assume that only voters in $Min_P \cup Maj_P$ can vote at the primaries of party P.*

FPTP without primaries. *Party leaders select the candidates in order to maximize the probability of winning the seat.*

The timing of the events is such that politicians decide their position on x before candidate selection takes place. Our first result for FPTP shows that, if no primary is held, there always exist equilibria with $MR > 0$.

Proposition 6. *If candidate selection is done by party leaders, $MR \in \{0, 1, \ldots, S\}$ in equilibrium. That is, there can exist situation where the minority is strongly over-represented.*

Proof. Consider a generic district s. At the voting stage, only two candidates a_s and b_s compete for the seat. By our assumptions on voters' preferences, Maj_A and Min_A voters will vote for a_s, while Maj_B and Min_B voters will vote for b_s. Thus, each candidate wins the seat with probability one-half, independently of their position on x. Now consider candidate selection stage. Given that the seat is won with probability one-half, independently of candidate's type, party leaders are indifferent about which politician they should select. A strategy that selects a politician supporting x_m (whenever such a politician exists) is therefore a legitimate equilibrium strategy. At the first stage, if politicians expect that party leaders will select a candidate supporting x_m, they will both choose that position. Thus, a candidate supporting the minority could always be elected in equilibrium. Extending this result to all S districts, it is easy to see that MR can take any value in between 0 and S. □

If primaries are held, instead, there can be no equilibrium where $MR > 0$.

Proposition 7. *If candidate selection is done through primaries, no minority candidate will be elected in equilibrium (i.e. $MR = 0$).*

---

[10]Our conclusions hold under more general assumptions about groups' sizes and preferences. The same results were obtained using a probabilistic voting model. Furthermore, these are in line with Hirano et al. (2013)'s analysis of public good provision and targeting in presence of primaries. The modeling strategy we present here was chosen for consistency with the rest of the paper.

Proof. Consider a generic district s. As before, independently of their position on x, candidates a_s and b_s win with probability one-half. In this specific district, voters' expected utility is therefore

$$EU_i(x_{a_s}, x_{b_s}) = \frac{1}{2}u_i(x_{a_s}, y_A) + \frac{1}{2}u_i(x_{b_s}, y_B)$$

At the primaries stage, all voters in the majority strictly prefer candidates supporting x_M. More precisely,

$$EU_i(x_M, x_{b_s}) > EU_i(x_m, x_{b_s}), \text{ for all } x_{b_s}$$

for all $i \in Maj_A$, and

$$EU_i(x_{a_s}, x_M) > EU_i(x_{a_s}, x_m), \text{ for all } x_{a_s}$$

for all $i \in Maj_B$. Thus, in equilibrium, Maj_A and Maj_B voters will always vote for a candidate supporting x_M (if there exists one). Since they are the majority in the population, they will always be able to elect their favorite candidate. At the first stage, choosing x_M is a dominant strategy for candidates. Therefore, no candidate will ever choose x_m in equilibrium. Since the same reasoning applies to all districts, $MR = 0$. □

The striking implication of Proposition 6 and 7 is that increasing voters' control on candidate selection can lead to opposite results under first-past-the-post and proportional representation. Under the former, minority representation decreases when primaries are introduced. Under proportional representation, instead, as long as the number of preferential votes that can be expressed by voters is not too large, minority representation increases. More generally, our analysis clearly shows that the simple distinction between first-past-the-post and proportional representation is not enough to capture the patterns of minority representation in Parliaments. Minorities can definitely be better represented in proportional representation systems than in first-past-the-post with primaries; however, first-past-the-post without primaries could guarantee a potentially much higher minority representation than any form of proportional representation considered here[11].

1.7 Conclusions

The paper provides the first theoretical study of proportional representation system that considers how seats are allocated within party-lists and the behavior of candidates belonging to them. We compare closed list proportional representation, where voters can only cast a vote for a party, to open list proportional representation, where instead they are also allowed to express preferential votes for candidates. Within this second system, we also investigate the effect of increasing the number of preferential votes that can be cast.

Our final goal is to examine the effect of these electoral systems on representation of minorities within Parliaments. We show that allowing voters to cast preferential votes for candidates can have different effects on minority representation. Under closed list, where

[11] Our comparison between first-past-the-post and proportional representation systems is related to Huber (2012)'s analysis of ethnic voting in the two systems. Contrary to common beliefs, the author finds that ethnic voting is higher in first-past-the-post than in proportional representation.

no vote can be cast, no candidate representing the minority will be elected in Parliament. Perfectly proportional representation of minorities is achieved under open list, when voters can cast a limited number of preferential votes. Minorities, however, can never be over-represented: increasing the number of preferential votes that can be cast only harms these voters. Most surprisingly, when this is too high, open list does not perform better than closed list in terms of minority representation.

The possibility to cast preferential votes for individual candidates can be interpreted as a form of control for voters on the selection of elected politicians. Our conclusions indicate that such control has a non-monotonic effect on minority representation. No candidate supporting the minority is elected when either the control is too low (closed list) or it is too high (open list with many preferential votes). Moderate degrees of control are instead those that maximize the representation of minorities in Parliaments.

Our results provide clear indications about the electoral system a country should choose, depending on whether its goal is to protect or prevent the representation of a minority. For instance, if a country wants to foster the representation of an ethno-linguistic group, it should choose an open list proportional system and set the number of preferences that can be cast by voters at most equal to the relative size of the group in the population. If, on the contrary, the goal was to prevent politicians from targeting specific groups in the population with resources subtracted from public good provision, then the system to be chosen should be either closed list or open list with many preferential votes. Our model can also be adapted to fit other applications. The main one we have in mind is women representation in Parliaments. In particular, future research could address the effectiveness of gender quotas under the different forms of proportional representation (see Esteve-Volart and Bagues (2012) for an empirical study of this in Spain).

Quite interestingly, higher levels of party discipline can only reduce minority representation. In an extension of the paper, we analyze a situation where party leaders can force their candidates to support a particular policy position. We show that, under strong party discipline, new equilibria emerge under open list with few preferential votes. These equilibria involve strictly lower levels of minority representation.

In addition to the main results, the paper provides other interesting conclusions. We compare open and closed list proportional representation with first-past-the-post with and without primaries. Primaries can be considered as a way to increase voters' control on candidate selection under first-past-the-post and are therefore the equivalent of open list under this electoral rule. However, the effect of such control is different under the two systems. In the paper, we show that primaries have a clear negative effect on minority representation, as no candidate supporting the minority will be elected in Parliament when they are held. If instead candidate selection is done by party leaders, minority representation can be very high, much higher than its perfectly proportional level. First-past-the-post without primaries could therefore guarantee a much higher degree of minority representation than any form of proportional representation. The result strongly contradicts standard beliefs about the two electoral systems.

1.A Appendix

Proof of Proposition 2

Proof. By our discussion in the main part of the paper, we know that if $n_m^P \geq k_m$ for some party P, then at least k_m candidates supporting the minority will be elected in equilibrium. Similarly, if $n_m^{P'} \leq k_m$ for some party P' (i.e. there are at least $S - k_m$ candidates supporting x_M in party P'), then at least $S - k_m$ majority candidates will be elected in equilibrium. As a direct consequence, if $n_m^P \geq k_m$ for some $P \in \{A, B\}$ and $n_m^{P'} \leq k_m$ for $P' \neq P$, exactly k_m minority candidates and $S - k_m$ majority ones will be elected. More precisely, the voting equilibrium will be such that all minority voters vote for P and all majority voters vote for P'. Voting for a different party can never increase the number of favorite candidates that are elected, for any of the two groups. For example, if minority voters decided to vote for P', they could elect *at most* $n_m^{P'} \leq k_m$ candidates.

To complete the proof, we need to show that no subgame where either $n_m^P < k_m$ or $n_m^{P'} > k_m$ for both $P \in \{A, B\}$ can be reached in equilibrium. Assume $n_m^P < k_m$ for both $P \in \{A, B\}$. Notice that in equilibrium, there must be a party where (weakly) more than $S/2$ candidates are not elected. Let this party be A. Among the non-elected candidates in it, consider those supporting the majority and notice that there are at least $S/2 - n_m^A$ of them. If $k_m - n_m^A < S/2 - n_m^A$ deviate to x_m a new subgame will arise where $\tilde{n}_m^A = n_m^A + (k_m - n_m^A) = k_m$. By our discussion above, all these candidates will be elected in equilibrium. By a symmetric reasoning, if $n_m^P > k_m$ for both $P \in \{A, B\}$ then a profitable deviation would exist for some candidates supporting x_m that are not elected. $\square$

Proof of Proposition 4

For expositional convenience, we divide the proof of Proposition 3 in different Lemmas. Lemma 3 shows that if there are sufficiently many candidates supporting the majority in the two parties, no minority candidate will be elected. Lemmas 4, 5 and 6 describe the equilibria in the other voting subgames and prove how, in each of them, a profitable deviation for some candidates exists. Figure 1.3 provides a graphical representation of the different subgames.

For all the Lemmas, we assume that $\pi \geq 2(k_m + 1)$ and $k_m < S/3 - 1$. As before, n_M^P and n_m^P denote the number of candidates in party P who choose to support x_M and x_m, respectively. We denote by v_M^P and v_m^P the number of majority and minority voters, respectively, who vote for party P and set $v_M = (v_M^A, v_M^B)$, $v_m = (v_m^A, v_m^B)$. To avoid confusion, it will turn useful to denote the number of seats party P obtains under voting profile (v_M, v_m) by $S^P(v_M, v_m)$.

Lemma 3. *If a subgame is such that* $n_M^P \geq 2k_m$ *for all* $P \in \{A, B\}$ *and* $n_M^A + n_M^B \geq S + k_m$, *then* $MR = 0$ *in equilibrium.*

Proof. Let $n_M^P \geq 2k_m$ for all $P \in \{A, B\}$ and $n_M^A + n_M^B \geq S + k_m$. The following voting profile is the only equilibrium of the game.

(a) If $n_M^A = n_M^B$, then the two groups equally split their votes between the two parties;

(b) If $n_M^P > n_M^{P'}$ for $P \in \{A, B\}$, $P \neq P'$, then (v_M, v_m) is such that

$$\begin{cases} v_M^P = \alpha \min\{n_M^P, (S - k_m)\} - \alpha k_m - 1 \\ v_M^{P'} = \alpha(S - k_m) - v_M^P \end{cases} \tag{1.3}$$

and

$$\begin{cases} v_m^P = 0 \\ v_m^{P'} = \alpha k_m \end{cases}$$

Start by assuming that $n_M^A = n_M^B$ and let majority voters divide their votes across the two parties. Notice that since $n_M^A + n_M^B \geq S + k_m$, it must be that

$$n_M^P \geq \frac{S + k_m}{2} = \frac{S - k_m}{2} + k_m \tag{1.4}$$

for both $P \in \{A, B\}$. Each of the $v_M^P = \alpha(S - k_m)/2$ majority voters voting for P can cast a preferential vote for $\min\{\pi, n_M^P\}$ majority candidates in the party. Thus, these voters can cast a total amount of preferential votes equal to $\alpha \min\{\pi, n_M^P\}(S - k_m)/2$. Denote by ν_M^P the number of majority candidates in P that can obtain at least $\alpha k_{min} + 1$ preferential votes. Then, by our assumptions on π and k_{maj},

$$\nu_M^P = \left\lfloor \frac{\alpha(S - k_m) \min\{\pi, n_M^P\}}{2(\alpha k_m + 1)} \right\rfloor \geq \frac{S - k_m}{2} + k_m \tag{1.5}$$

Now assume the minority votes for any of the two parties $P \in \{A, B\}$, so that $v_m^P = \alpha k_m$ and

$$S^P(v_M, v_m) = \frac{S - k_m}{2} + k_m$$

The highest number of preferential votes that a minority candidate can obtain is αk_m. By (1.4) and (1.5), there are at least $S^P(v_M, v_m)$ majority candidates in party P that obtain at least $\alpha k_m + 1$ preferential votes, so that no minority candidate will be elected. As this holds for both $P \in \{A, B\}$, the minority will never be able to elect any of its favorite candidates. Since $n_M^A = n_M^B$, our tie-breaking rule then implies that they will equally split their votes between the two parties. Each of the parties will therefore obtain $S/2$ seats and only majority candidates will be elected. Uniqueness of this equilibrium is a direct consequence of our tie-breaking rule.

Now assume without loss of generality that $n_M^B > n_M^A$ and let voters vote as described in (b), i.e.

$$\begin{cases} (v_M^A, v_M^B) = (\alpha(S - k_m) - v_M^B, \alpha \min\{n_M^B, (S - k_m)\} - \alpha k_m - 1) \\ (v_m^A, v_m^B) = (\alpha k_m, 0) \end{cases}$$

The number of votes obtained by each party P, denoted by v^P, is therefore equal to

$$\begin{aligned} v^A &= v_M^A + v_m^A = \alpha(S - \min\{n_M^B, (S - k_m)\} + k_m) + 1 \\ v^B &= v_M^B + v_m^B = \alpha(\min\{n_M^B, (S - k_m)\} - k_m - 1) + \alpha - 1 \end{aligned}$$

And, by the Largest Remainder Method, we have

$$S^A(v_M, v_m) = S - \min\{n_M^B, (S - k_m)\} + k_m$$
$$S^B(v_M, v_m) = \min\{n_M^B, (S - k_m)\} - k_m$$

Now notice that $n_M^B > S^B(v_M, v_m)$ and, since $n_M^A > S + k_m - n_M^B$, it must also be that $n_M^A > S^A(v_M, v_m)$. Furthermore, the number of majority candidates in each party that can be assigned at least $\alpha k_m + 1$ votes is

$$\begin{aligned}\nu_M^A &= \left\lfloor \frac{[\alpha(S - \min\{n_M^B, (S - k_m)\}) + 1] \min\{\pi, n_M^A\}}{\alpha k_m + 1} \right\rfloor \\ &\geq S - \min\{n_M^B, (S - k_m)\} + k_m \end{aligned} \tag{1.6}$$

$$\begin{aligned}\nu_M^B &= \left\lfloor \frac{[\alpha \min\{n_M^B, (S - k_m)\} - \alpha k_m - 1] \min\{\pi, n_M^B\}}{\alpha k_{min} + 1} \right\rfloor \\ &\geq \min\{n_M^B, (S - k_m)\} \end{aligned} \tag{1.7}$$

As before, then, $\nu_M^P \geq S^P(v_M, v_m)$ for both $P \in \{A, B\}$ and no minority candidate will be assigned a seat.

Consider a deviation by the minority now and assume they all vote for party B. Denote this strategy by $\tilde{v}_m = (0, \alpha k_m)$. Under the new strategy profile,

$$S^A(v_M, \tilde{v}_m) = S - \min\{n_M^B, (S - k_m)\}$$
$$S^B(v_M, \tilde{v}_m) = \min\{n_M^B, (S - k_m)\}$$

As before, $n_M^A > S^A(v_M, \tilde{v}_m)$ and it must also be that $n_M^B > S^B(v_M, \tilde{v}_m)$. Furthermore, (1.6) and (1.7) imply $\nu_M^A > S^A(v_M, \tilde{v}_m)$ and $\nu_M^B > S^B(v_M, \tilde{v}_m)$. As no minority candidate would be elected under $\tilde{v}_m$ and this strategy prescribes to vote for the party with the largest number of majority candidates, the deviation is not profitable.

The strategy for the majority is the one that assigns the minimum amount of votes to A that is sufficient to prevent the election of any minority candidate and it is therefore the unique equilibrium strategy under our tie-breaking rule. □

Before turning to Lemma 4, let us consider one implication of Lemma 3.

Corollary 1. *If $n_M^P > n_M^{P'}$ and*

(i) $n_M^{P'} = S + k_m - \min\{n_M^P, (S - k_m)\}$, then all $n_M^{P'}$ majority candidates in P' win a seat with probability one.

(ii) $n_M^{P'} \geq 2k_m$ and $n_M^A + n_M^B = S + k_m$, then all n_M^P majority candidates in P win a seat with probability

$$\lambda(n_M^{P'}) = \frac{S - n_M^{P'}}{S + k_m - n_M^{P'}}$$

Proof. Consider point (i) first. The assumptions on n_M^P and $n_M^{P'}$ imply that $n_M^P + n_M^{P'} \geq S + k_m$, so that Lemma 3 applies. In particular, as $n_M^P > n_M^{P'}$, equilibrium voting behavior is such that $S^{P'} = S - \min\{n_M^B, (S - k_m)\} + k_m = n_M^{P'}$. Thus, all the $n_M^{P'}$ majority candidates in P' win a seat for sure.

For point (ii), notice that $n_M^{P'} \geq 2k_m$ and $n_M^A + n_M^B = S + k_m$ imply that $n_M^P \leq S - k_m$. As $n_M^P > n_M^{P'}$, therefore, $S^P = n_M^P - k_m$ in equilibrium. Thus, each of the n_M^P majority candidates in P wins with probability

$$\lambda(n_M^{P'}) = \frac{S^P}{n_M^P} = \frac{n_M^P - k_m}{n_M^P} = \frac{S - n_M^{P'}}{S + k_m - n_M^{P'}}$$

□

Lemma 4. *If a subgame is such that $n_M^P \geq \max\{S - 2k_m, S - 3k_m + n_M^{P'}\}$ for some $P \in \{A, B\}$ and $n_M^{P'} < 2k_{min}$ for $P' \neq P$, the subgame is never reached in equilibrium.*

Proof. In this subgame, party P' will obtain at most $2k_m$ seats. Assume $v_m^{P'} = \alpha k_m$ and $v_m^P = 0$. The best response by the majority must be one of the two following strategies:

$$\begin{cases} v_M^{P'} = \alpha k_m + 1 \\ v_M^P = \alpha(S - 2k_m) - 1 \end{cases} \qquad \begin{cases} \tilde{v}_M^{P'} = 0 \\ \tilde{v}_M^P = \alpha(S - k_m) \end{cases}$$

Giving more than $\alpha k_m + 1$ votes to P' would imply that $S^{P'} > 2k_m$ and, as there are at most $n_M^{P'} < 2k_m$ majority candidates in P', some minority candidates would be elected for sure. Assigning to P' an amount of votes strictly in between zero and $\alpha k_m + 1$ would not allow the majority to give enough preferential votes to their favorite candidates. Under v_M, $S^{P'}(v_M, v_m) = 2k_m$ and $S^P(v_M, v_m) = S - 2k_m$. The number of elected majority candidates will therefore be

$$n_M^{P'} + \min\{n_M^P, S - 2k_m\}$$

Under $\tilde{v}_M$, instead $S^{P'}(\tilde{v}_M, v_m) = k_m$ and $S^P(\tilde{v}_M, v_m) = S - k_m$ and $\min\{n_B^P, S - k_m\}$ majority candidates will be elected in P. Thus, v_M is a best response to the strategy adopted by minority voters if and only if

$$n_M^{P'} + \min\{n_M^P, S - 2k_m\} \geq \min\{n_B^P, S - k_m\} \tag{1.8}$$

Under the assumptions on n_M^P, (3.2) is verified if and only if $n_M^{P'} \geq k_m$. Thus, party P' will win $2k_m$ seats if $n_M^{P'} \geq k_m$, it will only win k_m seats otherwise. Furthermore, the number of elected minority candidates cannot be greater than k_m: as soon as the majority votes for P', only $2k_m - n_M^{P'} \leq k_m$ minority candidates will be elected. Incidentally, notice that if the minority did not vote for P', no minority candidate would be elected in this party, so k_m is indeed the upper bound for the equilibrium number of elected minority candidates. Now notice that

$$n_m^{P'} = S - n_M^{P'} \geq S - 2k_m$$

Among these $n_m^{P'}$ candidates, there must be at least $S + k_m - \min\{n_M^P, (S - k_m)\} - n_M^{P'}$ of them that win with probability lower than one. If these candidates deviate to x_M, a new subgame will arise where $\tilde{n}_M^{P'} = S + k_m - \min\{n_M^P, (S - k_m)\}$ and $n_M^P > n_M^{P'}$. By Corollary 1 (point (i)), these candidates will be elected for sure. □

Corollary 2. *If $n_M^P = S - 3k_m + n_M^{P'}$ and $k_m \leq n_M^{P'} \leq 2k_m$, then all n_M^P majority candidates in P win with probability*

$$\tilde{\lambda}(n_M^{P'}) \geq \frac{S - 2k_m}{S - 3k_m + n_M^{P'}}$$

Proof. By Lemma 4, the assumptions on n_M^P and $n_M^{P'}$ are such that at least $\alpha(S - 2k_m)$ majority voters will vote for party P. Thus, this party will obtain at least $S - 2k_m$ seats and each of the $n_M^P = S - 3k_m + n_M^{P'}$ majority candidates in P receiving a positive amount of preferential votes will win with probability $\tilde{\lambda}_M^P$. □

Lemma 5. *If a subgame is such that $n_M^P < S - 2k_m$ for both $P \in \{A, B\}$ and $n_M^A + n_M^B \leq S - k_m$, then the subgame will never be reached in equilibrium.*

Proof. Assume $n_M^P \leq S - 2k_m$ for both $P \in \{A, B\}$ and $n_M^A + n_M^B \leq S - k_m$. In this subgame, exactly $n_M^A + n_M^P$ majority candidates are elected in equilibrium. To see why, let us begin by assuming that $n_M^B > n_m^A$. The case where $n_M^B < n_m^A$ is perfectly symmetric. Consider a voting profile by the majority such that

$$v_M^A = \max\{\alpha n_M^A, \alpha k_m + 1\}$$

and $v_M^B = \alpha(S - k_m) - v_M^A$. Under this strategy,

$$\begin{aligned} S^A(v_M, v_m) &\geq \max\{n_M^A, k_m\} \geq n_M^A \\ S^B(v_M, v_m) &\geq (S - k_m) - \max\{n_M^A, k_m\} \geq n_M^B \end{aligned}$$

for any strategy v_m played by the minority. Furthermore, one can check that the number of majority candidates that can be given at least $\alpha k_m + 1$ votes in the two parties is

$$\begin{aligned} \nu_M^A &= \left\lfloor \frac{\max\{\alpha n_M^A, \alpha k_m + 1\} \min\{\pi, n_M^A\}}{\alpha k_m + 1} \right\rfloor \geq n_M^A \\ \nu_M^B &= \left\lfloor \frac{[\alpha(S - k_m) - \max\{\alpha n_M^A, \alpha k_m + 1\}] \min\{\pi, n_M^B\}}{\alpha k_m + 1} \right\rfloor \geq n_M^B \end{aligned}$$

Thus, the majority is able to assign enough seats to each party and enough votes to each of the $n_M^A + n_M^B$ majority candidates to elect them independently of the behavior of the minority.

Now assume $n_M^A = n_M^B \leq (S - k_m)/2$. If the majority equally divides their votes between A and B, they will be able to assign $(S - k_m)/2$ seats to each of them. Furthermore,

$$\begin{aligned} \nu_M^A &= \left\lfloor \frac{\alpha(S - k_m) \min\{\pi, n_M^A\}}{2(\alpha k_m + 1)} \right\rfloor \geq n_M^A \\ \nu_M^B &= \left\lfloor \frac{\alpha(S - k_m) \min\{\pi, n_M^B\}}{2(\alpha k_m + 1)} \right\rfloor \geq n_M^B \end{aligned}$$

and the $n_M^A + n_M^B$ candidates will be able to receive enough preferential votes to be elected for sure.

An immediate implication of our discussion is the following: in equilibrium, it cannot be that, for some $P \in \{A, B\}$, $n_M^P < S - 2k_m$ when $n_M^A + n_M^B \leq S - k_m$. In this case, indeed, a minority candidate in P could always deviate to x_M and be elected for sure. The same reasoning holds for the case where $n_M^P \leq S - 2k_m$ and $n_M^A + n_M^B < S - k_m$.

Notice that if $n_M^P = S - 2k_m$, then $n_M^{P'} \leq k_m$ and a profitable deviation would exist for a minority candidate because of Lemma 4. To conclude the proof, we only need to show that no subgame where $n_M^P \leq S - 2k_m$ and $n_M^A + n_M^B = S - k_m$ can be reached in equilibrium. Notice that the maximum number of minority candidates elected in equilibrium is $S - n_M^A + n_M^B = k_m$. We distinguish three possible cases:

- *Case 1*: $n_M^P \geq (S + k_m)/2$: by Corolloary 1, if $S + k_m - n_M^{P'}$ minority candidates in P' deviate to x_M, they will win with probability one;

- *Case 2*: $n_M^P < (S + k_m)/2$ and $n_M^{P'} \leq 2k_m$: by Corollary 2, if $S - 3k_m + n_M^{P'} - n_M^P$ minority candidates in P deviate to x_M, they will win with probability $\tilde{\lambda}(n_M^{P'})$. By Corollary 1, if $S + k_m - n_M^{P'}$ minority candidates in P' deviate to x_M, they will win with probability $\lambda(n_M^P)$. Denote by ν_m^P and $\nu_m^{P'}$ the number of minority candidates in P and P' that are assigned a positive amount of preferential votes and let S_m^P and $S_m^{P'}$ be the number of seats won by party P and P' that is assigned to minority candidates. A profitable deviation does not exist if the following conditions are all satisfied:

$$\begin{aligned}
&\frac{S_m^P}{\nu_m^P} \geq \tilde{\lambda}(n_M^{P'}) \quad \text{and} \quad n_m^P - \nu_m^P < S - 3k_m + n_M^{P'} - n_M^P \\
&\frac{S_m^{P'}}{\nu_m^{P'}} \geq \lambda(n_M^P) \quad \text{and} \quad n_m^{P'} - \nu_m^{P'} < S + k_m - n_M^{P'} \\
&S_m^P + S_m^{P'} \leq k_m
\end{aligned}$$

  The first condition says that the number of seats won by minority candidates in party P is enough to guarantee to some candidates a probability of winning at least equal to $\tilde{\lambda}(n_M^{P'})$ and that those who win with lower probability are not enough to deviate. The second condition imposes the equivalent requirements for minority candidates in P'. Finally, the last condition directly follows from the fact that the total number of seats won by minority candidates in the two parties cannot be greater than k_m. Substituting for $\tilde{\lambda}(n_M^{P'})$ and $\lambda(n_M^P)$, one can check that the three conditions are incompatible.

- *Case 3*: $n_M^P < (S + k_m)/2$ and $n_M^{P'} > 2k_m$: by Corollary 1, if $S + k_m - n_M^{P'}$ minority candidates in P' deviate to x_M, they will win with probability $\lambda(n_M^P)$. Similarly, if $S + k_m - n_M^P$ minority candidates in P deviate to x_M, they will win with probability $\lambda(n_M^{P'})$. As before, a profitable deviation does not exist if the following conditions are all satisfied:

$$\begin{aligned}
&\frac{S_m^{P'}}{\nu_m^{P'}} \geq \lambda(n_M^P) \quad \text{and} \quad n_m^{P'} - \nu_m^{P'} < S + k_m - n_M^{P'} \\
&\frac{S_m^P}{\nu_m^P} \geq \lambda(n_M^{P'}) \quad \text{and} \quad n_m^P - \nu_m^P < S + k_m - n_M^P \\
&S_m^P + S_m^{P'} \leq k_m
\end{aligned}$$

Substituting for $\lambda(n_M^P)$ and $\lambda(n_M^{P'})$, one can check that the three conditions are incompatible.

□

Lemma 6. *If a subgame is such that $n_M^P \leq S - 3k_m + n_M^{P'}$ for both $P \in \{A, B\}$ and $S - k_m < n_M^A + n_M^B < S + k_m$, then the subgame will never be reached in equilibrium.*

Proof. A Strong Nash equilibrium in pure strategy does not exist in this subgame. Without loss of generality, let $n_M^A \leq n_M^B$. Notice that $n_M^B \leq S - 3k_m + n_M^A$ and $S - k_m < n_M^A + n_M^B$ imply that $n_M^A > k_m$. Assume $v_m = (\alpha k_m, 0)$ and consider the best response by the majority. If

$$\begin{cases} v_M^A &= \max\{\alpha k_m + 1, \alpha(n_M^A - k_m)\} \\ v_M^B &= \alpha(S - k_m) - v_M^A \end{cases} \tag{1.9}$$

then

$$S^A(v_M, v_m) = \max\{2k_m, n_M^A\} \geq n_M^A$$
$$S^B(v_M, v_m) = S - \max\{2k_m, n_M^A\}$$

This strategy guarantees the election of all n_M^A majority candidates in A and of

$$\min\{n_M^B, S^B(v_M, v_m)\}$$

majority candidates in B. Notice that $v_M^A = \alpha k_m + 1$ if $n_M^A \leq 2k_m$. Our discussion in the proof of Lemma 4 has already proven that v_M is the best response to v_m if $n_M^A > k_m$. When $n_M^A \geq 2k_m$, this strategy is the only one that prevents the election of any minority candidate in A.

Consider the best response of the minority to v_M now. If these voters vote according to v_m, the number of elected minority candidates will be

$$\max\{S^A(v_M, v_m) - n_M^A, 0\} + \max\{S^B(v_M, v_m) - n_M^B, 0\}$$

Consider the alternative strategy $\tilde{v}_m = (0, \alpha k_m)$. Under this strategy,

$$S^A(v_M, \tilde{v}_m) = \max\{k_m, n_M^A - k_m\}$$
$$S^B(v_M, \tilde{v}_m) = S - \max\{k_m, n_M^A - k_m\}$$

and $\max\{S^B(v_M, \tilde{v}_m) - n_M^B, 0\}$ minority candidates will be elected in B. Minority voters will therefore play v_m if and only if

$$\max\{S^A(v_M, v_m) - n_M^A, 0\} + \max\{S^B(v_M, v_m) - n_M^B, 0\} > \max\{S^B(v_M, \tilde{v}_m) - n_M^B\} \tag{1.10}$$

If $n_M^A < 2k_m$, (3.15) is equivalent to

$$2k_m - n_m^A + \max\{S - 2k_m - n_M^B, 0\} > S - k_m - n_M^B$$

where the right hand side of the inequality is a direct consequence of $n_M^B \leq S - 3k_m + n_M^A$ and $n_M^A < 2k_m$. One can check that this condition is never satisfied. If $n_M^A \geq 2k_m$, then (3.15) becomes

$$\max\{S - n_M^A - n_M^B, 0\} > S + k_m - n_M^A - n_M^B$$

which, again, is never satisfied. The best response of the minority to v_M is therefore $\tilde{v}_m$.

Let us now consider the best response of the majority to $\tilde{v}_m$. Under $(v_M, \tilde{v}_m)$, the number of elected majority candidates is equal to

$$S^A(v_M, \tilde{v}_m) + \min\{S^B(v_M, \tilde{v}_m), n_M^B\}$$

Consider a strategy that assigns exactly αn_M^A votes to A, that is $\tilde{v}_M = (\alpha n_M^A, \alpha(S - k_m - n_M^A))$. Under this strategy and given the behavior of the minority,

$$S^A(\tilde{v}_M, \tilde{v}_m) = n_M^A$$
$$S^B(\tilde{v}_M, \tilde{v}_m) = S - n_M^A$$

and the number of elected majority candidates will be $n_M^A + \min\{S - n_M^A, n_M^B\}$. Then, $\tilde{v}_M$ will be preferred to v_M if

$$n_M^A + \min\{S - n_M^A, n_M^B\} > \max\{k_m, n_M^A - k_m\} + \min\{S - \max\{2k_m, n_M^A\}, n_M^B\} \quad (1.11)$$

If $n_M^A < 2k_m$, (3.16) becomes

$$n_M^A + \min\{S - n_M^A, n_M^B\} > k_m + \min\{S - 2k_m, n_M^B\}$$

which is always satisfied as $n_M^A > k_m$. If $n_M^A \geq 2k_m$ instead, we have

$$n_M^A + \min\{S - n_M^A, n_M^B\} > n_M^A - k_m + \min\{S - n_M^A, n_M^B\}$$

which is again always true. The best response to $\tilde{v}_m$ is therefore $\tilde{v}_M$. The non-existence of a Strong Nash equilibrium in pure strategy is now proven by noticing that the best response of the minority to $\tilde{v}_M$ is v_m. Under this strategy, indeed

$$S^A(\tilde{v}_M, \tilde{v}_m) = n_M^A + k_m$$
$$S^B(\tilde{v}_M, \tilde{v}_m) = S - n_M^A - k_m$$

and at least k_m minority candidates will be elected in A. If the minority played according to $\tilde{v}_m$ instead, they would elect

$$S - n_M^A + \min\{S - n_M^A, n_M^B\} < k_m$$

minority candidates.

Thus, if an equilibrium exists, it must be in mixed strategies. Now notice that if $n_M^P \geq (S + k_m)/2$ for some P, Corollary 1 guarantees that a profitable deviation exists for $S + k_m - n_M^{P'} - n_M^P$ minority candidates. If these candidates deviated to x_M, indeed, the new subgame would be one where $\tilde{n}_M^P = S + k_m - n_M^{P'}$ and $n_M^P > n_M^{P'}$ and all these candidates would be elected with probability one.

For all other cases, notice that, as $n_M^A + n_M > S - k_m$ the majority can always guarantee the election of $S - k_m$ candidates. This is a direct consequence of our discussion in Lemma 5. The type of deviation is identical to the one described in that Lemma. □

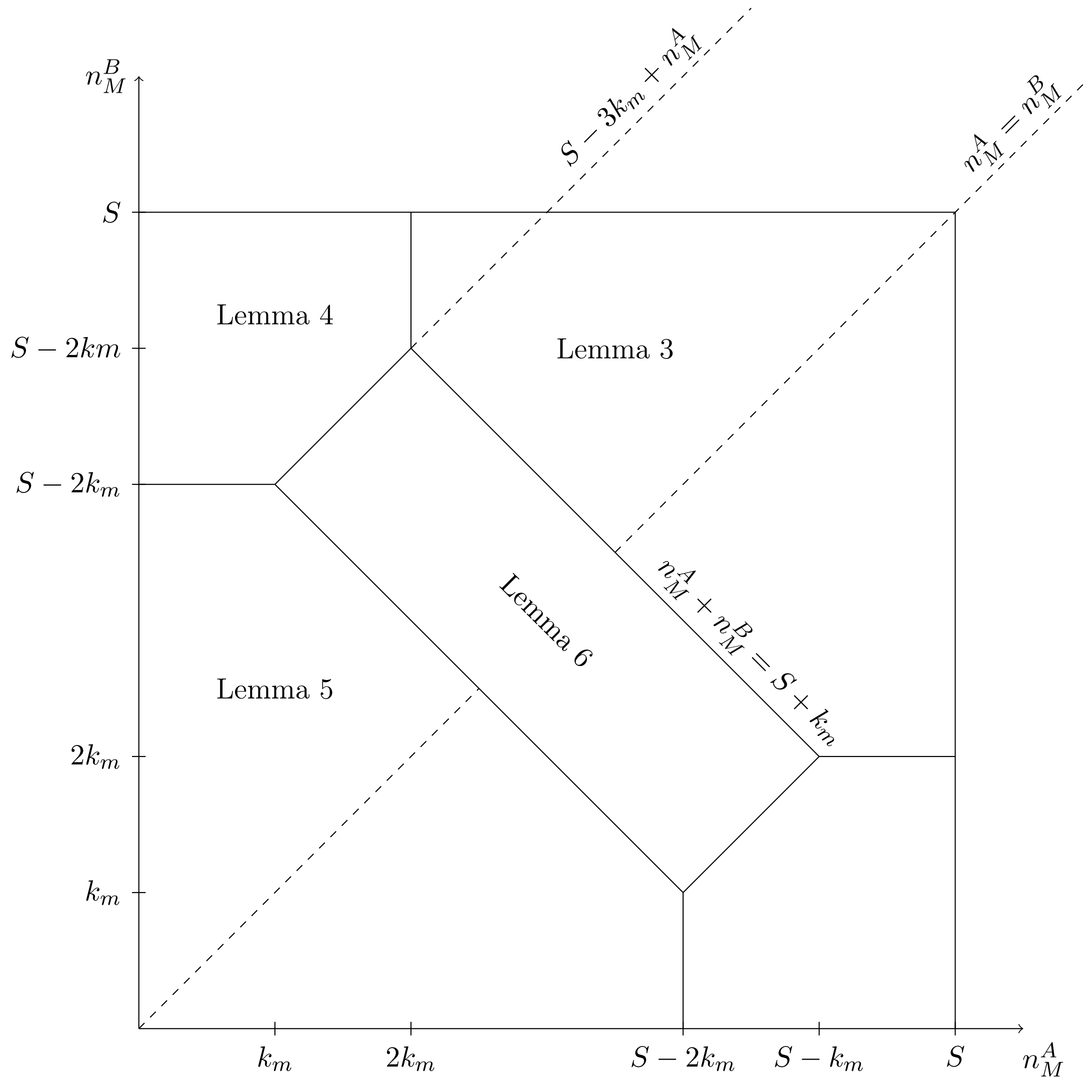

Figure 1.3: Subgames that can arise under open list proportional representation with $\pi \geq 2(k_m + 1)$ and $k_m < S/3 - 1$ as a function of the number of majority candidates in A and B, n_M^A and n_M^B.

Chapter 2

Why do Good Politicians Take Sub-Optimal Decisions?

Abstract: The paper provides an explanation for sub-optimal policy making decisions by incumbents. We show that electoral incentives can induce them to address less relevant issues, disregarding more important ones. This happens even if voters are perfectly informed about issues relevance and politicians are policy oriented. Our explanation relies on the negative correlation between issue importance and probability of solving it: for a given effort exerted by incumbents, less relevant issues guarantee higher probability of success. In equilibrium, the solution of these issues is anyway a positive signal about incumbents' type. Thus, whenever re-election is sufficiently valuable, politicians prefer not to take the risk of addressing more relevant issues.

Keywords: political accountability, elections, term-limits.
JEL Classification: D02, D72, D78.

2.1 Introduction

In all representative democracies, elections have a double objective: first, they allow to aggregate preferences to facilitate policy making decisions; second, they improve politicians' performances by making them accountable for their actions. In this paper, our attention is on the second goal. More precisely, we focus on the fact that, in many situations, this goal seems far to be reached. The observation of politicians' behavior, indeed, shows how they sometimes take decisions that are not optimal for the electorate. We here refer, in particular, to situations where they address issues that clearly look less relevant with respect to others.

For example, consider the case of Belgium. The central topic during the last three federal elections in the country was the design of the electoral district of Brussels: extra political rights of a minority of 150,000 French speaking citizens (over a total population of almost 11 million people) living in Flanders were constantly presented by politicians as a major issue, while, for example, the problem of a very low employment rate was almost ignored.

The Roma case in France constitutes another good example. During summer 2010, the country's political debate has been monopolized by the decision of dismantling 51 Roma camps and repatriate most of the Roma citizens living in them. A roaring campaign has been carried on by the Elysee Palace to explain how camps where "sources of illegal trafficking, profoundly shocking living standards, the exploitation of children for begging, prostitution and crime"[1]. French people positively responded to the initiative (around 80% of the citizens interviewed in a survey were in favor of it)[2] and president Sarkozy saw his poll ratings edge up. A closer look at the case of French Roma, however, shows how the gravity of the situation was clearly over-emphasized: Roma constitute between the 0.5% and 1% of French population and many of them had been peacefully and lawfully living in the country for many years. More importantly, although "the latest unemployment figure remained stubbornly high at 9.5%"[3], unemployment was completely absent from political debate.

In this paper, we provide an answer to the question of why, despite the presence of elections, politicians take decisions that are not optimal for voters. Our explanation relies on the negative correlation between importance of an issue and how difficult it is to solve it. We prove that less relevant issues will be addressed whenever they are *easier* to solve than more important ones, provided that re-election is *sufficiently* valuable for politicians. Indeed, if re-election is conditional on issue resolution, addressing problems that are more likely to be solved maximizes incumbents' probability of remaining in power. Most importantly, we show that this holds even if politicians care about issue resolution *per se*.

We provide a two-period political agency model that combines moral hazard and adverse selection. Before election, an incumbent politician is assigned the task of solving two issues. One of these issues is at the same time more important and more difficult to solve. We model these two characteristics by assuming that i) the solution of the issue guarantees higher utility to the electorate, but ii) for any given level of effort exerted by the politician, the probability of success is lower for this issue than for the other. The incumbent can be of two types: a *bad* one only wants to be elected for a second period; a *good* one cares about issue resolution as well. When elections are held, voters must decide whether to re-elect the incumbent or not. Their goal is to maximize the probability of having a good type in power in the second period. To do so, they must infer incumbent's quality from his actions in the past term. Effort is not observable and voters can only see whether an issue was solved or not.

We assume the existence of three states of the world: in one of them, both issues can be addressed and solved with positive probability; in the other two states, the probability of resolution for one of the issues is equal to zero, independently of the effort provided by the incumbent. The assumption aims at capturing the existence of external limits to the resolution of the issues, which are beyond incumbent's control. Politicians are informed about the state of the world, while voters only know the probability distribution over the possible states.

[1]The Roma repatriation, Gavin Hewitt, *BBC News*, Aug. 19th 2010.

[2]France Prepares to Deport Foreign Gypsies, Sebastian Moffet and Max Colchester, *Wall Street Journal*, Aug. 18th 2010

[3]Ibid.

In our benchmark model, we assume that the incumbent must choose on which issue she wants to exert effort. Exerting effort on both is impossible. Our first proposition examines effort decisions in the second term. Since no election takes place at the end of this term, the proposition allows us to characterize incumbents' behavior in absence of electoral incentives. The most important conclusion is that, absent any electoral pressure, a good incumbent would always choose to exert effort on the most important issue, even though this is more difficult to solve.

Our main contribution shows how re-electoral incentives change incumbent's effort allocation decisions. We show that, if being in power in the second period is sufficiently important for the politician, she will always address the less relevant issue in the states where it can be solved. An interesting and important observation to understand the result is the following: it is optimal for voters to re-elect the incumbent even if she solved only the less important issue. Intuitively, they recognize that good types will always exert higher effort than bad ones, as they also care about issue resolution. Since this implies that good incumbents always have higher chances of being successful, issue resolution is always a positive signal for voters. Given voters' re-election strategy, a good incumbent faces a trade-off between solving the issue she considers more important and increasing her chances of being re-elected by tackling the easier one. When re-election is sufficiently valuable, politicians will always prefer the latter.

The striking feature about our result is the negative effect of electoral competition on politicians' behavior. Elections are supposed to benefit voters in two ways: first, observing incumbents' performances allows for a better selection of politicians; second, the threat of punishment (i.e., no re-election) should force incumbents to work harder. Our model shows that this might not be the full picture. The same fear of losing her seat, indeed, can induce an incumbent to act against public interest, tackling issues of lower importance.

In an extension of the paper, we consider the more general case where effort can be exerted on both issues. We show that the conclusions depend on the degree of voters' ignorance about incumbent's decisions. Whenever they have no information about effort decisions, good types will exert positive effort on both issues. This case is compared to the one where voters acquire a limited amount of information about effort decisions. More precisely, we assume they can observe whether effort was exerted or not on one issue, but they are unable to find out the exact amount. Surprisingly, allowing voters to be better informed about politicians' behavior brings back the bad equilibrium we found in the benchmark case. This is supported by the out-of-equilibrium belief that exerting effort on an issue and not solving it constitutes a negative signal about incumbent's type. For example, voters might think that effort was sub-optimally allocated, or that the incumbent just pretended to address one issue. Under such a re-election strategy, exerting effort on both issues reduces the probability of being re-elected. As for our benchmark case, whenever re-election is more valuable than first period utility, both types of politician will only address the less important issue.

In the discussion of our results, we also consider the case in which one unique state of the world exists, where both issues can be solved with positive probability when effort is exerted on them. We show that the equilibrium found in the main part of the paper is robust to this alternative assumption, but it is no more unique. Most importantly for our purposes, an alternative equilibrium arises where both types of incumbent exert positive effort on the important issue. This happens whenever they know they will be re-elected if and only if this issue was solved.

The remainder of the paper is organized as follows: Section 2.2 reviews the literature related to our work. Section 2.3 and 2.4 present the formal model and the main results, respectively. In Section 2.5, we discuss the robustness of our conclusions. In Section 2.6, we develop a welfare analysis where we compare one-term and two-term limits for incumbent. Section 2.7 concludes.

2.2 Literature

Failures in the political process of policy making are not a new concept in political economics literature (for a general review of political agency models addressing the topic, see Besley (2007)). Our work is related to the branch of the literature studying how these failures can arise from the combined presence of asymmetric information, uncertainty, career concerns by politicians and term limits.

Political budget cycle literature (Rogoff (1990)) already highlighted how the need to separate from incompetent types could force competent incumbents to inefficiently cut taxes and increase government consumption spendings right before elections. Similar effects are present in Majumdar and Mukand (2004), who focus on the introduction of experimental policies by incumbents. They show that, because of electoral pressures, politicians tend to have inefficiently low degrees of flexibility in their policy making choices: they either persist in the adoption of policies that resulted to be sub-optimal from previous experimentation or they are over-conservative in protecting the status quo. Dewan and Hortala-Vallve (2012) show that electoral accountability can induce inefficient risk taking behavior by incumbents, as compared to situations where politicians are appointed. Non competent incumbents tend to underinvest in risky policies, while competent ones always over-invest. The authors also prove that allowing the challenger to play an active role, thus increasing the competition for election, does not restore efficiency.

Combining moral-hazard and adverse selection considerations, Maskin and Tirole (2004) and Canes-Wrone et al. (2001) prove that, whenever gains from being in power for one additional period are sufficiently high, reputation concerns can also lead to *pandering*, defined as a situation in which politicians take the "popular" action (i.e., the action ex ante preferred by the voters) even if their private information suggests it is not the optimal one. In our model, the possibility of pandering is ruled out by the assumption of voters' perfect information about the characteristics of the issues. In particular, we prove that politicians will sometimes take actions that are not in citizens' interests *even if* voters perfectly know what is best for them. Clearly, adding some uncertainty in the model would only go in the direction of our result.

Morelli and Van Weelden (2013) examine the case where only one issue is a common value one, while voters' opinions are divided about the second. The authors prove that reelectoral incentives can induce politicians to provide sub-optimally high effort on the divisive issue, rather than focusing on the solution of the common value one. This strategy helps signaling to the majority of the electorate that incumbent's preferences are aligned with theirs. All these results are obtained under the assumption that voters can observe the amount of effort exerted by the incumbent on an issue. In the second part of the paper, the authors also examine the effect of no transparency in effort allocation decisions. They show that the conclusions depend on the value attached to re-election by politicians and on the total amount of effort that they can exert on the issues.

Our paper is also related to the literature on secondary issues and single issue voters. Secondary issues are those that only affect a small subset of the population. Environmental regulations (List and Sturm (2006)) or gun control (Bouton et al. (2014)) are typical examples. Single issue voters are characterized by a strong intensity in their preferences over one specific secondary issue. The intensity is so strong that candidates' choices on all other dimensions of the policy space are completely irrelevant to them. Politicians can therefore attract their votes by implementing their favorite policy on the secondary issue, even though this might not be the one favored by the majority. As long as they please the majority on front-line issues (i.e. those of general interest), they will not lose their votes. In Bouton et al. (2014), the argument is used to explain why no bill in favor of gun control is passed by the U.S. Congress.

The positive effects of making incumbents accountable by allowing them to run for re-election had already been highlighted by Besley and Smart (2007). The authors identify a *selection effect*, by which politicians that are kept in power in the second period have an higher average quality of first period ones, and a *discipline effect*, which induces politicians to work harder in the first period if they want to be re-elected. The same effects are identified in a more general principal-agent model by Banks and Sundaram (1998). Despite recognizing the existence of these two effects, our model produces partially different conclusions. Key for this divergence is the assumption on a good politician's behavior: we depart from Besley and Smart's formalization by allowing the good type to act strategically. As noticed above, this induces a *re-election effect*, which consists in the possibility that the good type takes the bad action. The introduction of one-term limits could therefore be beneficial for voters. Maskin and Tirole (2004) also recognize this possibility when they prove that judicial power can dominate representative democracy if benefits form re-election are sufficiently high. According to Smart and Sturm (2006), instead, two-term limits would be better than one-term limits: the impossibility to run for re-election in the second period would induce a truthful behavior on first period incumbent, which, in turn, would increase voters' ability to screen politicians.

2.3 The Model

We consider a two-period model with an election in between. In each period $t \in \{1,2\}$, two issues are present in the country. To keep notation simple, we denote them by a and b in both periods. Voters care about the resolution of the issues and obtain utility u_i if issue $i \in \{a,b\}$ is solved. We assume they all have identical preferences over the issues and, without loss of generality, we set $u_a > u_b$. In the remainder of the paper, we will refer to a as to the *more important* issue.

In each period, an incumbent politician can exert effort to solve the issues. Let $e_i^t \in [0,\infty)$ denote the effort exerted on issue $i \in \{a,b\}$ in period t and define $\mathbf{e^t} = (e_a^t, e_b^t)$. The probability of solving issue i depends on e_i^t, but we assume it is also affected by events beyond politician's control. In particular, we assume the existence of two states of the world where one of the two issues cannot be solved, independently of the amount of effort exerted by the politician on it. Let σ_i be the state of the world where only issue i can be solved and denote by σ_{ab} the remaining one. Formally, for a given effort $\mathbf{e}^t$ exerted in period t, the probability of solving issue $i \in \{a, b\}$ in a given state $\sigma^t \in \{\sigma_a, \sigma_b, \sigma_{ab}\}$ is

$$\pi_i(\mathbf{e}^t|\sigma^t) = \begin{cases} 0 & \text{if } \sigma^t = \sigma_j, j \neq i \\ \pi_i(e_i^t) & \text{otherwise} \end{cases}$$

We assume the incumbent knows the state of the world, while voters are only informed about the probability distribution over the three states. For the states where issue i can be solved, we make the following assumptions on $\pi_i(\cdot)$:

Assumption 3. *For all* $i \in \{a, b\}$,

i) $\pi_i(e) \in [0, 1)$ *for all* $e \in [0, \infty)$;

ii) $\lim_{e\to\infty} \pi_i(e) = 1$;

iii) $\pi_i(0) = 0$;

iv) $\pi_i'(e) > 0, \pi_i''(e) < 0$ *for all* $e \in [0, \infty)$;

In words, Assumption 3 states that no issue can be solved with certainty (i and ii), that it is impossible to solve an issue without exerting effort on it (iii) and that the probability of resolution is an increasing and concave function of effort (iv). While these characteristics are common to the two probabilities, some asymmetry is introduced in Assumption 4:

Assumption 4. *For all* $e \in [0, \infty)$,

i) $\pi_b(e) > \pi_a(e)$;

ii) $\pi_a(e)u_a > \pi_b(e)u_b$;

Assumption 4 is the key of our model: it states that the more important issue is also *more difficult* to solve, in the sense that its probability of resolution is always lower than the probability of solving the other one (i); the second part of the assumption, however, imposes that this issue guarantees a higher expected utility for any level of exerted effort.

The incumbent politician wants to be re-elected and obtains ego rents R from being in power for a second period. With probability $(1 - \gamma)$, she is a *bad* type (B) and only cares about re-election. With probability γ, instead, the incumbent is *good* (G) and cares as much as voters about issue resolution. More precisely, in each period $t \in \{1, 2\}$ and for each state of the world $\sigma^t \in \{\sigma_a, \sigma_b, \sigma_{ab}\}$, her period-$t$ expected utility from issue resolution is

$$V(\mathbf{e}^t|\sigma^t) = EU_v(\mathbf{e}^t|\sigma^t) - c(e_a^t + e_b^t) \tag{2.1}$$

where

$$EU_v(\mathbf{e}^t|\sigma^t) = \pi_a(\mathbf{e}^t|\sigma^t)u_a + \pi_b(\mathbf{e}^t|\sigma^t)u_b$$

represents voters' expected utility in that period and state of the world[4] and $c(e_a^t + e_b^t)$ is her cost of effort in period t. We assume it increasing, convex and such that $c(0) = 0$[5]. To sum up, conditional on being in state of the world σ^1 in period $t = 1$, the expected utility for each type can be written as

$$EU_G(\mathbf{e}^1, \mathbf{e}^2|\sigma^1) = V(\mathbf{e}^1|\sigma^1) + q\left[R + \sum_{\sigma^2} \text{Prob}(\sigma^2)V(\mathbf{e}^2|\sigma^2)\right]$$

$$EU_B(\mathbf{e}^1, \mathbf{e}^2|\sigma^1) = q\left[R - c(e_a^2 + e_b^2)\right] - c(e_a^1 + e_b^1)$$

where q represents incumbent's (endogenous) probability of re-election[6]. The timing of the events is the following:

(0) Nature draws the state of the world $\sigma^1 \in \{\sigma_a, \sigma_b, \sigma_{ab}\}$ and the type of incumbent $P \in \{G, B\}$. Both σ^1 and P are only observed by the politician.

(1) The incumbent decides how much effort to exert to solve the issue in the first period, $\mathbf{e}^1 = (e_a^1, e_b^1)$. Effort cannot be observed by voters.

(2) Issue resolution is realized and observed by the electorate.

(3) Voters decide whether to re-elect the incumbent or to elect a challenger. The latter is good with probability γ.

(4) Point (0), (1) and (2) are repeated for period 2.

The next section studies the Perfect Bayesian equilibria of the model. To begin with, we will assume that the incumbent can only exert effort on one issue in each period (i.e. if $e_i^t > 0$ then $e_j^t = 0$, $j \neq i$, for all $t \in \{1, 2\}$). In Section 2.5.1, we will extend the analysis to more general assumptions. As we will show, our results will not change as long as politicians must announce which issue they plan to solve before exerting effort. Let us denote by $r \in \{a, b, \emptyset\}$ the outcome of issue resolution: $r = a$ ($r = b$) if issue a (b) was solved and $r = \emptyset$ otherwise. Voters' re-election decision will be denoted by $v \in \{1, 0\}$, with $v = 1$ if the incumbent is re-elected, $v = 0$ otherwise.

2.4 Results

Proceeding backward, we first consider incumbent's behavior in period 2. This will also give us the opportunity to analyze each type's behavior in absence of any electoral pressure. Clearly, a bad incumbent will never exert effort on any issue, in any state of the world. This is because exerting a positive amount of effort would increase costs without producing any

[4] Clearly, $\mathbf{e}^t = \mathbf{e}^t(\sigma^t)$, but we preferred to omit this specification for the sake of the exposition.

[5] Our results would not change if, instead of assuming different benefits from issue resolution, we assumed different costs of effort for the two types.

[6] For type G utility function, we omitted a term representing the utility this type would receive if not elected. This is positive if a good challenger is elected. Introducing such term makes analysis and notation more cumbersome without altering our results and we therefore preferred to omit it.

benefit. Let us turn to the other type of incumbent now. Conditional on being in state of the world σ^2, her maximization problem is

$$\max_{\mathbf{e^2}} V(\mathbf{e^2}|\sigma^2) \tag{2.2}$$

where $V(\mathbf{e^2}|\sigma^2)$ was defined in (2.1). In those states where only issue $i \in \{a, b\}$ can be solved, problem (2.2) is equivalent to maximizing the expected utility of solving this issue. For each issue i, define $\hat{e}_i$ as

$$\hat{e}_i = \arg\max_{e} \pi_i(e)u_i - c(e)$$

Since the politician can address only one issue in each period, in state σ_{ab} the incumbent will exert effort on the one that guarantees higher expected utility. In other words, she will provide effort $\hat{e}_i$ instead of $\hat{e}_j$ if and only if

$$V(\hat{e}_i|\sigma_{ab}) \geq V(\hat{e}_j|\sigma_{ab})$$

(since one of the two arguments of $\mathbf{e}$ is always zero, from now on notation will be simplified by specifying only the positive component of the effort vector). Assumption 4 (point *ii*) implies that the incumbent will always exert effort on issue a in σ_{ab}. Indeed, we have

$$\begin{aligned} V(\hat{e}_a|\sigma_{ab}) &= \pi_a(\hat{e}_a)u_a - c(\hat{e}_a) \\ &\geq \pi_a(\hat{e}_b)u_a - c(\hat{e}_b) \\ &> \pi_b(\hat{e}_b)u_b - c(\hat{e}_b) \\ &= V(\hat{e}_b|\sigma_{ab}) \end{aligned}$$

The first inequality is a direct consequence of $\hat{e}_a$ being the result of a maximization problem; Assumption 4 is used to deduce the second one. Figure 2.1 provides a graphical interpretation of the result.

The following proposition summarizes incumbent behavior in period 2, i.e. in absence of any electoral pressure. Denote the equilibrium effort provided by incumbent P in state of the world σ^2 by $\hat{\mathbf{e}}^2_P(\sigma^2)$.

Proposition 8. *Incumbent equilibrium effort provision in period $t = 2$ is*

$$\hat{\mathbf{e}}^2_B(\sigma^2) = (0, 0) \textit{ for all } \sigma^2 \in \{\sigma_a, \sigma_b, \sigma_{ab}\}$$

$$\hat{\mathbf{e}}^2_G(\sigma^2) = \begin{cases} (\hat{e}_a, 0) & \textit{if } \sigma^2 \in \{\sigma_a, \sigma_{ab}\} \\ (0, \hat{e}_b) & \textit{otherwise} \end{cases}$$

Before turning to incumbent behavior in the first period, let us analyze re-election decisions by voters. Given that a bad incumbent exerts no effort in $t = 2$, voters' expected utility in this period is always higher if a good type is in power. Thus, when deciding whether to re-elect the incumbent or not they will try to maximize the probability of electing this type of politician. Let $\tilde{\gamma}(r)$ be their updated beliefs about incumbent's type after having observed issue resolution $r \in \{a, b, \emptyset\}$. Since the challenger is of type G with probability γ, re-electing

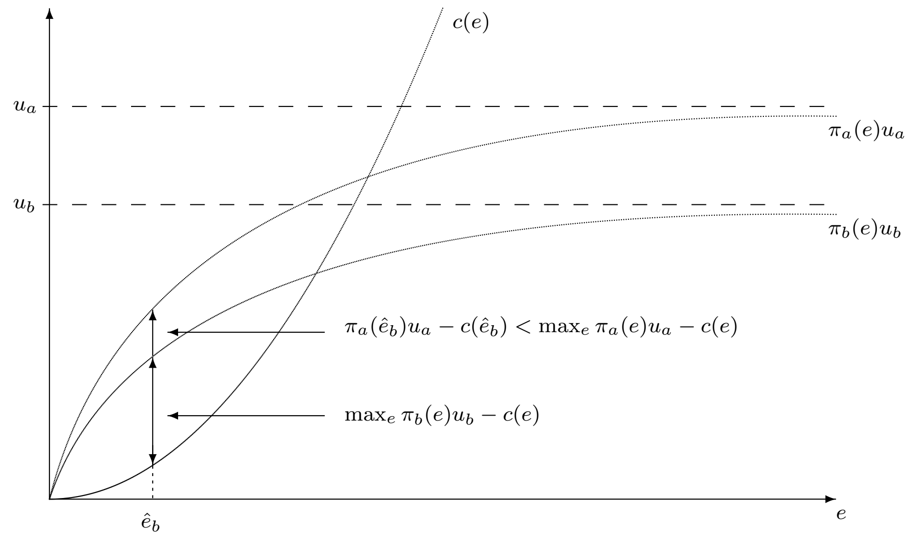

Figure 2.1: Second period issue choice for a good politician in state σ_{ab}

the incumbent ($v = 1$) is optimal if and only if $\tilde{\gamma}(r) > \gamma$. We assume each politician wins with probability one-half when $\tilde{\gamma}(r) = \gamma$.

Our main result concerns incumbent equilibrium behavior in $t = 1$. Proposition 9 shows that, whenever R is sufficiently high, both types of incumbent will exert effort on issue b in all states of the world where it can be solved (i.e., σ_b and σ_{ab}). The result is based on two important observations. First, the only equilibrium strategy for voters is the one that re-elects the incumbent if and only if *any* of the two issues was solved, i.e.

$$\hat{v} = \begin{cases} 1 & \text{if } r \in \{a, b\} \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

To see why the strategy is indeed optimal in equilibrium, one should look at the incentives for each type of politician to exert effort on an issue. Assume the incumbent knows that solving an issue is sufficient to be re-elected. For a bad politician, effort only increases the probability of obtaining R; for a good politician, it also raises the probability of obtaining a positive first period utility and of being able to tackle the more important issue in the second period. Given that the cost of effort is identical for the two types, this implies that the good politician will always exert higher effort than the bad one. Since both $\pi_a(\cdot)$ and $\pi_b(\cdot)$ are increasing functions of e, solving an issue is always a signal in favor of a type G incumbent, i.e. $\tilde{\gamma}(r) > \gamma$ for $r \in \{a, b\}$. To see that $\hat{v}$ is the unique equilibrium strategy for voters, consider the following re-election rule:

$$v' = \begin{cases} 1 & \text{if } r = a \\ 0 & \text{otherwise} \end{cases}$$

In other words, voters re-elect the incumbent if and only if issue a was solved. This clearly induces both types of politician to exert positive effort on a (and none on b) in state σ_a and

σ_{ab}. When only issue b can be solved (state σ_b), however, a bad type will never provide effort: given the impossibility of solving a and voters' re-election strategy, re-election probability is always zero and exerting effort would only increase her cost. Since a good type also cares about solving issue b *per se*, instead, her exerted effort will never be equal to zero. This implies that $\tilde{\gamma}(b) > \gamma$. Thus, not re-electing the incumbent when issue b was solved cannot be an equilibrium behavior.

Our second important point concerns incumbent behavior given voters' re-election strategy $\hat{v}$. In particular, we want to focus on type G (it seems straightforward to conclude that type B will always exert effort on issue b). If solving any issue is sufficient to be re-elected, this type will face the following choice: on the one hand, exerting positive effort on a increases her first period expected utility by Assumption 4 (*ii*); on the other hand, however, since b is *easier* to solve (Assumption 4 (*i*)), it guarantees higher probability of re-election. Whenever being in power in the second period is sufficiently important for the incumbent, exerting effort on b will always be the optimal strategy.

Before stating our main proposition, let us introduce some useful notation. Define $\tilde{e}_{iB}$ and $\tilde{e}_{iG}$ as

$$\tilde{e}_{iB} = \arg\max_e \pi_i(e)R - c(e) \tag{2.4}$$

$$\tilde{e}_{iG} = \arg\max_e \pi_i(e)[u_i + R + \hat{V}_2] - c(e) \tag{2.5}$$

where

$$\hat{V}_2 = \sum_\sigma \text{Prob}(\sigma^2)V(\hat{\mathbf{e}}_G^2(\sigma^2)|\sigma^2)$$

represents equilibrium second period utility from issue resolution. The two terms represent optimal effort exerted on issue $i \in \{a, b\}$ by a B and G type, respectively. Finally, denote by $\hat{\mathbf{s}} = [\hat{\mathbf{e}}_B^1(\sigma^1), \hat{\mathbf{e}}_G^1(\sigma^1), \hat{v}, \hat{\mathbf{e}}_B^2(\sigma^2), \hat{\mathbf{e}}_G^2(\sigma^2)]$ the strategy profile where a type P politician exerts effort $\hat{\mathbf{e}}_P^t(\sigma^t)$ in period t, with

$$\hat{\mathbf{e}}_B^1(\sigma^1) = \begin{cases} (0, \tilde{e}_{bB}) & \text{if } \sigma^1 \in \{\sigma_b, \sigma_{ab}\} \\ (\tilde{e}_{aB}, 0) & \text{if } \sigma^1 = \sigma_a \end{cases}$$

and

$$\hat{\mathbf{e}}_G^1(\sigma^1) = \begin{cases} (0, \tilde{e}_{bG}) & \text{if } \sigma^1 \in \{\sigma_b, \sigma_{ab}\} \\ (\tilde{e}_{aG}, 0) & \text{if } \sigma^1 = \sigma_a \end{cases}$$

and $\hat{\mathbf{e}}_B^2(\sigma^2), \hat{\mathbf{e}}_G^2(\sigma^2)$ as defined in Proposition 8. Voters adopt re-election decision $\hat{v}$, which we specified in (2.3).

Proposition 9. *There exists $\underline{R} \in [0, \infty)$ such that, whenever $R > \underline{R}$, strategy profile $\hat{\mathbf{s}}$ is the unique equilibrium of the game.*

A formal proof of Proposition 9 can be found in the Appendix.

2.5 Extensions and Robustness

In this section, we prove that our main result is robust to changes in the assumptions of the model.

2.5.1 Addressing both Issues

In the benchmark model, we imposed that the incumbent could only exert effort on one issue in each period. Let us now relax this assumption and consider the case where both issues can be addressed. In particular, let us distinguish between two different cases. In the *complete ignorance* case, voters do not observe anything about incumbent's effort decisions. In the *partial ignorance* case, instead, they only observe whether the amount of effort exerted on an issue was positive or not[7]. We begin the analysis from the first one.

Complete Ignorance Case

Assume the incumbent can exert positive effort on both issues, but voters receive no information at all about her choices. The possible issue resolution outcomes are now $\{a, b, (a, b), \emptyset\}$, where (a, b) denotes the case where both issues were solved. By the same reasoning used in the main section of the paper, a re-election strategy that elects the challenger if the incumbent solved only one issue is never optimal. This is because, in the state of the world where that issue cannot be solved, only the good incumbent will exert a positive effort (see the end of the proof of Proposition 9). Thus, the only optimal re-election strategy must now be

$$\hat{v} = \begin{cases} 1 & \text{if } r \in \{a, b, (a, b)\} \\ 0 & \text{otherwise} \end{cases}$$

Clearly, for states σ_a and σ_b, effort decisions by the two types are still those found in the benchmark case. The same holds for a bad incumbent in state σ_{ab}, as she would never exert effort on both issues when solving the easier one is sufficient to be re-elected. Consider the behavior of type G now. If she decides to provide effort on both issues, then her new maximization program is

$$\max_{e_a, e_b} \pi_a(e_a)u_a + \pi_b(e_b)u_b + (\pi_a(e_a) + \pi_b(e_b))[R + \hat{V}_2] - c(e_a + e_b)$$

which is equivalent to solving maximization problem (2.5) for both issues. Then, the politician will exert positive effort on both a and b.

Proposition 10. *Whenever effort can be exerted on both issues and voters are completely ignorant about incumbent's decisions, a good type will exert positive effort on both issue a and b in state σ_{ab}.*

We now consider the case where voters are able to see whether some positive effort was exerted on an issue, but remain ignorant about its exact amount.

Partial Ignorance Case

Assume that before election, voters can observe which issue the incumbent tried to solve and whether she was successful or not. What we have in mind is a situation where, because of

[7]In the benchmark model, there is no need to distinguish between the two cases as the results would coincide.

politician's announcements, voters can infer which issues she addressed, but cannot really determine how much effort she provided on each of them. Voters therefore observe the vector, (d, r), with $d \in \{a, b, (a, b), \emptyset\}$ and such that

$$d = \begin{cases} i & \text{if } e_i > 0, e_j = 0, i, j \in \{a, b\}, j \neq i \\ (a, b) & \text{if } e_i > 0, \forall i \in \{a, b\} \\ \emptyset & \text{otherwise} \end{cases}$$

As for the previous case, $r \in \{a, b, (a, b), \emptyset\}$. A re-election strategy like

$$\hat{v} = \begin{cases} 1 & \text{if } r \in \{a, b, (a, b)\}, \forall d \\ 0 & \text{otherwise} \end{cases}$$

(i.e., a strategy that re-elects if any issue was solved, independently of the announcement made by the politician) clearly induces the same equilibrium effort decisions as in the complete ignorance case. However, it is interesting to see what happens when voters assign more importance to d. Consider the alternative strategy

$$v' = \begin{cases} 1 & \text{if } d \in \{a, b, (a, b)\}, r = d \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

Under v', the incumbent is not elected if voters observe that she put effort on both issues, but failed at solving one of them (or both). For example, they might think that the incumbent only pretended to address one issue or they might simply think that effort could (and should) have been better allocated. Equilibrium effort provision for a B type does not change under this re-election rule, and neither does the one of a good politician in states σ_a and σ_b. In state σ_{ab}, however, exerting effort on both issues implies the following maximization program for G

$$\max_{e_a, e_b} \pi_a(e_a)u_a + \pi_b(e_b)u_b + \pi_a(e_a)\pi_b(e_b)[R + \tilde{V}_2] - c(e_a) - c(e_b) \tag{2.7}$$

Assume $R > \underline{R}$, so that, if effort has to be exerted on a unique issue, it is optimal to exert it on b. When comparing program (2.7) with program (2.5), it is immediate to see that providing positive effort on a produces two opposite effects: on the one hand, it increases first period utility of issue resolution; on the other hand, however, it decreases the probability of being elected, since now both issues have to be solved. As for our main result, there must therefore exist a value of rents from re-election $\underline{R}'$ such that, if $R > \max\{\underline{R}, \underline{R}'\}$, no effort will be exerted on the more important issue.

Proposition 11. *When effort can be exerted on both issues and voters can observe the vector (d, r), there always exist $\underline{R}'$ such that, if $R > \max\{\underline{R}, \underline{R}'\}$, equilibrium effort exerted on a in state σ_{ab} is equal to zero.*

A formal proof of Proposition 11 can be found in the Appendix.

2.5.2 Unique State of the World

In this section, we assume that the only possible state of the world is σ_{ab}, so that both issues can be solved with positive probability if some effort is exerted on them. The equilibrium found in the main section still exists, and the proof is identical to the one provided for (the existence part of) Proposition 9. However, the equilibrium is no more unique. More precisely, two other equilibria arise where the incumbent is elected if and only if she solved one specific issue.

Assume the incumbent is re-elected if and only if i was solved, i.e.

$$v^i = \begin{cases} 1 & \text{if } r = i \\ 0 & \text{otherwise} \end{cases}$$

Under this re-election strategy, the maximization program for type B and G are those reported in (2.4) and (2.5), respectively. That is, both politicians exert effort on the only issue that guarantees re-election if solved. Re-electing if $r = i$ is clearly optimal, since, as proven for Proposition 9, $\tilde{e}_{iG} > \tilde{e}_{iB}$ implies that $\tilde{\gamma}(i) > \gamma$. Finally, not re-electing the incumbent if $j \neq i$ was solved can be supported by out-of-equilibrium beliefs $\tilde{\gamma}(j) < \gamma$. This proves the following proposition.

Proposition 12. *Whenever σ_{ab} is the only possible state of the world, the strategy profiles $\hat{\mathbf{s}}^{\mathbf{i}} = (\mathbf{e}^1_B, \mathbf{e}^1_G, v', \mathbf{e}^2_B, \mathbf{e}^2_G) = ((\tilde{e}_{iG}, 0), (\tilde{e}_{iB}, 0), v^i, (\hat{e}_a, 0), (0, 0))$, $i \in \{a, b\}$, are equilibria of the game.*

Thus, when only one state of the world exist, there might be cases where politicians take the optimal decision and address issue a. The equilibria shown in Proposition 12 survive standard refinements such as the Intuitive Criterion or the D1 Criterion (Cho and Kreps (1987) and Banks and Sobel (1987)). These refinements require voters' out-of-equilibrium beliefs to assign more probability weight to the type of incumbent that is *more likely* to deviate, in the sense that, given voters' best response under the restricted beliefs, this type would benefit more than the other from the deviation. Notice that a bad incumbent wants to mimic a good one. In other words, this type will always have incentives to choose the issue that voters' believe a good incumbent would choose. Consequently, a good incumbent will never be the one who is *more likely* to deviate.

2.6 Term Limits: a Welfare Analysis

Given our conclusions on the distorting effect of elections, it is interesting to consider the consequences of imposing a one-term limit for incumbents. To do so, we compare voters' expected welfare when i) period one incumbent is subject to re-election and ii) a new politician is elected in each period. Clearly, this is an interesting comparison only for the states of the world where both issues can be solved with positive probability. To avoid useless complications and keep the analysis tractable, we here consider the limit case where the probability of state σ_{ab} is almost one in both periods. When doing so, voters' welfare in the other two states becomes negligible. Define

$$\tilde{\pi}_{bG} = \pi_b(\tilde{e}_{bG}) \quad \tilde{\pi}_{bB} = \pi_b(\tilde{e}_{bB}) \quad \hat{\pi}_{aG} = \pi_a(\hat{e}_a)$$

In words, $\tilde{\pi}_{bG}$ and $\tilde{\pi}_{bB}$ represent the probability that issue b is solved in the first period given optimal effort decision by the two types of incumbent. Similarly, $\hat{\pi}_{aG}$ represents the probability of solving a when a good politician is in power for one term.

Welfare in presence of a two-term limit, W^{2T}, and for a one-term limit, W^{1T}, can be computed as

$$W^{2T} = [\gamma\tilde{\pi}_{bG} + (1-\gamma)\tilde{\pi}_{bB}]u_b + \\ + [\gamma(\tilde{\pi}_{bG} + (1-\tilde{\pi}_{bG})\gamma) + (1-\gamma)(1-\tilde{\pi}_{bB})\gamma]\hat{\pi}_{aG}u_a \quad (2.8)$$

$$W^{1T} = 2\gamma\hat{\pi}_{aG}u_a \quad (2.9)$$

The first line in (2.8) represents first period expected utility for voters: it is the probability that each of the two types manages to solve issue b. In the second period (second line of (2.8)), voters obtain utility u_a only if a good incumbent solves issue a, which happens with probability $\hat{\pi}_{aG}$. The probability of having a good type in power in the second period coincides with the probability that i) a good incumbent solves the issue in the first period and is re-elected or ii) a (good or bad) incumbent fails to solve the issue but the elected challenger is a good type. If politicians can only be elected for one term, welfare is simply the expected utility of having a good type in power in each period.

Taking the difference between the two, we obtain

$$\Delta W = (1-\gamma)\tilde{\pi}_{bB}u_b + \gamma(\tilde{\pi}_{bG}u_b - \hat{\pi}_{aG}u_a) + \gamma(1-\gamma)(\tilde{\pi}_{bG} - \tilde{\pi}_{bB})\hat{\pi}_{aG}u_a \quad (2.10)$$

The first term in condition (2.10) corresponds to the *discipline effect* described by Besley and Smart [9]: if a bad period-one incumbent wants to be re-elected, she will be willing to exert a positive level of effort to try to appear as a good type. With respect to the case where a one-term limit is imposed, this increases votersÕ welfare. Furthermore, as a good incumbent is always more likely to solve an issue, the presence of elections raises voters (second period) welfare by increasing the ability of screening politicians. The third term in condition (2.10) represents this *selection effect.* The condition we provide, however, includes an additional term, which is related to the possibility that a good incumbent acts against public interest in order to remain in power. We call it the *re-election effect.* Whenever a is sufficiently important (i.e., u_a is sufficiently high),

$$\tilde{\pi}_{bG}u_b - \hat{\pi}_{aG}u_a < 0$$

and the re-election effect is negative. Most importantly, using (2.10), one can see that $\Delta W < 0$ if and only if

$$\hat{\pi}_{aG}u_a > \frac{(1-\gamma)\tilde{\pi}_{bB}u_b + \gamma\tilde{\pi}_{bG}u_b}{1-\gamma(1-\gamma)(\tilde{\pi}_{bG} - \tilde{\pi}_{bB})}$$

Notice that, for any value of u_a, there exist rents from re-election R that induce a good politician to address b. When u_a is very large and R satisfies the condition stated in Proposition 9, $\Delta W < 0$ and one-term limits guarantee higher welfare to the electorate.

2.7 Conclusions

The paper provides an answer to the question of why politicians sometimes address issues that are not the most relevant ones in their costituency. The explanation we propose is based on the negative correlation between importance of the issue and how difficult it is to solve it. Whenever the relevant issues in a country are also very complicated, politicians might try to signal their quality to voters by engaging in projects of lower value, that, however, ensure higher chances of success. Voters, for their part, re-elect successful incumbents even if they did not address the most important problems. Their behavior is justified by the idea that managing to successfully address an issue, even if not the main one, is always a signal of quality of the politician.

The examples mentioned in the introduction fit quite well our explanation. Roma case in France was clearly a less relevant issue in the country. Nevertheless, compared, for example, to unemployment, it constituted a much easier problem to solve. The same holds for the design of the electoral district of Brussels.

Two observations are important about our results. First, the politicians we consider are not all purely office-seeking ones: the Òbad actionÓ is also taken by those who care about citizensÕ welfare. Second, in the paper we prove that, absent the incentive to seek for re-election, good politicians would always address the relevant issues. These two considerations suggest a distortive effect of elections, which, by modifying good politicianÕs objectives, induces them to act against public interest. Allowing politicians to run for re-election, therefore, might not necessarily be as beneficial for voters as one might hope. Long term limits for sure favor the selection of candidates and provide incentives to act in votersÕ interest for purely office seeking politicians. On the other hand, however, they could also induce a welfare loss for citizens, by distorting a good incumbent's incentives.

2.A Appendix

Proof of Proposition 9

Proof. We start by proving that $\hat{\mathbf{s}}$ is an equilibrium. Let voters re-elect according to $\hat{v}$ and consider a B type in $t=1$. In state $\sigma^1 = \sigma_i$, $i \in \{a, b\}$, her optimal effort provision will be the one that solves problem (2.4), i.e. $\tilde{e}_{iB}$. In state σ_{ab}, B will always exert effort on issue b, as by Assumption 4 (i),

$$\begin{aligned}\max_e \pi_b(e)R - c(e) &= \pi_b(\tilde{e}_{bB})R - c(\tilde{e}_{bB}) \\ &\geq \pi_b(\tilde{e}_{aB})R - c(\tilde{e}_{aB}) \\ &> \pi_a(\tilde{e}_{aB})R - c(\tilde{e}_{aB})\end{aligned}$$

Consider a good incumbent now. If $\sigma^1 = \sigma_i$, she will exert effort $\tilde{e}_{iG}$ as defined in problem (2.5). Now assume $\sigma^1 = \sigma_{ab}$. Define the two functions

$$f_a(e) = \pi_a(e)[u_a + R + \hat{V}_2]$$

and

$$f_b(e) = \pi_b(e)[u_b + R + \hat{V}_2]$$

Clearly, $f_a(0) = f_b(0) = 0$. Furthermore, we have that

$$f_b'(0) > f_a'(0) \Leftrightarrow [\pi_b'(0) - \pi_a'(0)]R > \pi_a'(0)[u_a + \hat{V}_2] + \pi_b'(0)[u_b + \hat{V}_2] \tag{2.11}$$

Because of Assumption 4, $\pi_b'(0) - \pi_a'(0) > 0$. Then, (2.11) holds if and only if

$$R > \frac{\pi_a'[u_a + \hat{V}_2] + \pi_b'(0)[u_b + \hat{V}_2]}{\pi_b'(0) - \pi_a'(0)} \equiv R_0 \tag{2.12}$$

so that $f_b(e) > f_a(e)$ for all $R > R_0$ and e sufficiently small. Furthermore, since $\lim_{e\to\infty} f_a(e) > \lim_{e\to\infty} f_b(e)$, there must exist a point $\bar{e} > 0$ such that $f_b(\bar{e}) = f_a(\bar{e})$ and $f_b(e) > f_a(e)$ for all $e \in (0, \bar{e})$. Notice that $\bar{e} = \bar{e}(R)$ and, by the implicit function theorem

$$\frac{\partial \bar{e}(R)}{\partial R} = \frac{\pi_b(\bar{e}) - \pi_a(\bar{e})}{f_a'(\bar{e}) - f_b'(\bar{e})} > 0$$

Define $\underline{R}$ as the value of R such that

$$f_a(\bar{e}(\underline{R})) = f_b(\bar{e}(\underline{R})) = c(\bar{e}(\underline{R}))$$

Then, for all $R > \underline{R}$,

$$f_b(e) - c(e) > f_a(e) - c(e) \tag{2.13}$$

for all $e \leq \bar{e}(R)$, which implies

$$\begin{aligned} \max_e \pi_b(e)[u_b + R + \hat{V}_2] - c(e) &= \pi_b(\tilde{e}_{bG})[u_b + R + \hat{V}_2] - c(\tilde{e}_{bG}) \\ &\geq \pi_b(\tilde{e}_{aG})[u_b + R + \hat{V}_2] - c(\tilde{e}_{aG}) \\ &> \pi_a(\tilde{e}_{aG})[u_a + R + \hat{V}_2] - c(\tilde{e}_{aG}) \\ &= \max_e \pi_a(e)[u_a + R + \hat{V}_2] - c(e) \end{aligned}$$

and b will always be chosen in state σ_{ab}.

We now prove that $\hat{v}$ is voters' best response to incumbent's behavior. If issue $i \in \{a, b\}$ was solved, then voters' update about incumbents' type is

$$\tilde{\gamma}(i) = \frac{\gamma \pi_i(\tilde{e}_{iG})}{\gamma \pi_i(\tilde{e}_{iG}) + (1-\gamma)\pi_i(\tilde{e}_{iB})} \tag{2.14}$$

For each issue i, first order conditions of incumbent optimization problem imply that $\tilde{e}_{iB}$ and $\tilde{e}_{iG}$ must satisfy

$$\begin{aligned} \frac{c'(\tilde{e}_{iB})}{\pi_i'(\tilde{e}_{iB})} &= R \\ \frac{c'(\tilde{e}_{iG})}{\pi_i'(\tilde{e}_{iG})} &= u_i + R + \hat{V}_2 \end{aligned}$$

and since $c'(e)/\pi_i'(e)$ is an increasing function of e for all $i \in \{a, b\}$, it must be that $\tilde{e}_{iB} < \tilde{e}_{iG}$. This immediately implies that $\tilde{\gamma}(i) > \gamma$ for all $i \in \{a, b\}$, concluding the first part of the proof.

To prove that $\hat{\mathbf{s}}$ is the only equilibrium strategy profile, we show that no other re-election strategy can be supported in equilibrium. Since each re-election rule is associated to one and only one optimal effort decision by the incumbent, this will be enough to conclude the proof. Consider the alternative strategy

$$v' = \begin{cases} 1 & \text{if } r = a \\ 0 & \text{otherwise} \end{cases}$$

Optimal effort decisions in state $\sigma^1 \in \{\sigma_a, \sigma_{ab}\}$ will be $\tilde{e}_{aP}$, for $P \in \{B, G\}$. Now let $\sigma^1 = \sigma_b$. The maximization problems for type B and G are

$$\max_e -c(e)$$

and

$$\max_e \pi_b(e)u_b - c(e)$$

respectively. These are solved by $e'_{bB} = 0$ and $e'_{bG} > 0$. Then, $\tilde{\gamma}(b) = 1 > \gamma$ and not re-electing the incumbent when she solved issue b is not optimal for voters. A symmetric reasoning can be adopted to prove that

$$v'' = \begin{cases} 1 & \text{if } r = b \\ 0 & \text{otherwise} \end{cases}$$

can never be sustained as an equilibrium strategy for voters. □

Proof of Proposition 11

Proof. We here show that there must exist $\underline{R}'$ such that, if $R > \underline{R}'$ and voters re-elect according to v' as defined in (2.6), a good incumbent will choose to exert effort only on issue b in state $\sigma^1 = \sigma_{ab}$. To see that, define e^*_{aG} and e^*_{bG} as the effort levels that solve problem (2.7) and denote by $\mathcal{V}^{ab}(u_a, u_b, R, \hat{V}_2)$ its value function,

$$\mathcal{V}^{ab}(u_a, u_b, R, \hat{V}_2) = \pi_a(e^*_{aG})u_a + \pi_b(e^*_{bG})u_b + \pi_a(e^*_{aG})\pi_b(e^*_{bG})[R + \tilde{V}_2] - c(e^*_{aG}) - c(e^*_{bG})$$

Then, by the envelope theorem,

$$\frac{\partial \mathcal{V}^{ab}(u_a, u_b, R, \hat{V}_2)}{\partial R} = \pi_a(e^*_{aG})\pi_b(e^*_{bG})$$

Similarly, let $\mathcal{V}^b(u_b, R, \hat{V}_2)$ be the value function of problem (2.5) with $i = b$,

$$\mathcal{V}^b(u_b, R, \hat{V}_2) = \pi_b(\tilde{e}_{bG})[u_b + R + \tilde{V}_2] - c(\tilde{e}_{bG})$$

Then,

$$\frac{\partial \mathcal{V}^b(u_b, R, \hat{V}_2)}{\partial R} = \pi_b(\tilde{e}_{bG})$$

Now notice that e^*_{bG} and $\tilde{e}_{bG}$ must satisfy the following first order conditions

$$\begin{aligned}\frac{c'(e^*_{bG})}{\pi'_b(e^*_{bG})} &= u_b + \pi_a(e^*_{aG})[R + \hat{V}_2] \\ \frac{c'(\tilde{e}_{bG})}{\pi'_b(\tilde{e}_{bG})} &= u_b + R + \hat{V}_2\end{aligned}$$

Since $\pi_a(e^*_{aG}) < 1$ and the function $c'(e)/\pi'_b(e)$ is increasing in e, it must be that $\tilde{e}_{bG} > e^*_{bG}$. But then,

$$\frac{\partial \mathcal{V}^b(u_b, R, \hat{V}_2)}{\partial R} > \frac{\partial \mathcal{V}^{ab}(u_a, u_b, R, \hat{V}_2)}{\partial R}$$

for all R. This implies that there must exist an $\underline{R}'$ such that, if $R > \underline{R}'$,

$$\mathcal{V}^{ab}(u_a, u_b, R, \hat{V}_2) < \mathcal{V}^b(u_b, R, \hat{V}_2)$$

□

Chapter 3

An Axiomatic Approach to the Measurement of Ethnic Voting

Join work with Yves Sprumont, Université de Montreal

Abstract: The paper proposes the first axiomatic approach to the measurement of ethnic voting. Our central axiom is a *Size Neutral* property: ethnic voting should not depend on the total number of people belonging to an ethnic group, as long as the distribution of this group's support across different parties remains constant; likewise, ethnic voting should not depend on the total number of individuals voting for a party, if the share of this party's support coming from each group does not vary. The main result of the paper is a characterization of the class of size neutral measures. The rest of analysis is centered on the identification of the strongest monotonicity axiom that is compatible with Size Neutrality. We are able to identify a weak form of monotonicity, for which compatibility is guaranteed.

Keywords: ethnic voting, categorical variables, scale invariance.
JEL Classification: D71, D72, J15.

3.1 Introduction

Population ethnic composition is very diversified in many countries around the world. As Fearon (2003) shows, the median value of the Ethnic Fractionalization Index (ELF)[1] at the world level is around 0.5. The highest levels of ELF are found in Sub-Saharian Africa, where lower and upper quartiles of distribution of the measure across countries are located at around 0.65 and 0.85, respectively, with a median value of more than 0.75. Even if average ethnic fractionalization is lower in Western Europe, high values can still be found in countries like Belgium (0.567), Spain (0.502) or the UK (0.324). In the USA and Canada, ELF is equal to

[1]ELF is a measure of ethnic diversity within a country, defined as the probability that two randomly selected individuals will belong to different ethnic groups.

0.491 and 0.596, respectively. Many questions have been raised concerning the relationship between ethnic diversity and other characteristics of a country, such as its level of democracy (Dowd and Driessen (2008)) or economic development (Alesina et al. (2003), Easterly and Levine (1997), Posner (2004)).

In this paper, we focus on the relationship between ethnic diversity and voting behavior within countries. Political science literature has long focused on the way ethnic diversity affects politics and on the role of electoral systems in shaping this tie. In this respect, large consent seems to be reached on the idea that *ethnic voting,* roughly defined as the extent to which ethnic considerations affect voting decisions, is stronger under proportional representation. As already highlighted by Huber (2012), however, this has been taken more as a premise than an actual result of research, and a careful analysis of the relationship between electoral systems and ethnic voting is not present in the literature. Even more importantly for our purposes, very little research has been conducted about the appropriate way to measure this phenomenon.

Using an appropriate measure of ethnic voting is clearly fundamental if one wants to correctly evaluate the impact of different electoral systems. The need for such measure, however, is not limited to this exercise. If ethnicity is the main driving force behind people's voting behavior, the electorate will be poorly responsive to other important policy dimensions, such as economic ones. Furthermore, the screening and selection of politicians on the basis of their ability might not be as effective as it should be. Being able to measure the extent of ethnic voting behavior can therefore be important to assess the well functioning of a democracy.

To the best of our knowledge, the only existing measures of ethnic voting are those proposed in Huber (2012). The author defines four different measures of *politicization of ethnicity*, depending on the approach one wants to have on two different dimensions of the problem. The first dimension refers to whether the focus of the analysis is on ethnic groups or parties. Under the *group-based* approach, politicization of ethnicity corresponds to the possibility of predicting individuals' voting decision on the basis of the ethnic group they belong to. Under the *party-based* approach, instead, the measure is increasing in the ability of inferring individuals' ethnicity from their vote choice. The second dimension focuses on the degree of governance problems created by politicization of ethnicity. The so called *fractionalization* approach claims that governance problems are maximized in presence of many different ethnic groups of small size. On the contrary, the *polarization* approach relates higher governance problems to situations in which only two parties of equal size exist. Combining the different approaches on each dimension, Huber (2012) defines four measures: the *Group Voting Fractionalization*, the *Group Voting Polarization*, the *Party Voting Fractionalization* and the *Party Voting Polarization*.

By their definition, fractionalization and polarization approaches assign a fundamental importance to the size of ethnic groups and parties. After all, if only a very small ethnic group voted ethnically, the impact on governance and policy making would be very limited. At the same time, however, there might be cases where allowing a measure to be too dependent on the sizes of groups and parties would produce misleading conclusions. Consider a country were an initially small ethnic group strongly supports one specific party. Now assume that, after a wave of immigration, many other people of this group move to the country and all vote for that party as well. If ethnic voting was assessed through a measure that strongly depends on group sizes, one would claim that the phenomenon has tremendously increased in the country. What this measure would be capturing, however, is the pure effect of the change in the group size and not a *real* change in its voting behavior. Similarly, assume that one party becomes particularly popular in the country and doubles its electoral support. If this happens because each group exactly doubles the support it gives to the party, why should a measure of ethnic voting be influenced by this change?

Finally, consider the evaluation of the impact of electoral systems on ethnic voting behavior. An empirical analysis of such effect necessarily requires a cross-country comparison. Ideally, one would like to consider countries that only differ in the electoral system they use. Since this is clearly not possible, however, the ethnic voting index to be used should be the one that is as neutral as possible with respect to all other characteristics of a country.

Invariance with respect to changes in groups' and parties' size is the main focus of our paper. Our goal is to construct an index of ethnic voting that is affected by these changes if and only if they are associated to an actual variation in the voting patterns of the different ethnic groups. To do so, we let the voting profile of a country be represented by a matrix. Each cell of the matrix contains the number of individuals belonging to a given group and voting for a given party. A change in the size of a group or party is then formally captured by a rescaling of one row or column of the matrix. Our *Size Neutrality* axiom requires the measure of ethnic voting to be invariant to such rescaling.

A fundamental property to be specified for an ethnic voting index is *monotonicity*. This should identify the changes in the cells of a matrix that are associated to higher degrees of ethnic voting. Consider a country where every group assigns the same share of votes to each party. Clearly, ethnicity is not a strong driving force behind people voting behavior in this country. A measure of ethnic voting should therefore take its lowest value. The opposite case is a country where each group votes for one and only one party and each party receives votes from one and only one group. Monotonicity axioms help to define the value of the measure for voting profiles that are in between these two polar cases. Intuitively, a measure should be lower when a voting profile is closer to the first case, it should be higher when the voting profile is closer to second. Many different axioms of this type can be formulated, from stronger to weaker versions. Our analysis in the paper is centered on the definition of the strongest version of monotonicity that is compatible with Size Neutrality.

From a technical point of view, the problem of measuring ethnic voting is identical to the one of measuring the degree of association between two categorical variables (Goodman and Kruskal (1954)). A variable is defined as categorical if its values are not endowed with an intrinsic ordering. For example, if the variable under consideration is *ethnic group*, there would be no reason to list Hispanics before Blacks. Similarly, considering a leftist or a rightist party first, would not alter the conclusions of the analysis. The most used statistical measure of association between categorical variables is the Chi-square index (χ^2), first proposed by Pearson (1900). The index measures the deviation between the observed value of each cell in the matrix and its expected one under the assumption of independence between the two variables. The lower the deviation, the higher the degree of independence. On the contrary, if the deviation is large, then the variables are strongly associated.

The Chi-square index is not size neutral. In the paper, we are able to identify the strongest version of monotonicity that is satisfied by this index. The axiom we define is based on the concept of *bias* between a group and a party. We say that a *positive bias* exists between a group and a party if the electoral support assigned by the group to the party is greater than the party's average support in the population. We say that the bias is *negative* in the opposite case. Now consider two ethnic groups, Green and Red, and two parties, Left and Right. Assume there exists a positive bias between Green and Left and a negative bias between Green and Right. Furthermore, assume there exists a positive bias between Red and Right and a negative bias between Red and Left. If, at the following elections, more Green people vote for Left and less vote for Right while the opposite happens within the Red group, then our *Monotonicity* axiom imposes that ethnic voting is increased in the country.

The strongest monotonicity axiom for which we are able to prove compatibility with Size Neutrality is based on a weaker concept of bias between a group and a party. We say that there exists a *universally favorable bias* between group Green and party Left if the ratio between the number of supporters of Left and the number of supporters of any other party is greater in Green than in any other group. Our *Weak Bias Monotonicity* axiom then states that, if the number of Green people voting for Left increases, ethnic voting increases as well.

The problem of measuring ethnic voting is similar to the one of measuring racial segregation of districts within a city (or schools within a district). Indices of segregation have been grouped in five different classes (Massey and Denton (1988)): evenness, exposure, concentration, centralization, and clustering. Indices of evenness measure the differential distribution of groups across areas in a city. Exposure refers to the degree of potential contact among different groups. The last three classes are based on spatial dimensions of segregations: concentration measures the relative amount of physical space occupied by each group; centralization is to the extent to which a group is located near the center of a urban area. Finally, clustering is the extent to which residential units inhabited by one group adjoin one another or are spatially clustered. Our analysis is most closely related to indices of evenness (James and Taeuber (1985)). The most known measures belonging to this class are the Dissimilarity Index, the Gini Index and the Entropy Index. The Dissimilarity Index (Jahn et al. (1947), Morgan (1975), Sakoda (1981), Reardon (1998)) measures the proportion of population that would have to move within the city in order to equalize group proportions across its districts. The Gini Index (Jahn et al. (1947), Reardon (1998)) is a function of the weighted sum, over all groups, of the absolute difference in group proportions between all possible pairs of districts. The Entropy Index (Theil (1972), Theil and Finizza (1971)) measures each district's departure from the city's *entropy*, defined as the extent of the city's ethnic diversity. None of these measures is size neutral. Furthermore, they can be proven to be closely related to measures of association between categorical variables as the Chi-square (Reardon and Firebaugh (2002)). All these measures, together with the Chi-Square index and Huber (2012)'s indices of ethnic voting, are defined in the Appendix. For an axiomatic characterization of the indices of segregation, see Frankel and Volij (2007).

3.2 Notation

Let $M = \{1, ..., m\}$ be a set of ethnic groups, $m \geq 2$, and let $N = \{1, ..., n\}$ be a set of political parties, $n \geq 2$. A voting profile is described by a positive $m \times n$ matrix A,

$$A = \begin{pmatrix} a_1^1 & \dots & a_1^j & \dots & a_1^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_i^1 & \dots & a_i^j & \dots & a_i^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_m^1 & \dots & a_m^j & \dots & a_m^n \end{pmatrix}$$

where the number a_i^j at the intersection of row i and column j records the number of agents from group i who vote for party j. We denote the ith row of A by A_i and its jth column by A^j. We also write

$$a_i = \sum_{j \in N} a_i^j \qquad a^j = \sum_{i \in M} a_i^j \qquad a = \sum_{i \in M} \sum_{j \in N} a_i^j$$

Thus, a_i is the total population in group $i \in M$, a^j is the total support of party $j \in N$ and a is the population in the whole country. We let $\mathcal{A}(m, n)$ denote the set of positive $m \times n$ matrices and define an *ethnic voting index* to be a function $F : \mathcal{A}(m, n) \to \mathbb{R}_+$.

3.3 Axioms and Results

Our central axiom is a *size neutrality* property. If $A \in \mathcal{A}(m, n)$, $i \in N$, $j \in M$ and $\lambda \in \mathbb{R}_{++}$, we denote by $(\lambda A_i, A_{-i})$ the matrix obtained by multiplying each entry of the ith row of A by λ and leaving all other entries unchanged,

$$(\lambda A_i, A_{-i}) = \begin{pmatrix} a_1^1 & \dots & a_1^j & \dots & a_1^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda a_i^1 & \dots & \lambda a_i^j & \dots & \lambda a_i^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_m^1 & \dots & a_m^j & \dots & a_m^n \end{pmatrix}$$

Likewise, $(\lambda A^j, A^{-j})$ is the matrix obtained by multiplying each entry of the jth column of A by λ and leaving all other entries unchanged,

$$(\lambda A^j, A^{-j}) = \begin{pmatrix} a_1^1 & \dots & \lambda a_1^j & \dots & a_1^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_i^1 & \dots & \lambda a_i^j & \dots & a_i^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_m^1 & \dots & \lambda a_m^j & \dots & a_m^n \end{pmatrix}$$

Matrices $(\lambda A_i, A_{-i})$ and $(\lambda A^j, A^{-j})$ therefore represent situations in which the total size of a group or a party in a country (represented by matrix A) has been multiplied by λ. Our *Size Neutrality* axiom states that the measure should not be affected by this transformation.

Size Neutrality. *For all $A \in \mathcal{A}(m, n)$, $i \in N$, $j \in M$ and $\lambda \in \mathbb{R}_{++}$, $F(\lambda A_i, A_{-i}) = F(A) = F(\lambda A^j, A^{-j})$.*

The standard statistical index of association between two categorical variables is the Chi-square index, defined as

$$\chi^2(A) = \sum_{i\in M}\sum_{j\in N} \frac{\left(a_i^j - \frac{a_i a^j}{a}\right)^2}{\frac{a_i a^j}{a}} \tag{3.1}$$

For each $(i,j) \in M \times N$, the index is a function of the distance between the observed cell count and its expected value under independence, $(a_i a^j)/a$. The lower the distance between the two values, the higher the independence between the two variables. In Appendix 3.A, we show that this index is not size neutral. In the same Appendix, we also show that none of the measures discussed in the introduction satisfies the axiom.

The class of size neutral indices can be described explicitly. An *elementary submatrix* of $A \in \mathcal{A}(m,n)$ is any 2×2 submatrix whose entries belong to adjacent rows and columns of A. The *cross-product ratio* of an elementary submatrix (or an *elementary cross-product ratio* for short) is the ratio between the product of the diagonal entries of this submatrix and the product of its off-diagonal entries. More formally: for all matrices $A \in \mathcal{A}(m,n)$, all groups $i \in M \setminus m$ and all parties $j \in N \setminus n$, define

$$r_i^j(A) = \frac{a_i^j a_{i+1}^{j+1}}{a_i^{j+1} a_{i+1}^j}$$

and let $r(A) \in \mathcal{A}(m-1, n-1)$ denote the matrix

$$r(A) = (r_i^j(A))_{i\in M\setminus m}^{j\in N\setminus n}$$

Theorem 1. *An index is size neutral if and only if it can be expressed as a function of all elementary cross-products of a matrix.*

Formally, Theorem 1 states that the function $F : \mathcal{A}(m,n) \rightarrow \mathbb{R}_+$ satisfies Size Neutrality if an only if there exists a function $f : \mathcal{A}(m-1, n-1) \rightarrow \mathbb{R}_+$ such that $F(A) = f(r(A))$ for all $A \in \mathcal{A}(m,n)$. The reader can find the proof in Appendix 3.B.

Of course, the cross-product ratio can be defined for every 2×2 (not necessarily elementary) submatrix of A : for all $i, k \in M$ such that $i < k$ and all $j, l \in N$ such that $j < l$, let

$$r_{i,k}^{j,l}(A) = \frac{a_i^j a_k^l}{a_i^l a_k^j}.$$

One checks that

$$r_{i,k}^{j,l}(A) = \prod_{\substack{i'=i,\ldots,k-1 \\ j'=j,\ldots,l-1}} r_{i'}^{j'}(A).$$

It follows that any function of the cross-product ratios of all 2×2 submatrices of a matrix can be rewritten as a function of its elementary cross-product ratios only. We may therefore reformulate Theorem 1 as follows. Denote by μ the number of 2×2 submatrices of a matrix $A \in \mathcal{A}(m,n)$. Let

$$\bar{r}(A) = (r_{i,k}^{j,l}(A))_{1\leq i<k\leq m}^{1\leq j<l\leq n} \in \mathbb{R}_{++}^{\mu}$$

be the vector of cross-product ratios of these submatrices. An index $F : A(m, n) \to \mathbb{R}_+$ is size neutral if and only if there exists a function $\bar{f} : \mathbb{R}^{\mu}_{++} \to \mathbb{R}_+$ such that $F(A) = \bar{f}(\bar{r}(A))$ for all $A \in \mathcal{A}(m, n)$.

In order to construct a good ethnic voting index, one must identify which changes in a matrix definitely increase the degree of ethnic voting. Consider a matrix A and fix a group $i \in M$ and a party $j \in N$. We say that there is a *favorable bias between i and j (in A)* if

$$a_i^j a > a_i a^j. \tag{3.2}$$

We say that there is an *unfavorable bias between i and j* if the opposite inequality holds. Notice that (3.2) is equivalent to

$$\frac{a_i^j}{a_i} > \frac{a^j}{a} \quad \text{or} \quad \frac{a_i^j}{a^j} > \frac{a_i}{a}$$

The first inequality says that the proportion of members of group i voting for party j is larger than the proportion of agents voting for j in the total population. The second says that the proportion of party j's supporters coming from group i is larger than the relative size of group i in the population.

Bias Monotonicity. *Let $i \in M$, $j \in N$ and $A, B \in \mathcal{A}(m, n)$. Suppose there is a favorable bias between i and j in A and suppose $b_i^j > a_i^j$ and $b_k^l = a_k^l$ whenever $(k, l) \neq (i, j)$. Then $F(B) > F(A)$. Likewise, if there is an unfavorable bias between i and j in A and if $b_i^j < a_i^j$ and $b_k^l = a_k^l$ whenever $(k, l) \neq (i, j)$, then again $F(B) > F(A)$.*

Fact 1. If $m \geq 3$ *or* $n \geq 3$, no index satisfies Bias Monotonicity and Size Neutrality.

Proof. If F satisfies Bias Monotonicity and Size Neutrality, then

$$\begin{aligned} F\begin{pmatrix} 3 & 1 \\ 2 & \frac{9}{10} \\ 1 & 2 \end{pmatrix} &= F\begin{pmatrix} 30 & 10 \\ \frac{20}{9} & 1 \\ 1 & 2 \end{pmatrix} \\ &< F\begin{pmatrix} 30 & 10 \\ 2 & 1 \\ 1 & 2 \end{pmatrix} \\ &= F\begin{pmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 2 \end{pmatrix} \\ &< F\begin{pmatrix} 3 & 1 \\ 2 & \frac{9}{10} \\ 1 & 2 \end{pmatrix}, \end{aligned}$$

which is impossible. □

The source of the incompatibility is clear: the very notion of a bias between a group and a party is not size neutral. In the argument above, the bias between group 2 and party 1 is unfavorable in the third matrix and favorable in the fourth although the two matrices are related by simply rescaling row 1. Note that the proof only uses the fact that F is insensitive to any rescaling of rows: this weaker version is incompatible with Bias Monotonicity when $m \geq 3$.

Note also that the incompatibility persists under various weakened formulations of Bias Monotonicity. Consider for instance the following variant.

Group Bias Monotonicity. *Let $i \in M$, let $j, j' \in N$ be such that $j \neq j'$, and let $A, B \in \mathcal{A}(m, n)$. Suppose there is a favorable bias between i and j in A and an unfavorable bias between i and j' in A. Suppose there exists a positive number δ such that $b_i^j = a_i^j + \delta$, $b_i^{j'} = a_i^{j'} - \delta$, and $b_k^l = a_k^l$ for all $(k, l) \neq (i, j), (i, j')$. Then $F(B) > F(A)$.*

This is a fixed-population implication of Bias Monotonicity. A dual axiom of Party Bias Monotonicity can also be defined. If F satisfies Group Bias Monotonicity and Size Neutrality, then

$$
\begin{aligned}
F\begin{pmatrix} 3 & 1 \\ \frac{21}{10} & \frac{9}{10} \\ 1 & 2 \end{pmatrix} &= F\begin{pmatrix} 30 & 10 \\ \frac{21}{10} & \frac{9}{10} \\ 1 & 2 \end{pmatrix} \\
&< F\begin{pmatrix} 30 & 10 \\ 2 & 1 \\ 1 & 2 \end{pmatrix} \\
&= F\begin{pmatrix} 3 & 1 \\ 2 & 1 \\ 1 & 2 \end{pmatrix} \\
&< F\begin{pmatrix} 3 & 1 \\ \frac{21}{10} & \frac{9}{10} \\ 1 & 2 \end{pmatrix},
\end{aligned}
$$

an impossibility.

In spite of the above incompatibility, Size Neutrality does make sense. Following our discussion above (see (3.2)), define the *bias between i and j (in A)* to be the quantity

$$\beta_i^j(A) = \frac{a_i^j a}{a_i a^j}.$$

Thus, a bias is favorable if it is greater than one; it is unfavorable if it is smaller than one. Multiplying rows or columns may change biases but does not alter the *relative biases*

$$r_{i,k}^j(A) := \frac{\beta_i^j(A)}{\beta_k^j(A)} = \frac{a_i^j a_k}{a_i a_k^j} \quad \text{and} \quad r_i^{j,l}(A) := \frac{\beta_i^j(A)}{\beta_i^l(A)} = \frac{a_i^j a^l}{a^j a_i^l}$$

for any $i, i' \in M$ and $j, j' \in N$. Size Neutrality amounts to requiring that only relative biases matter.

One may also argue that Bias Monotonicity, while certainly reasonable, is perhaps not as compelling as it looks. Observe that (3.2) can be rewritten as

$$\frac{a_i^j}{a_i} > \sum_{k \in M} \frac{a_k}{a} \frac{a_k^j}{a_k}.$$

Thus, we consider that there is a favorable bias between i and j when the proportion of members of group i voting for j exceeds the (weighted) arithmetic average of such proportions over all ethnic groups. The choice of the *arithmetic* average is arbitrary. More fundamentally, focusing on averages may be misleading. For instance, Bias Monotonicity implies that

$$F\begin{pmatrix} 2 & 1 & \varepsilon \\ 1 & 2 & \varepsilon \\ \varepsilon & \varepsilon & 4 \end{pmatrix} < F\begin{pmatrix} 2 & 2 & \varepsilon \\ 1 & 2 & \varepsilon \\ \varepsilon & \varepsilon & 4 \end{pmatrix} < F\begin{pmatrix} 2 & 2 & \varepsilon \\ 2 & 2 & \varepsilon \\ \varepsilon & \varepsilon & 4 \end{pmatrix}$$

for ε small enough, which seems to be questionable.

Finally, Group Bias Monotonicity is not satisfied by the Chi-square index either. We show it in Appendix 3.A. Since the index seems to be regarded as a valid one for the fixed population case, this suggests that Group Bias Monotonicity and Party Bias Monotonicity might be too strong, even when Size Neutrality is not imposed.

A property that is implied by the conjunction of Group Bias Monotonicity and Party Bias Monotonicity is the following:

Monotonicity. *Let $i, i' \in M,$, $i \neq i'$ let $j, j' \in N$ be such that $j \neq j'$, and let $A, B \in \mathcal{A}(m,n)$. Suppose that, in A, there is (i) a favorable bias between i and j and between i' and j' and (ii) an unfavorable bias between i and j' and between i' and j. Suppose there exists a positive number δ such that $b_i^j = a_i^j + \delta$, $b_i^{j'} = a_i^{j'} - \delta$, $b_{i'}^{j'} = a_{i'}^{j'} + \delta$, $b_{i'}^j = a_{i'}^j - \delta$ and $b_k^l = a_k^l$ for all $(k,l) \neq (i,j), (i,j'), (i',j), (i',j')$. Then $F(B) > F(A)$.*

Fact 2. The Chi-square index satisfies Monotonicity.

Proof. See Appendix 3.A. □

The main question to ask now is whether Size Neutrality and Monotonicity are compatible. Unfortunately, we are not able to answer it yet. Our intuition suggests that this should not be the case. Indeed, Monotonicity applies to an environment that is not size neutral to begin with. Furthermore, two of the most natural size neutral indices fail to satisfy the axiom. Consider a matrix $A \in \mathcal{A}(m,n)$ and, for any pair $\{i,k\} \in 2^M$ and any pair $\{j,l\} \in 2^N$, define

$$\rho_{\{i,k\}}^{\{j,l\}}(A) = \max\left\{ r_{i,k}^{j,l}(A), \frac{1}{r_{i,k}^{j,l}(A)} \right\}$$

By Theorem 1, the following two indices are size neutral,

$$F_1(A) = \prod_{i \in M} \prod_{j \in N} \rho_{\{i,k\}}^{\{j,l\}}(A) \tag{3.3}$$

$$F_2(A) = \sum_{i \in M} \sum_{j \in N} \rho_{\{i,k\}}^{\{j,l\}}(A) \tag{3.4}$$

Now consider the two matrices $A, B \in \mathcal{A}(2,5)$, such that

$$A = \begin{pmatrix} 130 & 2 & 2 & 2 & 2 \\ 100 & 10 & 12 & 12 & 12 \end{pmatrix} \quad B = \begin{pmatrix} 130.1 & 1.9 & 2 & 2 & 2 \\ 99.9 & 10.1 & 12 & 12 & 12 \end{pmatrix}$$

Since $\beta_1^1(A) \approx 1.16$, $\beta_2^2(A) \approx 1.62$, $\beta_1^2(A) \approx 0.34$ and $\beta_2^1(A) \approx 0.84$, there is a favorable bias between $i = 1$ and $j = 1$ and between $i = 2$ and $j = 2$ and an unfavorable bias between $i = 1$ and $j = 2$ and between $i = 2$ and $j = 1$ in matrix A. Furthermore, $b_1^1 = a_1^1 + 0.1$, $b_1^2 = a_1^2 - 0.1$, $b_2^2 = a_2^2 + 0.1$ and $b_2^1 = a_2^1 - 0.1$. By Monotonicity, $F(B) > F(A)$. Yet one can check that $F_1(A) \approx 21320.67 > F_1(B) \approx 17144.56$ and $F_2(A) = 56 > F_2(B) \approx 54.82$. The source of incompatibility is the following. Notice that $r_{1,2}^{j,l}(A) > 1$ for all $j, l \in N$, $j \neq l$. The transformation implied by Monotonicity affects all $r_{1,2}^{j,l}$ with $j \in \{1, 2\}$. More precisely, it is such that

$$\rho_{\{1,2\}}^{\{1,l\}}(B) = r_{1,2}^{1,l}(B) > \rho_{\{1,2\}}^{\{1,l\}}(A) = r_{1,2}^{1,l}(A)$$
$$\rho_{\{1,2\}}^{\{2,l\}}(B) = r_{1,2}^{2,l}(B) < \rho_{\{1,2\}}^{\{2,l\}}(A) = r_{1,2}^{2,l}(A)$$

for all $l \in N \setminus \{1, 2\}$. As the negative effect on the $r_{1,2}^{2,l}(B)$'s is stronger than the positive effect on the $r_{1,2}^{1,l}(B)$'s, $F_1(A) > F_1(B)$ and $F_2(A) > F_2(B)$.

Clearly, the impossibility of finding an index that satisfies Monotonicity and Size Neutrality - if proven - does not necessarily imply the incompatibility between the two axioms. Indeed, there might exist a preorder satisfying both of them, that is however not representable. This turns out to be the case when one restricts the analysis to the class of $2 \times n$ matrices. Let $A \in \mathcal{A}(2, n)$ and, for each pair $\{j, k\} \in 2^N$, define

$$\rho^{\{j,k\}}(A) = \max\left\{ r_{1,2}^{j,k}(A), \frac{1}{r_{1,2}^{j,k}(A)} \right\}$$

Note that $\rho^{\{j,k\}}(A) \geq 1$ for all $\{j, k\} \in 2^N$. Denote by $\rho(A) \in \mathbb{R}^{\frac{(n-1)n}{2}}$ the vector

$$\rho(A) = \left(\rho^{\{j,k\}}(A)\right)_{\{j,k\} \in 2^N}$$

Let $\succsim^L$ represent the leximax ordering on $\mathbb{R}^{\frac{(n-1)n}{2}}$ and define the preordering $\succsim$ on $\mathcal{A}(2, n)$ as

$$A \succsim B \;\; \Leftrightarrow \;\; \rho(A) \succsim^L \rho(B) \tag{3.5}$$

Fact 3. The preordering $\succsim$ on $\mathcal{A}(2, n)$ satisfies Size Neutrality and Monotonicity.

Proof. See Appendix 3.B. □

Clearly, an equivalent statement can be formulated for the $\mathcal{A}(m, 2)$ class. Unfortunately, however, the argument cannot be extended to the general $\mathcal{A}(m, n)$ class. Consider the matrices $A, B \in \mathcal{A}(3, 3)$

$$A = \begin{pmatrix} 30 & 10 & 8 \\ 10 & 30 & 25 \\ 8 & 25 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 30+\delta & 10-\delta & 8 \\ 10-\delta & 30+\delta & 25 \\ 8 & 25 & 1 \end{pmatrix}$$

In matrix A, there is a favorable bias between $i = 1$ and $j = 1$ and between $i = 2$ and $j = 2$ and an unfavorable bias between $i = 1$ and $j = 2$ and $i = 2$ and $j = 1$, as $\beta_1^1(A) \approx 1.91$, $\beta_2^2(A) \approx 1.04$, $\beta_1^2(A) \approx 0.47$ and $\beta_2^1(A) \approx 0.47$. By Monotonicity, $B \succ A$. One can however check that, for δ small enough,

$$\max_{\{i,k\},\{j,l\}} \rho_{\{i,k\}}^{\{j,l\}}(B) = \rho_{\{2,3\}}^{\{2,3\}}(B) < \rho_{\{2,3\}}^{\{2,3\}}(A) = \max_{\{i,k\},\{j,l\}} \rho_{\{i,k\}}^{\{j,l\}}(A)$$

and $\rho(A) \succ^L \rho(B)$, violating Monotonicity.

We now introduce a weaker form of monotonicity, which is compatible with Size Neutrality. Let $i \in M$, $j \in N$ and $A \in \mathcal{A}(m,n)$. Say that there is a *universally favorable bias between i and j (in A)* if

$$a_i^j a_k^l > a_i^l a_k^j \text{ for all } k \in M \setminus i \text{ and all } l \in N \setminus j. \tag{3.6}$$

Notice that (3.6) is equivalent to

$$\frac{a_i^j}{a_i^l} > \frac{a_k^j}{a_k^l} \quad \text{and} \quad \frac{a_i^j}{a_k^j} > \frac{a_i^l}{a_k^l}$$

for all $k \in M \setminus i$ and all $l \in N \setminus j$. This means that the ratio between the number of supporters of party j and the number of supporters of any other party is larger in ethnic group i than in any other group. Equivalently, the ratio between the number of supporters from ethnic group i and the number of supporters from any other group is larger in party j than in any other party. We say that there is a *universally unfavorable bias between i and j (in A)* if all strict inequalities in (3.6) are replaced with the opposite strict inequalities.

For instance, there is a universally favorable bias between group 1 and party 1 in the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 6 & 4 & 4 \\ 2 & 8 & 7 \end{pmatrix}$$

since $a_1^1 a_2^2 > a_1^2 a_2^1$, $a_1^1 a_3^2 > a_3^1 a_1^2$, $a_1^1 a_2^3 > a_2^1 a_1^3$, and $a_1^1 a_3^3 > a_1^3 a_3^1$. On the other hand, there is a favorable bias between group 1 and party 1 but this bias is not *universally* favorable in the matrix

$$B = \begin{pmatrix} 4 & 3 & 2 \\ 6 & 4 & 4 \\ 2 & 8 & 7 \end{pmatrix}$$

since $b_1^1 b > b_1 b^1$ but $b_1^1 b_2^2 < b_1^2 b_2^1$. Observe also that when a universally favorable bias exists, it need not be unique. In A for instance, such a bias exists between group 1 and party 1 and also beween group 3 and party 2.

Weak Bias Monotonicity. *Let $i \in M$, $j \in N$ and $A, B \in \mathcal{A}(m,n)$. Suppose there is a universally favorable bias between i and j in A and suppose $b_i^j > a_i^j$ and $b_k^l = a_k^l$ if $(k,l) \neq (i,j)$. Then $F(B) > F(A)$. Likewise, if there is a universally unfavorable bias between i and j in A and if $b_i^j < a_i^j$ and $b_k^l = a_k^l$ whenever $(k,l) \neq (i,j)$, then again $F(B) > F(A)$.*

The functions F_1 and F_2 defined in (3.3) and (3.4) satisfy Weak Bias Monotonicity. Suppose the conditions stated in the axiom hold for two matrices $A, B \in \mathcal{A}(m,n)$ and consider the cross-product ratios for their 2×2 submatrices. All cross-products that do not depend on the (i,j)th cell will not be affected by the transformation. If there is a universally favorable bias between $i \in M$ and $j \in N$ in A, then the cross-product ratio in A containing a_i^j will be greater than one if a_i^j appears at its numerator and it will be smaller than one if a_i^j appears at its denominator. Formally,

$$
\begin{aligned}
r_{i,k}^{j,l}(A) &= \frac{a_i^j a_k^l}{a_i^l a_k^j} > 1 && \text{for all } (k,l) \in M \times N : k > i, l > j \\
r_{k,i}^{l,j}(A) &= \frac{a_k^l a_i^j}{a_k^j a_i^l} > 1 && \text{for all } (k,l) \in M \times N : k < i, l < j \\
r_{i,k}^{l,j}(A) &= \frac{a_i^l a_k^j}{a_i^j a_k^l} < 1 && \text{for all } (k,l) \in M \times N : k > i, l < j \\
r_{k,i}^{j,l}(A) &= \frac{a_k^j a_i^l}{a_k^l a_i^j} < 1 && \text{for all } (k,l) \in M \times N : k < i, l > j
\end{aligned}
$$

The following figure might help the interpretation of the conditions just stated.

$$
\begin{array}{ccccc}
a_1^1 & \cdots & a_1^j & \cdots & a_1^n \\
 & & & & \\
\vdots & \begin{matrix} k < i, l < j \\ r_{k,i}^{l,j}(A) > 1 \end{matrix} & & \begin{matrix} k < i, l > j \\ r_{k,i}^{j,l}(A) < 1 \end{matrix} & \vdots \\
 & & & & \\
a_i^1 & & a_i^j & & a_i^n \\
 & & & & \\
\vdots & \begin{matrix} k > i, l < j \\ r_{i,k}^{l,j}(A) < 1 \end{matrix} & & \begin{matrix} k > i, l > j \\ r_{i,k}^{j,l}(A) > 1 \end{matrix} & \vdots \\
 & & & & \\
a_m^1 & \cdots & a_m^j & \cdots & a_m^n
\end{array}
$$

The fact that F_1 and F_2 satisfy Weak Bias Monotonicity now immediately follows from

$$\rho^{j,l}_{i,k}(B) = \frac{b^j_i a^l_k}{a^l_i a^j_k} > \rho^{j,l}_{i,k}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l > j$$
$$\rho^{l,j}_{k,i}(B) = \frac{a^l_k b^j_i}{a^j_k a^l_i} > \rho^{l,j}_{k,i}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l < j$$
$$\rho^{l,j}_{i,k}(B) = \frac{b^j_i a^l_k}{a^l_i a^j_k} > \rho^{l,j}_{i,k}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l < j$$
$$\rho^{j,l}_{k,i}(B) = \frac{a^l_k b^j_i}{a^j_k a^l_i} > \rho^{j,l}_{k,i}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l > j$$

By a symmetric reasoning, whenever the bias between i and j is universally unfavorable in A, then the conditions on the cross-product ratios become

$$r^{j,l}_{i,k}(A) = \frac{a^j_i a^l_k}{a^l_i a^j_k} < 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l > j$$
$$r^{l,j}_{k,i}(A) = \frac{a^l_k a^j_i}{a^j_k a^l_i} < 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l < j$$
$$r^{l,j}_{i,k}(A) = \frac{a^l_i a^j_k}{a^j_i a^l_k} > 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l < j$$
$$r^{j,l}_{k,i}(A) = \frac{a^j_k a^l_i}{a^l_k a^j_i} > 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l > j$$

Consequently,

$$\rho^{j,l}_{i,k}(B) = \frac{a^l_i a^j_k}{b^j_i a^l_k} > \rho^{j,l}_{i,k}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l > j$$
$$\rho^{l,j}_{k,i}(B) = \frac{a^j_k a^l_i}{a^l_k b^j_i} > \rho^{l,j}_{k,i}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l < j$$
$$\rho^{l,j}_{i,k}(B) = \frac{a^l_i a^j_k}{b^j_i a^l_k} > \rho^{l,j}_{i,k}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k > i, l < j$$
$$\rho^{j,l}_{k,i}(B) = \frac{a^j_k a^l_i}{a^l_k b^j_i} > \rho^{j,l}_{k,i}(A) > 1 \qquad \text{for all } (k,l) \in M \times N : k < i, l > j$$

and $F_1(B) > F_1(A)$ and $F_2(B) > F_2(A)$.

The class of size neutral and weakly monotonic measures can be esplicitly identified. Fix a size neutral index $F : \mathcal{A}(m,n) \to \mathbb{R}_+$ and let $f : \mathcal{A}(m-1, n-1) \to \mathbb{R}_+$ be the function that represents it according to Theorem 1. Theorem 2 below identifies the restrictions imposed on f by Weak Bias Monotonicity under the additional assumption that F, hence also f, is differentiable. Define

$$Q(M,N) = \left\{((i,j),(k,l)) \in (M \times N)^2 \mid i < k \text{ and } j < l\right\}$$

Given $z = (z_i^j)_{i \in M \setminus m}^{j \in N \setminus n} \in \mathcal{A}(m-1, n-1)$, define, for each $((i,j),(k,l)) \in Q(M,N)$,

$$Z_{i,k}^{j,l} = \prod_{\substack{i'=i,\dots,k-1 \\ j'=j,\dots,l-1}} z_{i'}^{j'}. \tag{3.7}$$

Let $B^+(z)$ denote the set of all $(i,j) \in (M \setminus m) \times (N \setminus n)$ satisfying the following inequalities

$$\left.\begin{array}{ll} Z_{i,k}^{j,l} > 1 & \text{for all } (k,l) \in M \times N : k > i, l > j \\ Z_{k,i}^{l,j} > 1 & \text{for all } (k,l) \in M \times N : k < i, l < j \\ Z_{k,i}^{j,l} < 1 & \text{for all } (k,l) \in M \times N : k < i, l > j \\ Z_{i,k}^{l,j} < 1 & \text{for all } (k,l) \in M \times N : k > i, l < j \end{array}\right\} \tag{3.8}$$

and let $B^-(z)$ denote the set of all $(i,j) \in (M \setminus m) \times (N \setminus n)$ satisfying the opposite strict inequalities.

Theorem 2. *Let F be a differentiable size neutral index represented by f. The index F satisfies Weak Bias Monotonicity if and only if for all $z \in \mathcal{A}(m-1, n-1)$,*

$$\frac{\partial f(z)}{\partial z_{i-1}^{j-1}} z_{i-1}^{j-1} + \frac{\partial f(z)}{\partial z_i^j} z_i^j > \frac{\partial f(z)}{\partial z_{i-1}^j} z_{i-1}^j + \frac{\partial f(z)}{\partial z_i^{j-1}} z_i^{j-1} \tag{3.9}$$

whenever $(i,j) \in B^+(z)$ and the opposite strict inequality holds whenever $(i,j) \in B^-(z)$ with the convention that expressions involving undefined terms are zero.

The reader can find the proof in Appendix 3.B. Simple sufficient conditions guaranteeing Weak Bias Monotonicity are stated in the following corollary. Notice that these conditions are satisfied by F_1 and F_2.

Corollary 3. *A differentiable size neutral index F represented by f satisfies Weak Bias Monotonicity if for all $z \in \mathcal{A}(m-1, n-1)$ and all $(i,j) \in (M \setminus m) \times (N \setminus n)$, we have*

$$\frac{\partial f(z)}{\partial z_i^j} > 0 \textit{ whenever } z_i^j > 1;$$
$$\frac{\partial f(z)}{\partial z_i^j} < 0 \textit{ whenever } z_i^j < 1.$$

Proof. Let f satisfy the conditions stated in the corollary. If there is a universally favorable bias between i and j in A, then $(i,j) \in B^+(z)$, $z_i^j > 1$, $z_{i-1}^{j-1} > 1$, $z_{i-1}^j < 1$ and $z_i^{j-1} < 1$ and

$$\frac{\partial f(z)}{\partial z_{i-1}^{j-1}} z_{i-1}^{j-1} + \frac{\partial f(z)}{\partial z_i^j} z_i^j > 0 > \frac{\partial f(z)}{\partial z_{i-1}^j} z_{i-1}^j + \frac{\partial f(z)}{\partial z_i^{j-1}} z_i^{j-1}$$

Similarly, if there is a universally unfavorable bias between i and j, then $(i,j) \in B^-(z)$, $z_i^j < 1$, $z_{i-1}^{j-1} < 1$, $z_{i-1}^j > 1$ and $z_i^{j-1} > 1$ and

$$\frac{\partial f(z)}{\partial z_{i-1}^{j-1}} z_{i-1}^{j-1} + \frac{\partial f(z)}{\partial z_i^j} z_i^j < 0 < \frac{\partial f(z)}{\partial z_{i-1}^j} z_{i-1}^j + \frac{\partial f(z)}{\partial z_i^{j-1}} z_i^{j-1}$$

Therefore, condition (3.9) is satisfies, and by Theorem 2, F satisfies Weak Bias Monotonicity. $\square$

3.4 Conclusions

The paper proposes the first axiomatic approach to the measurement of ethnic voting, defined as the extent to which individuals' voting decisions depend on their ethnicity. The problem is related to other two widely studied questions. The first is the measurement of the degree of association between categorical variables. The second consists in the assessment of the level of segregation among districts in a city. Despite this close relation, however, none of the measures proposed for the two problems is suitable for our purposes. The reason relies on a size neutral property that we impose on the index of ethnic voting. Size Neutrality states that the index should not be affected by the rescaling of any row or any column of the matrix representing the ethnic voting profile of a country. In other words, we require that the measure depends on the behavior of the different groups and on the distribution of support of the different parties, but not on their sizes. The most used index of association between categorical variables - the Chi-square index - and the existing measures of segregation do not satisfy this property. Nor do previously proposed measures of ethnic voting.

The central problem of the paper is to identify the strongest monotonicity axiom that is compatible with Size Neutrality. Monotonicity is a property that specifies which changes in the cells of a matrix are associated to higher degrees of ethnic voting. We propose a *Weak Bias Monotonicity* axiom that is for sure compatible with Size Neutrality. The axiom is based on the concept of *universally favorable or unfavorable bias* between a group and a party. A universally favorable bias exists between a party and a group if the ratio between the number of supporters of that party and the number of supporters of any other party is largest for that specific ethnic group. The axiom requires that any increase in the support that the party receives from that group raises the degree of ethnic voting in the country.

In the paper, we also propose a stronger version of monotonicity, which is satisfied by the Chi-square index. This is based on a concept of bias defined in terms of averages rather than individual comparisons. A *favorable bias* exists between a group and a party if the share of support that the group assigns to the party is greater than the average support the party receives at the country level. The bias is *unfavorable* in the opposite case. Our *Monotonicity* axiom requires that all changes of the matrix that reinforce the biases are associated to higher degrees of ethnic voting. A question that remains unsolved is whether this version of monotonicity is compatible with Size Neutrality. In the paper, we show that the most natural size neutral indices fail to satisfy Monotonicity. We also show that a particular preordering based on the leximax criterion satisfies Size Neutrality and Monotonicity only within a particular class of matrices.

3.A Appendix A: Existing Measures

Huber (2012)

For any matrix $A \in \mathcal{A}(m,n)$, let the *distance between two groups* $i, k \in M$, $i \neq k$, be defined as

$$d_{ik}(A) = \sqrt{\frac{1}{2}\sum_{j\in N}\left(\frac{a_i^j}{a_i} - \frac{a_k^j}{a_k}\right)^2}$$

Each component of $d_{ik}(A)$ measures the difference between the share of people in groups i and k that vote for party $j \in N$. The distance between two groups is equal to zero whenever they have the same voting behavior and reaches its maximum when they concentrate all their support on two different parties. Similarly, for any matrix $A \in \mathcal{A}(m,n)$, define the *distance between two parties* $j, l \in N$, $j \neq l$, as

$$d^{jl}(A) = \sqrt{\frac{1}{2}\sum_{i \in M}\left(\frac{a_i^j}{a^j} - \frac{a_i^l}{a^l}\right)^2}$$

Each component of $d^{jl}(A)$ measures the difference between the share of electoral support coming from i in party j and k. The distance is equal to zero when the distribution of electoral support across ethnic groups is identical for the two parties, while it is maximized when the two parties only receive votes from two different ethnic groups.

The measures proposed in Huber (2012) are based on two different ways of aggregating all $d_{ik}(A)$ and $d^{jl}(A)$. The *Group Voting Fractionalization (GVF)* and *Party Voting Fractionalization (PVF)* measures are defined as

$$GVF(A) = \sum_{i \in M}\sum_{k \in M}\left(\frac{a_i}{a}\right)\left(\frac{a_k}{a}\right) d_{ik}(A) \tag{3.10}$$

$$PVF(A) = \sum_{j \in N}\sum_{l \in N}\left(\frac{a^j}{a}\right)\left(\frac{a^l}{a}\right) d^{jl}(A) \tag{3.11}$$

while the *Group Voting Polarization (GVP)* and *Party Voting Polarization (PVP)* measures are defined as

$$GVP(A) = 4\sum_{i \in M}\sum_{k \in M}\left(\frac{a_i}{a}\right)\left(\frac{a_k}{a}\right)^2 d_{ik}(A) \tag{3.12}$$

$$PVP(A) = 4\sum_{j \in N}\sum_{l \in N}\left(\frac{a^j}{a}\right)\left(\frac{a^l}{a}\right)^2 d^{jl}(A) \tag{3.13}$$

The difference between the two pairs of measures relies on the weights used to aggregate the distances between parties or groups. Under the *fractionalization* approach, ethnic voting is maximized in presence of many different groups or many different parties of equal size. Under the *polarization approach*, instead, the measure is highest in presence of two big groups or two big parties.

Measures of Evenness in Segregation Literature

Districts in segregation literature are equivalent to political parties in the measurement of ethnic voting. For any matrix $A \in \mathcal{A}(m,n)$, the Dissimilarity ($D(A)$), Gini ($G(A)$) and Entropy ($H(A)$) indices are defined as[2]

[2] The Dissimilarity and Gini Indices were originally formulated under the assumption of only two groups (Jahn et al. (1947)). We here report their generalized versions (Morgan (1975), Sakoda (1981), Reardon (1998))

$$D(A) = \frac{1}{2I} \sum_{i \in M} \frac{a_i}{a} \sum_{j \in N} \frac{a^j}{a} \left| \frac{a_i^j a}{a_i a^j} - 1 \right|$$

$$G(A) = \frac{1}{2I} \sum_{i \in M} \sum_{j \in N} \sum_{l \in N} \frac{a^j}{a} \frac{a^l}{a} \left| \frac{a_i^j}{a^j} - \frac{a_i^l}{a^l} \right|$$

$$H(A) = \frac{1}{E} \sum_{i \in M} \frac{a_i}{a} \sum_{j \in N} \frac{a^j}{a} \left(\frac{a_i^j a}{a_i a^j} \right) \ln \left(\frac{a_i^j a}{a_i a^j} \right)$$

where I and E are defined as

$$I = \sum_{i \in M} \frac{a_i}{a} \left(1 - \frac{a_i}{a} \right)$$

$$E = \sum_{i \in M} \frac{a_i}{a} \ln \left(\frac{a}{a_i} \right)$$

Reardon and Firebaugh (2002) define the class of segregation indices that measure the *disproportionality* in group proportions across districts. For any matrix $A \in \mathcal{A}(m.n)$, let $W(A)$ be the quantity

$$W(A) = \sum_{i \in M} \frac{a_i}{a} \sum_{j \in N} \frac{a^j}{a} g \left(\frac{a_i^j a}{a_i a^j} \right)$$

for some function $g : \mathcal{A}(m,n) \to \mathbb{R}_{++}$. The class of dissimilarity indices can be defined as

$$S(A) = \frac{W(A)}{W^*}$$

where

$$W^* = \max_{A \in \mathcal{A}(m,n)} W(A)$$

The authors show that the three main indices reported above can be derived from S through different specifications of the function g. Notice that

$$S(A) = \frac{a}{W^*} \sum_{i \in M} \sum_{j \in N} \frac{a_i a^j}{a} g \left(\frac{a_i^j a}{a_i a^j} \right)$$

which is a normalized general formulation of an index of association between two categorical variables. For example, if

$$g \left(\frac{a_i^j a}{a_i a^j} \right) = \left(\frac{a_i^j a}{a_i a^j} - 1 \right)^2$$

then

$$S(A) = \frac{a}{W^*} \chi^2(A)$$

For a fixed-population size, then, segregation measures are measures of association.

Size Neutrality

One can show that none of the measures introduced above satisfies Size Neutrality. Consider the matrices $A, B \in \mathcal{A}(2,2)$ such that

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & 1 \\ 2 & 2 \end{pmatrix}$$

Notice that B is obtained by multiplying the first column of A by two, i.e. $B = (2A^1, A^2)$. By Size Neutrality it must be that $F(A) = F(B)$. As shown in the table below, however, all the measures introduced above compute different values for the two matrices.

	GVF	PVF	GVP	PVP	χ^2	D	G	H
A	0.17	0.17	0.33	0.33	0.67	0.67	0.33	0.08
B	0.15	0.15	0.3	0.3	0.9	0.6	0.3	0.07

Group Bias Monotonicity and Monotonicity

We here show that the Chi-Square index does not satisfy Group Bias Monotonicity. Consider the following matrices

$$A = \begin{pmatrix} 0.7 & 2.7 & 1.6 \\ 1.3 & 1.3 & 2.4 \\ 3 & 1 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 0.7 & 2.7 & 1.6 \\ 1.3 & 1.4 & 2.3 \\ 3 & 1 & 6 \end{pmatrix}$$

There is a favorable bias between $i = 2$ and $j = 2$ and an unfavorable bias between $i = 2$ and $j' = 3$ in A, since $\beta_2^2(A) = 1,04$ and $\beta_2^3(A) = 0.96$. Furthermore, $b_2^2 = a_2^2 + 0.1$ and $b_2^3 = a_2^3 - 0.1$. Group Bias Monotonicity therefore implies $F(B) > F(A)$. Yet one computes $\chi^2(B) \approx 3.436 < \chi^2(A) \approx 3.456$. The reason for the incompatibility is that, when applying Group Bias Monotonicity, we do not only affect the terms in the index that correspond to elements a_i^j and $a_i^{j'}$. Indeed, $b^j = a^j + \delta$ and $b^{j'} = a^{j'} - \delta$. This might decrease some terms corresponding to elements a_k^j and $a_{k'}^{j'}$, $k, k' \neq i$. In the example above,

$$\frac{\left(b_1^2 - \frac{b_1 b^2}{b}\right)}{\frac{b_1 b^2}{b}} \approx 1.59 < \frac{\left(a_1^2 - \frac{a_1 a^2}{a}\right)}{\frac{a_1 a^2}{a}} \approx 1.68$$

and

$$\frac{\left(b_1^3 - \frac{b_1 b^3}{b}\right)}{\frac{b_1 b^3}{b}} \approx 0.31 < \frac{\left(a_1^3 - \frac{a_1 a^3}{a}\right)}{\frac{a_1 a^3}{a}} \approx 0.32$$

As the decrease in these terms exceeds the increase in the other terms, $\chi^2(B) < \chi^2(A)$.

The index however satisfies our Monotonicity axiom. Notice that for any given $A \in \mathcal{A}(m,n)$, $\chi^2(A)$ can be rewritten as

$$\chi^2(A) = \sum_{(i,j)\in M\times N} \left[(a_i^j)^2 \frac{a}{a_i a^j} - 2a_i^j + \frac{a_i a^j}{a} \right]$$

Let A and B be such that the conditions in the Monotonicity axiom are satisfied. Then,

$$\begin{aligned}\chi^2(B) = &\sum_{(k,l)\notin\{(i,j),(i',j),(i,j'),(i',j')\}} \left[(a_k^l)^2\frac{a}{a_k a^l} - 2a_k^l + \frac{a_k a^l}{a}\right]\\ &+ \left[(a_i^j+\delta)^2\frac{a}{a_i a^j} - 2(a_i^j+\delta) + \frac{a_i a^j}{a}\right]\\ &+ \left[(a_i^{j'}-\delta)^2\frac{a}{a_i a^{j'}} - 2(a_i^{j'}-\delta) + \frac{a_i a^{j'}}{a}\right]\\ &+ \left[(a_{i'}^{j'}+\delta)^2\frac{a}{a_{i'} a^{j'}} - 2(a_{i'}^{j'}+\delta) + \frac{a_{i'} a^{j'}}{a}\right]\\ &+ \left[(a_{i'}^{j}-\delta)^2\frac{a}{a_{i'} a^{j}} - 2(a_{i'}^{j}-\delta) + \frac{a_{i'} a^{j}}{a}\right]\end{aligned}$$

and

$$\frac{\partial\chi^2(B)}{\partial\delta} = 2a\left(\frac{a_i^j+\delta}{a_i a^j} - \frac{a_i^{j'}-\delta}{a_i a^{j'}} + \frac{a_{i'}^{j'}+\delta}{a_{i'} a^{j'}} - \frac{a_{i'}^{j}-\delta}{a_{i'} a^{j}}\right) > 0$$

for any δ.

3.B Appendix B: Proofs

Proof of Theorem 1

Proof. Sufficiency follows from the fact that for any $\lambda > 0$, $i \in M$ and $j \in N$, we have $r(\lambda A_i, A_{-i}) = r(A) = r(\lambda A^j, A^{-j})$.

Now let F be a size neutral index and let $A, B \in \mathcal{A}(m,n)$ be such that $r(A) = r(B)$. We show that $F(A) = F(B)$. We use the following notation: if $C, D \in \mathcal{A}(p,q)$, then $C * D$ is the matrix in $\mathcal{A}(p,q)$ whose (i,j)th entry is the product $c_i^j d_i^j$ of the (i,j)th entries of C and D.

Step 1. There exist $\lambda, \lambda_2, \lambda^2 > 0$ such that

$$\begin{pmatrix} b_1^1 & b_1^2 \\ b_2^1 & b_2^2 \end{pmatrix} = \begin{pmatrix} \lambda & \lambda^2 \\ \lambda_2 & \frac{\lambda_2\lambda^2}{\lambda} \end{pmatrix} * \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix}. \tag{3.14}$$

To see this, simply define

$$\lambda = \frac{b_1^1}{a_1^1} \;\; \lambda_2 = \frac{b_2^1}{a_2^1} \;\; \lambda^2 = \frac{b_1^2}{a_1^2}$$

Trivially, $b_1^1 = \lambda a_1^1$, $b_2^1 = \lambda_2 a_2^1$, and $b_1^2 = \lambda^2 a_1^2$. And since $r(A) = r(B)$ implies

$$\frac{a_1^1 a_2^2}{a_1^2 a_2^1} = \frac{b_1^1 b_2^2}{b_1^2 b_2^1},$$

we have

$$b_2^2 = \left(\frac{b_1^2 b_2^1 a_1^1}{a_1^2 a_2^1 b_1^1}\right) a_2^2 = \frac{\lambda_2\lambda^2}{\lambda} a_2^2,$$

proving (3.14).

Step 2. There exist λ; $\lambda_2, ..., \lambda_m$; $\lambda^2, ..., \lambda^n > 0$ such that

$$B = \begin{pmatrix} \lambda & \lambda^2 & \lambda^3 & \cdots & \lambda^n \\ \lambda_2 & \frac{\lambda_2\lambda^2}{\lambda} & \frac{\lambda_2\lambda^3}{\lambda} & \cdots & \frac{\lambda_2\lambda^n}{\lambda} \\ \vdots & & & & \vdots \\ \lambda_m & \frac{\lambda_m\lambda^2}{\lambda} & \frac{\lambda_m\lambda^3}{\lambda} & \cdots & \frac{\lambda_m\lambda^n}{\lambda} \end{pmatrix} * A. \tag{3.15}$$

For ease of exposition we assume $m \leq n$; the case $n \leq m$ is treated similarly. From Step 1 we know that there exist $\lambda, \lambda_2, \lambda^2 > 0$ such that $b_i^j = \left(\frac{\lambda_i\lambda^j}{\lambda}\right) a_i^j$ for all $(i,j) \in M \times N$ such that $\max(i,j) \leq 2$. Here and below, $\lambda_1 = \lambda^1 = \lambda$. The rest of the proof proceeds by induction on $\max(i,j)$.

Fix an integer k such that $2 < k < n$ and make the following *induction hypothesis*: there exist λ; $\lambda_2, ..., \lambda_{\min(k-1,m)}$; $\lambda^2, ..., \lambda^{k-1} > 0$ such that $b_i^j = \left(\frac{\lambda_i\lambda^j}{\lambda}\right) a_i^j$ for all $(i,j) \in M \times N$ such that $\max(i,j) \leq k-1$. We make two claims. The first is that there exists a number $\lambda^k > 0$ such that

$$b_i^k = \left(\frac{\lambda_i\lambda^k}{\lambda}\right) a_i^k \text{ for } i = 1, ..., \min(k-1, m). \tag{3.16}$$

The second claim is that, if $k-1 < m$, there exists a number $\lambda_k > 0$ such that

$$b_k^j = \left(\frac{\lambda_k\lambda^j}{\lambda}\right) a_k^j \text{ for } j = 1, ..., k. \tag{3.17}$$

In order to prove the first claim, define $\lambda^k = \frac{b_1^k}{a_1^k}$. By construction,

$$b_1^k = \lambda^k a_1^k = \frac{\lambda_1\lambda^k}{\lambda} a_1^k$$

Next, fix an integer q such that $1 < q < k$ and make the *auxiliary induction hypothesis* that $b_i^k = \left(\frac{\lambda_i\lambda^k}{\lambda}\right) a_i^k$ for $i = 1, ..., q-1$. By the assumption $r(A) = r(B)$ we have

$$\frac{a_{q-1}^{k-1}a_q^k}{a_{q-1}^k a_q^{k-1}} = \frac{b_{q-1}^{k-1}b_q^k}{b_{q-1}^k b_q^{k-1}},$$

hence,

$$b_q^k = \left(\frac{a_{q-1}^{k-1}b_{q-1}^k b_q^{k-1}}{a_{q-1}^k a_q^{k-1} b_{q-1}^{k-1}}\right) a_q^k. \tag{3.18}$$

By the (main) induction hypothesis,

$$b_q^{k-1} = \left(\frac{\lambda_q\lambda^{k-1}}{\lambda}\right) a_q^{k-1} \quad \text{and} \quad b_{q-1}^{k-1} = \left(\frac{\lambda_{q-1}\lambda^{k-1}}{\lambda}\right) a_{q-1}^{k-1}$$

whereas by the auxiliary induction hypothesis

$$b_{q-1}^{k-1} = \left(\frac{\lambda_{q-1}\lambda^k}{\lambda}\right) a_{q-1}^k$$

Therefore (3.18) implies

$$b_q^k = \left(\frac{\lambda_q \lambda^k}{\lambda}\right) a_q^k,$$

proving our first claim.

The proof of the second claim is similar. Suppose $k-1<m$. Define $\lambda_k = \frac{b_k^1}{a_k^1}$. Then $b_k^1 = \lambda_k a_k^1 = \frac{\lambda_k \lambda^1}{\lambda} a_k^1$. Fix an integer q such that $1 < q \leq k$ (notice the weak inequality this time) and make the *second auxiliary hypothesis* that $b_k^j = \left(\frac{\lambda_k \lambda^j}{\lambda}\right) a_k^j$ for $j = 1, ..., q-1$. By assumption,

$$\frac{a_{k-1}^{q-1} a_k^q}{a_k^{q-1} a_{k-1}^q} = \frac{b_{k-1}^{q-1} b_k^q}{b_k^{q-1} b_{k-1}^q},$$

hence

$$b_k^q = \left(\frac{a_{k-1}^{q-1} b_k^{q-1} b_{k-1}^q}{a_k^{q-1} a_{k-1}^q b_{k-1}^{q-1}}\right) a_k^q. \tag{3.19}$$

By the main induction hypothesis,

$$b_{k-1}^{q-1} = \left(\frac{\lambda_{k-1} \lambda^{q-1}}{\lambda}\right) a_{k-1}^{q-1}$$

By the second auxiliary induction hypothesis,

$$b_k^{q-1} = \left(\frac{\lambda_k \lambda^{q-1}}{\lambda}\right) a_k^{q-1}$$

Finally,

$$b_{k-1}^q = \left(\frac{\lambda_{k-1} \lambda^q}{\lambda}\right) a_k^{q-1}$$

This holds by virtue of the main induction hypothesis if $q < k$ and by virtue of the first claim we proved if $q = k$. Therefore (3.19) implies

$$b_k^q = \left(\frac{\lambda_k \lambda^q}{\lambda}\right) a_k^q,$$

proving our second claim.

Step 3. Because of (3.15), the matrix B can be obtained by (i) multiplying A_1 and A^1 by $\sqrt{\lambda}$, (ii) multiplying A_i by $\frac{\lambda_i}{\sqrt{\lambda}}$ for each $i = 2, ..., m$, and (iii) multiplying A^j by $\frac{\lambda^j}{\sqrt{\lambda}}$ for each $j = 2, ..., n$. Size Neutrality now implies that $F(A) = F(B)$. □

Proof of Fact 3

In this section, we show that preordering $\succsim$ on $\mathcal{A}(2,n)$ defined in 3.5 satisfies Size Neutrality and Monotonicity.

Proof. Since $\succsim^L$ is antisymmetric, $A \sim B \Leftrightarrow \rho(A) = \rho(B)$. If $B = (\lambda A_i, A_{-i})$ or $B = (\lambda A^j, A^{-j})$ for some $\lambda \in \mathbb{R}_{++}$, $i \in \{1,2\}$, $j \in N$, then $\rho(A) = \rho(B)$ and $A \sim B$. This proves that $\succsim$ is size neutral.

We now show that $\succsim$ is also monotonic. Let $\{j,l\} \in 2^N$ and consider two matrices $A, B \in \mathcal{A}(2,n)$. Without loss of generality, suppose that in A there is a favorable bias between $i=1$ and $j=1$ and between $i=2$ and $j=2$ and an unfavorable bias between $i=1$ and $j=2$ and between $i=2$ and $j=1$. Also, suppose that there exists a $\delta > 0$ such that $b_1^1 = a_1^1 + \delta$, $b_2^2 = a_2^2 + \delta$, $b_1^2 = a_1^2 - \delta$ and $b_2^1 = a_2^1 - \delta$. We claim that $B \succ A$ or, equivalently, $\rho(B) \succ^L \rho(A)$.

To prove the claim, we will show that if $\rho^{\{j,l\}}(B) < \rho^{\{j,l\}}(A)$ for some $\{j,l\} \in 2^N$, then there must exist $\{j',l'\} \in 2^N$, $\{j',l'\} \neq \{j',l'\}$, such that

$$\rho^{\{j,l\}}(B) < \rho^{\{j',l'\}}(B) \quad \text{and} \quad \rho^{\{j',l'\}}(A) < \rho^{\{j',l'\}}(B)$$

The leximax ordering then implies that the comparison between $\rho^{\{j,l\}}(A)$ and $\rho^{\{j,l\}}(B)$ is irrelevant.

We begin the argument by noticing that $\rho^{\{j,l\}}(B) = \rho^{\{j,l\}}(A)$ whenever $j,l \notin \{1,2\}$. Moreover, our assumptions about biases imply that $a_1^1 a_2^2 > a_1^2 a_2^1$, so that

$$\rho^{\{1,2\}}(A) = \frac{a_1^1 a_2^2}{a_1^2 a_2^1} < \frac{(a_1^1+\delta)(a_2^2+\delta)}{(a_1^2-\delta)(a_2^1-\delta)} = \rho^{\{1,2\}}(B)$$

Thus, $\rho^{\{j,l\}}(B) < \rho^{\{j,l\}}(A)$ only if i) $\{j,l\} = \{1,l\}$, $l \neq 2$ or ii) $\{j,l\} = \{2,l\}$, $l \neq 1$. Suppose first that there exists $l \neq 2$ such that

$$\rho^{\{1,l\}}(B) < \rho^{\{1,l\}}(A)$$

This can only occur if

$$\rho^{\{1,l\}}(A) = \frac{a_2^1 a_1^l}{a_1^1 a_2^l} \quad \text{and} \quad \rho^{\{1,l\}}(B) = \frac{(a_2^1-\delta)a_1^l}{(a_1^1+\delta)a_2^l}$$

If this is the case, then $\rho^{\{1,l\}}(A) \geq 1$ and $\rho^{\{1,2\}}(A) > 1$ imply that the product

$$\rho^{\{1,l\}}(A)\rho^{\{1,2\}}(A) = \frac{a_2^1 a_1^l}{a_1^1 a_2^l}\frac{a_1^1 a_2^2}{a_1^2 a_2^1} = \frac{a_1^l a_2^2}{a_2^l a_1^2} > 1$$

so that

$$\rho^{\{2,l\}}(A) = \frac{a_1^l a_2^2}{a_2^l a_1^2} > \rho^{\{1,l\}}(A)$$

But then,

$$\rho^{\{2,l\}}(B) = \frac{a_1^l(a_2^2+\delta)}{a_2^l(a_1^2-\delta)} > \rho^{\{2,l\}}(A) > \rho^{\{1,l\}}(A) > \rho^{\{1,l\}}(B)$$

Suppose next that there exists $l \neq 1$ such that

$$\rho^{\{2,l\}}(B) < r^{\{2,l\}}(A)$$

By the same argument used above, one can show that

$$\rho^{\{1,l\}}(B) > \rho^{\{1,l\}}(A) > \rho^{\{2,l\}}(A) > \rho^{\{2,l\}}(B)$$

□

Proof of Theorem 2

Proof. For any matrix $A \in \mathcal{A}(m,n)$, let $z_i^j = r_{i,i+1}^{j,j+1}(A)$, so that $z = r(A)$ and

$$Z_{i,k}^{j,l} = \prod_{\substack{i'=i,\dots,k-1\\ j'=j,\dots,l-1}} r_{i',i'+1}^{j',j'+1}(A) = r_{i,k}^{j,l}(A)$$

for all $((i,j),(k,l)) \in Q(M,N)$. Let F satisfy Weak Bias Monotonicity and assume that, for some $i \in M$ and $j \in N$, there is a universally favorable bias between i and j in A. By our discussion above, then

$$\begin{aligned}
Z_{i,k}^{j,l} &= r_{i,k}^{j,l}(A) > 1 &\quad& \text{for all } (k,l) \in M \times N : k > i, l > j\\
Z_{k,i}^{l,j} &= r_{k,i}^{l,j}(A) > 1 && \text{for all } (k,l) \in M \times N : k < i, l < j\\
Z_{i,k}^{l,j} &= r_{i,k}^{l,j}(A) < 1 && \text{for all } (k,l) \in M \times N : k > i, l < j\\
Z_{k,i}^{j,l} &= r_{k,i}^{j,l}(A) < 1 && \text{for all } (k,l) \in M \times N : k < i, l > j
\end{aligned}$$

and $(i,j) \in B^+(z)$. Let $i \geq 2$ and $j \geq 2$. If F satisfies Weak Bias Monotonicity, then $dF(A)/da_i^j > 0$. That is,

$$\begin{aligned}
\frac{dF(A)}{da_i^j} =& \frac{\partial f(r(A))}{\partial r_{i-1,i}^{j-1,j}} \frac{\partial r_{i-1,i}^{j-1,j}}{\partial a_i^j} + \frac{\partial f(r(A))}{\partial r_{i,i+1}^{j,j+1}} \frac{\partial r_{i,i+1}^{j,j+1}}{\partial a_i^j}\\
&+ \frac{\partial f(r(A))}{\partial r_{i-1,i}^{j,j+1}} \frac{\partial r_{i-1,i}^{j,j+1}}{\partial a_i^j} + \frac{\partial f(r(A))}{\partial r_{i,i+1}^{j-1,j}} \frac{\partial r_{i,i+1}^{j-1,j}}{\partial a_i^j}\\
=& \frac{\partial f(r(A))}{\partial r_{i-1,i}^{j-1,j}} \frac{a_{i-1}^{j-1}}{a_{i-1}^{j} a_i^{j-1}} + \frac{\partial f(r(A))}{\partial r_{i,i+1}^{j,j+1}} \frac{a_{i+1}^{j+1}}{a_i^{j+1} a_{i+1}^{j}}\\
&- \frac{\partial f(r(A))}{\partial r_{i-1,i}^{j,j+1}} \frac{a_{i-1}^{j} a_i^{j+1}}{a_{i-1}^{j+1} (a_i^j)^2} - \frac{\partial f(r(A))}{\partial r_{i,i+1}^{j-1,j}} \frac{a_i^{j-1} a_{i+1}^{j}}{a_{i+1}^{j-1} (a_i^j)^2}\\
=& \left(\frac{\partial f(z)}{z_{i-1}^{j-1}} z_{i-1}^{j-1} + \frac{\partial f(z)}{\partial z_i^j} z_i^j - \frac{\partial f(z)}{\partial z_{i-1}^{j}} z_{i-1}^{j} - \frac{\partial f(z)}{\partial z_i^{j-1}} z_i^{j-1} \right) a_i^j > 0
\end{aligned}$$

which directly implies (3.9). It follows immediately that, if $i = 1$, $j = 1$ or both, then

$$\begin{aligned}
\frac{dF(A)}{da_1^j} &= \left(\frac{\partial f(z)}{\partial z_1^j} z_1^j - \frac{\partial f(z)}{\partial z_1^{j-1}} z_1^{j-1} \right) a_1^j > 0 &\qquad& (j \geq 2)\\
\frac{dF(A)}{da_i^1} &= \left(\frac{\partial f(z)}{\partial z_i^1} z_i^1 - \frac{\partial f(z)}{\partial z_{i-1}^1} z_{i-1}^1 \right) a_i^1 > 0 && (i \geq 2)\\
\frac{dF(A)}{da_1^1} &= \left(\frac{\partial f(z)}{\partial z_1^1} z_1^1 \right) a_1^1 > 0
\end{aligned}$$

The case of universally unfavorable bias between i and j is proven in the same way, just by noticing that now

$$
\begin{aligned}
&Z_{i,k}^{j,l} = r_{i,k}^{j,l}(A) < 1 && \text{for all } (k,l) \in M \times N : k > i, l > j \\
&Z_{k,i}^{l,j} = r_{k,i}^{l,j}(A) < 1 && \text{for all } (k,l) \in M \times N : k < i, l < j \\
&Z_{i,k}^{l,j} = r_{i,k}^{l,j}(A) > 1 && \text{for all } (k,l) \in M \times N : k > i, l < j \\
&Z_{k,i}^{j,l} = r_{k,i}^{j,l}(A) > 1 && \text{for all } (k,l) \in M \times N : k < i, l > j
\end{aligned}
$$

so that $(i,j) \in B^-(z)$, and $dF(A)/da_i^j < 0$.

To prove that (3.9) is also sufficient, consider a matrix $A \in \mathcal{A}(m,n)$ such that for some $i \in M$ and $j \in N$, there exists a universally favorable bias between i and j. Then $(i,j) \in B^+(z)$. Furthermore, for any matrix $B \in \mathcal{A}(m,n)$ such that $b_i^j > a_i^j$ and $b_k^l = a_k^l$, for all $(k,l) \neq (i,j)$, $(i,j) \in B^+(\tilde{z})$, with $\tilde{z} \equiv r(B)$. That is, whenever there is a universally favorable bias between i and j in A, any increase in a_i^j does not change the direction of the inequality in (3.9). This immediately implies that $F(B) > F(A)$. Similarly, if there exists a universally unfavorable bias between i and j in A, then $(i,j) \in B^-(z)$ and $(i,j) \in B^-(\tilde{z})$ for all $B \in \mathcal{A}(m,n)$ such that $b_i^j < a_i^j$ and $b_k^l = a_k^l$, for all $(k,l) \neq (i,j)$. Then again, $F(B) > F(A)$. $\square$

Conclusions

This thesis has addressed questions related to two main goals of elections: the aggregation of the different preferences in the population and the improvement of politicians' performances through accountability.

In the first chapter, we showed how common beliefs about minority representation under proportional representation systems might be completely reversed when seat allocation within party-lists is included in the analysis. We distinguished between closed and open list proportional representation and focused on the behavior of single candidates rather than parties. Our results show that minority representation strongly depends on the degree of voters' control on the selection of their representatives, measured by the number of preferential votes that can be cast for candidates within a party-list. Furthermore, this dependence is non-monotonic. The two extreme cases where voters can only vote for party-lists (closed list systems) and where they are allowed to express many preferential votes both lead to Parliaments that only represent the majority. Minorities are perfectly represented under open list when the number of preferential votes that can be cast is at most equal to their relative size. Quite surprisingly, there cannot be cases where minorities are over-represented in Parliament.

In addition to the main results, our first chapter produced interesting conclusions on the comparison between proportional representation and first-past-the-post systems. We analyzed first-past-the-post with and without primaries, which were interpreted as a way to increase voters' control on politicians' selection under this system. Two main results emerged from the analysis: first, first-past-the-post without primaries could potentially guarantee much higher levels of minority representation than any form of proportional representation. Second, increasing voters' control on candidates' selection has a clear negative effect under first-past-the-post. This system indeed produces Parliaments where no minority candidate is elected.

Finally, the chapter opens the way for future research on closed and open list proportional representation systems. Endogenous party formation and party discipline are elements that we only partially considered as extensions of our main model. Both of them deserve a deeper and more articulated analysis.

The focus of our second chapter was on the impact of re-electoral concerns on politicians' behavior. We there showed that the possibility of being re-elected for a second term might induce incumbents to address issues that are not the most relevant ones for the electorate. This happens because relevant issues are too difficult to solve and failing to successfully address them would constitute a bad signal about incumbent's quality. Under these circumstances, politicians prefer to address issues that, despite being less important, guarantee a higher probability of success.

The chapter therefore highlights a possible drawback of electoral competition. Incumbents that, absent any re-electoral concern, would act in the interest of the electorate might modify their behavior and take sub-optimal decisions when they have to stand for re-election. In the analysis, we compared voters' welfare under the assumptions that politicians are subject to a one- or two-term limit. We showed that, in addition to the positive effects of elections already stressed by the literature, a negative one could arise, which is related to the induced change in politicians' incentives during the first term. Most importantly, there exist cases in which the negative effect overcomes the positive ones, so that the very presence of elections is detrimental to voters' welfare.

In the third chapter, joint work with Yves Sprumont, we returned to the role of elections as a way to aggregate the diverse preferences in the population. The focus of the chapter was on the driving forces behind people's voting decision. More precisely, we there analyzed the problem of constructing a measure of *ethnic voting*. A population votes ethnically if there is a strong association between individuals' voting decisions and their ethnic belonging, or, in other words, when people belonging to the same ethnic group tend to vote for the same party (and each party's support primarily comes from a different ethnic group).

The main property we imposed on the measure of ethnic voting was neutrality with respect to the size of groups and parties. In our view, changes in the total number of people belonging to a group or voting for a party should be captured by the measure only if they are associated to changes in voting patterns by the different groups. Whenever the size of a group increases, but individuals within it still assign the same share of support to each party, ethnic voting should not increase. The same holds if, for example, a party doubles its support because each ethnic group gives to it twice as many votes as before.

The main contribution of our third chapter was the characterization of the class of size neutral measures of ethnic voting. Together with it, we also proposed a monotonicity axiom that is compatible with our main property. Monotonicity defines the changes in people's voting behavior that increase or decrease the level of ethnic voting in a country. Future research will focus on the identification of stronger monotonicity axioms that can still be compatible with size neutrality.

Bibliography

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.

Ames, B. (1995a). Electoral rules, constituency pressures, and pork barrel: Bases of voting in the brazilian congress. *The Journal of Politics*, 57(2):324–343.

Ames, B. (1995b). Electoral strategy under open-list proportional representation. *American Journal of Political Science*, 39(2):406–433.

Aumann, R. J. (1959). Acceptable points in general cooperative n-person games. In *Contributions to the Theory of Games IV*. Princeton University Press, Princeton NJ.

Austen-Smith, D. and Banks, J. (1988). Elections, coalitions, and legislative outcomes. *American Political Science Review*, 82(2):405–422.

Banks, J. S. and Sobel, J. (1987). Equilibrium selection in signaling games. *Econometrica*, 55(3):647–661.

Banks, J. S. and Sundaram, R. K. (1998). Optimal retention in agency problems. *Journal of Economic Theory*, 82(2):293–323.

Baron, D. P. and Diermeier, D. (2001). Elections, governments, and parliaments in proportional representation systems. *The Quarterly Journal of Economics*, 116(3):933–967.

Bernheim, B. D., Peleg, B., and Whinston, M. D. (1987). Coalition-proof nash equilibria i. concepts. *Journal of Economic Theory*, 42(1):1–12.

Bernheim, B. D. and Whinston, M. D. (1987). Coalition-proof nash equilibria ii. applications. *Journal of Economic Theory*, 42(1):13–29.

Besley, T. (2007). *Principled agents? The political economy of good government, The Lindhal Lectures*. Oxford University Press.

Besley, T. and Smart, M. (2007). Fiscal restraints and voter welfare. *Journal of Public Economics*, 91(3):755–773.

Bouton, L., Conconi, P., Pino, F., and Zanardi, M. (2014). Guns and votes.

Canes-Wrone, B., Herron, M. C., and Shotts, K. W. (2001). Leadership and pandering: A theory of executive policymaking. *American Journal of Political Science*, 45(3):532–550.

Carey, J. M. and Hix, S. (2011). The electoral sweet spot: Low-magnitude proportional electoral systems. *American Journal of Political Science*, 55(2):383–397.

Carey, J. M. and Hix, S. (2013). District magnitude and representation of the majority's preferences: a comment and reinterpretation. *Public choice*, 154(1-2):139–148.

Carey, J. M. and Shugart, M. S. (1995). Incentives to cultivate a personal vote: a rank ordering of electoral formulas. *Electoral studies*, 14(4):417–439.

Chang, E. C. (2005). Electoral incentives for political corruption under open-list proportional representation. *Journal of Politics*, 67(3):716–730.

Chang, E. C. and Golden, M. A. (2007). Electoral systems, district magnitude and corruption. *Journal of Political Science*, 37(1):115–37.

Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

Cox, G. W. and Shugart, M. S. (1996). Strategic voting under proportional representation. *Journal of Law, Economics, and Organization*, 12(2):299–324.

Crutzen, B. (2013). Keeping politicians on their toes: Does the way parties select candidates matter?

Dewan, T. and Hortala-Vallve, R. (2012). Policy learning and elections.

Dowd, R. A. and Driessen, M. (2008). Ethnically dominated party systems and the quality of democracy: Evidence from Sub-Saharan Africa. *Afrobarometer Working Paper*, 92.

Dutta, B. and Sen, A. (2012). Nash implementation with partially honest individuals. *Games and Economic Behavior*, 74(1):154–169.

Duverger, M. (1959). *Political parties: Their organization and activity in the modern state*. London: Methuen.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.

Esteve-Volart, B. and 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.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth*, 8(2):195–222.

Feddersen, T. J. (1992). A voting model implying duverger's law and positive turnout. *American journal of political science*, 36(4):938–962.

Fey, M. (1997). Stability and coordination in duverger's law: A formal model of preelection polls and strategic voting. *American Political Science Review*, 91(1):135–147.

Frankel, D. M. and Volij, O. (2007). Measuring segregation. *Iowa State University*.

Goodman, L. A. and Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, 49(268):732–764.

Hirano, S., Snyder, J., and Ting, M. M. (2013). Primary elections and the provision of public goods.

Huber, J. D. (2012). Measuring ethnic voting: Do proportional electoral laws politicize ethnicity? *American Journal of Political Science*, 56(4):986–1001.

Jahn, J., Schmid, C. F., and Schrag, C. (1947). The measurement of ecological segregation. *American Sociological Review*, 12(3):293–303.

James, D. R. and Taeuber, K. E. (1985). Measures of segregation. *Social Methodology*, 15:1–32.

Kelley, J. and McAllister, I. (1984). Ballot paper cues and the vote in australia and britain: Alphabetic voting, sex, and title. *Public Opinion Quarterly*, 48(2):452–466.

Lijphart, A. and Aitkin, D. (1994). *Electoral systems and party systems: A study of twenty-seven democracies, 1945-1990*, volume 159. Oxford University Press Oxford.

Lijphart, A. and Pintor, R. L. (1988). Alphabetic bias in partisan elections: Patterns of voting for the spanish senate, 1982 and 1986. *Electoral Studies*, 7(3):225 – 231.

List, J. A. and Sturm, D. M. (2006). How elections matter: Theory and evidence from environmental policy. *The Quarterly Journal of Economics*, 121(4):1249–1281.

Lizzeri, A. and Persico, N. (2001). The provision of public goods under alternative electoral incentives. *American Economic Review*, 91(1):225–239.

Majumdar, S. and Mukand, S. W. (2004). Policy gambles. *The American Economic Review*, 94(4):1207–1222.

Maniquet, F. (2011). Strategic voting in large elections under proportional representation: Why vote for center parties?

Maskin, E. and Tirole, J. (2004). The politician and the judge: Accountability in government. *The American Economic Review*, 94(4):1034–1054.

Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, 67(2):281–315.

Milesi-Ferretti, G. M., Perotti, R., and Rostagno, M. (2002). Electoral systems and public spending. *The Quarterly Journal of Economics*, 117(2):609–657.

Morelli, M. (2004). Party formation and policy outcomes under different electoral systems. *The Review of Economic Studies*, 71(3):829–853.

Morelli, M. and Van Weelden, R. (2013). Re-election through division.

Morgan, B. S. (1975). The segregation of socio-economic groups in urban areas: a comparative analysis. *Urban Studies*, 12(1):47–60.

Myerson, R. B. (1993a). Effectiveness of electoral systems for reducing government corruption: a game-theoretic analysis. *Games and Economic Behavior*, 5(1):118–132.

Myerson, R. B. (1993b). Incentives to cultivate favored minorities under alternative electoral systems. *American Political Science Review*, 87(4):856–869.

Palfrey, T. (1989). A mathematical proof of duverger's law. In Ordeshook, P., editor, *Models of Strategic Choice in Politics*, pages 69–91. Ann Arbor: University of Michigan Press.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

Persson, T. and Tabellini, G. (1999). The size and scope of government: Comparative politics with rational politicians. *European Economic Review*, 43(4):699–735.

Persson, T., Tabellini, G., and Trebbi, F. (2003). Electoral rules and corruption. *Journal of the European Economic Association*, 1(4):958–989.

Portmann, M., Stadelmann, D., and Eichenberger, R. (2012). District magnitude and representation of the majority's preferences: Evidence from popular and parliamentary votes. *Public Choice*, 151(3-4):585–610.

Posner, D. N. (2004). Measuring ethnic fractionalization in Africa. *American Journal of Political Science*, 48(4):849–862.

Reardon, S. F. (1998). Measures of racial diversity and segregation in multigroup and hierarchically structured populations. *Presented at the Annual meeting of the Eastern Sociological Society, Philadelphia, PA*.

Reardon, S. F. and Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological methodology*, 32(1):33–67.

Rogoff, K. (1990). Equilibrium political budget cycles. *The American Economic Review*, 80(1):21–36.

Sakoda, J. M. (1981). A generalized index of dissimilarity. *Demography*, 18(2):245–250.

Shugart, M. S., Valdini, M. E., and Suominen, K. (2005). Looking for locals: Voter information demands and personal vote-earning attributes of legislators under proportional representation. *American Journal of Political Science*, 49(2):437–449.

Smart, M. and Sturm, D. M. (2006). *Term limits and electoral accountability*. Centre for Economic Performance, London School of Economics and Political Science.

Taagepera, R. and Shugart, M. S. (1989). *Seats and votes: The effects and determinants of electoral systems*. Yale University Press New Haven.

Theil, H. (1972). *Statistical decomposition analysis: with applications in the social and administrative sciences*. North-Holland Publishing Company Amsterdam.

Theil, H. and Finizza, A. (1971). A note on the measurement of racial segregation in schools by means of informational concepts. *Journal of Mathematical Sociology*, 1(2):187–94.