Spring 5-15-2021

Essays on Diversity and Public Policy

Saumya Deojain
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee
Marcus Berliant, Chair
Randall Calvert
Valerio Dotti
Justin Fox
John Nachbar
Brian Rogers

Essays on Diversity and Public Policy

by

Saumya Deojain

A dissertation presented to
The Graduate School
of Washington University in
partial fullfillment of the
requirements for the degree
of Doctor of Philosophy

May, 2021
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgements

This work has been made possible due to the love and support of many people over the years. First, I would like to express my deepest appreciation to my committee members. This work was possible because my advisor, Prof. Marcus Berliant, gave me the complete freedom to pursue whatever I wanted and carefully guided me towards the elusive path of doing meaningful research. This work would not have found its shape without Prof. John Nachbar's compassion, guidance and unwavering support. His pragmatism and sensitivity to my needs as a young student and researcher has helped me immensely to anchor this work. I'm deeply indebted to Prof. Randall Calvert who was integral to the evolution of this research. His insights and pointers have greatly influenced the way I tackle research questions. My sincere thanks go out to Prof. Brian Rogers and Prof. Justin Fox. Both of them generously took out time in the past several years to discuss my papers in detail. I am extremely fortunate and grateful to have found my committee member, Valerio Dotti in the Department of Economics at WashU. It would have been difficult for this work to come together without the support of such a wonderful co-author, mentor, and friend.

Too often, research feels like a Sisyphean task. Friends, however, have the awesome power to tip the balance against the impossible. I am very lucky and grateful to have made so many meaningful relationships during the economics Ph.D. program at WashU. First among them, is my ex-officemate, co-author, and confidant David Lindequist. Our shared teas, walks and subway lunches have, among other things, made me a stronger and more independent thinker and explorer. My friends Alper Sükrü Gençer, Dominic Jarkey, Rodrigo Moser, Hyewon Jeong, and Keh-Kwan Sun have had a big hand in helping me realize the potential of my work. They are peers I look up to that have indulged my many musings and helped me uncover many missed nuances. I am also thankful to my friends in the Washington University Graduate Workers Union. Their commitment to their convictions not only inspired me personally but also some of the work here.

I am forever grateful to my long-distance friends Dave Carlson, Prakrati Thakur, and Abhishek Sekhri who have stuck by me all this time. They have supported and cared for me through everything that mattered. I want to thank Palak Dudani for being the rock in my life whose unshakable belief in me has made it possible to navigate through perpetually uncertain waters. Lastly, but perhaps most importantly, I am grateful to my parents who have made me who I am and, at every step of the way, done everything in their power to help me realize the best version of myself.

Saumya Deojain

Washington University in St. Louis

May 2021

Dedicated to Ma and Babu

ABSTRACT OF THE DISSERTATION

Essays on Diversity and Public Policy

for Arts and Sciences Graduate Students

by

Saumya Deojain

Doctor of Philosophy in Economics

Washington University in St. Louis, 2021

Professor Marcus Berliant, Chair

This dissertation is divided into three chapters, each of which explores a particular aspect of the effect of socio-political institutions on cooperation and conflict between diverse individuals. The first chapter shows how the relationship between diversity and cooperation changes with social norms. The second chapter investigates the relationship between cultural conflict and government regulation. The third chapter describes conditions on the distribution of preferences of legislators that allow legislative coalitions to induce a legislative gridlock and block reform.

In the first chapter the notion of a 'norm of compromise' is introduced and its importance for cooperation among agents with diverse preferences over a public policy is demonstrated. Agents choose to cooperate when they join a coalition and agree to support a commonly proposed policy that could be far away from their preferred policy. A norm of compromise is an exogenous protocol used by a coalition to arrive at this commonly proposed or 'compromise' policy. I consider a

parameterized class of norms in which the compromise policy's relative sensitivity to moderates and extremists in the coalition can be dialed. I study the effect of these norms on the stability of the grand coalition (or full cooperation) in a model where an agent faces a trade-off between compromise if she joins a coalition and increased risk if she does not. I find that polarization does not always reduce cooperation: it destabilizes the grand coalition under norms with low relative sensitivity to moderates but stabilizes it under norms with high relative sensitivity to moderates. This can lead to a situation where norms enabling cooperation in a polarized society do not enable cooperation in a homogeneous one. I also find the counterintuitive result that under some norms extremists are less willing to cooperate when moderates' preferences get closer to these extremists. This work sheds light on the emergence of cooperation within social movements like Black Lives Matter and the Arab Spring, where political actors compromise in the relative absence of formal institutional structures.

The second chapter explores how social conflict generated through cultural diversity affects public policy. In the model, social conflict arises when culturally differentiated groups impose negative cultural consumption externalities on each other. These externalities can be mitigated by a government that uses taxation to reduce disposable income, thereby transforming cultural consumption into public good consumption. We link the size of equilibrium taxation to characteristics of the underlying distribution of cultural groups as well as to the local government's social welfare function. We test the predictions from our theoretical framework using U.S. city and county data from 1990. Controlling for a variety of socioeconomic and demographic indicators, we verify a key prediction: local taxes per capita are increasing in diversity as measured by ethnic fractionalization. We further document that other characteristics of the group size distribution have effects on local taxes per capita that are in line with our theoretical analysis.

Finally, the third chapter looks at possible conditions for legislative gridlock to occur within a majoritarian assembly with endogenous coalitions. Legislators possess Euclidean preferences over multiple policy dimensions and are ordered according to a multidimensional notion of the left vs. right divide. Gridlock occurs if no coalition can propose a reform that gains stable support.

As a result, the status quo is maintained. We derive both necessary and sufficient conditions for the existence of a stable gridlock. These conditions depend on two measures characterizing a legislator's ideology, which we call conservatism and extremism. Conservatism characterizes the attractiveness of the status quo relative to the median legislator's preferred policy. Extremism measures the distance of an agent's preference from the preference of the median legislator along the left-right divide. We show that gridlock can be a stable outcome only if a majority of legislators are sufficiently conservative and extremist. Moreover, gridlock occurs if a coalition that includes relatively conservative legislators located on both sides of the political spectrum tactically proposes a more conservative reform than the one that is preferred by the median legislator.

Chapter 1

Unity in Diversity

AUTHOR: SAUMYA DEOJAIN

1.1 Introduction

Any collective effort requires a compromise between a diverse set of actors. This is true for people who come together to write a constitution, form cohesive coalitions of protesters/lobbyists, or form international alliances. These compromises are often arrived at using pre-existing norms of compromise, such as majority voting, contest games, or bargaining. Sometimes these norms are successful in enabling cooperation, and sometimes they are not. Sometimes, the same set of people that can cooperate under one norm fragment in another. In this work, I analyze the conditions where a norm of compromise enables cooperation between a diverse set of agents and where it does not.

This work extends the literature concerning the effect of diversity and institutions on cooperation (Alesina et al., 2001; Easterly and Levine, 1997; Gorodnienko and Roland, 2015; Stichnoth and Van der Straeten, 2013) and, specifically, on the prevention of fragmentation in diverse communities by institutions (Alesina and Reich, 2015, Reynal-Querol, 2002, Spolaore and Wacziarg, 2017). Unlike most of the existing literature which has been concerned with the distribution of public goods and conflict, this chapter is focused on the trade-offs between the cost of compromise and fragmentation for individuals within a society. I show how an individual's willingness to compromise and cooperate depends upon the distribution of preferences and norms of compromise within a society. This willingness depends on how sensitive a norm of compromise is to the preferences of members of a coalition. The main contribution of this chapter is identifying this channel of sensitivity to individual preferences that explicitly links the emergence of cooperation to the underlying norms of compromise and the nature of diversity within a society.

In the model, policies are defined over the real line and each agent has a 'most preferred' policy. The closer a policy is implemented to an agent's preferred policy, the better off she is. All agents know each other's preferences. A group of agents whose preferences need not be the same may form a coalition and agree on a compromise policy that will be proposed by the coalition. A norm of compromise determines which policy is chosen by the coalition as the compromise policy, taking into account the preferences of each agent in the coalition[1]. Each norm is characterized by a real number between zero and unity. When this parameter is close to unity, the compromise policy within the coalition is more sensitive to moderate positions than it is to extreme positions. When the parameter is close to zero, the compromise policy is more sensitive to the extreme positions than it is to moderate positions. Therefore, the compromise policy of the coalition non-trivially depends on the interplay between the norm of compromise and the distribution of preferences of the agents in the coalition.

I apply these norms of compromise in a simple game in which an agent can either join a coalition or back her individual preferred policy. If the agent chooses to join the coalition, she must support its compromise policy (which depends on the norms of compromise as discussed in the previous paragraph). The policy that is finally implemented has a probability distribution over the policies that are backed. The greater the number of people backing a policy, the higher the probability that that policy is implemented. An agent's utility depends on the distance between her preferred policy and the finally implemented policy. The risk aversion of an agent characterizes how much she prefers an uncertain implementation of policy to a certain one. Every agent chooses to join a coalition or back her individually preferred policy based on the assessment of which

[1] This means the compromise policy is endogenous to the membership of the coalition even though the norm of compromise, as will be defined, is exogenously specified.

decision increases her expected utility. (Full) Cooperation emerges in society if all agents join a grand coalition and agree to compromise with each other.

Joining a coalition generates a cost of compromise for an agent since the compromise policy of the coalition need not coincide with the agent's own preferred policy. At the same time, breaking off the coalition creates a cost of fragmentation. When an agent leaves the coalition the compromise policy of the new fragmented coalition is determined by the same norms of compromise without consideration of her preferences. This creates the risk that a policy worse for her is implemented with a positive probability. This results in a trade-off between the cost of compromise and the cost of fragmentation for an agent. If the compromise policy of a coalition is not sensitive to an agent, then the agent faces no risk of leaving the coalition. Cooperation by joining a coalition only reduces the risk for an agent when the compromise policy is sensitive to her preferences. Norms of compromise used by the coalition determine this sensitivity to her preference, therefore also determining her willingness to cooperate.

One of the main results concerns the emergence of cooperation when policy preferences become polarized. Polarization occurs when agents' preferred policies start concentrating on the preferred policy of the extremists in a society. A polarized society is characterized by a bimodal distribution of preferences at the extreme political position. I find when norms of compromise are relatively more sensitive to extremists than moderates, polarization results in a smaller willingness to cooperate. The compromise policies determined by the set of norms with more sensitivity to the extremists are the same as the equilibrium policies chosen in the contest game described in Duggan and Gao (2020). Therefore, this relationship between cooperation and polarization is in line with Esteban and Ray (1999) who find increased polarization results in a greater conflict when agents are in a contest game. When norms of compromise are relatively more sensitive to moderates, polarization results in a greater willingness to cooperate. The more sensitive norms of compromise are to moderates, the closer the compromise policy gets to the outcome of majority voting: the median preferred policy. This relationship between cooperation, polarization, and sensitivity to moderates is mirrored in the result in Reynal-Querol (2002) who finds that conflict is reduced in a polarized society in the presence of democracy.

By applying these results about polarization to a simple example, I find that in a polarized society cooperation does not emerge under norms that are too sensitive to extremes because moderates leave. Conversely, cooperation does not emerge in a more homogeneous society under norms that are too sensitive to moderates because extremes leave. For a given level of polarization, there is a window of norms that depends on the level of polarization under which cooperation will emerge.

Therefore, societies with the same diversity of preferences may have vastly different capabilities to form large coalitions if their norms of compromise are different. This could explain why in the Arab Spring different countries had different levels of cooperation between protesters. Many of these countries had very polarized protesters, but some of these countries were more reliant than others on tribal ties to galvanize the opposition against the status quo. One can argue that when tribal groups cooperate the norm of compromise is to play a contest game. Compromise policies of contest games have high relative sensitivity to extremist positions. As tribes are more differentiated in a society, the compromise is more sensitive to extreme positions. This could explain why Libya had more fractured protests compared to Tunisia even though both were highly polarized. Libya used its more entrenched tribal ties to collectivize in the wake of the Arab Spring - a norm of compromise that could not enable cooperation between its protesters (Anderson, 2011, Tufecki, 2017).

The last main result in the model looks at how the willingness of an extremist to cooperate changes as a moderate comes closer to her preferences. This willingness depends on the norms of compromise too and leads to some counter-intuitive results. When norms of compromise are sensitive to moderate positions and the moderate moves towards an extremist the cost of fragmentation for that extremist increases. Therefore, an extremist is more willing to cooperate within the grand coalition when the moderate's preferred policy gets closer to the extremist's preferred policy. This result is what one would expect. I get counter-intuitive results when compromise policies are sensitive to extremists. In this case, the cost of fragmentation for the extremist decreases as the moderate's preferred policy gets closer to the extremist's preferred policy. So, the extremist is less willing to cooperate when a moderate's preferences move closer to hers. In other words, this extremist could move out of the coalition and back her own policy as other agents' preferred policy move closer to hers. This could explain an interesting phenomenon in the Black Lives Matter movement. After many moderates within the left changed their preferences towards greater police regulation, sections of the movement switched to a more radical ask of defunding the police. This created a fissure within the movement.

The rest of the chapter is divided as follows. The next section is a literature review of the theoretical apparatus used in this chapter and other related works. Section 2.2.1 introduces the theoretical framework discussed in the chapter. This is divided into subsection 1.3.1, where our set of norms of compromise is defined, and subsection 1.3.2, which describes the non-cooperative game and stability conditions that drive the results discussed in the body of the chapter. The results are discussed in section 1.4. I discuss the effect of polarization and radicalization of moderates in section 1.4.1 and 1.4.2 respectively. I conclude with section 1.5.

1.2 Related Literature

This chapter uses the framework provided by non-cooperative games with partition functions (Ray and Vohra, 2015, Yi, 1997, Diamantoudi and Xue, 2007). The tension between cooperation and competition has been studied in many papers like Finus and McGinty (2019), Levy (2004) and Dotti (2020). Of these Finus and McGinty (2019) has focused on the specific relationship between diversity and cooperation the most. The authors show that diversity can sometimes be good for enabling cooperation by using a Cournot model of competition and cartel formation. They also discuss the multidimensionality of diversity and show that diversity in some dimensions may not be effective in enabling cooperation. My contribution is an added layer of 'norms of compromise' that interacts with diversity for cooperation to emerge. These norms of compromise provide a very specific structure on decision making within a coalition. This structure is different from that of Demange (2004) who models this as an exogenous hierarchy/network of decision making. In the present work, it is imposed by an exogenous function that determines how the cost of compromise within a coalition is distributed among its members through a single compromise policy. Specifically, this structure/function is parameterized by a variable that measures the relative sensitivity of a compromise policy to preferences of the moderate with respect to the extremist. To the best of my knowledge, this is a new way of teasing out concrete relationships between distribution of preferences and structural differences among societies.

One structure on cooperation and compromise that has been extensively studied is legislative bargaining, reviewed in Eraslan et al. (2020) and Eraslan and Evdokimov (2019). This literature investigates which policies are chosen by heterogeneous players that are playing a bargaining game (Baron, 1991; Cho and Duggan, 2009; Calvert and Dietz, 2005; Battaglini, 2020). In the context of the present work, the bargaining literature is important in providing micro-foundations for policy outcomes achieved by a norm of compromise. This chapter assumes the existence of certain kinds of effective norms of compromise without going into their micro-foundations. This allows me to focus on explicitly studying how different norms of compromise interact with preferences to make cooperation successful.

Games that model information frictions (Barbera and Jackson, 2020; Dai and Yang, 2019, Battaglini, 2017) are generally very concerned with the coordination problems that arise because of diversity. Dai and Yang (2019) directly address the tension between joining a coalition and staying independent when preferences are heterogeneous. They model a coordination game in which an agent forgoes her independence when she joins an organization that aggregates information and chooses the strategy of the agent with median preferences within the organization. This fundamental trade-off between cost of compromise and cost of fragmentation is very similar to the one I use in this chapter. The main difference is between their focus on information aggregation, and my focus on norms of compromise. This shift away from information aggregation allows me to talk of a system of compromise within organizations that could be separate from their system of information aggregation. This distinction effects the results of the model where I find that cooperation emerges only if the compromise policy is sufficiently sensitive to *both* extremists *and* moderates. Information aggregation models on the other hand are generally looking for solutions to cooperation by worrying solely about those who are at the extremes.

There are many ways to model social interaction or norms or culture [2]. Norms of compromise in the present work are modeled as exogenously determined protocols used by coalitions to arrive at compromise policy. They are slightly different from 'norms' understood as outcomes reached in a repeated game succinctly reviewed in Bisin and Verdier (2017). Instead, one can think of these norms as historically (or culturally) determined rules of an underlying strategic game agents play to coordinate which compromise policy is chosen by a coalition. This is akin to the concept of 'core equilibrium beliefs' discussed in Schofield (2006) in which it is hypothesized that 'core equilibrium beliefs' determine which game agents believe they are playing at a given point of time. A natural way to fit the norms of compromise in the present chapter to this literature is to endogenize these norms by studying their evolution over time. This is a promising direction of further study that is beyond the scope of this chapter.

Lastly, these norms of compromise can also be interpreted as a class of preference aggregators used in social choice theory. In the context of social choice theory, the class of preference aggregators considered in the chapter violate the assumption of neutrality and are not complete. This is why the aggregators in question can be uniquely distinguished by a continuous parameter and violate the assumptions in May's theorem and Arrow's theorem[3] respectively. Despite being in a world of single peaked preferences these norms are not strategy-proof when there is imperfect information [4]. To simplify the analysis I assume complete information. While this is a strong assumption it is not out of place given the motivation behind the norms of compromise[5].

[2] Alesina and Giuliano (2015), Young (2015), Schofield (2006), Bisin and Verdier (2017), Gorodnienko and Roland (2015): These papers conceive of norms, culture and social interaction very differently from each other

[3] However, these aggregators satisfy many of the other assumptions in Arrow's theorem like Weak Pareto Optimality and Anonymity.

[4] See Barbera et al. (1993) for discussion on which types of social choice functions are strategy-proof.

[5] For example the equilibrium policy arrived at by playing contest games as modeled in Duggan and Gao (2020) can also be interpreted in terms of social choice functions. These social choice functions will also not be strategy-proof and require the assumption of perfect information for equilibrium to be reached. These contest games give us the same compromise policies for a subclass of the norms of compromise considered in the present work

1.3 Setup

This section is split in two parts major parts. The first part, subsection 1.3.1, defines and investigates the properties of the class of norms of compromise considered. The second part, subsection 1.3.2, describes the model for which these norms of compromise are applied. I use the example of this model to show that norms of compromise and diversity work together for cooperation to emerge.

Preferences

The model is a one stage game with complete information and a finite set of agents, $\alpha \in \mathscr{N} = \{1,...,N\}$. Agents have Euclidean preferences over a single policy dimension with their ideal policy denoted by $y_\alpha \in \mathbb{R}$. In other words, if the distance between some policy y and y_α, $d_\alpha(y) \equiv |y - y_\alpha|$, is lesser than the distance between another policy y' and y_α agent α prefers y over y'.

Without loss of generality I assume $y_1 \leq ... \leq y_N$, and to avoid trivial solutions assume $y_1 < y_N$. For the purposes of this analysis I modify the ideal policy space to $x = \frac{y - y_1}{y_N - y_1}$. This will not change the qualitative results of the chapter and will simplify analysis of the effect of the distribution of preferences. For the rest of the chapter $x_\alpha = \frac{y_\alpha - y_1}{y_N - y_1}$ will be referred to as the preferred policy of α, thus $x_1 = 0$, $x_N = 1$ and $x_\alpha \in [0,1]$. The set of ideal policies of all agents, $\mathscr{A} = \{x_1,...,x_N\}$, fully describes the distribution of preferences within a society. I consider cases where $N > 2$ to eliminate trivial cases.

1.3.1 Norms

'Norms of compromise' determine the proposed or 'compromise' policy of coalitions. A norm of compromise is essentially a protocol used by a coalition to account for the preferences of all agents in it and minimize the collective cost of compromise. In the present work, a norm is a rule characterized by a number $\rho \in (0,1)$ which determines the proposed policy of a coalition given the ideal policies of all the agents belonging to the coalition. Consider a coalition $C \subset \mathscr{N}$ containing

m agents. Let $A \in [0,1]^m$ be the list or distribution of ideal policies of all agents in C. Then the proposed policy or 'compromise policy' of the coalition C is given by:

$$f^{\rho}(A) = \arg\min_{x \in \mathbb{R}} \Big(\sum_{\alpha \in C} \big(d_{\alpha}(x)\big)^{\frac{1}{\rho}} \Big)^{\rho} \tag{1.1}$$

The distance from a compromise policy, $d_{\alpha}(f^{\rho}(A))$, describes the cost of compromise that the coalition C imposes on agent α. At $x = f^{\rho}(A)$ the socially determined cost of compromise the coalition imposes on all its members of is minimized. This socially determined cost is parametrized by ρ. Moderates of a coalition are defined as agents closest to the compromise policy, $f^{\rho}(A)$, and extremists of a coalition are agents that are furthest from the compromise policy[6]. We will see the exact relationship between ρ and the sensitivity to moderates when we investigate the properties of the norms of compromise. An alternative way of thinking about a norm of compromise, is that it is a measure of centrality given by ρ which a coalition uses to choose its compromise policy[7].

The solution of the optimization problem in (1.1) always exists and is unique [8], for $\rho \in (0,1)$. This means $\forall\ \rho \in (0,1)$, the function f^{ρ} is well-defined for any distribution of preferences $A \in [0,1]^m$. As this chapter focuses only on policy proposals generated by the function f^{ρ}, $d^{\rho}_{\alpha}(A) \equiv d_{\alpha}(f^{\rho}(A))$ denotes distance of α's preferred policy from the coalition's compromise policy $f^{\rho}(A)$ where A is the distribution of preferences and ρ is the norm of compromise. To get a sense of how the compromise policy changes with ρ let us explore the following example that has a closed form solution of f^{ρ}.

***Example* 1.3.1.** *Let C be a coalition that uses the norm of compromise ρ. Suppose $\forall \alpha \in C$, $x_{\alpha} \in \{x, y\}$ with $x < y$. Let l denote the total number of agents with $x_{\alpha} = x$, and $r \equiv m - l$ the*

[6]In the grand coalition, either 1 or N will be an extremist depending on which policy is chosen.

[7]Another interpretation of ρ comes from the use of a constant elasticity of substitution objective function in equation 1.1. One can think of this objective function as the cooperation compromise objective function, or the specific way this coalition minimizes the cost of compromise of each agent. With this interpretation, ρ determines of elasticity of substitution of costs of compromise of agents in the social compromise objective function. The exact measure of elasticity of substitution of a moderate's cost of compromise with respect to an extremist's cost of compromise is $\frac{\rho}{1-\rho}$. Note, when $\rho = 0$ the elasticity of substitution is 0 and when $\rho \to 1$ then elasticity of substitution tends to arbitrarily large values.

[8]See appendix for all proofs.

number of agents with $x_\alpha = y$. Then,

$$f^\rho(A) = \frac{x + y\left(\frac{r}{l}\right)^{\frac{\rho}{1-\rho}}}{1 + \left(\frac{r}{l}\right)^{\frac{\rho}{1-\rho}}}$$

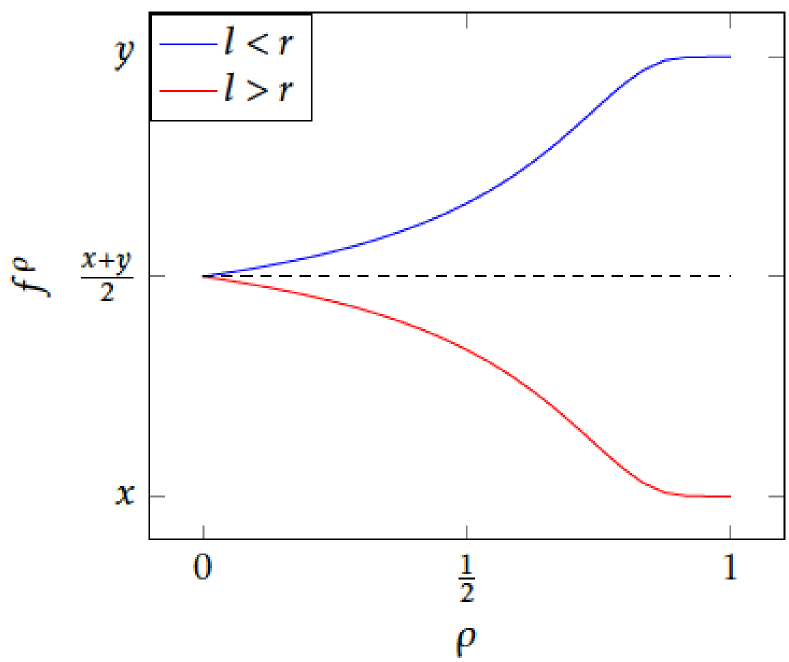


Figure 1.1: Example of the relationship between compromise policy and norms of compromise

Figure 1.1 shows how f^ρ changes if there are only two types of agents within a coalition. We see that f^ρ converges to $(x+y)/2$ as $\rho \to 0$ whether $l > r$ or $l < r$. When $r > l$ then f^ρ increases as ρ increases converging to the median agent's preferred policy, y. When $r < l$ then f^ρ decreases as ρ increases, finally converging to the median agent's preferred policy, x. This example illustrates how dependent policy proposals are on the norm of compromise *and* the distribution of preferences within a coalition. It follows that costs of compromise imposed by the coalition on each agent also depends on this interaction between norms of compromise and the distribution of preferences within a coalition.

Norms of compromise at $\rho \to 0$

In the example above, at the limit values of ρ, one can see how the sensitivity of the compromise policy changes from the median to the extremists. Even for a general distribution of preference this change in sensitivity is most stark at the limiting values of ρ.

$$f^0(A) = \lim_{\rho \to 0} f^\rho(A) = \arg\min_x \max_{\alpha \in C} d_\alpha(x) = \frac{\max_{\alpha \in C} x_\alpha + \min_{\alpha \in S} x_\alpha}{2} \tag{1.2}$$

Expression (1.2) states that when sensitivity to the cost of risk imposed by the coalition becomes arbitrarily large (or $\rho \to 0$) the norms of compromise will only consider minimizing the cost of risk of the extremists within a coalition. This is because these agents bear the highest cost of compromise within the coalition from the compromise policy.

Norms of compromise at $\rho \to 1$

When $\rho = 1$ then the solution to (1.1) is no longer a function for all possible coalitions in $\mathscr{N}$. This is because the solution to (1.1) at $\rho = 1$ is the set of values in $[0,1]$ that are a median to the distribution of ideal points A. From here on out assume for the sake of simplicity at $\rho = 1$,

$$f^1(A) = \begin{cases} x_\alpha & m \text{ is odd; } x_\alpha = med\,\langle A \rangle \\ \\ \frac{x_\alpha + x_\beta}{2} & m \text{ is even; } x_\alpha, x_\beta \in med\langle A \rangle;\ \alpha \neq \beta \end{cases} \tag{1.3}$$

where $med\langle A \rangle$ is a set of all median values in A. This means that at $\rho = 1$, either the unique median ideal policy of agents in C is chosen or the midpoint of two distinct median ideal policies in A is chosen by the coalition. These functional forms of f^ρ at the limits give us an idea of which norms are more sensitive to extremists and which to moderates. When ρ is close to 0 then the compromise policy is more sensitive to the preferences of the extremists relative to the moderates. At $\rho = 0$, the compromise policy ignores the moderates' preferences; it depends only on the extremists. When ρ is close to 1, the compromise policy is sensitive to preferences of the moderates relative to the extremists. At $\rho = 1$, it ignores the extremists' preferences; it only depends on the median(s).

Norms of compromise at $\rho = 1/2$

At $\rho = 1/2$, the compromise policy is simply mean of the ideal points of all the agents in a coalition. In other words,

$$f^{\frac{1}{2}}(A) = \frac{\sum\limits_{\alpha \in C} x_\alpha}{m}, \tag{1.4}$$

where m is the number of agents within coalition C. The norm of compromise represented by $\rho = 1/2$ is very special. The compromise policy induced by the norm $\rho = 1/2$ is equally sensitive to all members of a coalition. Additionally, under $\rho > 1/2$, the compromise policy is sensitive to moderates and under $\rho < 1/2$ the compromise policy is sensitive to extremists. This discussed in greater detail when we explore the properties of the norm.

Before we continue to a more detailed discussion on the sensitivity of a compromise policy to extremists or moderates in the coalition, it is important to note that the compromise policy by coalition need not be monotonic in ρ. Consider the following example:

***Example* 1.3.2.** *$C \subset \mathcal{N}$ has a distribution of preferences given by $A = (0, 0, 2/5, 2/5, 1)$.*

Here $f^1(A) = 2/5$, $f^{\frac{1}{2}}(A) = 9/25 < 2/5$ and $f^0(A) = 1/2 > 2/5$. The non-monotonicity is shown on a schematic diagram in figure 1.2.

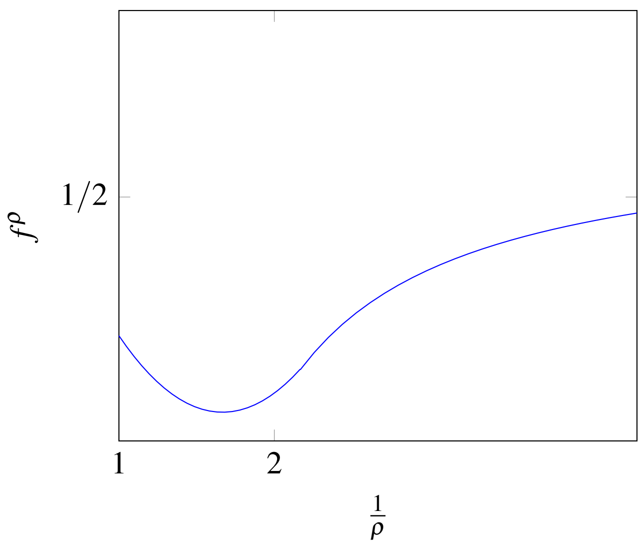


Figure 1.2: Relationship between the compromise policy and the norm of compromise

In this example, the minimum value of f^ρ, the compromise policy, is outside the range $[med\langle A\rangle, 1/2]$. It is possible to numerically see that the minimum value of f^ρ is reached for this distribution when $\rho \in (1/2, 1)$. If one were to slightly change the median value in A, we would find that when $\rho \in (0, 1/2)$ as ρ increases, the compromise policy does not change very much with the median value i.e. the compromise policy is not sensitive to preferences close the median. On the other hand changing $med\langle A\rangle$ when $\rho \in (1/2, 1)$, the compromise policy is much more sensitive to preferences of the median. This already alludes to an interesting property of the norm: increasing sensitivity of the compromise policy to the moderate by increasing ρ is not equivalent to saying there is a monotonic change in the compromise policy as ρ increases.

Properties of the Norm

To better understand the effect of norms and the distribution of preferences on cooperation, let us see how perturbations of preferences within a coalition affect policy proposals under a norm. The following expression gives us the relationship between a social norm and the preferences of an agent α.

***Lemma* 1.3.1.** *Let $\rho \in (0,1)$, and $\alpha \in A \subseteq \mathcal{N}$ with $x_\alpha = x$ then,*

$$\left.\frac{\partial f^\rho(A)}{\partial x_\alpha}\right|_{x_\alpha = x} = \left.\frac{\left(d_\alpha^\rho(A)\right)^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\beta \in C}\left(d_\beta^\rho(A)\right)^{\frac{1-2\rho}{\rho}}}\right|_{x_\alpha = x} \tag{1.5}$$

The right hand side of expression (1.3.1) is an increasing function of $d_\alpha^\rho(A)$ for $\rho < 1/2$ and a decreasing function of $d_\alpha^\rho(A)$ for $\rho > 1/2$. Consequently for $\rho < 1/2$ is most sensitive to the preference of the extremist (whose $d_\alpha^\rho(A)$ is largest) while for $\rho > 1/2$ it is most sensitive to the preferences of the moderate (whose $d_\alpha^\rho(A)$ is the smallest). In other words, when $\rho < 1/2$, the extremists will affect the compromise policy the most. When $\rho > 1/2$, the extremists will affect the compromise policy the least. Further, when $\rho = 1/2$, the right hand side of expression (4) is independent of $d_\alpha^\rho(A)$ and hence of α. This means $f^\rho(A)$ is equally sensitive to the preferences of all agents.

This sensitivity of a compromise or proposed policy to preferences is important in understanding the trade off for α. Sensitivity towards an agent's preferences is defined as follows:

***Definition* 1.3.1.** *Sensitivity of a compromise policy to an agent α's preference at a given distribution of preferences, A, within a coalition, is defined as $\left|\frac{\partial f^\rho(A)}{\partial x_\alpha}\right|$ and denoted by $S_\alpha(\rho, A)$.*

The following figure plots the sensitivity of the compromise policy to α's preference ρ, against the distance of x_α from f^ρ to see how this sensitivity qualitatively changes for moderates and extremists for different ranges of ρ.

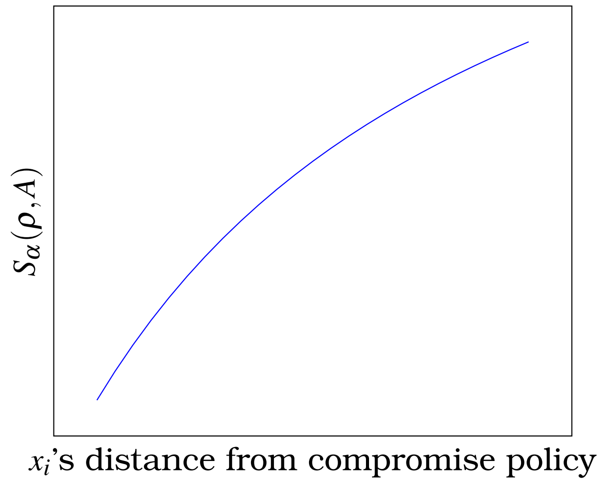


Figure 1.3: Sensitivity of compromise policy to agent's preferences under norm of compromise, $\rho \in (0, 1/2)$

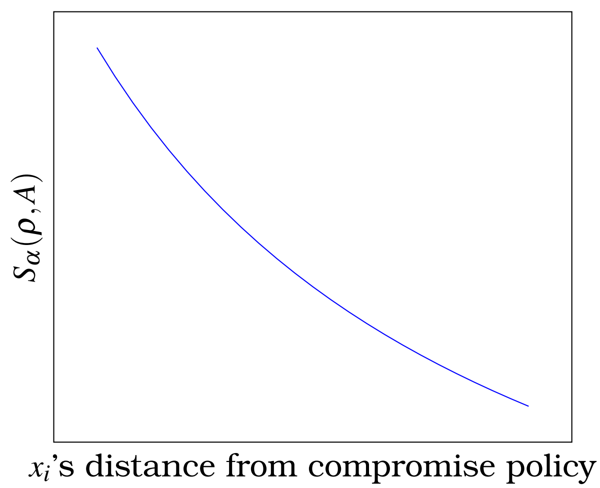


Figure 1.4: Sensitivity of compromise policy to agent's preferences under norm of compromise, $(\rho \in (1/2, 1))$

To get into a better sense of how sensitivity changes for moderates and extremists, recall moderates within a coalition are agents with ideal policies closest to the compromise policy, and extremists within a coalition are agents with the ideal policy furthest away from the compromise policy. Notationally, $mod \in C$ minimizes $d_\alpha(f^\rho(A))\ \forall \alpha \in C$ and $ext \in C$ maximizes $d_\alpha(f^\rho(A))\ \forall \alpha \in C$. Therefore, the relative sensitivity, given by

$$\frac{S_{mod}(\rho, A)}{S_{ext}(\rho, A)},$$

increases as ρ increases. The elasticity of the ratio of the relative distances of the moderates and the extremist with respect to the relative sensitivity of the compromise policy under a given norm is given by

$$\varepsilon_{d,S} \equiv \frac{\text{\% change in } \frac{d_{mod}(f^{\rho}(A))}{d_{ext}(f^{\rho}(A))}}{\text{\% change in } \frac{S_{mod}(\rho,A)}{S_{ext}(\rho,A)}} = \frac{\rho}{1-2\rho}$$

Note, the absolute value of this elasticity is the lowest when $\rho = 0$. In the interval $(1/2, 1]$, $\rho = 1$ also has the lowest elasticity. This elasticity measures how responsive relative distances are to changes in relative sensitivity. This change in relative distances is precisely what agents factor in when considering whether or not it is worthwhile to stay within the grand coalition. In the next section a game of coalition formation is discussed which fully describes the cost of an agent fragmenting from C. This will allow us to determine the stability of a coalition once the notion of equilibrium is also defined.

One can think of norms of compromise, as defined above, as the pre-existing rules that agents tacitly agree on when they join a coalition.This makes the compromise policy endogenous to membership but exogenously constrained by the process of arriving at a compromise policy encapsulated in the norm ρ. For example, if lobbyists come together to form a coalition they may all choose a policy between them that depends on their financial contributions. In this casethe compromise policy depends on the distribution of preferences within the coalition but it also depends on the method of agreement: financial contributions. This rule of agreement that uses financial contributions to determine policy of a coalition would mean ρ is somewhere between 0 and $1/2$, as shown in Duggan and Gao (2020). Alternatively, the rules of agreement within the coalition may involve majority voting; in this case $\rho \to 1$; where the compromise policy tends to the median preferred policy. This way, ρ acts as a representation of the underlying rules of compromise within a coalition. This chapter is silent about the *strategic process* through which the compromise policy is arrived at. While a discussion of these micro-foundations of norms of compromise is out of the scope of this work, it is important to underscore that each norm of compromise encapsulated by ρ can capture a special underlying exogenous social process through which agreement is reached within a coalition[9].

1.3.2 Coalition Structure

An agent, α, has two choices: she can propose and back her most preferred policy, x_α, or she can back the policy of a coalition, $C \subset \mathscr{N}$, that proposes z. This is essentially an announcement game with two possible strategies: all players that announce 1 are part of the coalition C and back its policy and all those that announce 0 are not part of C and back their own policy. This is similar to the non-cooperative coalition game structure used in Finus and McGinty (2019) and Yi (1997) where agents can join a coalition as long as they adhere to the rules within the coalition. It is important to note that this chapter focuses solely on the stability of the grand coalition[10].

The probability a public policy, $x \in \mathbb{R}$, is implemented depends on the number of agents backing x. This probability is denoted by $p(b_x, \mathscr{P})$ where b_x is the number agents backing policy x and $\mathscr{P}$ is the partition of the set of agents $\mathscr{N}$. To determine the expected utility of an agent the functional form of the utility of agent α when policy x is implemented is given by:

$$u(x) = -|x_\alpha - x|^\theta \equiv -d_\alpha(x)^\theta. \tag{1.6}$$

Here $\theta \geq 1$ captures the level of relative risk aversion of an agent [11]. This means that an increase in θ will decrease the attractiveness of a gamble of policy positions with respect to a safe policy position. One way to interpret θ is the dislike agents have for uncertainty of public policy imple-

[9]It is important to clarify the distinction between a norm of compromise and a social choice function. While both are preference aggregators,unlike the social choice function in the literature, in this model agents can opt in and out of engaging in an aggregation of preferences. The norm of compromise only applies if agents decide to form a coalition and use the preference agreggator to come to a compromise with each other.

[10]One can generalize the concept of equilibrium that emerges from this type of game structure to one in which any coalition can block the grand coalition. For the purposes of exposition the discussion in the main body of the chapter restricts the equilibrium discussion to individual deviation from the grand coalition. The proofs in the appendix solve for more general definitions of the stability of the grand coalition

[11]As distance of a policy from the implemented policy increases the risky behavior decreases (increasing absolute risk aversion of policy distance). The risky behavior with respect to the current distance of policy remains constant (constant relative risk aversion of policy distance given by θ).

mentation. This distaste for uncertainty about which public policy is implemented would likely increase when there are sudden exogenous shocks like a pandemic or war[12].

Finally, let us denote $\mathscr{A} \in [0,1]^N$ as the full distribution of ideal points of the agents in $\mathscr{N}$. The expected utility of an agent α for a given partition $\mathscr{P}$ that describes whether or not agents choose to be in a coalition C under norm ρ is given by

$$\mathbb{E}_{\alpha}^{\theta}(\mathscr{A}, z, \mathscr{P}) = -p(b_z, \mathscr{P})(d_\alpha(z))^\theta - \sum_{\beta \notin S} p(b_{x_\beta}, \mathscr{P})(d_\alpha(x_\beta))^\theta \tag{1.7}$$

where the expected utility of α depends on the distribution of preferences in society, $\mathscr{A}$, the compromise policy within the coalition, z and the partition $\mathscr{P}$ which generates the probability distribution over the backed policies given by $p(b_x, \mathscr{P})$. Note that there can only be a single coalition, given by C, with more than one agent backing its compromise policy. This assumption will allow us to clearly identify the trade-off agents face for joining or leaving a coalition.

Given that $z = f^\rho(A)$ in this model, both the probability distribution and the compromise policy is determined by which agents choose to join the coalition. In fact, agent α's expected utility from a partition can be re-written as:

$$\mathbb{E}_{\alpha}^{\rho,\theta}(\mathscr{A}, C) = -p(m, \mathscr{P})(d_\alpha^\rho(A))^\theta - \sum_{\beta \notin S} p(1, \mathscr{P})(d_\alpha(x_\beta))^\theta \tag{1.8}$$

where the expected utility of an agent from a partition depends on the norm of compromise[13], ρ, the members of the coalition, C, and the complete distribution of preferences in the society $\mathscr{A}$. An agent chooses to join C if and only if $\mathbb{E}_{\alpha}^{\rho,\theta}(\mathscr{A}, C \cup \{\alpha\}) > \mathbb{E}_{\alpha}^{\rho,\theta}(\mathscr{A}, C)$ and $C \cap \{\alpha\} = \emptyset$[14].

The coalition structure C is said to be stable under ρ if and only if all agents in C prefer to re-

[12] Consider again the related policy spaces $y \in \mathbb{R}$ and $x = \frac{y - y_1}{y_N - y_1}$. The absolute risk aversion for an agent $\alpha \in \mathscr{N}$ in both policy spaces is the same, and the relative risk aversion of an agent in the policy space determined by y is given by $\theta(y_N - y_1) + 1$.

[13] Remember $d_\alpha^\rho(A) \equiv d_\alpha(f^\rho(A))$.

[14] If an agent is indifferent between joining a coalition and being independent, it is assumed she joins the coalition.

main in C and all agents not in C prefer to stay outside of C given the norm of compromise ρ. Formally, C is stable under $\mathscr{A}$ if and only if internal and external stability are satisfied, where

Internal stability: $\mathbb{E}_{\alpha}^{\rho,\theta}(C,\mathscr{A}) \geq \mathbb{E}_{\alpha}^{\rho,\theta}(C_{-\alpha},\mathscr{A}), \forall\alpha \in C$

External stability: $\mathbb{E}_{\alpha}^{\rho,\theta}(C,\mathscr{A}) > \mathbb{E}_{\alpha}^{\rho,\theta}(C \cup \{\alpha\},\mathscr{A}), \forall\alpha \notin C$

where $C_{-\alpha} \equiv \{\beta\}_{\{\beta\in C:\beta\neq\alpha\}}$. Full cooperation is considered stable under ρ when the grand coalition, $C = \mathscr{N}$, is stable under ρ.

The formulation of this model may elicit more than one equilibrium but for the purposes of this chapter we will focus our attention to the case when full cooperation is stable. The next section discusses the cost of fragmenting from this grand coalition.

1.3.3 Cost of fragmentation

Stability of the grand coalition depends on agents' trade-off between compromising by staying in the grand coalition and increasing their risk by leaving it. This trade-off is encapsulated by the cost of fragmentation given by the expression

$$R_{\alpha}^{\rho,\theta}(\mathscr{A}) = p^{\frac{1}{\theta}} d_{\alpha}^{\rho}(\mathscr{A}_{-\alpha}) - d_{\alpha}^{\rho}(\mathscr{A}) \tag{1.9}$$

where $p = p(N-1, \langle\mathscr{N}_{-\alpha}, \{\alpha\}\rangle)$ and $\mathscr{A}_{-\alpha}$ denotes the distribution of preferences of agents in $\mathscr{N}_{-\alpha}$. The cost of leaving the grand coalition is given by $p^{\frac{1}{\theta}} d_{\alpha}^{\rho}(\mathscr{A}_{-\alpha})$ which depends on how far away the fragmented coalition's compromise policy will be from α's preferred policy, and on the probability that that policy is chosen. The cost of staying in it is given by the cost of compromise $d_{\alpha}^{\rho}(\mathscr{A})$. The cost of fragmentation is essentially the cost of leaving normalized by the cost of compromise under ρ. When the cost of fragmentation $R_{\alpha}^{\rho,\theta}(\mathscr{A})$, is negative agent α will fragment the grand coalition by leaving it, and when cost of fragmentation is positive α will not. Therefore, when the cost of fragmentation is non-negative $\forall \alpha \in \mathscr{N}$ then full cooperation is stable, otherwise there exists at least one agent that prefers to fragment the grand coalition by leaving it.

The riskiness of leaving the grand coalition increases with $p^{\frac{1}{\theta}} d_\alpha^\rho(\mathscr{A}_{-\alpha})$ and decreases with $d_\alpha^\rho(\mathscr{A})$ for α. Conversely, the willingness to leave the grand coalition increases in the cost of compromise, $d_\alpha^\rho(\mathscr{A})$, and decreases with $p^{\frac{1}{\theta}} d_\alpha^\rho(\mathscr{A}_{-\alpha})$ for α. If $d_\alpha^\rho(\mathscr{A}_{-\alpha}) = d_\alpha^\rho(\mathscr{A})$ then agent α has no incentive to stay in the coalition. On the other hand, if $p = 1$ then full-cooperation is always stable since leaving the coalition always creates more risk for any agent α since $d_\alpha^\rho(\mathscr{A}_{-\alpha}) > d_\alpha^\rho(\mathscr{A})$ for $\rho \in (0,1)$.

1.4 Results

The first relationship established is between risk aversion and the stability of the grand coalition. Stability of the grand coalition is achieved when creating divisions becomes very costly for the agents, or cost of fragmentation is very high. The most straight forward way that these costs rise is when risk aversion, θ, increases. Risk aversion captures an agent's aversion to political uncertainty created by divisions along a specific public policy dimension. For example, people investing in higher education who are risk averse to income losses (Jung, 2015) would also be more risk averse about public infrastructure being allocated away from their preferences. This could be an explanation as to why there is cooperation between diverse social groups within college campuses that choose to protest (Dahlum and Wig, 2017).

The first lemma states the relationship between risk aversion and cost of risk for an agent α. The larger the risk aversion to divisions, the more attractive a safe option offered by the grand coalition becomes for an agent.

***Lemma* 1.4.1.** *Fix* $\rho \in (0,1)$ *and distribution of preferences* $\mathscr{A}$*, then* $\exists \bar{\theta}_\alpha^\rho(\mathscr{A}) \geq 1$ *such that* $\forall \theta >$

$\bar{\theta}^{\rho}_{\alpha}(\mathscr{A})$ *and* $\forall\, i \in \mathscr{N}$

$$R^{\rho,\theta}_{\alpha}(\mathscr{A}) > 0 \tag{1.10}$$

Lemma 1.4.1 states that, under a fixed norm, for every α there exists a threshold of risk aversion that makes the grand coalition more attractive than any other partition. As there are a finite number of agents, there must exist a maximum threshold level of risk aversion. Any risk aversion larger than this maximum threshold of risk aversion makes the grand coalition attractive for *all* agents in the population. This leads us to the first proposition.

***Proposition* 1.4.1.** *Fix* $\rho \in (0,1)$*, then* $\exists \bar{\theta}^{\rho}(\mathscr{A}) \geq 1$ *such that* $\forall \theta > \bar{\theta}^{\rho}(\mathscr{A})$ *the grand coalition is stable.*

For large enough risk aversion, under a norm ρ, the grand coalition is the most preferred partition for any agent in $\mathscr{N}$. It is important to note, that for low θ the grand coalition need *not* be either stable nor efficient under a specific norm. Proposition 1.4.1 merely states that high enough risk aversion overrides the instability of grand coalition for a given norm. This relationship between θ and ρ is illustrated in figure 1.5. The shaded region represents the θ, ρ combinations where full cooperation is stable. Outside the shaded region risk aversion is not high enough to induce cooperation for a given ρ and θ.

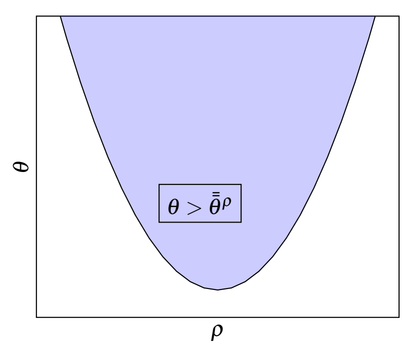


Figure 1.5: Full-cooperation emerges with high enough risk aversion (shaded).

Proposition 1.4.1 alludes to the fact that when risk-aversion is high enough, the distribution of preferences do not matter. Therefore, when risk aversion is high enough there is no relationship of diversity of preferences and norms of compromise. This result is here so that when we begin exploring the relationship between norms of compromise, distribution of preferences and cooperation it is clear that the societies investigated has low enough risk aversion, θ so that variation is possible. The next two sections that explore this relationship with the assumption that risk aversion is low.

1.4.1 Polarization and Norms of Compromise

In this section, we will discuss how different norms of compromise enable cooperation when agents' policy preferences polarize. For simplicity the analysis is restricted to symmetric distribution of preferences. Before defining what polarization is let us consider two examples that describe two societies with different levels of polarization. One society is fully polarized, i.e. agents only prefer extreme policies. The other society is 'minimally' polarized, i.e. all agents are concentrated in the middle of the spectrum[15]. For the same norm, agents in a polarized society have different costs of fragmentation than agents in an unpolarized society. This is because the cost of fragmentation for any agent, determined by a fragmented coalition's compromise policy, also crucially depends on the polarization within the fragmented coalition.

In an unpolarized society, when norms are too sensitive to moderates of the grand coalition, $\rho \to 1$, then extremists have low costs of fragmentation. On the other hand in the same society when norms are sensitive to extremists, $\rho \to 0$, the extremists' costs of fragmentation is too high for them to fragment from the coalition.

***Example* 1.4.1.** *Let $\theta \approx 1$, consider $\mathscr{A}$ which is symmetric such that $\forall \alpha \in \mathscr{N} \setminus \{1,N\}$ $x_\alpha = 1/2$ and $N \geq 4$.*

This example describes a minimally polarized or maximally homogeneous society where all agents except 1 and N are have preferred policies at $x = 1/2$. Therefore, all moderate agents of the grand coalition will be indifferent about fragmenting the grand coalition. The threat of fragmentation comes from the extremists. Extremist agents face very small costs of fragmentation when ρ is

[15] This excludes agents 1 and N who have preferred policies at 0 and 1 respectively

close to 1 because compromise policy of the fragmented coalition, $f^{\rho}(\mathscr{A}_{-1})$, is very close to compromise policy by the grand coalition, $f^{\rho}(\mathscr{A})$. This is because for high ρ the compromise policy is most sensitive to moderates all of whom are concentrated at $1/2$. Alternatively, when ρ is low $f^{\rho}(\mathscr{A})$ is sensitive to the extreme positions. If an extremist fragments the grand coalition the fragmented coalition's compromise policy will shift significantly to the opposite political extreme. This will increase the risk of fragmentation so high that neither extremist will be willing to leave the grand coalition.

In a fully polarized society, this relationship between the stability of the grand coalition and norms of compromise flips.

***Example* 1.4.2.** *Fix θ, consider $\mathscr{A}$ which is symmetric such that $\forall \alpha \in \mathscr{N}$, $x_{\alpha} \in \{0,1\}$. Let l be the number of agents with ideal point x_{α}. If $l = N - l$, $\exists \hat{\rho}$ such that $\forall \rho > \hat{\rho}$ the grand coalition is stable and $\forall \rho < \hat{\rho}$ the grand coalition is not stable.*

This example describes a society which is 'maximally' or fully polarized i.e. agents' preferred policy are either 0 or 1. When $l = N - l$, the society has a symmetric distribution of preferences. This implies everyone is a moderate of the grand coalition. If the extremist agent N fragments from the grand coalition then agent i with $x^i = 0$ will become the moderate of the fragmented coalition, and any agent j with $x^j = 1$ will become the extremist[16]. Similarly, if 1 fragments from the grand coalition then agent j with $x^j = 1$ becomes the moderate and any agent i with $x^i = 0$ will become moderates. Therefore, when norms of compromise are sensitive to moderates (ρ close to 1) the cost of fragmentation for an extremist is extremely high. This incentive to stay in the grand coalition gets reversed for very low ρ. The compromise policy of the fragmented coalition will not change much from the compromise policy in the grand coalition because everyone in the society is an extremist. This makes cost of fragmentation small and full cooperation breaks down when $\rho \to 0$ in a fully polarized society.

These simple examples show us, as we expect, that different norms can lead to different levels of cooperation for a distribution of preferences. More interestingly, they are also able to illustrate

[16]This follows by the centralizing tendency of all norms of compromise considered.

that there can be societies with the *same* norms of compromise and have *different* levels of co-operation if the level of polarization within the society is different. Through these examples we can see that the level of cooperation cannot solely be determined by polarization or the norms of compromise. In the following example, we will consider the relationship between polarization and cooperation under different norms of compromise.

***Example* 1.4.3.** *Consider a population with $N = 4$, $p = \frac{m}{N}$ and $\mathscr{A} = \{0, 1-y, y, 1\}$, where $y \in [1/2, 1]$ measures the level of polarization in the population.*

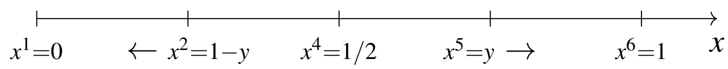


Figure 1.6: Pictorial depiction of a polarizing society.

In this example, full cooperation would give us a compromise policy $f^{\rho}(\mathscr{A}) = 1/2$ for all values of ρ. The cost of fragmentation for an agent i is determined by comparing the distance of her preferred policy to the policy of the fragmented coalition, $f^{\rho}(\mathscr{A}_{-i})$ and $1/2$. Formally, agent i with ideal policy $x_1 = 0$ will remain in the grand coalition if and only if

$$\frac{3}{4}|f^{\rho}(\mathscr{A}_{-i}) - x^i| - \frac{1}{4}|x^i - x^i| \geq |\frac{1}{2} - x^i| \tag{1.11}$$

This means, that if the threat of leaving the grand coalition, $|f^{\rho}(\mathscr{A}_{-i}) - x^i|$, is large enough for i she will have the incentive to remain in the grand coalition. For example, agent 1 will remain in the grand coalition if and only if

$$f^{\rho}(\mathscr{A}_{-i} = \{1-y, y, 1\}) \geq \frac{2}{3} \tag{1.12}$$

When $f^{\rho}(\mathscr{A}_{-i})$ is high then the threat of leaving the coalition for agent 1. We can be even more specific with this example.

When $\rho = 1$ then $f^{\rho}(\mathscr{A}_{-1}) = y$ and agent 1 stays in the grand coalition if and only if $y \geq 2/3$. So, polarization slackens this constraint and incentivizes agent 1 to stay in the grand coalition because the threat of leaving increases. The intuition behind this is that under $\rho = 1$, it is the moderates of the fragmented coalition that affect its compromise policy. So when agent 4 moves further away from agent 1 with polarization the threat of leaving the grand coalition increases for agent 1.

When $\rho = 0$ then $f^{\rho}(\mathscr{A}_{-1}) = 1 - \frac{y}{2}$ and agent 1 stays in the grand coalition if and only if $y \leq \frac{2}{3}$. So, polarization binds the constraint and incentivizes agent 1 to leave the coalition. The intuition behind this is that under $\rho = 0$, it is the extremists within the fragmented group whose policy positions affect polarization. This means, the threat of leaving the coalition decreases as polarization increases and agent 2 gets closer to agent 1.

The example of agent 1's change of preferences for fragmenting the grand coalition shows that the effect of polarization on cooperation depends on which norm of compromise exists in the population. One can qualitatively see the relationship between ρ and the willingness of agent 1 staying in the coalition.

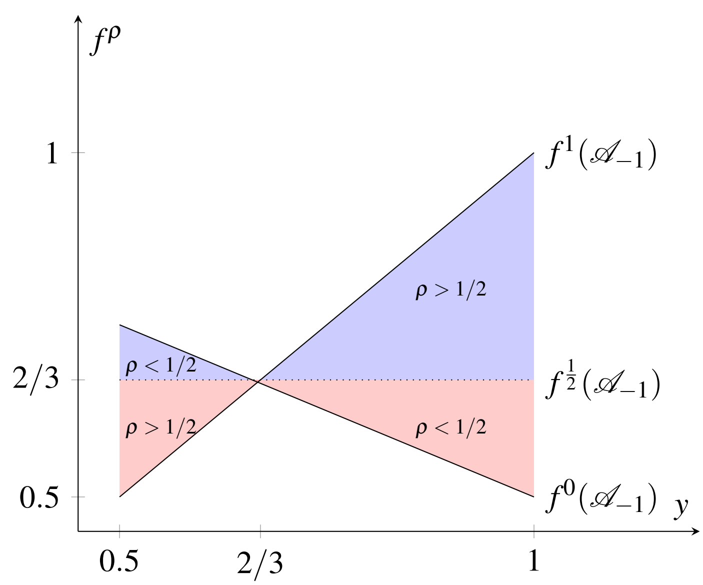


Figure 1.7: Norms of compromise under which agent 1 will leave (below dotted line) or stay (above dotted line) the grand coalition for a given level of polarization.

From figure 1.7 we see that under $\rho > 1/2$ polarization is makes the grand coalition stable to fragmentation from agent 1 and under $\rho < 1/2$ polarization makes the grand coalition unstable to fragmentation agent 1 [17]. The intuition is the following: when agent 1 fragments from the grand coalition agent 2 ($x_2 = 1 - y$) becomes the extremist agent of the fragmented coalition and agent 3 ($x_3 = y$) becomes the moderate in the fragmented coalition. When polarization increases, the moderate of the fragmented coalition moves away from agent 1 and the extremist moves toward her. So when polarization is increasing under $\rho > 1/2$, norms with high sensitivity to moderates of a coalition, the fragmented coalition's compromise policy will move further away from agent 1's ideal policy. On the other hand, under $\rho < 1/2$, norms with high sensitivity to extremists of a coalition, polarization makes the fragmented coalition's compromise policy move closer to agent 1's ideal policy. This example lays down the basic intuition for the general result about polarization, norms of compromise and cooperation.

A more general definition of polarization is as follows:

***Definition* 1.4.1.** *Consider a symmetric distribution $\mathscr{A} = (x_\alpha)_{\alpha \in \mathscr{N}} \in [0,1]^N$. $\mathscr{A}$ is symmetrically polarized to another symmetric distribution of preferences $\mathscr{A}' \in \mathbb{R}^N$ if and only if $v = \mathscr{A}' - \mathscr{A}$ is also symmetric and $v_\alpha < 0$ if $x_\alpha < 1/2$ and $v_\alpha > 0$ if $x_\alpha > 1/2$.*

This definition of polarization leads from one symmetric distribution for another. This is done to simplify the basic mechanics of the effect of polarization on the distribution of preferences. Note, the value $||v||$ gives us the magnitude of polarization.

The following proposition illustrates the effect of polarization on cooperation depends on the norms of compromise. When norms are sensitive to the moderates, $\rho > 1/2$, then the willingness to be part of a coalition increases with polarization for agents who are not polarized. However, when norms are sensitive to extremists, $\rho < 1/2$ then the willingness to be part of a coalition decreases with polarization for agents who are not polarized.

***Proposition* 1.4.2.** *Consider a symmetric distribution of preferences given by $\mathscr{A}$ that is symmetrically polarized to $\mathscr{A}'$ with $v = \mathscr{A}' - \mathscr{A}$. Let $\beta \in \mathscr{N}$ be such that $v_\beta = 0$. Then,*

- *$\rho < 1/2$,*

$$v.\nabla R_\beta^{\rho,\theta} < 0$$

[17] In fact, a symmetric argument can be given for the trade-offs for agent 4 with $x_4 = 1$

In words, if full cooperation is not stable to deviations from β at $\mathscr{A}$ under ρ then full cooperation is not stable to deviations from β at $\mathscr{A}'$ under ρ. increases.

- $\rho > 1/2$,

$$v.\nabla R_{\beta}^{\rho,\theta} > 0$$

In words, if full cooperation is not stable to deviations from β at $\mathscr{A}'$ under ρ then full cooperation is not stable to deviations from β at $\mathscr{A}$ under ρ.

where ∇R_{β} is the partial of the cost of fragmentation of β with respect to the distribution of preferences $\mathscr{A}$.

Proposition (1.4.2) states that for agents that are not getting polarized, polarization of preferences makes the grand coalition less attractive under norms that are more sensitive to the extremists. Therefore, polarization is 'destabilizing' for full cooperation when $\rho \in (0,1/2)$. Conversely, under norms that are more sensitive to moderates, polarization makes full cooperation more attractive to those who are not getting polarized. In other words, polarization is more 'stabilizing' to full cooperation when $\rho \in (1/2,1)$.

The intuition lies in how sensitivity towards extremists and moderates play into the cost of fragmentation. A fragmented coalition's compromise policy is always further away from the agent who fragments away from the grand coalition. In a symmetric distribution, fragmentation of the grand coalition by some $\alpha \in \mathscr{N}$ results in agents at the opposite side of the political spectrum (other side of $x = 1/2$) to become moderates in the fragmented coalition and agents on the same side of the political spectrum of α closer to the extremist. This means, when $\rho \in (0,1/2)$, extremists of the fragmented coalition pull policies closer to them. If an agent who is not polarized leaves the grand coalition, polarization brings the compromise policy closer to them than it did before. This results in a lowering of a cost of fragmentation and a decreasing willingness to cooperate. Conversely, when $\rho \in (1/2,1)$ the compromise policy is sensitive to moderates. This means the compromise policy is pulled in the direction of the new moderates in the fragmented coalition. With this, cost of fragmentation goes up and agents have a greater willingness to stay in the grand coalition.

The following example illustrates the power of this result. It shows when the grand coalition can be vulnerable to fragmentation from extremists, and when it can be vulnerable to fragmentation from moderates.

***Example* 1.4.4.** *Consider a symmetric distribution of preferences given by $\mathscr{A}(y)$ where $x_\alpha \in \{0,(1-y),(y),1\}$ for an even number of agents in a population. Let the number of agents with $x_\alpha = \{(1-y)$ or $x_\alpha = y$ be the same and equal to $m = (N/2-2)$. Further, assume that $p \geq \frac{N-1}{N}$.*

- *For extremist agents, $\alpha \in \{0,N\}$,*

  - *fix $\rho \in (0,1/2)$, $\exists y^*_{ext}(\rho) \in (0,1)$ such that,*

    $\forall y < y^*_{ext}(\rho)$, full-cooperation is stable to deviations from extremists.

    $\forall y > y^*_{ext}(\rho)$, full-cooperation is not stable to deviations from extremists.

  - *fix $\rho \in (1/2,1)$, $\exists y^{**}_{ext}(\rho) \in (0,1)$ such that,*

    $\forall y > y^{**}_{ext}(\rho)$, full-cooperation is stable to deviations from extremists.

    $\forall y < y^{**}_{ext}(\rho)$, full-cooperation is not stable to deviations from extremists.

- *For moderate agents $\alpha \in \mathscr{N} \setminus \{1,N\}$,*

  - *$\exists \bar{\rho} \in (0,1/2)$ such that $\forall \rho < \bar{\rho}$ $\exists y^*_{mod}(\rho) \in (0,1)$ such that*

    $\forall y > y^*_{mod}(\rho)$, full-cooperation is stable to deviations from moderates

    $\forall y > y^*_{mod}(\rho)$, full-cooperation is not stable to deviations from moderates

  - *$\exists\ \underline{\rho} < 1/2$ such that if $\rho > \underline{\rho}$, $\forall y \in (0,1)$*

    full-cooperation is stable to deviations from moderates.

This result is illustrated in figure 1.8. When polarization is high and ρ is low moderates will leave the grand coalition. At high polarization, moderates do not have much to lose from fragmentation as norms of compromise are very sensitive to extremists that remain in the coalition. As moderates are themselves close to extremists fragmentation is no longer as costly. When norms of compromise are sensitive to moderates, $\rho > 1/2$, polarized moderates will always find fragmentation costly.

Any fragmentation of the grand coalition by moderates of one side of the political spectrum will result in agents on the other side of the political spectrum become the sole moderates in the fragmented coalition. When $\rho > 1/2$, this creates bigger risks of fragmentation for moderates. Therefore, the willingness of moderates to cooperate always increases as polarization increases. when $\rho > 1/2$.

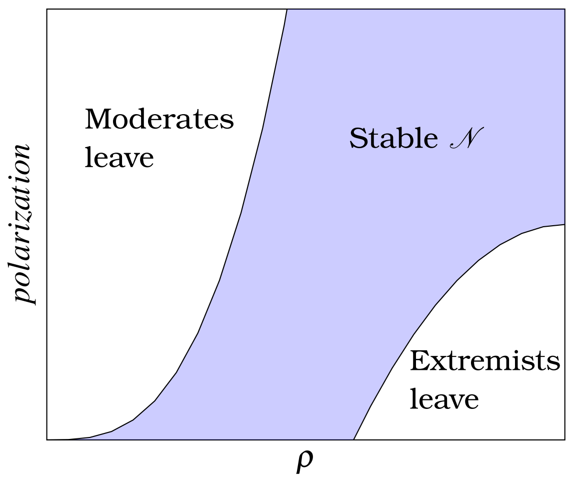


Figure 1.8: Schematic diagram of the relationship between polarization and norms of compromise

For extremists the intuition for staying or leaving a coalition is very similar. When polarization increases and norms of compromise become more sensitive to the extremists, extremists find it costlier to fragment. This is why when polarization is high enough extremists do not fragment from the grand coalition. In contrast, when polarization increases and norms of compromise are more sensitive to extremists ($\rho < 1/2$), the incentive to stay in the grand coalition decreases. When $\rho < 1/2$, the extremists within the fragmented coalition are closer to the extremists. When an extremist of the grand coalition fragments it, the fragmented coalition's compromise policy is most sensitive to its extremists which are close to the extremists of the grand coalition. Therefore, fragmentation for the extremists becomes less costly as polarization of the moderates increases. This is why for sufficiently large polarization extremists are no longer willing to be part of the grand coalition.

The examples in this section show that the same distribution of preferences may not support full cooperation because norms of compromise may decrease. This may explain why countries in the Arab spring saw such many types of protests. Some protests were fragmented like Libya and Yemen where moderates within the protesting group were fragmented into extremists and moderates. Some were more cohesive protests like Tunisa and Egypt which had a a single coalition of protesters that were able to get to the point of drafting constitutions after overthrowing the dictatorship government. These differences in the character of protests may have come from the differences in the way norms of compromise work in these societies. People in Yemen and Libya are have stronger tribal ties (Khan and Mezran (2013), Holm (2013)) and so compromise between these tribal groups may involve a contest game with lower ρ.

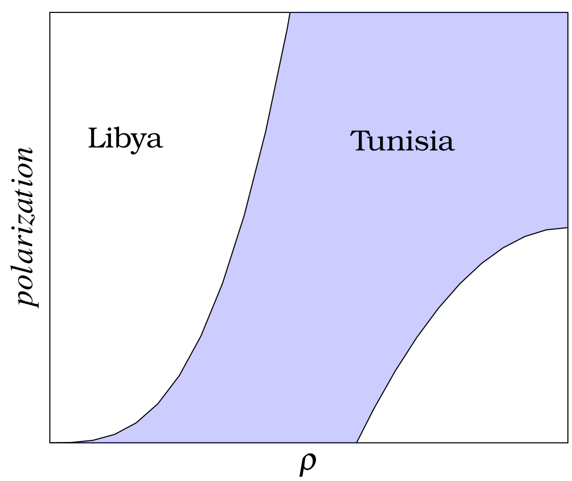


Figure 1.9: Schematic diagram of where countries in the Arab Spring were placed in the polarization/norms axis

A natural question to ask is which norms enable cooperation independent of polarization. This is discussed in the next section.

***Proposition* 1.4.3.** *Let $p \geq \frac{N-1}{N}$ then, $\forall \theta \geq 1$ full cooperation is always stable under $\rho = 1/2$.*

For $\theta \neq 2$, $\rho = 1/2$ does not propose the most efficient policy in the grand coalition, but under the norm $\rho = 1/2$ induces the most efficient partition it can, which is the grand coalition. Thus,

$\rho = 1/2$ is able to make agents sufficiently internalize the costs of risk and compromise so that the most efficient partition under the norm is stable to deviations.

The relationship of norms and diversity of preferences studied through polarization has been by comparing symmetric distributions. However, it is worth considering what happens when symmetric distributions skew towards one extreme position because the change of preferences of the moderate. In the next section, we will see how the willingness to cooperate of the extremist changes for different types of norms when moderates change their preferences towards them.

1.4.2 Effect of radicalization of moderates on extremists willingness to cooperate

Even small changes to the moderates preferred policy position has significant implications on the cost of fragmentation for extremists. Consider a distribution given by $\mathscr{A}$ which is symmetric and $1/2 = x_\alpha \in \mathscr{A}$. Then, for $\rho < 1/2$, the effect of a distortion of α's ideal point makes the grand coalition more attractive to the agent who is furthest away from the distortion. Where as, for $\rho > 1/2$, the agent who is in the same direction of the distortion is finds the grand coalition more attractive.

***Proposition* 1.4.4.** *Let $\rho \in (1/2, 1)$ and $\mathscr{A}$ be the distribution of preferences such that $\exists \alpha \in \mathscr{N}$ where $x_\alpha = f^\rho(\mathscr{A})$. Then*

$$\left.\frac{\partial R_N^{\rho\theta}(\mathscr{P})}{\partial x_\alpha}\right|_{x_\alpha = f^\rho} > 0 \tag{1.13}$$

$$\tag{1.14}$$

In words, if at $\mathscr{A}$ the grand coalition is stable to deviations from N under N, then when a moderate shifts her preferences towards N the grand coalition will still be stable to deviations from N under ρ.

The intuition behind proposition 1.4.4 comes from the sensitivity of norms of compromise towards moderates. When $\rho > 1/2$ the compromise policy is sensitive to the moderate's position. This means when a distribution is skewed towards $x_N = 1$, agent N's influence within the grand coalition will increase relative to its influence on a fragmented partition. This makes the grand coalition more attractive to this extremist, in other words the cost of fragmentation increases with this perturbation. A similar argument can be made to agent 1's willingness to stay in the grand coalition when the moderate's preferred policy comes closer to hers.

***Proposition* 1.4.5.** *Let $\rho \in (0, 1/2)$ and $\mathscr{A}$ be the distribution of preferences. If $x_\alpha = f^\rho(\mathscr{A})$ then*

$$\left.\frac{\partial R_N^{\rho\theta}(\mathscr{P})}{\partial x_\alpha}\right|_{x_\alpha = f^\rho} < 0 \tag{1.15}$$

In words, if at $\mathscr{A}$ the grand coalition is not stable to deviatitions from N under ρ, then when moderate shifts her preferences towards N the grand coalition will still not be stable to deviations from N under ρ.

The intuition behind proposition 1.4.5 given by the sensitivity of norms towards the extremists. When $\rho < 1/2$, policy proposals are most sensitive to the extremists. Therefore, when distribution is skewed toward an extreme the influence of the extremist is diminished within the grand coalition compared to a fragmented partition. This is why, even when the distribution of preferences is skewed toward an extreme, the grand coalition becomes less attractive for the corresponding extremist when $\rho < 1/2$. This could explain when the Black Lives Matter movement gained momentum and moderates were changing their positions towards more police regulation, many in the black lives matter movement took the conversation further towards an extreme position demanding 'defund the police'. It can be argued that this stance created fissures within the movement, particularly online.

Proposition 1.4.5 and 1.4.4 illustrate that skewness of a distribution of preferences may affect the cost of fragmentation very differently for different social norms of decision making. Thus, the effect of moderates moving closer to the extremists on the stability of cooperation may entirely depend on what norms of compromise are employed within the coalition. We are agnostic as to which norms are better or worse.

1.5 Conclusion

This chapter provides a theoretical framework to explore the combined effect of norms of compromise and distribution of preferences on the emergence of cooperation. I find the relationship between cooperation and diversity depends on norms of compromise within the society. These norms of compromise are defined as protocols used by members of a society to come to a compromise if they choose to work in a coalition together. Unlike much of the literature these norms of compromise are not limited to democratic versus autocratic political systems. They can be used to describe more norms that emerge from more cultural methods of compromise like the use of Twitter hashtags (whose compromise policies are more like outcomes of contest games) or voting within councils (whose compromise policies are more like outcomes of majority voting). I find that the relationship between cooperation and and diversity can flip if we consider different norms of compromise.

In the model, Full cooperation under a norm for a given distribution emerges when the grand coalition is stable to individual deviations under a norm. I find that high risk aversion end up stabilizing the grand coalition for a given norm and preference distribution. When risk aversion is low enough, the norms of compromise play a crucial role in explaining whether or not polarization will enable cooperation. It is also important whether the radicalization of the moderate makes the extremist more willing to stay in the coalition. The key driving force of these results is the sensitivity a norm of compromise has towards a moderate with respect to the extremist. When the norm is more sensitive to moderates then polarization is stabilizes the grand coalition whereas when it is more sensitive to extremists polarization is 'destabilizes' the grand coalition.

Further work would include extending the model to multiple dimensions, which would strengthen the generality of these results. Another important aspect to explore would be the effect of the position of the status quo on cooperation. This model has abstracted away from the status quo in order to focus solely on the relationship between cooperation and social norms. Once status quo is included to the model, the political process that chooses policy proposals will become important in determining cooperation. Introducing a status quo will allow us to look at cooperation across an entire society as opposed to a group of individuals already opposed to the status quo. This might allow for analysis on the more general question of when revolution and counter revolutions emerge. Another very interesting direction to expand this work would be to find what norms of compromise survive if agents have a choice to join coalitions with different norms. This is particularly interesting in the context of a dynamical framework where norms of compromise may 'compete' each with other by virtue of how many people adopt their norm at any given time period. This formulation may be able say something about which norms of compromise are more stable in the long run than others.

In the real world, change of risk aversion to public good provision probably comes hand in hand with changes in social norms. This would make studying the partial effect of norms on cooperation difficult. For example, the BLM movement happened after a large change in life style but also saw a large shock on income for many citizens. The change in lifestyle of staying home and shock in income were probably highly correlated. However, it may be possible to study one hypothesis of the model in the data: the positive effect of risk aversion to public good provision on collective action. In the case of the US, the two of the largest movements in the past two centuries (the Civil Rights movement and the BLM movement) have escalated either during war (Vietnam) or a pandemic. It would be worthwhile investigating time-series data for changing risk aversion and its affect on collective action for which this model has clear predictions. To the best of my knowledge, the literature on link between risk aversion the fragmentation of protests is scarce.

Chapter 2

Diversity Taxes

AUTHOR(S): SAUMYA DEOJAIN & DAVID LINDEQUIST[1]

2.1 Introduction

The record-breaking volume of migration in the past few years has made the question of how to deal with conflict within culturally diverse communities a pressing issue for governments all over the world. Elements of this conflict are often embodied in discomfort with differentiated cultural consumption. A Christian may dislike the type of consumption that is specific to Ramadan, and a Muslim may dislike the expression of religiosity celebrated during Lent. Similarly, racial tensions within communities might manifest themselves by one ethnicity disliking the consumption of another ethnicity. By framing conflict between divided groups as negative consumption externalities we study how governments use taxation and public spending to mitigate conflict created by social divisions.

The vast literature on diversity and public policy focuses primarily on how governments allocate public funds in the face of diversity, often modeled as heterogeneous preferences for a public good. In most models, this heterogeneity in preferences leads to an overall smaller size of the government compared to an economy without diverse social groups. We depart from this view and

[1]Assistant Professor, Department of Economics, Miami University

focus on how public spending and taxation regulate conflict between socially divided groups. We argue that governments use taxation as a tool to limit the consumption externalities created in the cultural sphere, and invest instead in secular celebrations, such as the 4th of July, which do not create cultural externalities. In contrast to much of the literature, we conclude that more cultural diversity leads to a bigger size of the government.

In our model, agents consume both private and cultural consumption goods. The type of cultural good an agent consumes is determined by the social group she belongs to. A key assumption is that cultural consumption creates positive externalities within a social group, and negative externalities across social groups.[2] The (local) government taxes labor income and spends all proceeds on public goods which are equally enjoyed by members of all social groups. We assume that the tax rate is the same for all social groups. This assumption implies that we abstract from government interventions that are targeted to cultural consumption directly (Pigovian taxes) or to specific social groups. We think of this assumption as being reflective of a society that values freedom of cultural expression by any of its members. The government can effectively use income taxes to limit the consumption externalities created in the cultural sphere, and invest instead in public goods which do not create cultural externalities. We refer to local taxes per capita that result from the government's desire to regulate the production of negative externalities in diverse communities as 'diversity taxes'.

Within this framework, the government's chosen tax rate balances two opposing effects. On the one hand, higher income taxes reduce total cultural externalities in the economy and allow for the provision of the public good. On the other hand, higher income taxes reduce the utility agents receive from private and cultural consumption. We show that in the equilibrium of the model, the chosen income tax rate depends on both the government's social welfare function as well as the characteristics of the cultural group distribution. The group distribution governs the total amount

[2]This assumption does not rule out that there are also positive cultural consumption externalities across social groups. What matters is that the *net* externality is negative. We note that our model is flexible enough to allow us to study the case of net negative externalities as well as the case of net positive externalities. Positive externalities would manifest themselves in negative 'diversity taxes' for which we find no evidence in the data.

of externalities produced in the economy, while the government's social welfare function determines which externalities the government prioritizies when setting tax rates. We define a majority government as a government social welfare function that only considers the total externalities imposed on the majority group. Our setup implies that cultural consumption of a group is concave in its size, i.e. smaller groups create more externalities per capita but less externalities in total compared to bigger groups. This concavity of externality production in group size implies that a majority government sets lower tax rates the larger the majority group. It further follows that a majority government sets higher tax rates the more fractionalized the set of minority groups is since total externalities imposed on the majority group are increasing in the fractionalization of minority groups.

One advantage of our setup is that it allows us to endogenize the government's social welfare function by considering specific political processes such as majority voting. Majority voting in the presence of a majority group that includes the median voter will endogenously generate a majority government. We also consider the case in which the median voter is in a minority group. The resulting minority government decreases tax rates as the majority group grows bigger, and increases tax rates as the other minority groups become more fractionalized. Both results are again driven by the fact that the amount of negative externalities each cultural group produces is concave in group size.

Another theoretical contribution of our work is to disentangle the impact of different dimensions of diversity on government regulation of social conflict through income taxation. In the literature on ethnic diversity, polarization and fractionalization indices are commonly used to measure diversity. We incorporate these measures into our model by distinguishing increases in diversity due to the increasing size of an already existing group (intensive margin) and due to the addition of a new group (extensive margin). While fractionalization increases both at the extensive and the intensive margin, polarization increases at the intensive margin and decreases at the extensive margin. This theoretical distinction allows us to study the impact of finer definitions of diversity on public policy than typically considered in the literature. More generally, our framework allows us to link any given distribution of social groups to total negative externalities created within the economy. Given a specific government social welfare function, the production of negative externalities translates into an equilibrium income tax rate.

Assuming that the median voter is in the majority group, we obtain a set of key predictions from our theoretical analysis. First, local taxes per capita are increasing in cultural fractionalization. Second, total local government taxes per capita are negatively correlated with the size of the majority group. Third, total local government taxes per capita are negatively correlated with the size of the biggest minority group. Fourth, total local government taxes per capita are positively correlated with the fractionalization of minority groups. We test these predictions using U.S. city and county data on ethnic diversity and local public finances from 1990 provided by Alesina et al. (1999). The data set contains a breakdown of US population into five ethnic groups (White, Black, Asian and Pacific Islander, American Indian, and Other) on both city and county level.[3] It further entails data on total local government taxes per capita (comprised of income, property, and sales taxes) as well as local government expenditure categories.

The empirical analysis confirms all predictions from the theoretical analysis, both on city and county level. Total local taxes per capita are significantly and positively influenced by ethnic fractionalization while controlling for a variety of socioeconomic and demographic indicators, such as population size, education, age distribution, violence per capita, income per capita, and income inequality. We find that the average U.S. city in 1990 would have experienced a decrease in local taxes per capita of nearly 16% if the population had been completely homogenized. We address the potential endogeneity of ethnic fractionalization through an instrumental variable approach. Using time-lagged values of ethnic fractionalization as an instrument, we find that our main results are robust to this modification.[4] Concerning our other three predictions, we find that (i) ethnic fractionalization of the minority groups significantly and positively influences taxes per capita within U.S. cities in 1990, (ii) local taxes per capita are negatively correlated with the size of the majority

[3]After dropping cities and counties in which the median voter is not in the majority ethnicity, we are left with data on 1045 cities and 1394 counties.

[4]In our setup, ethnic fractionalization from 1980 is a valid instrument if it is not impacted by taxes per capita in 1990.

group, and (iii) local taxes per capita are negatively correlated with the size of the biggest minority group. We interpret these last three findings as evidence for the existence of sizeable 'diversity taxes': local governments set local tax rates to mitigate the total amount of negative externalities imposed on the majority group. We conclude that the US evidence corroborates our hypothesized link between cultural diversity and local taxes per capita through negative cultural consumption externalities.

The rest of this chapter is organized as follows. Section 2.1.1 contrasts our setup and findings to the literature on diversity, public policy, and conflict. Section 2.2 presents our theoretical model which we use to study the relationship between diversity, government type, and taxation. In this section, 2.2.1 sets up the model and 2.2.2 establishes and discusses the key results of the theoretical analysis. In 2.2.3, we discuss the inclusion of a political process like majority voting. We further discuss the relationship between majority voting, minority fragmentation and equilibrium taxation. Section 2.3 modifies our general model to derive predictions which we test using U.S. city and county data. In this section, we discuss multiple robustness checks and address potential endogeneity concerns. Section 2.4 concludes.

2.1.1 Related Literature

We contribute to a large literature that studies social conflict arising within diverse communities. Esteban and Ray (Esteban and Ray, 1999; Esteban and Ray, 2011a; Esteban and Ray, 2011b) provide formal models of ethnic conflict in which social tensions are greatest when two equally powerful groups fight over the control of economc resources. Novta (2016) studies the impact of ethnic diversity on the spread of civil war. She finds that greater ethnic diversity is associated with costlier conflict. Similarly, both Amodio and Chiovelli (2018) and Caselli and Coleman (2013) show that ethnic diversity intensifies social conflict and violence. We add to this line of research by framing ethnic diversity as social conflict arising from negative consumption externalities. The focus of this chapter, however, is not on the violence ensuing this social conflict, but rather the local government's fiscal response to it.

Previous work mostly finds negative effects of ethnic diversity on economic growth and public good provision. Lane and Tornell (1999), Azzimonti (2011), Woo (2005), Hodler (2006), and Ager and Brückner (2013) all find negative effects from ethnic diversity on economic growth and fiscal stability. Stichnoth and Van der Straeten (2013) provide a comprehensive survey of recent empirical work that illustrates the evidence of the impact diversity has on government expenditure and public good provision. The channel through which ethnic diversity impacts public good provision remains contested. When diversity is negatively correlated with public good provision (Alesina et al., 1999; Hopkins, 2009; Spolaore and Wacziarg, 2017; Miguel and Gugerty, 2005), it has been attributed to coordination failures due to heterogeneous public good preferences. When diversity has been positively correlated to public good provision (Gisselquist et al., 2016; Gisselquist, 2014; Banerjee and Somanathan, 2007), it has been either attributed to 'diversity dividends' that arise when politically competing social groups keep each other in check for the provision of public goods, or to a rebalancing of local public spending (Lee et al., 2016). We remain neutral about these findings as our results do not depend on the public good provided per se. While our model incorporates secular good provision, it is not central to our analysis. We do not find any conclusive evidence about the effect of diversity on public good provision like education and hospitals. Our findings highlight that if public goods are a means to reduce conflict between divided groups, then the provision of public goods increases with diversity as we find in this chapter. This finding does not contradict the notion that if public goods are solely a productive public good, then miscoordination within a society can result in a negative relationship between diversity and public good provision as other authors have argued (Alesina et al., 2001; Alesina and Glaeser, 2004; Alesina et al., 2019). Hence, we see our findings as complementary to the findings by previous work about the relationship between public good provision and diversity. The main contribution of this chapter is to introduce the concept of 'diversity taxes' imposed by governments to regulate consumption externalities created between groups.

Measuring diversity has been a contentious issue when using it as an indicator for conflict (Somanathan, 2018). This is because diversity has several dimension that can affect conflict and coordination between groups. Esteban et al. (2012) have shown that polarization, a proxy for group competition, may be a better measurement of ethnic friction than ethnic fractionalization which solely measures the relative sizes of groups. We incorporate these distinctions in our empirical analysis and are able to predict the effect of these different measurements on taxes. A contribution of this chapter is to identify the effects of finer definitions of diversity, such as the size of the majority group and fractionalization within the minority, on local taxes per capita.

Our work also contributes to the theoretical literature that explores how governments allocate resources between redistributive policies and ethnic goods (see Fernández and Levy, 2008 and Ghosh and Mitra, 2016). Specifically, we explore how governments regulate disposable income if they cannot directly control cultural good consumption of culturally competitive groups. We theoretically show how the size of 'diversity taxes' depends on the type of government (utilitarian, majority and minority) in place.

2.2 Model

In this section, we set up and discuss the results of a model for government taxation of cultural consumption externalities. The setup is discussed in section 2.2.1 and the main results for the relationship between different dimensions of diversity and government type are discussed in section 2.2.2. In the latter section, we discuss how we interpret diversity changing at the intensive and the extensive margin. We define a government type by the weights it puts on different groups in the society. By focusing on three government types (majority, minority and utilitarian) we are able to distinguish three different functional forms of the equilibrium government tax. These functional forms depend on the cultural groups whose well-being enters a government's social welfare function. Section 2.2.3 discusses the impact of the political process of majority voting on equilibrium taxation, as well as the impact of minority fragmentation at the intensive margin.

2.2.1 Setup

The model is a two stage game. In the first stage, the government chooses the tax rate on labor income for the entire economy. In the second stage, agents allocate their post-tax labor income between private and cultural good consumption. The government uses tax revenue to provide a 'secular' public good. Hence, the relative consumption of private, cultural, and secular goods in the equilibrium of this two-stage game is ultimately regulated by the government tax rate.

We consider a continuum of agents in $[0,1]$ where each agent belongs to a social group $i \in M = \{1,2,...,m\}$. An agent from group i inelastically supplies one unit of labor to the labor market where he obtains a fixed wage rate w. The government sets a tax rate t which is applied to an agent's labor market income. Group i allocates its post-tax wage income $(1-t)w$ between per-capita cultural good consumption, e_i, and per-capita private goods consumption, c_i. Cultural good consumption by one group creates negative externalities on agents belonging to other groups. The government can mitigate these externalities by taxing labor market incomes and using the proceeds to provide a secular public good, g. This public good is equally enjoyed by all agents from all groups. The government is restricted to apply the same tax rate to all agents.

The government chooses its tax rate through backward induction. It is constrained by the optimization problem of the social groups and a budget constraint for the expenditure on the secular good, g. We present the second stage first.

Second Stage: Cultural Group Optimization Problem

An agent from group i has the following utility function

$$u^i(E_i, E_{-i}, c_i, g) = \alpha\left(\beta E_i^{\rho} + (1-\beta)c_i^{\rho}\right)^{\frac{\nu}{\rho}} + \gamma g - \delta\sum_{j\neq i} E_j \tag{2.1}$$

with $\alpha, \beta, \gamma, \delta > 0$. For the rest of this chapter we normalize δ to 1. $E_i = \phi_i e_i$ is the total cultural good consumption by group $i \in M$ where $\phi_i \in (0,1)$ is the size of group i, c_i is the private good consumption by an individual in group i, and $E_{-i} = \sum_{j \neq i} E_j$ is the total amount of negative externalities imposed on group i. We assume that $\nu \in (0,1)$ and $\rho < 0$. That is, returns to total consumption are decreasing, and private and cultural consumption are complements.[5] Our setup implies that there are two types of externalities between agents in this economy. First, there is a negative externality that is created by the exposure to the cultural good consumption of other groups. Second, there is positive externality created by same-group members who all contribute to their common cultural good consumption.[6]

For a given tax rate t, group i optimizes cultural and private good consumption facing the following budget constraint:

$$e_i + c_i = w(1-t) \tag{2.2}$$

Thus, the optimization problem for group i is given by:

$$\left.\begin{array}{ll} \max\limits_{c_i, E_i} & u^i(E_i, E_{-i}, c_i, g) \\ s.t. & e_i + c_i = w(1-t) \end{array}\right\} C_i \tag{2.3}$$

The c_i^* and $E_i^* = \phi_i e_i^*$ that solve C_i will be functions of the exogenous variables w, $\boldsymbol{\phi}_m = (\phi_1, ..., \phi_m)$, and the tax rate, t, set by the government. As $\boldsymbol{\phi}_m$ is the distribution vector of group sizes, it holds that $\sum_{i \in M} \phi_i = 1$.

First Stage: Government Optimization Problem

Through backward induction the government sets a tax rate which maximizes its objective function.

[5]The assumption of complementarity between cultural and private consumption is essential for our main results. If cultural and private consumption were substitutes ($\rho > 0$), then some of our results would be reversed. We are confident in the assumption of complementarity since the predictions derived from the theory align with the empirical evidence we present later, and since private (secular) and cultural consumption seem sufficiently interdependent to render complementarity a natural assumption.

[6]For the sake of simplicity, we assume away the problem of voluntary contributions of agents to total cultural consumption and the resulting free-riding problem. Instead, we assume that each cultural group has a government body that is able to impose a fixed per-capita consumption of cultural good on all members of the group. This is akin to a situation in which the per-capita consumption of cultural good is decided by a vote by all members of the group. As members within a group have the same utility and budget constraints, the desired per-capita cultural consumption is the same for all group members.

A government's social welfare function depends on the social weights it puts on the different cultural groups given by vector $\boldsymbol{\lambda}_m = (\lambda_1, ..., \lambda_m)$ and the size of each group given by the distribution vector $\boldsymbol{\phi}_m$. The government's budget constraint states that total spending equals total tax revenue:

$$g = wt \tag{2.4}$$

The government maximizes its social welfare function based on the consumption decisions $c_i^*(t)$ and $E_i^*(t)$ of each cultural group. Any government at $\boldsymbol{\phi}_m$ defined by $\boldsymbol{\lambda}_m$ solves the following optimization problem:

$$\left.\begin{array}{lll} \max\limits_{t} & U_{\boldsymbol{\lambda}_m}(\boldsymbol{\phi}_m) = \sum\limits_{i \in M} \lambda_i \phi_i u^i(.) & \\ & & \\ s.t. & g = tw & \\ & & \\ and & \forall i \in M \quad c_i^*(t), E_i^*(t) \text{solve} & \\ & & \\ & \left.\begin{array}{ll} \max\limits_{c_i, E_i} & u^i(.) \\ s.t. & c_i + \frac{E_i}{\phi_i} = w(1-t) \end{array}\right\} C_i & \end{array}\right\} G_{\boldsymbol{\lambda}_m}(\boldsymbol{\phi}_m)$$

where $\boldsymbol{\lambda}_m$ is normalized such that $\sum_{i \in M} \lambda_i = 1$.

2.2.2 Results

Second stage solution:

Solving for C_i for group $i \in M$ gives the solutions for private good consumption, $c_i^*(t)$, and cultural good consumption, $E_i^*(t)$. We obtain

$$c_i^*(t) = (1-t)\kappa_c(\phi_i) \tag{2.5}$$

$$E_i^*(t) = (1-t)\kappa_E(\phi_i) \tag{2.6}$$

$$e_i^*(t) = \frac{1-t}{\phi_i}\kappa_E(\phi_i) \tag{2.7}$$

where $\kappa_c(\phi_i) \equiv w\left[\frac{\phi_i^{\frac{-\rho}{1-\rho}}}{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}}+\phi_i^{\frac{-\rho}{1-\rho}}}\right]$ and $\kappa_E(\phi_i) \equiv w\left(\frac{\beta}{1-\beta}\right)^{\frac{1}{1-\rho}}\left[\frac{\phi_i}{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}}+\phi_i^{\frac{-\rho}{1-\rho}}}\right]$ are exogenous coefficients which are a function of the group distribution. It then holds that

$$E_{-i}^* = \sum_{j\neq i} E_j^* = (1-t)\sum_{j\neq i}\kappa_E(\phi_j) \tag{2.8}$$

Note that the total cultural good consumption by group i, $E_i^*(t)$, as well as private good consumption per capita, $c_i^*(t)$, are increasing and concave in group size ϕ_i. The concavity of total cultural consumption in group size implies that negative externalities increase by a larger amount when smaller groups become bigger compared to already large groups becoming even larger.

The government can reduce consumption externalities created in the economy by increasing t. It prioritizes the externalities imposed on different groups based on the social weights it applies to each group.

First stage solution: equilibrium government tax rate

Using the solutions from the maximization problem at the group level to solve $G_{\lambda_m}(\phi_m)$, we get a closed form solution for the equilibrium tax rate, $t^*_{\lambda_m}(\phi_m)$.

***Lemma* 2.2.1.** *For a given number of groups m and distribution of size over groups,* $\boldsymbol{\phi_m} = (\phi_1, ..., \phi_m)$, *the sub-game perfect equilibrium tax rate set by the government with social weight distribution over groups,* $\boldsymbol{\lambda_m} = (\lambda_1, ..., \lambda_m)$*, is given by*

$$t_{\lambda_m}(\boldsymbol{\phi_m}) = 1 - (\Omega_{\lambda_m}(\boldsymbol{\phi_m}))^{\frac{-1}{1-\nu}} \tag{2.9}$$

where

$$\Omega_{\lambda_m}(\boldsymbol{\phi_m}) \equiv \frac{\sum_{i \in M} \lambda_i \phi_i \{\gamma w + \sum_{j \neq i} \kappa_E(\phi_j)\}}{\nu\alpha \sum_{i \in M} \lambda_i \phi_i \{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}}} \tag{2.10}$$

with $\Omega_{\lambda_m}(\boldsymbol{\phi_m}) > 0$*.*

Proof. All proofs are in the Appendix. ■

Lemma 2.2.1 tells us that the relationship between the equilibrium tax rate, diversity (ϕ_m) and government (λ_m) parameters is fully described by the variable $\Omega_{\lambda_m}(\phi_m)$. $\Omega_{\lambda_m}(\phi_m)$ represents the *ratio between the total welfare gain from increasing taxes*, through reduction of total externalities and provision of the public good, *and the total welfare loss of increasing taxes*, through the reduction of consumption of cultural and private consumption goods. We call $\Omega_{\lambda_m}(\phi_m)$ the *government benefit to loss ratio of cultural regulation*. Note that when $\phi_i = 1$, i.e. when there is only one group in the economy, then $\sum_{j \neq i} \kappa_E(\phi_j) = 0$ as there are no negative externalities created. This lowers the benefit of taxation compared to the scenario with other groups present in the economy ($\phi_i < 1$). As a result, a fully homogeneous society sets a lower tax rate than a heterogeneous society. Thus, ceteris paribus, the presence of negative consumption externalities created through cultural diversity increases tax rates. This is our notion of 'diversity taxes'.

The formulation in lemma 2.2.1 allows us to neatly study how the interaction between different dimensions of diversity and the government type impact taxation, all through the variable $\Omega_{\lambda_m}(\phi_m)$.

***Corollary* 2.2.1.** *The equilibrium tax rate* $t^*_{\lambda_m}(\phi_m)$ *is monotonically increasing in* $\Omega_{\lambda_m}(\phi_m)$*.*

We proceed to define three types of government, at a given ϕ_m, which have different functional forms of the government benefit to loss ratio of cultural regulation, $\Omega_{\lambda_m}(\phi_m)$. We assume that group 1 is the biggest group in the economy with m groups and that all other $m-1$ groups are of equal size:

***Assumption* 2.2.1.** *For a given* $\boldsymbol{\phi}_m = (\phi_1, ..., \phi_m)$

1. $\phi_1 > \frac{1}{m}$

2. $\phi_j = \frac{1-\phi_1}{m-1} \; \forall \; j \neq 1$

***Definition* 2.2.1.** *For a given m and* $\boldsymbol{\phi}_m = (\phi_1, ..., \phi_m)$

1. *A utilitarian government has a social welfare function with* $\lambda_i = \lambda_j = \frac{1}{m} \; \forall i, j \in M$.

2. *A majority government has a social welfare function with* $\lambda_1 = 1$.

3. *A minority government i has a social welfare function with* $\lambda_i = 1$ *where* $i \neq 1$.

The distribution of $\boldsymbol{\lambda}_m$ characterizes which groups the government cares about. A majority government prioritizes the majority group which puts all the social weight on its members, the minority government prioritizes a minority group, and a utilitarian government is controlled by a benevolent dictator prioritizing all agents. The groups that are prioritized by the government through $\boldsymbol{\lambda}_m$ determine the externalities that are prioritized by the government when taxes are imposed. As a result, the vector $\boldsymbol{\lambda}_m$ determines the government benefit to loss ratio of cultural regulation, $\Omega_{\lambda_m}(\boldsymbol{\phi}_m)$.

***Corollary* 2.2.2.** *Take m as given and* $\boldsymbol{\phi_m} = (\phi_1, ..., \phi_m)$.

- *For a utilitarian government,* $\Omega_u(\boldsymbol{\phi}_m) = \frac{\sum\limits_{i \in M} \phi_i \{\gamma w + \sum\limits_{j \neq i} \kappa_E(\phi_j)\}}{\nu\alpha \sum\limits_{i \in M} \phi_i \{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}}}$

- *For a majority government,* $\Omega_{maj}(\boldsymbol{\phi_m}) = \frac{\{\gamma w + \sum\limits_{j \neq 1} \kappa_E(\phi_j)\}}{\nu\alpha \{\beta(\kappa_E(\phi_1))^\rho + (1-\beta)(\kappa_c(\phi_1))^\rho\}^{\frac{\nu}{\rho}}}$

- *For a minority government* $i \neq 1$, $\Omega_{min_i}(\boldsymbol{\phi_m}) = \frac{\{\gamma w + \sum\limits_{j \neq i} \kappa_E(\phi_j)\}}{\nu\alpha \{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}}}$

While a utilitarian government takes into account the total externalities created in the entire economy, a minority or majority government only takes into account the negative externalities faced by the group they prioritize.

From our definitions, a majority government focuses solely on the negative externalities imposed on the majority group. As the majority group faces the least cultural externalities among the groups in the economy, a majority government imposes a smaller tax than either the minority group or the utilitarian group. Conversely, the smallest group faces the largest externalities. Hence, a minority government which prioritizes the smallest group imposes the highest taxes compared to any other minority or majority government. This result is summarized in the following Proposition.

***Proposition* 2.2.1.** *For a given* $\boldsymbol{\phi}_m = (\phi_1, ..., \phi_m)$, $t^*_{min_i}(\boldsymbol{\phi}_m) \geq t^*_u(\boldsymbol{\phi}_m) \geq t^*_{maj}(\boldsymbol{\phi}_m)$.

Here, $t^*_u(\boldsymbol{\phi}_m)$, $t^*_{maj}(\boldsymbol{\phi}_m)$ and $t^*_{min_i}(\boldsymbol{\phi}_m)$ are the equilibrium tax rates imposed by a utilitarian government, a majoritarian and a minority government respectively for a given distribution of groups $\boldsymbol{\phi}_m$.

Comparative statics of tax rates with respect to diversity

In our model, the distribution vector $\boldsymbol{\phi}_m$ captures two different dimensions of diversity. First, the length of $\boldsymbol{\phi}_m$ describes the number of groups within a society. Second, the elements of $\boldsymbol{\phi}_m$ describe how agents are distributed into different groups. This allows us to separate two dimensions of diversity through our distribution vector $\boldsymbol{\phi}_m$. One dimension is how agents are distributed between groups for a fixed m. We call this dimension of diversity the *intensive margin*. A change along the intensive margin implies a change in the elements of $\boldsymbol{\phi_m}$. Given assumption 2.2.1, ϕ_1 is a proxy for the change of diversity along the intensive margin. As group 1 is assumed to be the majority, for a fixed m, an increase in ϕ_1 is a decrease in diversity along the intensive margin. The other dimension of diversity that we define is the *extensive margin*. A change along this dimension implies that the number of groups, m, changes. When considering changes along the extensive margin, we fix the size of the majority group. That is, diversity increases at the extensive margin when we increase m, for a fixed ϕ_1.

Taxation and intensive margin diversity

Higher diversity at the intensive margin means a lower value of ϕ_1 with m being fixed. For a given $\phi_1 > 1/m$, the majority group faces lower externalities than the minority groups. Hence, total externalities in the society decrease as intensive margin diversity decreases.[7]

***Proposition* 2.2.2.** *Let* $\phi_1 > \frac{1}{m}$ *then:*

- *For the majority government:*
$$\frac{\partial t^*_{maj}(\phi_1, m)}{\partial \phi_1} < 0$$

- *For the minority government* $i \neq 1$ *with* ϕ_m *such that* $\forall j \notin \{i, 1\}$ $\phi_j = (1-\phi_1)/(m-2)$*:*

$$\frac{\partial t^*_{min_i}(\phi_1, \phi'_i, m)}{\partial \phi_1} < 0$$

Proposition 2.2.2 shows that there is a positive relationship between diversity along the intensive margin (ϕ_1) and equilibrium taxation. Remember that the government sets tax rates so as to reduce the amount of negative externalities faced by the group it cares about. The majority government only cares about the welfare of the majority group. When the majority group gets larger (ϕ_1 increases), then all other $m-1$ minority groups decrease in size equally. Since the total externalities produced by any cultural group is increasing in its group size, this implies that the total amount of externalities of the $m-1$ minority groups decreases. Thus, the majority group faces lower negative externalities and, as a result, the majority government decreases taxes in response to an increase in the size of the majority group.

Similarly, the minority government only internalizes the negative externalities imposed on a given minority group. Fixing the size of this minority group, an increase in the size of the majority group (increase in ϕ_1) leads to an increase in the externalities created through the majority group, and a decrease in the externalities created through the other minority groups. Due to the concavity

[7] We omit the discussion of the utilitarian government since this case does not yield unambiguous predictions for general parameter values.

of the creation of total externalities in group size, the total externalities created in this economy go down. That is, the decrease in externalities by minority groups outweigh the increase in externalities by the majority group. As a result, the minority government reduces taxation in response to an increasing majority group size.

Taxation and extensive margin diversity

At the extensive margin, diversity increases as the number of groups, m, increases while holding ϕ_1 constant. When an additional small group comes into the society, we find that it increases externalities across the board for all previously existing groups.

***Proposition* 2.2.3.** *Given* ϕ_1,

- *for the majority government:*

$$t_{maj}(\phi_1, m+1) - t_{maj}(\phi_1, m) > 0$$

- *for a minority government* $i \in \{1, ..., m\}$*:*

$$t_{min_i}(\phi_1, m+1) - t_{min_i}(\phi_1, m) > 0$$

Proposition 2.2.3 implies that any majority or minority government that existed before the creation of a new group will increase taxes with the introduction of this new group. While a new group results in smaller minority groups, it increases the total amount of negative externalities imposed on all the old groups. This result is again driven by the concavity of a group's consumption of the cultural good. The introduction of a new group reduces the size of existing smaller groups which makes them consume less of their cultural good as a group, but more of their own cultural goods per capita. This implies that the decrease in total externalities from existing groups is smaller than the increase in externalities from the newly added group. As a result, all existing groups face more total negative externalities when a new group is introduced. This results in higher taxes for all governments that prioritize a previously existing social group.

2.2.3 Extensions

In this section we discuss two extensions to the baseline model. First, we explore the relationship between the equilibrium tax and diversity if we add the political process of majority voting. Second, we study the relationship between fragmentation of the minority group at the intensive margin. Both these results are important when we test the predictions from our model on the data.

Majority voting

So far we have made no assumptions on the political process of how a government chooses taxes. Suppose we add a stage to the game with majority voting over taxes before the government sets the taxes. Group i's individual tax rate, $t_i^*(\boldsymbol{\phi_m})$, that solves (2.9) is well-defined. This means that the median voter tax policy is well-defined. If the majority group is the median voter then, $t_{maj}^*(\phi_m)$ is imposed. If the minority group is the median voter, $t_{min}^*(\phi_m)$ is imposed. We know from proposition 2.2.1 that smaller groups prefer higher taxes than bigger groups. This means if $\phi_1 < 0.5$, the majority group is no longer the median voter and a minority group chooses it's most preferred taxes as the median voter. This gives us the following result:

***Proposition* 2.2.4.** *Suppose $m \geq 3$. In a political process of majority voting, the equilibrium regulatory tax will be $t_{med}^*(\phi_1, m) = t_{maj}^*(\phi_1, m)$ for $\phi_1 > 0.5$, and $t_{med}^*(\phi_1, m) = t_{min}^*(\phi_1, m)$ for $\phi_1 < 0.5$*

Proposition 2.2.4 says that there is a discontinuity in $t_{med}^*(\phi_1, m)$ at $\phi_1 = 0.5$ when $m \geq 3$. This discontinuity at $\phi_1 = 0.5$ is a result of majority voting. At $\phi_1 = 0.5$ the government switches from prioritizing the majority group to prioritizing a minority which means higher taxes than $t_{maj}^*(0.5, m)$.

Fragmentation of the minority group

In Proposition 2.2.3 we explored the relationship between taxes and fragmentation of the minority at the extensive margin as minorities become smaller through the introduction of a new group. We established that higher fragmentation at the extensive margin unambiguously leads to an increase in taxes. In this section, we explore how fragmentation of the minority groups affects taxation at the intensive margin.

Without loss of generality assume that group 2 is the largest minority group i.e. $\phi_2 \in (\frac{1}{m-1}, \phi_1)$. Again, for simplicity we assume that all other smaller minority groups are of the same size, specifically, $\phi_i = \frac{1-\phi_1-\phi_2}{m-2}$ $\forall i \in M \setminus \{1,2\}$. Similar to ϕ_1, ϕ_2 is a proxy of the fragmentation within the minority. Higher ϕ_2 implies lower fragmentation of the minority at the intensive margin.

Assume that the political process is majority voting. Then, for $\phi_1 > 0.5$ we have majority government taxes. For $\phi_1 < 0.5$ and $\phi_1 + \phi_2 > 0.5$ the biggest minority group, group 2, is the median voter.[8]

***Proposition* 2.2.5.** *Suppose $m \geq 3$. Assume that the political process is majority voting.*

1. *For $\phi_1 > 0.5$*
$$\frac{\partial t^*_{med}(\phi_1, \phi_2, m)}{\partial \phi_2} < 0 \qquad (2.11)$$

2. *For $\phi_1 < 0.5$ and $\phi_1 + \phi_2 > 0.5$*
$$\frac{\partial t^*_{med}(\phi_1, \phi_2, m)}{\partial \phi_2} < 0 \qquad (2.12)$$

The second part of the proposition states that if group 2 is the median voter, then taxes decline as ϕ_2 increases. Clearly, as group 2 grows larger smaller minority groups become smaller and externalities imposed on group 2 decline. The first part of the proposition reveals that if there is a big enough majority group then increases in ϕ_2 results in lower taxes. This is because as

[8]We omit the case of $\phi_1 + \phi_2 < 0.5$ in which a smaller minority group sets the tax rate since this case does not yield unambiguous predictions for general parameter values.

ϕ_2 becomes larger the increase in the negative externality on group 1 by group 2 is less than the decrease in negative externalities coming from all the other smaller groups. In other words, higher fragmentation of the minority group results in higher taxes imposed by the majority government. We use this particular result when testing our predictions using U.S. data.

2.3 Empirical Evaluation

In this section, we provide empirical evidence for 'diversity taxes' using U.S. data from Alesina et al. (1999). The theoretical analysis in the previous section makes predictions about how intensive and extensive margin variations in diversity affect government taxation and spending. In addition, it relates these predictions to the outcome of the political process (majority versus minority governments). Intensive and extensive margin variations have no direct counterpart in the data. Instead, empirical work has used fractionalization and polarization indices as measures of diversity to study the effect of diversity on conflict and public spending (Alesina et al., 2000; Esteban et al., 2012; Montalvo and Reynal-Querol, 2005). Fractionalization (Taylor and Hudson, 1972) is defined as the probability that members of two different groups meet one another. Specifically,

$$FRAC = 1 - \sum_{i \in M} \phi_i^2 \tag{2.13}$$

Polarization captures how far the distribution of groups is from a bipolar distribution which represents the highest level of polarization. We use the Reynal-Querol index (Reynal-Querol, 2002) to measure polarization within a given population:

$$POL = 1 - \sum_{i \in M} \left(\frac{0.5 - \phi_i}{0.5} \right)^2 \phi_i \tag{2.14}$$

Both indices are imperfect measures for our two-dimensional specification of diversity. More specifically, fractionalization and polarization move in opposite directions at the extensive margin and at the intensive margin fractionalization and polarization monotonically decreases with ϕ_1.

Lemma 2.3.1. *Assume* $\phi_m = (\phi_1, \frac{1-\phi_1}{m-1}, ..., \frac{1-\phi_1}{m-1})$

At the extensive margin:

- *If* $\phi_1 \in (1/m, 1)\ \forall\ m > 2$, $\frac{\Delta FRAC}{\Delta m} > 0$.
- *If* $\phi_1 \in (1/m, 1)\ \forall\ m > 2$, $\frac{\Delta POL}{\Delta m} < 0$.

At the intensive margin:

- *If* $m > 2\ \forall\ \phi_1 \in (1/m, 1)$, $\frac{\partial FRAC}{\partial \phi_1} < 0$.
- *If* $m > 2\ \forall\ \phi_1 \in (1/m, 1)$, $\frac{\partial POL}{\partial \phi_1} < 0$.

At the intensive margin, $FRAC$ decreases with ϕ_1 because the probability of meeting other groups decreases as small groups become smaller, whereas POL decreases with ϕ_1 as minority groups become smaller they move further away from the distribution $(1/2, 1/2, 0, ..., 0)$. At the extensive margin, more number of groups increase $FRAC$ because of increased probability of randomly meeting a member of different group, whereas more groups enlarge the difference between the sizes of the majority and the largest minority, decreasing POL.

The data set provided by Alesina et al. (1999) contains a fixed number of ethnic groups and hence, it rules out any analysis along the extensive margin of diversity. As a result, we will adjust our theoretical framework to derive predictions specific to the intensive margin variation of diversity. Before doing so, we will briefly describe the data we use to conduct our empirical analysis.

2.3.1 Data and Sources

Alesina et al. (1999) provide a comprehensive database on ethnic fractionalization and public finances for three levels of U.S. urban localities in the year 1990: cities, counties, and metropolitan areas.[9] They use an ethnic fractionalization index as their measure of ethnic fragmentation:

$$ETHNIC = 1 - \sum_i (Race_i)^2 \tag{2.15}$$

where $Race_i$ denotes the share of population self-identified as of race i with

$$i = \{\text{White, Black, Asian and Pacific Islander, American Indian, Other}\}$$

This racial classification is adopted from the U.S. Census. It is noteworthy that Hispanics as an ethnic group fall under the category "Other". Table B.1 gives a description of all the variables we employ from the data set.[10] To test our hypothesis about the existence of 'diversity taxes', we use the data on population distribution by race to construct various indices measuring the degree of ethnic fragmentation. More specifically, we construct an index of *ethnic polarization* as in (2.14) which captures how far the distribution of ethnic groups is from a bipolar distribution. The index *size of majority group* measures the dominance of one ethnic group. With *fractionalization of minority* and *size of biggest minority group* we try to capture the fragmentation of the minority groups. Below we will derive distinct theoretical predictions for how we expect these different ethnic diversity variables to impact local taxes per capita.

A second set of variables concerns government finances on the local level. For both levels of aggregation (city and county), we have data on general local government expenditures per capita as well as total local government taxes per capita. The data set also allows us to break down general expenditure into specific categories (health spending, education spending, police spending, welfare spending, and others), although sometimes only at county level. We also observe the local government debt outstanding per capita on the county and metro level. One potential concern is

[9] The data can be accessed through the World Bank: `https://datacatalog.worldbank.org/dataset/wps2108-public-goods-and-ethnic-divisions`. We drop metropolitan area data as the sample size is small compared to city and county data and, more importantly, because our channel of negative consumption externalities is operating on the local level which is better approximated by city and county data.

[10] A detailed description of the data set and data sources can be found in Alesina et al. (1999).

that total local government taxes per capita are comprised of income, property, and sales taxes, but the data set only contains total taxes per capita without the breakdown into these subcategories. We note that our theoretical analysis is focused on income taxation as a tool to address social conflict. The inclusion of property taxes in the taxes per capita data is problematic if property values are correlated with ethnic diversity. As it is likely the case that this correlation is negative (predominantly white communities are more propertied), the inclusion of property taxes works against our hypothesis of diversity taxes. That is, the negative correlation between property taxes and ethnic fractionalization will lead us to underestimate the effect of ethnic fractionalization on total local taxes per capita.

Finally, the Alesina et al. (1999) data set allows us to control for a variety of factors beyond ethnic diversity which might affect local government taxation and spending, such as population size, the percentage of people above 25 with a Bachelor's degree or higher, fraction of the population above 65, violent crimes per capita, median and mean income per capita, as well as mean-to-median income (as a measure of inequality). Table B.2 provides summary statistics for all variables in our data set (at city level).

2.3.2 Modified Model

Given the data availability described above, we make the following modifications to our general model. First, we assume that diversity changes only at the intensive margin and keep m fixed. Further, we assume that the majority group is the median voter i.e. $\phi_1 > 0.5$ and $\lambda_1 = 1$. This implies that the majority group is effectively setting its most preferred tax rate. The majority group will set a tax rate which is positively related to the total amount of externalities that the minority groups as a whole impose on the majority group.

***Corollary* 2.3.1.** *Assume m is fixed and $\phi_1 > 0.5$. Then the tax parameter t^* is chosen by the majority and positively correlated with the total amount of externalities imposed on the majority group.*

Thus, our model predicts that for localities in which there is a majority, local government taxes per capita are positively correlated with the amount of externalities imposed on the majority group. In the following, we present a set of testable predictions based on how ethnic composition impacts the negative externalities imposed on the majority group.

In the modified version of our model, all variation in diversity occurs along the intensive margin, i.e. through changes in the distribution of people over a fixed number of ethnic groups (i.e. through the vector ϕ_m). Proposition 2.2.2 showed that an increase in the size of the majority leads to a decrease in the total amount of externalities imposed on the majority group. As a result, our first prediction is that the majority government imposes a lower tax rate as the size of the majority increases.

***Prediction* 2.3.1.** *Total local government taxes per capita are negatively correlated with the size of the majority since*

$$\frac{\partial t^*_{maj}(\phi_1, \phi_2, m)}{\partial \phi_1} < 0 \tag{2.16}$$

Furthermore, lemma 2.3.1 showed that both the fractionalization index as well as the polarization index decrease when the majority group grows bigger. This implies that the total amount of negative externalities imposed on the majority group is increasing in fractionalization and polarization. This gives rise to our second prediction.

***Prediction* 2.3.2.** *Both the fractionalization index and the polarization index are positively correlated with local government taxes per capita.*

Next, consider variations in the composition of the minority groups. More specifically, assume that the biggest minority group gets bigger (i.e. ϕ_2 increases) while holding the size of the majority constant. From proposition 2.2.5 we know that this reduces the total amount of externalities imposed on the majority group. As a result, the tax rate set by the majority government will decrease.

***Prediction* 2.3.3.** *Total local government taxes per capita are negatively correlated with the size of the biggest minority group since*

$$\frac{\partial t^*_{maj}(\phi_1, \phi_2, m)}{\partial \phi_2} < 0 \tag{2.17}$$

In our model, ϕ_2 measures the size of the biggest minority group and as such is also a proxy for the relative sizes of the minority groups. This comes from the simplifying assumption that other minority groups are the same size. When looking at the data, the level of fractionalization within the minority is a related indicator of how cultural externalities of the minority are affecting the majority. Fractionalization of the minority is given by:

$$FRACMIN = 1 - \sum_{i \neq 1} \left(\frac{\phi_i}{1 - \phi_1} \right)^2 \tag{2.18}$$

As group 2 is the biggest minority, this implies that $\frac{\partial FRACMIN}{\partial \phi_2} < 0$. That is, as group 2 gets smaller, minority groups become more similar in size and the fractionalization of the minority groups increases. This increases the total amount of externalities imposed on the majority group. From this observation we obtain our final prediction from the theoretical analysis.

***Prediction* 2.3.4.** *Total local government taxes per capita are positively correlated with the fractionalization of minority groups.*

In the next subsection, we test our four theoretical predictions on the data set from Alesina et al. (1999). As our theoretical predictions are derived under the assumption that there is a majority government, we drop all cities and counties from the sample in which there is no majority, i.e. in which no ethnic group has more than 50% share of the population.[11]

2.3.3 Results

We find evidence in the data for all four of our theoretical predictions. We structure the discussion of our empirical results around these four predictions.

[11] This implies dropping 31 cities, reducing the sample of cities from 1076 to 1045, and dropping 6 counties, reducing the sample of counties from 1400 to 1394. Hence, this sub-sampling of the data comes with little reduction of sample size.

Documenting diversity taxes: taxes per capita and ethnic fractionalization

We first test Prediction 2.3.2 from above. We expect the fractionalization index of ethnic diversity to be positively correlated with taxes per capita. Table B.3 presents results from a regression analysis on city level data. We regress different fiscal variables on an ethnic fractionalization index as constructed in (2.15). The first two columns of the table present our results from the city-level analysis. In the first column of the table we regress taxes per capita on ethnic fractionalization and various city-level controls. We find that ethnic fractionalization is positively influencing taxes per capita. This finding is statistically significant after controlling for income per capita, population size, fraction of people above 25 with a Bachelor's degree or higher, inequality, fraction of the population above 65, violence per capita, and state-fixed effects. We conclude that the data confirms our prediction about the impact of fractionalization on local taxation.[12] To quantify the relevance of the impact of diversity on local government taxation, consider the average U.S. city in 1990. No diversity (fully homogeneous population) results in a hypothetical \$248.77 of taxes per capita for an otherwise average city, while maximum diversity would imply taxes per capita of \$415.64. Fully homogenizing U.S. cities in 1990 would reduce taxes per capita by an average of 15.96%. This suggest that 'diversity taxes' are significant and relevant drivers of both the level and the variation of city taxes per capita observed in the data.

The second column of table B.3 presents the results from regressing total city government expenditures per capita on ethnic fractionalization and the various controls. We find a positive and significant effect of fractionalization on expenditures. This is in line with our prediction that more diversity leads to more provision of (secular) public goods as a by-product of the government regulating negative externalities arising from diversity through taxation. The effect of ethnic diversity on local exenditures per capita is of considerable size (although less significant than the impact of diversity on taxation). The government of a fully homogenized city ($ETHNIC = 0$) in the U.S. in 1990 spends \$215.14 less than its fully heterogenized counterpart ($ETHNIC = 1$). Fully ho-

[12] We ran the same type of regression with ethnic polarization as in (2.14) and found a positive and significant effect as well.

mogenizing U.S. cities in 1990 would reduce government expenditures per capita by an average of 6.98%.

The third and fourth column of table B.3 repeat the analysis for county-level data. We find that ethnic fractionalization has a positive and significant impact on taxes per capita.[13] The magnitude of this impact is less than on the city level. Fully homogenizing U.S. counties in 1990 would reduce taxes per capita by an average of 4.9%. Similarly, ethnic fractionalization positively and significantly impacts expenditure per capita on county level. Fully homogenizing U.S. counties in 1990 would reduce expenditure per capita by an average of 6.46%. We conclude that 'diversity taxes' are impactful on the county level, although less so than on city level. This is in line with our hypothesis that negative externalities from cultural diversity mostly arise in interactions between people in small localities, i.e. neighborhoods or cities. County level data also includes a spending category 'Public Welfare'. Our proposed channel from diversity to taxation implies that the local government does not redistribute its tax revenues back to the different groups through public transfers, but rather spends the revenues on secular goods (roads, sewage, police, schools, hospitals etc.). Table B.5 reports the results from regressing several public expenditure categories on ethnic fractionalization. We find no significant relationship between the fractionalization and the share of welfare spending on county level. This strengthens our notion of a government which taxes the consumption of cultural goods and provides secular goods instead. Table B.5 shows that neither the spending share for education nor for hospitals is significantly correlated with ethnic diversity. Only the county level share of expenditures on roads is significant and negatively correlated with ethnic fractionalization.[14] We conclude that there is a significant and positive impact of ethnic fractionalization on both taxes per capita and expenditures per capita. However, there is a less significant relationship between ethnic fractionalization and specific expenditure categories. Importantly for us, ethnic fractionalization does not significantly increase welfare spending.

Regressing taxes per capita on ethnic fractionalization might be problematic if taxes per capita

[13] We ran regressions with our polarization index and obtained qualitatively similar and significant result.

[14] This observation is in line with the findings of Alesina et al. (1999). They hypothesize that more ethnic fractionalization leads to less 'productive' public provision (roads, education, sewerage) due to heterogeneity of preferences across ethnic groups.

have a causal impact on ethnic diversity. To address this endogeneity concern, we follow Alesina et al. (1999) and instrument ethnic fractionalization in 1990 by ethnic fractionalization in 1980 (at the county level). The results from the two-stage-least squares are reported in Table B.6. We find that the impact of ethnic fractionalization on taxes per capita and expenditures per capita remains positive and significant.[15]

To sum up, data on U.S. cities and counties in 1990 confirm our prediction that more ethnic diversity (as measured by fractionalization or polarization) leads to more government taxation and public expenditure. However, mutliple explanations are possible for this positive relationship between diversity and taxation. In the following, we try to identify the empirical importance of our proposed channel by testing predictions from our model which are specific to our notion of 'diversity taxes'.

Exploring the channel: taxes per capita, majority size and minority fragmentation

Our theoretical analysis predicts (i) a negative relation between taxes per capita and the size of the majority (prediction 2.3.1), (ii) a negative relation between taxes per capita and the size of the biggest minority group (prediction 2.3.3), and (iii) a positive relation between taxes per capita and the fragmentation of minority groups (prediction 2.3.4).

Table B.4 reports the results from regressing taxes per capita on the different variables of ethnic diversity on city and county level. We run two specifications per local level. Both specifications test for the impact of the size of the majority as well the fragmentation of the minority on local taxes per capita. In the first specification, we capture minority fragmentation by the fractionalization of the minority groups. In the second specification, we use the size of the biggest minority group as a measure of minority fragmentation. Consider columns (1) and (3) which show the results from regressing taxes per capita on the size of the majority and the fractionalization of the minority. We find that the size of majority is negatively impacting taxes per capita while fractionalization of the minority is positively influencing taxes per capita, both on the city and on the county level. Next,

[15] Note that our results may be biased if ethnic fractionalization in 1980 was already causally impacted by local taxes per capita, and if local taxes per capita are persistent over time.

consider columns (2) and (4). We regress taxes per capita on the size of the majority group and the size of the biggest minority group. Again, we obtain a significant negative impact from the size of the majority on taxes per capita. In addition, we find that the size of the biggest minority group is significantly and negatively affecting taxes per capita. We conclude that prediction 1, prediction 3, and prediction 4 are confirmed. We believe that the evidence presented in table B.4 further strengthens the case for the existence of 'diversity taxes' as the channel through which cultural diversity impacts local taxation. Our theoretical analysis establishes a positive link between taxes per capita and total negative externalities imposed on the majority group. As we discussed, this link is able to explain the direction and significance of the coefficients reported in table B.4 which might otherwise be puzzling.

We conclude that the empirical evidence for U.S. cities and counties in 1990 confirms the predictions from our theoretical analysis. More ethnic diversity (as measured by fractionalization or polarization indices) leads to more taxes per capita and local government expenditure per capita. We see this as evidence for 'diversity taxes'. Our model explains the existence of these diversity taxes in the context of diverse groups imposing negative consumption externalities on each other. We tested this theory by relating the total amount of negative externalities imposed on the majority group by minority groups to various characteristics of the distribution of groups. The empirical evidence confirms a significant positive relationship between taxes per capita observed in the data and the amount of negative externalities imposed on the majority as we infer it from the distribution of groups. We see this as evidence for our hypothesized channel which relates diversity to taxation.

2.4 Conclusion

We propose a model in which diverse cultural groups impose negative consumption externalities on each other. Assuming that governments are restricted to use income taxes to control the production of negative consumption externalities, we show that 'diversity taxes' arise as a tool to mitigate social conflict. We find this restriction to be a reasonable approximation of liberal democracies which value freedom of (cultural) expression by any of its members. We provide a theoretical framework to study the relationship between different government social welfare functions and the taxes they impose for a given level of diversity. We show that in the equilibrium of the model, local taxes per capita are determined by the interplay between the government's social welfare function and the total amount of negative externalities produced in the economy. We further endogenize the government's social welfare function by incorporating the political process of majority voting.

One of the main predictions from our theoretical analysis is that more diversity leads to a bigger size of the government as measured by taxes per capita. We test this prediction using the U.S. city and county data provided by Alesina et al. (1999). We find robust and significant evidence for the existence of 'diversity taxes' even after including a variety of socioeconomic and demographic controls and after instrumenting for ethnic fractionalization. We further document significant relationships between majority group size and taxes as well as between minority fractionalization and taxes in line with the predictions of our theory. These results lend credence to our notion of social conflict manifesting itself in negative consumption externalities between diverse groups, and its regulation through public policy.

Documenting 'diversity taxes' in countries other than the US is a fruitful avenue for future research. We expect that negative consumption externalities between culturally divided groups lead to higher local taxes per capita in other countries as well. It is worth pointing out that the fact that the relatively homogeneous Scandinavian countries have generally higher taxes per capita (and more public good provision) than the relatively heterogeneous US does not contradict our theory. Our notion of 'diversity taxes' is meant to explain within-country variation in local taxes per capita, not across-country variation. Our theory predicts that within a given country, more diverse communities impose higher local taxes per capita.

Chapter 3

Ideologically Radical, Tactically Conservative

AUTHORS: SAUMYA DEOJAIN & VALERIO DOTTI[1]

3.1 Introduction

Political coalitions play a crucial role in shaping legislative bargaining and policy making within majoritarian institutions (Baron, 1991). The restrictions within a coalition together with its internal composition determine the policy proposals that are brought to the floor for a vote. This is why legislative coalitions can potentially affect policy outcomes.

A successful stream of literature strongly supports the claim that parties *do matter* in shaping policy reforms, even if there is a Condorcet winner – i.e., a policy proposal that defeats all the possible alternatives by simple majority voting – within the choice set of legislators or voters (Austen-Smith, 1986; Austen-Smith and Banks, 1996; Jackson and Moselle, 2002; Levy, 2004, Morelli, 2004; Eguia, 2011). In contrast, there is little theoretical research examining the conditions on the characteristics of legislators which can create legislative coalitions that *prevent* or limit the extent of policy reforms.

In an assembly, a reform backed by a majority of legislators can be blocked if (i) the *status quo* is preferred to the reform by some agents, and (ii) no majority consensus as to which policy should

[1]Post-doctoral Fellow, Department of Economics, Washington University in St. Louis

pass – i.e., a *legislative gridlock*.

Political coalitions (or in some cases, political parties) are deemed to play a key role in this process. For instance, by widening the set of available alternatives, they can generate disagreement regarding which reform should be implemented. In turn, disagreement causes the set of feasible reforms to shrink, and in some cases a legislative gridlock may prevail.

In this chapter, we study the role of legislative coalitions in blocking policy reforms in a legislature. The goal is to identify the features of the ideological composition of a legislative body that can induce a gridlock.

To achieve this goal, we adopt a model of legislative bargaining among endogenous political parties that is based on the work of Levy (2004) and Dotti (2020). In detail, we propose a model of legislative bargaining within an assembly in which legislators are organized into parties. A partition of the set of legislators into parties constitutes a party structure. The policy space is a multidimensional partially ordered set. Parties engage in legislative bargaining with the aim of selecting a policy to implement from such a set.

Legislators' preferences over policies are Euclidean and satisfy some ordinal properties: namely, legislators can be ordered along a generalized left vs. right preference dimension. This assumption implies that a median legislator can be identified even if neither single-peakedness nor the unidimensional single-crossing condition (Gans and Smart, 1996) hold, and no Condorcet winner exists.

We find that two measures of a legislator's ideology play a key role in shaping the answer to the main question of this chapter. The first is *conservatism*, i.e. the extent to which a legislator finds the status quo desirable relative to median legislator. The second is *extremism*, which captures how far is a legislator is from the median legislator with respect to the left vs. right divide.

We show that a stable gridlock occurs only if legislators located on both side of the left-right political spectrum are sufficiently conservative and extremists. If that is the case, conservative-extremists on opposite sides can form a coalition aiming to tactically propose a less ambitious reform relative to the one preferred by the median legislator. The goal of such proposal is solely that of generating disagreement in the assembly, which leads to a gridlock. Thus, the coalition of conservative-extremists never actually propose any successful reform, but can block any reformist effort by the moderate legislators. This situation can occur even if a large majority in the assembly prefers the median legislator's ideal point to the status quo, preventing reforms that enjoy large support from being implemented. We also show that stable gridlock[2] can occur when a majority of legislators are moderately conservative and extremist. Particularly, if on one side of the political spectrum there are a large number of non-extremist conservatives and on the other side of the political spectrum there are large number of extremists and some non-extremists stable gridlock may occur. This result indicates that if extremists and conservatives, defined with respect to the median votes and the status quo, can be divided into the left and right political legislators gridlock can occur.

The key mechanisms that underpin the results of the chapter are illustrated in the following example.

3.1.1 Motivating Example

Consider an assembly consisting of 7 legislators bargaining under majority rule. The set of legislators is $\mathscr{N} = \{1,2,3,4,5,6,7\}$. The policy space is $X = \mathbb{R}^2$, and the *status quo* is x^0. Legislator i's preferences are Euclidean with ideal point x^i.[3] The ideal points of the legislators are: $x^1 = (-4,-5), x^2 = (-4,-4), x^3 = (0,-27/28), x^4 = (0,0), x^5 = (2,0), x^6 = (2,7/2), x^7 = (8,4)$ where x^i is the ideal point of legislature in $i \in \mathscr{N}$. Ideal points are totally ordered in X but not aligned, as illustrated in figure 3.1. The latter assumption is necessary to rule out the existence of a Condorcet winner Plott (1967).[4] The ideal point of the median legislator is $x^4 = (0,0)$ is denoted as x^v.

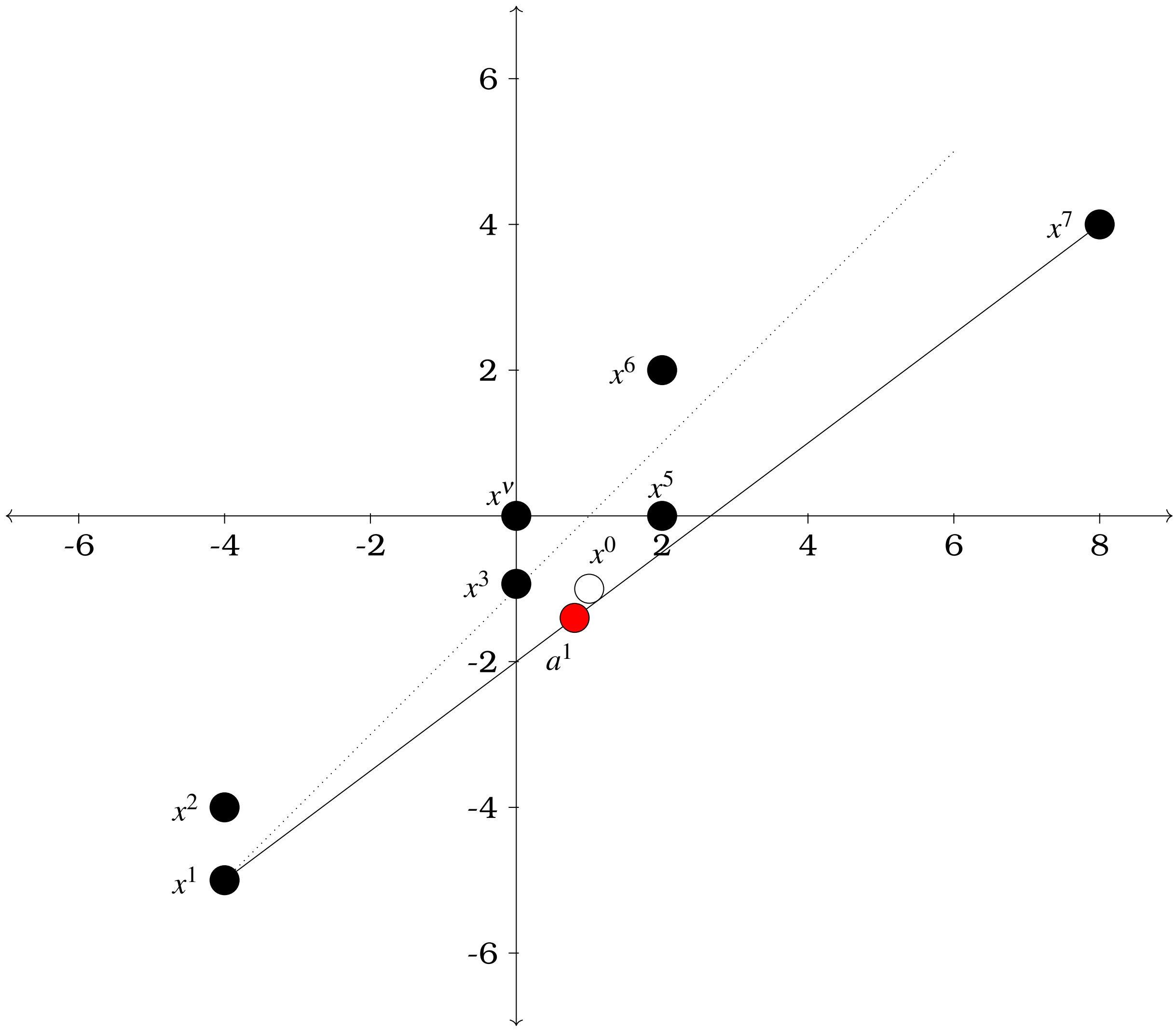
[2]For other definitions of gridlock used in the political literature see seminal works: Krehbiel (1998), Shepsle and Weingast (1987), Baron and Ferejohn (1989).

[3]I.e., they can be represented by the objective function $u^i(x) = -\left[x - x^i\right]^T \left[x - x^i\right]$.

[4]Formally $x^j \leq x^k$ for all $j \leq k$, and for any three distinct ideal points x^k, x^l, x^j, one gets $\left[x^k - x^l\right]^T \left[x^l - x^j\right] \neq 0$. This assumption ensures that the condition of *radial symmetry* in Plott (1967) is violated for any subset of three alternatives.

A partition $\mathbb{P}$ of $\mathscr{N}$ is a *party structure*. Each set $\mathscr{C}^j \in \mathbb{P}$ is a *party*. The political process is divided in two stages. In the first stage, each party $\mathscr{C}^j$ simultaneously make a *proposal* $\boldsymbol{a^j}$. A proposal can be any policy in the Pareto set of the party members, or $\boldsymbol{a^j} = \emptyset$ if party $\mathscr{C}^j$ is inactive. In the second stage, the legislators use the method of majority rule to select one proposal, called *policy outcome*, in the set $A = \{\boldsymbol{a^j}\}_{\mathscr{C}^j \in \mathbb{P}}$. If the majority rule does not deliver a stable outcome, then the policy outcome is the status quo x^0. In this example, the status quo is $x^0 = (1, -1)$.
Lastly, a party structure is required to be *stable*. Informally, a party structure $\mathbb{P}$ is stable if, given A, no subset of members has a strict incentive to quit his/her party in order to induce a different policy outcome.

Notice that the ideal point of the median legislator x^4 is strictly preferred to the status quo by a majority of legislators. Specifically, legislators 2, 3, 4, 6 strictly prefer x^4 to x^0, legislator 1 is indifferent between the two alternatives, and only legislators 5 and 7 strictly prefer x^0 to x^4. In this example even a policy reform that enjoys such a large support can be blocked.



In this figure, The black dots denote the position ideal policies of the legislators in the $\mathbb{R}^2$ policy space. Legislator 4 is the median voter with ideal policy $x^4 \equiv x^v$. The coalition $\{1,7\}$ is denoted by the solid line between ideal points of legislator 1 and 7. All other legislatures form singleton coalitions. In this example, coalition $\{1,7\}$ proposes a^1, legislators $3,2$ and 6 propose $\emptyset$ and legislator 4 proposed x^v and legislator 5 proposes x^5. The position of a^1, x^v and x^5 results in a Condorcet cycles which in turn results in a gridlock. The dotted line partitions the policy space into those legislators that prefer x^0 to x^v and those who do not. In this case, $1,5$ and 7 prefer the status quo while legislators $2,3,4$ and 6 prefer the median ideal point.

Figure 3.1: Stable legislative gridlock

Consider the party structure $\mathbb{P} = \{\{1,7\};\{2\};\{3\};\{4\};\{5\};\{6\}\}$ and the vector of policy proposals $A = \{a^1;\emptyset;\emptyset;x^4;x^5;\emptyset\}$. It is easy to verify that such proposals are optimal for all legislators and parties. Specifically, parties $\{2\}$, $\{3\}$, and $\{6\}$ are inactive because each of them cannot change the policy outcome by proposing the ideal policy of its unique member. Party $\{4\}$ optimally proposes x^4 because its unique member prefers x^0 to the policy outcome that would prevail if the party becomes inactive, which is x^5. Similarly, party $\{5\}$ optimally proposes x^5 because $u^5(x^0) > u^5(a^1)$. Lastly, party $\{1,7\}$ optimally proposes a^1 because both members prefer x^0 to x^4.

Moreover, neither 1 nor 7 has a strictly incentive to quit the party, because such a choice would lead to a policy outcome x^4, which both consider weakly worse than x^0. Thus, $\langle \mathbb{P}, \{a^1; x^4; x^5\}, x^0 \rangle$ is a stable legislative gridlock. This implies that the reform proposed by the median legislator does not gain a stable support and the status quo is maintained.

In the next section, we set up the model of this endogenous party formation when legislatures have Euclidean preferences in finite dimensional policy space. We use this model to characterize necessary conditions under which legislature gridlock is possible.

3.2 The Model

Consider an assembly that consists of an odd number $N \geq 3$ of legislators. The set of legislators is $\mathscr{N} = \{1, 2, ..., i, ..., N\}$. The median element of $\mathscr{N}$ is the *median legislator*, denoted by ν. The policy space is assumed to be $\mathbb{R}^m$. A party structure is a partition $\mathbb{P}$ of the set $\mathscr{N}$. A typical element $\mathscr{C}^j \in \mathbb{P}$ is called a *party*.

Each party $\mathscr{C}^j$ makes a proposal $\boldsymbol{a^j} \in X^j \cup \{\emptyset\}$, where $X^j \subseteq \mathbb{R}^m$, and $\boldsymbol{a^j} = \emptyset$ means that party $\mathscr{C}^j$ proposes no policy reform – i.e., it is *inactive*. A *policy profile* $A := \{\boldsymbol{a^j}\}_{\mathscr{C}^j \in \mathbb{P}}$ is the set of policy proposed by each party.

Given a party structure $\mathbb{P}$, the legislators play a two-stage game:

1. the members of each party $\mathscr{C}^j$ collectively make a proposal $\boldsymbol{a^j}$ (proposal stage);

2. each legislator votes for one of the available proposals $\boldsymbol{a^j} \in A$ (voting stage).

A voting rule selects the winning proposal (if any exists), which is called the *policy outcome* denoted by w. Let $\boldsymbol{x}$ be the policy implemented. If $w \neq \emptyset$ then $\boldsymbol{x} = w$ otherwise $\boldsymbol{x} = \boldsymbol{x_0}$.

3.2.1 Preferences

Preferences over policy outcomes. Each legislator $i \in \mathscr{N}$ has Euclidean preferences over their ideal point $\mathbf{x^i} \in \mathbb{R}^m$. Specifically, the utility function of legislator i is given by:

$$u^i(\boldsymbol{x}) = -||\boldsymbol{x^i} - \boldsymbol{x}||$$

where $||.||$ denotes the standard Euclidean distance in $\mathbb{R}^m$. As preferences of agents are uniquely defined by their ideal points we can define a preference profile as $\pi \equiv (\boldsymbol{x^1}, ..., \boldsymbol{x}^N)$. Let Π denote the $N-$ fold Cartesian product of the set of individual preferences. Thus, any collection of preference orderings of all $i \in \mathscr{N}$, π is such that $\pi \in \Pi$. The tuple $(\mathbb{R}^m, \pi)$ defines the preferences of all the legislators over the policy space.

Legislatures are ordered according to their preference type $\theta^i \in \Theta$, where Θ is a totally ordered set. Two legislatures $i, j \in \mathscr{N}$ have the same type $\theta^i \in \Theta$ if and only if they have the same preferences i.e., if $u^i(\boldsymbol{x}) \geq u^i(\boldsymbol{x}') \longleftrightarrow u^i(\boldsymbol{x}) \geq u^i(\boldsymbol{x}')$ for all $x \in X$. In other words, two legislatures $i, j \in \mathscr{N}$ have the same preferences if and only if $\boldsymbol{x^i} = \boldsymbol{x^j}$. Using the ordinal restrictions in Dotti (2020) and Milgrom (1994) [5], we assume that legislator's policy preferences have *Strict Single Crossing Property (SSC)*. Formally,

(SSC in x, θ*)* if, for all $\boldsymbol{x}', \boldsymbol{x}'' \in \mathbb{R}$ such that $\boldsymbol{x}' > \boldsymbol{x}''$:[6]

1. $u^i(\boldsymbol{x}') \geq u^j(\boldsymbol{x}'') \rightarrow u^i(\boldsymbol{x}') \geq u^j(\boldsymbol{x}'')$ for all $j \geq i$, and

2. $u^i(\boldsymbol{x}') > u^i(\boldsymbol{x}'') \rightarrow u^j(\boldsymbol{x}') > u^j(\boldsymbol{x}'')$ for all $j > i$ such that $\theta^j > \theta^i$.

The *SSC* property imposes a restriction on Euclidean preferences.

If Euclidean preferences have a Strict Single Crossing Property then there must be a total ordering of the ideal points in $\mathbb{R}^m$.

[5] For our results to hold we also need preferences to be *Quasisupermodular* (QSM), see Dotti (2020). We show that Euclidean preferences satisfy QSM in the appendix

[6] The strict order $\boldsymbol{x}' > \boldsymbol{x}''$ indicates that $\boldsymbol{x}' \geq \boldsymbol{x}''$ and $x'_k > x''_k$ for at least one element k.

The *SSC* property on Euclidean preferences restricts the ideal points of the legislators to be totally ordered in $\mathbb{R}^m$. This ensures the existence of a median type of legislator. We label this median voter type as $\theta^\nu \in \Theta$. Formally, median voter θ^ν is defined as the legislator $\nu \in \mathscr{N}$ that satisfies conditions $|\sum_{i \in \mathscr{N}} \boldsymbol{x}^i \geq \boldsymbol{x}^\nu| \geq \frac{1}{2}$ and $|\sum_{i \in \mathscr{N}} \boldsymbol{x}^i \leq \boldsymbol{x}^\nu| \geq \frac{1}{2}$.

Conditional preferences over policy proposals. Each legislator i member of party $\mathscr{C}^j \in \mathbb{P}$ also has preferences over X^j in relation to which proposal in X^j his/her party should put forward conditional on the proposals of other parties $\boldsymbol{a}^{-j}$.[7] We define for each legislator $i \in \mathscr{C}^j$ and each pair of proposals $\boldsymbol{x}', \boldsymbol{x}'' \in X^j$ the conditional utility $u^{j,i}(.|\boldsymbol{a}^{-j}, \pi)$ over X^j. For instance, $u^{j,i}(\boldsymbol{x}'|\boldsymbol{a}^{-j}) \geq u^{j,i}(\boldsymbol{x}''|\boldsymbol{a}^{-j})$ means that legislator i, member of party $\mathscr{C}^j \in \mathbb{P}$ prefers his/her party to propose $\boldsymbol{x}'$ rather than $\boldsymbol{x}''$ conditional on other parties proposing $\boldsymbol{a}^{-j}$. There are no further restrictions on $u^{j,i}(.|\boldsymbol{a}^{-j})$. In section 2.4, we restrict the type of $u^{j,i}(.|\boldsymbol{a}^{-j})$ that can be part of an equilibrium to those that are consistent with purely policy-motivated legislators.

Lastly, let $u^j(.|\boldsymbol{a}^{-j})$ denote the conditional utility profile of all $i \in \mathscr{C}^j$, and $\mathbf{R}^j$ be the $|\mathscr{C}^j|$-fold Cartesian product of the set of individual preferences over policy proposals in X^j.[8]

3.2.2 Proposal Stage

Following Levy (2004), we assume that the proposal of each party $\mathscr{C}^j$ must lie in the Pareto set of its members, such that the set of available proposals is $X^j_\pi := \{\boldsymbol{x} \in \mathbb{R}^m | \nexists \boldsymbol{x}' \in \mathbb{R}^m \, s.t. \, u^i(\boldsymbol{x}') \geq u^i(\boldsymbol{x}) \; \forall i \in \mathscr{C}^j \; and \; u^i(\boldsymbol{x}') \geq u^i(\boldsymbol{x}) \, for \, some \, k \in \mathscr{C}^j\}$.[9] This assumption is common in models of endogenous parties (e.g., Roemer (1999), Levy (2004)), and it is further discussed in Levy

[7]Notice that $\boldsymbol{a}^{-j}$ lists all the proposals of other parties, including null proposals $\emptyset$.
[8]Both $R^j(\boldsymbol{a}^{-j})$ and $\mathbf{R}^j$ do not depend on the party structure $\mathbb{P}$.
[9]Notice that if a party is a singleton, then X^j reduces to the set of ideal policies of its unique member (as in a citizen–candidate model).

(2004)[10]. For Euclidean preferences X^j is equivalent to the convex hull of the types' ideal points, $\pi^j = \{\boldsymbol{x}^i\}_{i \in \mathscr{C}^j}$.

Departing from Levy (2004) and following Dotti (2020), we assume that each party chooses its proposal $\boldsymbol{a}^j$ according to a *protocol* in the form of a surjective conditional group choice function:

$$g^j_{\boldsymbol{a}^{-j}} : \mathbf{R}^j \to \mathbb{R}^m \cup \{\emptyset\}$$

that must satisfy the following three conditions:

1. *Monotonicity (M)*: Consider any two conditional preference profiles $u^j(.|\boldsymbol{a}^{-j})$, $\hat{u}^j(.|\boldsymbol{a}^{-j}) \in \mathbf{R}^j$ such that for all $i \in \mathscr{C}^j$ and $\exists\ \boldsymbol{x}', \boldsymbol{x}'' \in X^j$, $\hat{u}^i(\boldsymbol{x}''|\boldsymbol{a}^{-j}) \geq \hat{u}^i(\boldsymbol{x}'|\boldsymbol{a}^{-j}) \to u^i(\boldsymbol{x}''|\boldsymbol{a}^{-j}) \geq u^i(\boldsymbol{x}'|\boldsymbol{a}^{-j})$. Then if $g^j_{\boldsymbol{a}^{-j}}(\hat{u}^j(.|\boldsymbol{a}^{-j})) = \boldsymbol{x}''$ implies $g^j_{\boldsymbol{a}^{-j}}(u^j(.|\boldsymbol{a}^{-j})) \neq \boldsymbol{x}'$.[11]

2. *Neutrality (N)*: For any conditional preference profile $u^j(.|\boldsymbol{a}^{-j}) \in \mathbf{R}^j$, let $\rho :\ X^j \to X^j$ be a permutation of X^j and let $u^j_\rho(.|\boldsymbol{a}^{-j}) \in \mathbf{R}^j$ be such that for all $i \in \mathscr{C}^j$ and all $\boldsymbol{x}, \boldsymbol{x}' \in X^j$ with $\boldsymbol{x} \neq \boldsymbol{x}'$, $u^i_\rho(\boldsymbol{x}|(\boldsymbol{a}^{-j})) \geq u^i_\rho(\boldsymbol{x}'|(\boldsymbol{a}^{-j}))$ if and only if $u^i(\rho(\boldsymbol{x})|(\boldsymbol{a}^{-j}) \geq u^i(\rho(\boldsymbol{x}')|(\boldsymbol{a}^{-j})$. The function $g^j_{\boldsymbol{a}^{-j}}$ is neutral if $\rho(g^j_{\boldsymbol{a}^{-j}}(u^j(.|\boldsymbol{a}^{-j}))) = g^j_{\boldsymbol{a}^{-j}}(u^j_\rho(.|\boldsymbol{a}^{-j}))$.

Let $w(A, \pi)$ denote the expected policy outcome given policy profile A and preference profile π, and define $\boldsymbol{a}^{-j} = A \setminus \{\boldsymbol{a}^j\}$. A tie-breaking rule on $g^j_{\boldsymbol{a}^{-j}}$ that satisfies the following conditions.

1. *TB1 (Inaction of irrelevant parties)*. If $w(\{\emptyset\} \cup \boldsymbol{a}^{-j}, \pi) = w(\{x\} \cup \boldsymbol{a}^{-j}, \pi)$ for some $\boldsymbol{x} \in X^j$, then $g^j_{\boldsymbol{a}^{-j}}(u^j(.|\boldsymbol{a}^{-j})) \neq \boldsymbol{x}$.

2. *TB2 (Office seeking)*. If $\exists x' \in X^j$ such that $w(\{\boldsymbol{x}'\} \cup \boldsymbol{a}^{-j}, \pi) \neq w(\{\emptyset\} \cup \boldsymbol{a}^{-j}, \pi)$ and $u^i(w(\{\boldsymbol{x}'\} \cup \boldsymbol{a}^{-j}, \pi)) \geq u^i(w(\{\emptyset\} \cup \boldsymbol{a}^{-j}, \pi))\ \forall i \in \mathscr{C}^j$, then $g^j_{\{\emptyset\}}(u^j(.|\boldsymbol{a}^{-j})) \neq \emptyset$.

The conditions *(M)* and *(N)* [12] discipline the choice of the party proposal. In particular, *(M)* ensures

[10] The author suggests some possible interpretations of this assumption, which are consistent with the setting of this chapter.

[11] Monotonicity implies the *weak Pareto principle* (with respect to the conditional preferences) because the function $g^j_{\boldsymbol{a}^{-j}}$ is surjective.

[12] Conditions *(M)* and *(N)* are widely adopted – and often deemed desirable – in social choice literature. They are satisfied by any weighted majority voting rule. See Dasgupta and Maskin (2008).

that (conditional) Pareto inferior proposals are never chosen.[13] Monotonicity is admittedly a strong assumption, but is deemed to be credible for this application. Specifically, it is typically satisfied by the choice methods adopted by most political parties in Western countries such as (i) voting within the party congress, (ii) the dictatorship of the leader, and (iii) many forms of bargaining within a small committee (Ceron, 2012).

Condition *(N)* implies that the choice protocol treats all alternatives in X^j symmetrically: if the alternatives are relabeled via ρ, then the chosen proposal is relabeled in the same way. Thus, *(N)* implies that the choice over policy proposals is solely driven by the party members' preferences over policy proposals, and by their relative political power within the party.

Condition *TB1* can be rationalized as the effect of a small cost of participating in the legislative process, which is not explicitly modeled in the proposed setting. Specifically, it states that a party chooses to be inactive rather than making a proposal that has no influence on the policy outcome. A very similar assumption is imposed in Levy (2004) for the same reason.

Condition *TB2* simply states that a party always makes a proposal if no other party is active and if it can secure an outcome that is, for all its members, at least as good as the one that would prevail if the party is inactive.

3.2.3 Voting Stage

In the second stage, legislators must select a proposal in A. The assembly ranks the alternatives in A using the *method of majority rule* whenever this method delivers an outcome, and selects the status quo $\boldsymbol{x}^0$ otherwise. Each legislator i's voting behavior is driven solely by his/her preferences over policy outcomes, and it is not affected by party membership.[14]

Formally, let $MV(\pi)$ be the complete social preference relation induced by the majority rule under

[13] *(M)* implies *conditional* weak Pareto efficiency, and it is motivated by standard social choice considerations. The assumption on the codomain of $g^j_{\boldsymbol{a}^{-j}}$ corresponds to *unconditional* Pareto efficiency, and it is justified by commitment issues.

[14] The results are unaffected if one restricts the attention to equilibria in which no legislator votes against the proposal made by his/her party.

a preference profile π.[15] Given this social preference relation $MV(\pi)$ and a subset $A \subseteq \mathbb{R}^m$, define $K(A,\pi) := \{\boldsymbol{x} \in A | \boldsymbol{x} MV(\pi) \boldsymbol{x}' \;\; \forall x' \in A\}$ as the set of *MV-maximal* alternatives in A, which corresponds to the core of the majority voting game over A. As legislators have convex preferences if $K(A,\pi)$ is non-empty then it is a singleton set.

The assembly chooses the outcome from a set of available alternatives $A \subseteq \mathbb{R}^m$ according to a social choice function:

$$W: \; \mathscr{P}(\mathbb{R}^m) \times \Pi \to \{\mathbb{R}^m \cup \boldsymbol{x}^0\}$$

Thus, the policy outcome w is the outcome of the function W. The social choice function W satisfies the following conditions.

1. *Majority Rule (MR)*: If $K(A,\pi)$ is nonempty and $A \neq \{\emptyset\}$, then $W(A,\pi) = K(A,\pi)$.

2. *Inertia (I)*. If either $K(A,\pi) = \{\emptyset\}$, or $A = \{\emptyset\}$, then $W(A,\pi) = \boldsymbol{x}^0$ – i.e., the *status quo* is maintained.

Assumptions *(MR)* and *(I)* correspond to the social choice procedure that results from the *method of majority rule*. In such a case, if there is a (weak) Condorcet winner among the set of alternatives, then the core of the voting game is nonempty[16] and the policy outcome lies in the core. Conversely, if no Condorcet winner exists, then no reform proposal can gain a stable support[17].

[15] Thus, $\boldsymbol{x}' MV(\pi) \boldsymbol{x}''$ if and only if $\sum_{i=1}^{N} \mathbf{1}[u^i(\boldsymbol{x}') \geq u^i(\boldsymbol{x}'')] \geq N/2$. This definition can be easily extended to the case in which the set of voters is a continuum. Notice that $MV(\pi)$ is complete, reflexive, but not necessarily transitive.

[16] The relationship between the concepts of (strong) *core* and *Condorcet winner* is described in Ordeshook (1986), pp 347-349.

[17] A potential shortcoming of this approach is that the status quo $\boldsymbol{x}^0$ may not be an element of A. This could lead to a strange scenario in which a the status quo wins the majority vote against alternatives that emerge in a stable equilibrium, but will never be the policy implemented. In order to get around this, one can assume that before the proposal stage agents vote whether or not to change the status quo. If they vote to keep the status quo then x^0 is implemented, if they vote to change the status quo then the proposal stage starts and the game proceeds exactly as before. With this assumption if majority of agents prefer the status quo to the proposals in A, the status quo will be implemented. This added assumption does not take away the core intuition behind our results for the emergence of gridlock.

3.2.4 Equilibrium of the Proposal Game

For a given party structure $\mathbb{P}$, we define an equilibrium of the proposal game as follows.

***Definition* 3.2.1.** *Equilibrium of the Proposal Game):. A collection* $\{\boldsymbol{a^j}\}_{\mathscr{C}^j\in\mathbb{P}} \equiv A(\mathbb{P})$ *is an equilibrium of the proposal game if there exists* $\{u^j(.|\boldsymbol{a}^{-j}), g^j_{\boldsymbol{a}^{-j}}\}_{\mathscr{C}^j\in\mathbb{P}}$*, such that for all* $\mathscr{C}^j \in \mathbb{P}$*, the following conditions hold: (i) for any* $i \in \mathscr{C}^j$*, and for any* $\boldsymbol{x}, \boldsymbol{x}' \in X^j$*,* $u^{j,i}(\boldsymbol{x}|\boldsymbol{a}^{-j}) \geq u^{j,i}(\boldsymbol{x}'|\boldsymbol{a}^{-j} \iff u^i(w(\{\boldsymbol{x}\}\cup\boldsymbol{a}^{-j},\pi)|\boldsymbol{a}^{-j}) \geq u^i(w(\{\boldsymbol{x}'\}\cup\boldsymbol{a}^{-j},\pi)|\boldsymbol{a}^{-j})$*; (ii)* $g^j_{\boldsymbol{a}^{-j}}(R^j(\boldsymbol{a}^{-j})) = \boldsymbol{a^j}$*. An equilibrium of the proposal game* $A(\mathbb{P})$ *is either (i) a* regular equilibrium *if* $K(A(\mathbb{P}),\pi) \neq \{\emptyset\}$*, or (ii) a* legislative gridlock *if* $K(A(\mathbb{P}),\pi) = \{\emptyset\}$.

The equilibrium condition on $R^j(\boldsymbol{a}^{-j})$ implies that each legislator i member of party $\mathscr{C}^j$, for the purpose of choosing the party proposal, ranks the policies in X^j solely with respect to the policy outcome they induce conditional on the proposals of other parties. For instance, the policy $\boldsymbol{x}'$ is (conditional) weakly preferred to $\boldsymbol{x}''$ by legislator i if and only if, given other party proposals $\boldsymbol{a}^{-j}$, he/she (unconditionally) prefers the policy outcome $w(\{\boldsymbol{x}'\}\cup A(\mathbb{P})\setminus\{\boldsymbol{a^j}\},\pi)$ to $w(\{\boldsymbol{x}''\}\cup A(\mathbb{P})\setminus\{\boldsymbol{a^j}\},\pi)$.

Notice that, in contrast to Levy (2004), in this setting an equilibrium of the proposal game may not always exist. This scenario occurs if for any $A \subseteq \mathbb{R}^m$ there is no collection $\{u^j(.|\boldsymbol{a}^{-j}), g^j_{\boldsymbol{a}^{-j}}\}_{\mathscr{C}^j\in\mathbb{P}}$, such that $g^j_{\boldsymbol{a}^{-j}}(u^j(.|\boldsymbol{a}^{-j})) = \boldsymbol{a^j}$ for all $\mathscr{C}^j \in \mathbb{P}$. In such a case, we assume $A(\mathbb{P}) = \{\emptyset\}$.[18] On the other hand, an equilibrium of the overall game, which we describe in the next section, always exists.

3.2.5 Stable Party Structures

The concept of stability is borrowed from Ray and Vohra (1997), and it has been implemented in a model of political parties by Levy (2004), Levy (2005) and Dotti (2020). Below, we briefly summarize the stability concept, highlighting the key differences with the framework in Levy (2004).

[18]This assumption is only relevant to evaluate the profitability of a deviation that induces a new proposal game in which no equilibrium exists. See next section.

Players start from some party structure $\mathbb{P}$, and are only allowed to break parties by inducing finer partitions. Let $\aleph(\mathbb{P})$ denote all the party structures that are refinements of $\mathbb{P}$. A partition $\mathbb{P}' \in \aleph(\mathbb{P})$ is induced from $\mathbb{P}$ if it is generated by breaking a party $\mathscr{C}^j \in \mathbb{P}$ into two. Consider a party $\mathscr{C}^j \in \mathbb{P}$. A party $\mathscr{C}^d \subseteq \mathscr{C}^j$ is a *deviator* if it can induce $\mathbb{P}' \in \aleph(\mathbb{P})$ from $\mathbb{P}$. The members of a deviator $\mathscr{C}^d \subseteq \mathscr{C}^j$ take into account future deviations, both by members of their own party $\mathscr{C}^d$ and by members of other parties $\mathscr{C}^s \in \mathbb{P}'$, $s \neq d$.[19] Credible threats are deviations to finer partitions which are stable – i.e., the definition of stability is recursive.

The only difference in the definition of stability relative to Levy (2004) is that we allow for the outcome of disagreement, which we refer to as a legislative gridlock, to be stable. Formally, consider a sequence of partitions $\{\mathbb{P}_k\}_{k=1}^{K}$ such that for every $k = 2, ..., K$, $\mathbb{P}_k$ is induced by a deviator $\mathscr{C}_k^d \subset \mathscr{C}_{k-1}^j$ for some $\mathscr{C}_{k-1}^j \in \mathbb{P}_{k-1}$. We define stability as follows.

***Definition* 3.2.2.** *(Stability):* $\langle \mathbb{P}, A(\mathbb{P}) \rangle$ *is* sequentially blocked *by* $\langle \mathbb{P}', A(\mathbb{P}') \rangle$ *for some* $\mathbb{P}' \in \aleph(\mathbb{P})$ *if there exists a sequence* $\{\langle \mathbb{P}_1, A(\mathbb{P}_1) \rangle, \langle \mathbb{P}_2, A(\mathbb{P}_2) \rangle, ..., \langle \mathbb{P}_K, A(\mathbb{P}_K) \rangle\}$*, such that:*

1. $\langle \mathbb{P}_1, A(\mathbb{P}_1) \rangle = \langle \mathbb{P}, A(\mathbb{P}) \rangle$, $\langle \mathbb{P}_K, A(\mathbb{P}_K) \rangle = \langle \mathbb{P}', A(\mathbb{P}') \rangle$ *and for every* $k = 2, ..., K$ *there is a deviator* $\mathscr{C}_k^j$ *that induces* $\mathbb{P}_k$ *from* $\mathbb{P}_{k-1}$.

2. $\langle \mathbb{P}', A(\mathbb{P}') \rangle$ *is stable with* $A(\mathbb{P}') = A(\mathbb{P}_K)$ *for some* $A(\mathbb{P}_K)$.

3. $\langle \mathbb{P}_k, A(\mathbb{P}_k) \rangle$ *is not stable for any* $A(\mathbb{P}_k)$ *and for* $1 < k < K$.

4. $u^i(w(A(\mathbb{P}'), \pi)) > u^i w(A(\mathbb{P}_{k-1}), \pi)$ *for all* $k = 2, ..., K$*, and* $i \in \mathscr{C}_k^j$.

Then, $\langle \mathbb{P}, A(\mathbb{P}) \rangle$ *is* stable *if there is no* $\langle \mathbb{P}', A(\mathbb{P}') \rangle$ *for* $\mathbb{P}' \in \aleph(\mathbb{P})$ *that sequentially blocks* $\langle \mathbb{P}, A(\mathbb{P}) \rangle$.

The definition states that a tuple $\langle \mathbb{P}, A \rangle$ is stable if there is no leading deviator $\mathscr{C}^d$ that can induce a sequence of deviations such that at each step of the sequence an equilibrium of the proposal game $\langle \mathbb{P}_k, A(\mathbb{P}_k) \rangle$ (if any exists) is played, and such that it is strictly profitable for each deviator to

[19]Because of that, this stability concept satisfies *consistency* and *farsightedness* in the sense of Ray and Vohra (2015). The main results of this chapter go through even if players do not exhibit such sophisticated behavior.

choose to deviate, with respect to the final outcome of the sequence. Notice that the finest partition of $\mathscr{N}$ – i.e., the one in which each party $\mathscr{C}^j \in \mathbb{P}$ is a singleton – is always stable.

We define an equilibrium for the overall game following Dotti (2020) as *stable party structure* in following manner:

***Definition* 3.2.3.** *(Stable Party Structure): A Stable Party Structure (*SPS*) is a tuple $\langle \mathbb{P}, A(\mathbb{P}), w \rangle$ such that (i) $A(\mathbb{P})$ is an equilibrium of the proposal game given partition $\mathbb{P}$; (ii) $\langle \mathbb{P}, A(\mathbb{P}) \rangle$ is stable; (iii) w is a policy outcome of the second-stage game – i.e., $W(A(\mathbb{P}), \pi) = w$. A* SPS *is* a party equilibrium *if $A(\mathbb{P})$ is a regular equilibrium of the proposal game, and it is a* stable legislative gridlock *otherwise.*

This definition completes the description of the political process.

3.3 Results

The main results of this chapter are stated in this section. Section 3.3.1 contains the results regarding the existence and characterization of a stable party structure. Section 3.3.2 states the necessary conditions under which there will be legislative gridlock.

3.3.1 Existence and Quasi-Median Voter Theorem

We use the result derived in Dotti (2020) for deriving existance of a stable party structure. We restate the result for Euclidiean preferences it in the following lemma.

***Lemma* 3.3.1.** (Existence). (i) A stable party structure $\langle \mathbb{P}, A(\mathbb{P}), w \rangle$ always exists for any preference profile π; (ii) a party equilibrium such that the policy outcome is an ideal point of the median legislator m – i.e., $\langle \mathbb{P}, \{w\}, w \rangle$ with $w = \boldsymbol{x}^{\nu}$ – always exists; (iii) a stable legislative gridlock $\left\langle \mathbb{P}, A(\mathbb{P}), \boldsymbol{x}^0 \right\rangle$ exists only if either $\boldsymbol{x}^0 = \boldsymbol{x}^{\nu}$ or if $\boldsymbol{x}^0 \ngtrsim x^{\nu}$and $\boldsymbol{x}^0 \nlesssim x^{\nu}$.

***Lemma* 3.3.2.** *(Quasi-median voter theorem) Any stable party structure* $\langle \mathbb{P}, A(\mathbb{P}), w \rangle$ *is such that either* $w = \boldsymbol{x^\nu}$ *or* $w = \boldsymbol{x^0}$*. If* $\left\langle \mathbb{P}, A(\mathbb{P}), \boldsymbol{x^0} \right\rangle$ *is stable party structure then there exists* $i, k \in \mathcal{N}$ *with* $\theta^i < \theta^\nu < \theta^k$ *and* $C^j \subset \mathbb{P}$ *such that* $i, k \in C^j$ *with* $\boldsymbol{a^j} \notin \{\emptyset, \boldsymbol{x^\nu}\}$*.*

Given lemma 3.3.1, lemma 3.3.2 states that if a policy that is not the status quo is passed in the legislature it must be the median voter's ideal policy. This is because of the assumptions of monotonicity and neutrality we place on the conditional group choice function $g^j_{\boldsymbol{a}^{-j}}(.)$.

However, this also a quasi-median voter theorem because under certain preference distributions there are stable legislative gridlocks in which the status quo will be passed instead of $\boldsymbol{x^\nu}$.

It is easy to show that for legislative gridlock, or a Condorcet cycle, to exist proposal $\boldsymbol{\alpha} \neq \boldsymbol{x^\nu}$ must beat $\boldsymbol{x^\nu}$ in majority voting. This happens only if $\boldsymbol{\alpha}$ is not ordered with respect to $\boldsymbol{x^\nu}$, i.e. $\boldsymbol{\alpha} \ngeq \boldsymbol{x^\nu}$ and $\boldsymbol{\alpha} \nleq \boldsymbol{x^\nu}$. But this type of proposal is only possible if there exists an active coalition with members from both sides of the median as Pareto optimality restricts the set of alternatives to be within the convex hull of the ideal points of members in a coalition. This is why a legislative gridlock only exists if members from opposite sides of the political spectrum to form an active coalition. This result informs the main focus of this chapter which lays down the necessary conditions under which these legislative gridlock arise.

3.3.2 Conditions for legislative gridlock

In order to characterize what happens when there is a legislative gridlock we decompose the preferences of a legislator into two intuitive components. This can be done by decomposing the distance vector $\boldsymbol{z^i} \equiv \boldsymbol{x^i} - \boldsymbol{x^\nu}$. One component is a measure of relative conservatism that we call intensity of conservatism, I^i,

$$I^i = \frac{(\boldsymbol{x^i} - \boldsymbol{x^\nu}).(\boldsymbol{x^0} - \boldsymbol{x^\nu})}{||\boldsymbol{x^0} - \boldsymbol{x^\nu}||^2} = \frac{\boldsymbol{z^i}.\boldsymbol{z^0}}{||\boldsymbol{z^0}||}.$$

The other component captures the degree of extremism of a legislator i from the median, E^i. It is derived from the plane perpendicular to the vector $\boldsymbol{z^0}$. For the purposes of this chapter we will focus on a two dimensional policy space. Therefore,

$$E^i = \frac{||\boldsymbol{z^i}||sin\theta^i}{||z^0||}$$

where θ^i is the angle between vectors $\boldsymbol{z^i}$ and $\boldsymbol{z^0} \equiv \boldsymbol{x^0} - \boldsymbol{x^\nu}$ measured as indicated in figure 3.2. It is important to note that the preferences of i are fully described by the components I^i, E^i and the vector $\boldsymbol{z^0}$.

This definition of intensity of conservatism and degree of extremism is zero for the median voter. Thus, If $I^i > I^\nu = 0$ then the intensity of conservatism of i is greater than the median legislator so we say that i is more conservative than the median. If $I^i < 0$ then the intensity of conservatism for i is less than the median voter so i is less conservative than the median. Similarly, if $E^i > E^\nu = 0$ we say that i is a right leaning legislator and if $E^i < 0$ we say that she is left leaning. $|E^i|$ describes how far legislature i's preferences is from the median on a one dimensional left-right spectrum as seen in figure 3.2. We call $|E^i|$ the intensity of extremism of legislator i.

This decomposition allows us to characterize the necessary conditions that are needed for a legislative gridlock. There are two steps required in order for a legislative gridlock to occur. The first is already indicated in lemma 3.3.2. This lemma indicates that there should be at least one coalition with members from both sides of the political spectrum that find the status quo more attractive than the median voter. Only if this is the case will members of the coalition propose an alternative policy, $\boldsymbol{\alpha}$, to the median voter's policy $\boldsymbol{x^\nu}$ in equilibrium. Second, by definition of a gridlock, a Condorcet cycle must emerge from the proposals on the table for legislators. This means that α must beat the median in a majority voting scenario. This imposes further restrictions on the distribution of the preferences of the legislators. These two restrictions for a stable existence of a Condorcet cycle give us the main results of this chapter on legislative gridlock that we write as proposition 3.3.1 and proposition 3.3.2.

***Proposition* 3.3.1.** *(Minimum conservatism) A stable legislative gridlock exists only if there exists* $i, j \in \mathscr{N}$ *with* $\theta^i < \theta^\nu < \theta^j$ *such that* $I^i, I^j > 1/2$.

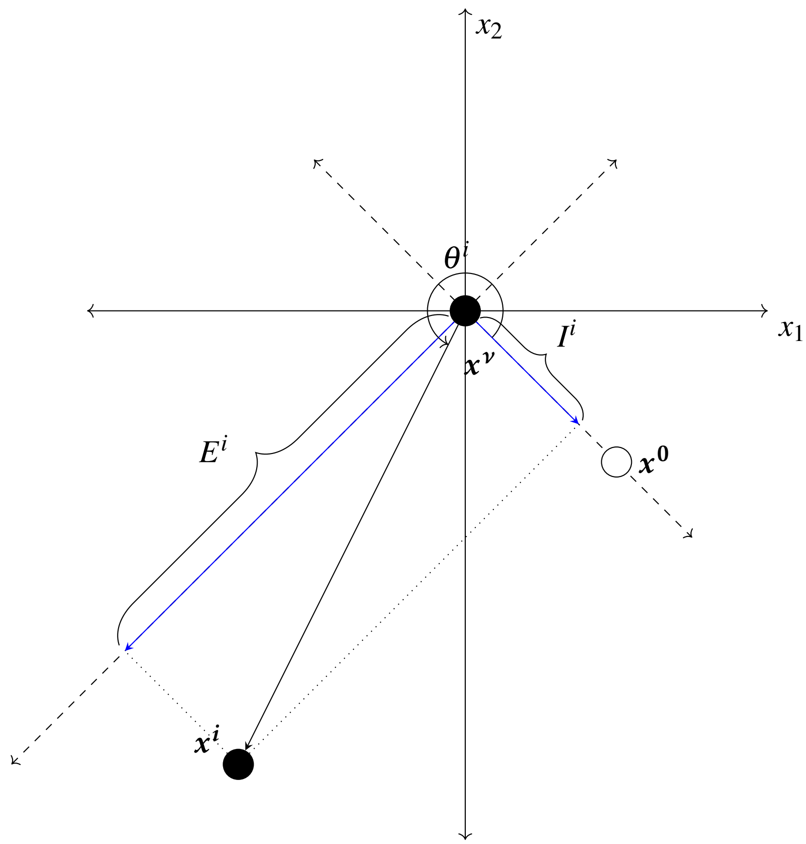


Figure 3.2: Decomposition of legislators' preferences into extremism with respect to median and conservatism with respect to the median

Proposition 3.3.1 states gridlock does not occur if legislators on both sides of the political spectrum are not conservative enough. If legislators are not conservative enough then they do not find the status quo attractive enough to even propose a platform. In fact, a legislator i prefers status quo to the median if and only if I^i is greater than $1/2$. This means, extremism plays no role in making the status quo more attractive to legislator i. Additionally, it is not enough if only one end of the political spectrum is conservative for legislative gridlocks to occur. This means if the legislative ideal points were aligned linearly no gridlock would have been possible because at least one side of the political spectrum would less conservative than the median.[20] Thus, our theory predicts that when a legislative gridlock occurs there must be some staggering of ideal points in a multidimensional space.

With a minimal level of conservatism required on both sides of the spectrum one can also easily derive the expression for a minimum level of distance from median required in order for legislative gridlock to be possible. We express it as *MD*.

$$MD \equiv \frac{\min\left\{\max_{i>v}\left\|\boldsymbol{x}^i - \boldsymbol{x}^v\right\|, \max_{i<v}\left\|\boldsymbol{x}^i - \boldsymbol{x}^v\right\|\right\}}{||\boldsymbol{z}^{\mathbf{0}}||}$$

***Corollary* 3.3.1.** *A stable legislative gridlock exists only if polarization is large enough, i.e.* $MD > 1/2$.

Using the triangle inequality we also get an expression for a minimum level of polarization required for legislative gridlock. Polarization is defined as the distance between the extreme most legislators. Formally, minimum polarization (*MP*) is defined as follows:

$$MP \equiv \frac{\max_{j,k}\left\|\boldsymbol{x}^j - \boldsymbol{x}^k\right\|}{||\boldsymbol{z}^{\mathbf{0}}||}$$

***Corollary* 3.3.2.** *A stable legislative gridlock exists only if* $MP > 1/\sqrt{2}$.

[20] This would mean that one side of political spectrum has negative intensity of conservatism which will always be less than $1/2$.

This corollary states that unless the legislators are sufficiently polarized legislative gridlocks will not occur. This is in line with some empirical literature on legislative gridlock.

The next proposition is a necessary condition for a Condorcet cycle to exist given a proposal $\boldsymbol{\alpha} \neq \boldsymbol{x^\nu}$ is proposed by a coalition, $\mathscr{C}^j$, that chooses $\boldsymbol{\alpha}$ specifically to block $\boldsymbol{x^\nu}$ from winning in order to retain status quo, $\boldsymbol{x^0}$. Proposition 3.3.1 imposes a restriction on preferences for members that will want to block $\boldsymbol{x^\nu}$ with $\boldsymbol{\alpha}$ in order to retain $\boldsymbol{x^0}$. As the set of policy alternatives for $\mathscr{C}^j$, X^j, is simply the convex hull of ideal points of members in $\mathscr{C}^j$, $\boldsymbol{\alpha} \in X^j$ is restricted by the preference restrictions given in Proposition 3.3.1. Suppose $\boldsymbol{\alpha}$ was not restricted by the members in a coalition with $\boldsymbol{x^0}\pi^i\boldsymbol{x^\nu}$. Then there is a member that is blocking $\boldsymbol{x^\nu}$ with $\boldsymbol{\alpha}$ even though it prefers $\boldsymbol{x^\nu}$ to $\boldsymbol{x^0}$. This is not equilibrium as this member will prefer to leave this coalition. Thus, $\boldsymbol{\alpha} \in \mathscr{X}_0^j$ where $\mathscr{X}_0^j$ is the convex hull of ideal points of the members in $\mathscr{C}^j$ who prefer the status quo to $\boldsymbol{x^\nu}$. This restriction on $\boldsymbol{\alpha}$ restricts the distribution of preferences that would induce a Condorcet cycle because they prefer $\boldsymbol{\alpha}$ to $\boldsymbol{x^\nu}$.

We define *conservative extremism* for a legislator i,

$$CE^i \equiv I^i + (E^i)^2$$

This captures the total intensity of extremism and conservatism. We derive the minimum level of conservative extremism required for majority legislators to vote for $\boldsymbol{\alpha}$ against a median voter's proposal.

***Proposition* 3.3.2.** *A stable legislative gridlock exists only if a majority of voters exhibit sufficiently high conservative extremism i.e.* $\sum_{i=1}^{n} \mathbf{1}\left[CE^i > 1/4\right] \geq n/2$.

Proposition 3.3.2 states that a gridlock occurs only if there exists only if a majority of voters exhibits sufficiently high conservative extremism. Moreover, it implies that the threshold on intensity of conservatism decreases with as intensity of extremism. This result suggests that a gridlock occurs only if a majority of voters have high intensity of conservatism relative to the median voter.

This condition leaves room for a situation where the median voter theorem fails. Even if majority of the voters are not conservative enough to like the status quo more than the median, we can still have a legislative gridlock. While conservatism is important for a Condorcet cycle to occur, the possibility of a Condorcet cycle emerging increases when extremism increases. According to our condition, the ones most likely to vote for a policy blocking the median policy are the extreme legislators.

Next, in order to state the necessary and sufficient conditions under which we have a legislative gridlock we introduce a little more notation. We first need to define a set of legislators which in equilibrium would block the median reform in the event of a legislative gridlock. To do this, we first define $\mathscr{L}(\pi) \equiv \{i \in \mathscr{N} | I^i > 1/2 \text{ and } \theta^i < \theta^\nu\}$ and $\mathscr{R}(\pi) \equiv \{i \in \mathscr{N} | I^i > 1/2 \text{ and } \theta^i > \theta^\nu\}$. $\mathscr{L}(\pi)$ and $\mathscr{R}(\pi)$ describe the right and left legislators that prefer the status quo to the median voter's platform. We denote a set of Pareto optimal policies of a coalition $C^j \in N$ as $O_{C^j} \in \mathbb{R}^m$.

***Proposition* 3.3.3.** *A stable legislative gridlock will exist in equilibrium for a preference profile* (π) *if* $\exists g \in N \backslash \nu$ *such that* $\exists \alpha \in O_S$ *where* $S \subset \mathscr{L}(\pi) \cup \mathscr{R}(\pi) \backslash g$ *and,*

1. $\sum_{i \in \mathscr{N}} \mathbf{1}.\Big\{I^i + E^i(\frac{E^a}{I^a}) > \frac{1}{2}\Big(\frac{(I^a)^2+(E^a)^2}{I^a}\Big)\Big\} > \frac{n}{2}$
2. $I^g + E^g\big(\frac{E^a}{I^a-1}\big) < \frac{1}{2}\Big[\frac{(I^a)^2+(E^a)^2-1}{I^a-1}\Big]$
3. $I^g < I^a$ *and* $\sum_{i \in \mathscr{N}} \mathbf{1}.\Big\{I^i + E^i\big(\frac{(E^a-E^g)}{I^a-I^g}\big) < \frac{1}{2}\Big[\frac{(I^a)^2+(E^a)^2-[(I^g)^2+(E^g)^2]}{I^a-I^g}\Big]\Big\} > \frac{n}{2}$

Proposition 3.3.3 describes a sufficient condition for a stable legislative gridlock. This gridlock is given by: $\langle\{\nu, g, S\}, \{x^\nu, x^g, a\}, x^0\rangle$. In other words, a coalition of left and right legislators, S, is created that proposes policy $a \in O_S$. Policy a is chosen so that the majority of legislators are sufficiently conservative and extremist with respect to a so they prefer a to the median's preferred policy x^ν, as given in condition 1. Agent g proposes her own ideal policy x^g to block a from winning against median policy x^ν. She does this because she prefers the implementation of the status quo to the implementation of policy a. This is given by condition 2 where x^g is not as conservative a policy as a. Note, if a and x^g are on the opposite side of the political spectrum, $E^g E^a < 0$, then g will be more extreme if she wants to implement x^0 instead of a. However, for the same to apply when x^g and a are on the same side of the political spectrum, $E^gE^a > 0$, g must be conservative. Finally, in order for a Condorcet cycle to emerge, it must be the case that a majority of the legislators prefer x^g to a. Therefore, as described by condition 3, majority of the legislators must have a bounded extremism and conservatism. This bound is defined by the extremism and conservatism of proposed policies a and x^g.

Conditions 1 and 3 together indicate: for gridlock to emerge it can be sufficient if majority of agents' preferences are neither too much nor too little conservative and extremist. Condition 1 also illustrates that not all those who vote for the more conservative policy a against x^v are those with extreme opinions. Those in the opposite side of the spectrum of a (with $E^iE^a < 0$) who vote for a must be very moderate with low E^i. Condition 3, on the other hand, indicates that there are at least some people in the same side of the spectrum of policy a, $E^iE^a > 0$, that are moderate or with low E^i. Therefore, sufficiency conditions may be met for a stable gridlock to occur if one side of the political spectrum are not extremists (low E^i) but fairly conservative and the other side of the political spectrum mostly has extremist legislators with exceptions. This indicates a type of skewness of legislators preferences with respect to the median legislator's preferred policy and the status quo to be sufficient for a gridlock to emerge.

3.4 Conclusion

This chapter represents a preliminary attempt of analyzing and characterizing the conditions that underpin a legislative gridlock.

The main novelty consists in identifying two main drivers of the legislators' ideological position that produce legislative gridlock, which we name conservatism and extremism. We show that the existence of a sufficiently high number of legislators that are both conservative and extremist is key to generate a legislative gridlock. Moreover, we show that a gridlock prevails when these conservative-extremist legislators are present both among left-wing and right-wing legislators, and a coalition is formed that includes conservatives from both ideological sides.

These results can help to shed light on the drivers of ineffective legislative barganing. Possible examples include voting on different forms of Brexit in the Parliament of the United Kingdom, and voting over economic and social reforms in the Italian Parliament.

References

Ager, P. and M. Brückner (2013). Cultural diversity and economic growth: Evidence from the us during the age of mass migration. *European Economic Review 64*, 76–97.

Alesina, A., R. Baqir, and W. Easterly (1999, November). Public goods and ethnic divisions. *The Quarterly Journal of Economics 114*, 1243–1284.

Alesina, A., R. Baqir, and W. Easterly (2000). Redistributive public employment. *Journal of Urban Economics 48*, 219–241.

Alesina, A., C. Gennaioli, and S. Lovo (2019). Public goods and ethnic diversity: Evidence from deforestation in indonesia. *Economica 86*(341), 32–66.

Alesina, A. and P. Giuliano (2015, December). Culture and institutions. *Journal of Economic Literature 53*(4), 898–944.

Alesina, A. and E. Glaeser (2004). *Fighting Poverty in the US and Europe: A World of Difference*. Oxford.

Alesina, A., E. Glaeser, and B. Sacerdote (2001, October). Why doesn't the us have a european-style welfare system? Working Paper 8524, National Bureau of Economic Research.

Alesina, A. and B. Reich (2015). Nation building. Working Paper 18839, National Bureau of Economic Research.

Amodio, F. and G. Chiovelli (2018). Ethnicity and violence during democratic transitions: Evidence from south africa. *Journal of the European Economic Association 16*(4), 1234–1280.

Anderson, L. (2011). Demystifying the arab spring: Parsing the differences between tunisia, egypt, and libya. *Foreign Affairs 90*(3), 2–7.

Austen-Smith, D. (1986). Legislative coalitions and electoral equilibrium. *Public Choice 50*, 185–210.

Austen-Smith, D. and J. S. Banks (1996). Information aggregation, rationality, and the condorcet jury theorem. *The American Political Science Review 90*(1), 34–45.

Azzimonti, M. (2011). Barriers to investment in polarized societies. *American Economic Review 101*(5), 2182–2204.

Banerjee, A. and R. Somanathan (2007). The political economy of public goods: Some evidence from india. *Journal of Development Economics 82*(2), 287–314.

Barbera, S., F. Gul, and E. Stacchetti (1993). Generalized median voter schemes and committees. *Journal of Economic Theory 61*(2), 262 – 289.

Barbera, S. and M. Jackson (2020, 01). A model of protests, revolution, and information. *Quarterly Journal of Political Science 15*, 297–335.

Baron, D. P. (1991). A spatial bargaining theory of government formation in parliamentary systems. *American Political Science Review 85*(1), 137–164.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in legislatures. *The American Political Science Review 83*(4), 1181–1206.

Battaglini, M. (2017). Public protests and policy making. *Quarterly Journal of Economics 132*(1), 485–549.

Battaglini, M. (2020, April). Coalition formation in legislative bargaining. *NBER Working Paper No. 25664*.

Bisin, A. and T. Verdier (2017). On the joint evolition of culture and institutions. *NBER working paper series*.

Calvert, R. L. and N. Dietz (2005). Legislative coalitions in a bargaining model with externalities. In D. Austen-Smith and J. Duggan (Eds.), *Social Choice and Strategic Decisions*. Studies in Choice and Welfare: Springer.

Caselli, F. and W. J. Coleman (2013). On the theory of ethnic conflict. *Journal of the European Economic Association 11*(suppl_1), 161–192.

Ceron, A. (2012). Bounded oligarchy: How and when factions constrain leaders in party position-taking. *Electoral Studies 31*, 689–701.

Cho, S.-j. and J. Duggan (2009). Bargaining foundations of the median voter theorem. *Journal of Economic Theory 144*(2), 851 – 868.

Dahlum, S. and T. Wig (2017, 13 September). Educating demonstrators: Education and mass protest in africa. *Journal of Conflict Resolution 63*.

Dai, L. and M. Yang (2019, 8 August). Organizations and coordination in a diverse population. *SSRN: https://ssrn.com/abstract=3051192*.

Demange, G. (2004). On group stability in hierarchies and networks. *Journal of Political Economy 112*(4), 754–778.

Diamantoudi, E. and L. Xue (2007). Coalitions, agreements and efficiency. *Journal of Economic Theory 136*, 105–125.

Dotti, V. (2020). Reaching across the aisle to block reforms. *Economic Theory*.

Duggan, J. and J. Gao (2020). Lobbying as a multidimensional tug of war. *Social Choice and Welfare 54*(1).

Easterly, W. and R. E. Levine (1997, November). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics 112*.

Eguia, J. X. (2011). Endogenous parties in an assembly. *American Journal of Political Science 55*(1), 16–26.

Eraslan, H. and K. S. Evdokimov (2019). Legislative and multilateral bargaining. *Annual Review of Economics 11*(1), 443–472.

Eraslan, H., K. S. Evdokimov, and J. Zapal (2020, June 10). Dynamic legislative bargaining. ISER DP No. 1090, Available at SSRN: https://ssrn.com/abstract=3634333 or http://dx.doi.org/10.2139/ssrn.3634333.

Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review 102*(4), 1310–42.

Esteban, J. and D. Ray (1999). Conflict and distribution. *Journal of Economic Theory 87*(2), 379 – 415.

Esteban, J. and D. Ray (2011a). Linking conflict to inequality and polarization. *American Economic Review 101*(4), 1345–74.

Esteban, J. and D. Ray (2011b). A model of ethnic conflict. *Journal of the European Economic Association 9*(3), 496–521.

Fernández, R. and G. Levy (2008, June). Diversity and redistribution. *Journal of Public Economics 92*, 925–943.

Finus, M. and M. McGinty (2019). The anti-paradox of cooperation: Diversity may pay! *Journal of Economic Behavior and Organization 157*, 541–559.

Gans, J. and M. Smart (1996). Majority voting with single-crossing preferences. *Journal of Public Economics 59*, 219–237.

Ghosh, S. and A. Mitra (2016). Ethnic Diversity, Public Spending and Political Regimes. MPRA Paper 75546, University Library of Munich, Germany.

Gisselquist, R. M. (2014). Ethnic divisions and public goods provision, revisited. *Ethnic and Racial Studies 37*(9), 1605–1627.

Gisselquist, R. M., S. Leiderer, and M. Niño-Zarazúa (2016). Ethnic heterogeneity and public goods provision in zambia: Evidence of a subnational diversity dividend. *World Development 78*, 308 – 323.

Gorodnienko, Y. and G. Roland (2015). Culture, institutions and democratization. Working Paper 21117, National Bureau of Economic Research.

Hodler, R. (2006). The curse of natural resources in fractionalized countries. *European Economic Review 50*(6), 1367–1386.

Holm, U. (2013). Libya in transition: the fragile and insecure relation between the local, the national and the regional. In L. R. Andersen (Ed.), *How the local matter: Democratization in Libya, Pakistan, Yemen and Palistine*. Danish Institute for International Studies Report.

Hopkins, D. J. (2009). The diversity discount: When increasing ethnic and racial diversity prevents tax increases. *The Journal of Politics 71*(1), 160–177.

Jackson, M. O. and B. Moselle (2002). Coalition and party formation in a legislative voting game. *Journal of Economic Theory 103*, 49–87.

Jung, S. (2015). Does education affect risk aversion? evidence from the british education reform. *Applied Economics 47*(28), 2924–2938.

Khan, M. and K. Mezran (2013). *The Libyan Economy After the Revolution: Still No Clear Vision*. Atlantic Council, Rafik Hariri Center for the Middle East.

Krehbiel, K. (1998). *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago, IL: University of Chicago Press.

Lane, P. and A. Tornell (1999). The voracity effect. *American Economic Review 89*(1), 22–46.

Lee, S., D. Lee, and T. E. Borcherding (2016). Ethnic diversity and public goods provision: Evidence from us municipalities and school districts. *Urban Affairs Review 52*(5), 685–713.

Levy, G. (2004). A model of political parties. *Journal of Economic Theory*, 250–277.

Levy, G. (2005). The politics of public provision of education. *The Quarterly Journal of Economics 120*(4), 1507–1534.

Miguel, E. and M. K. Gugerty (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of Public Economics 89*(11-12), 2325–2368.

Milgrom, Paul, S. C. (1994). Monotone comparative statics. *Econometrica 62*(1), 157–180.

Montalvo, J. G. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *The American Economic Review 95*(3), 796–816.

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

Novta, N. (2016). Ethnic diversity and the spread of civil war. *Journal of the European Economic Association 14*(5), 1074–1100.

Ordeshook, P. C. (1986). *The core*, pp. 339–386. Cambridge University Press.

Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Review 57*(4), 787–806.

Ray, D. and R. Vohra (1997). Equilibrium binding agreements. *Journal of Economic Theory 73*, 30.

Ray, D. and R. Vohra (2015). Coalition formation. *Handbook of Game Theory*.

Reynal-Querol, M. (2002). Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution 46*(1), 29–54.

Roemer, J. E. (1999). The democratic political economy of progressive taxation. *Econometrica 67*, 1–19.

Schofield, N. (2006). *Architects of Political Change*. Cambridge University Press.

Shepsle, K. A. and B. R. Weingast (1987). The institutional foundations of committee power. *The American Political Science Review 81*(1), 85–104.

Somanathan, R. (2018, October). The Measurement and Mismeasurement of Social Difference. Working papers 294, Centre for Development Economics, Delhi School of Economics.

Spolaore, E. and R. Wacziarg (2017, March). The political economy of heterogeneity and conflict. Working Paper 23278, National Bureau of Economic Research.

Stichnoth, H. and K. Van der Straeten (2013). Ethnic diversity, public spending, and individual support for the welfare state: A review of the empirical literature. *Journal of Economic Surveys 27*(2), 364–389.

Taylor, C. L. and M. C. W. Hudson (1972). *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Tufecki, Z. (2017). *Twitter and Tear Gas*. Yale University Press.

Woo, J. (2005). Social polarization, fiscal instability and growth. *European Economic Review 49*(6), 1451–1477.

Yi, S.-S. (1997). Stable coalition structures with externalities. *Games and Economic Behavior 20*(2), 201 – 237.

Young, P. H. (2015). The evolution of social norms. *Annual Review of Economics*, 359–87.

Appendix A

Unity in Diversity

A.1 Norms of Compromise

A.1.1 Existence and uniqueness of the compromise policy

Let $C \subset \mathscr{N}$, m be the number of agents in C, and A be the list of ideal policy positions of agents in C.

***Lemma* A.1.1.** *Let $\rho \in (0,1)$. Define $G^{\rho} : \mathbb{R}^m \to \mathbb{R}$,*

$$G^{\rho}(A,x) = \sum_{\alpha \in C} \big(d_{\alpha}(x)\big)^{\frac{1}{\rho}} \tag{A.1}$$

then $G^{\rho}(A,.)$ is strictly convex in x.

Proof. For $\rho \in (0,1)$, $G_{\rho}(A,.)$ is differentiable function of x.

$$D^1 G^{\rho}(A,x) = \frac{1}{\rho} \sum_{\alpha \in C} \big(d_{\alpha}(x)\big)^{\frac{1}{\rho}-1} \frac{\partial d_{\alpha}(x)}{\partial x} \tag{A.2}$$

Since $\frac{\partial^2 d_{\alpha}(x)}{\partial x^2}[d_{\alpha}^{1/\rho-1}] = 0$

$$D^2 G^{\rho}(A,x) = \left(\frac{1-\rho}{\rho^2}\right) \sum_{\alpha \in C_i} \big(d_{\alpha}(x)\big)^{\frac{1}{\rho}-2} \Big(\frac{\partial d_{\alpha}(x)}{\partial x}\Big)^2 > 0 \tag{A.3}$$

■

***Lemma* A.1.2.** *Fix* $A \in [0,1]^{|C|}$ *and let* $\rho \in (0,1)$. *Consider the optimization problem*

$$\min_x G^\rho(A,x) \tag{A.4}$$

Then policy proposal of $C \subset \mathcal{N}$, *generated by* $f^\rho : [0,1]^{|C|} \to \mathbb{R}$ *as defined in (1.1) is uniquely defined, always exists and solves (A.4). Specifically,*

$$D^1 G^\rho(A, f^\rho(A)) = 0 \tag{A.5}$$

Proof. Since objective function in (1.1) is a strictly increasing function in the value of G^ρ, if solution to (A.4) exists then solution to optimizing problem in (1.1) exists. Furthermore, any solution of (A.4) is a solution to (1.1).

Since $G^\rho(A^i,.)$ is a strictly convex differentiable function, if $x^* \in \mathscr{R}$ solves (A.4) then $D^1 G^\rho(A,x^*) = 0$.

To show x^* exists: Let $x < x_1$, then $D^1 G^\rho(A,x) < 0$ since $\frac{\partial d_\alpha(x)}{\partial x} < 0 \ \forall \ \alpha \in C$. Let $x > x_N$, then $D^1 G^\rho(A,x) > 0$ since $\frac{\partial d_\alpha(x)}{\partial x} > 0 \ \forall \ \alpha \in C$. Since at $\rho \in (0,1)$, $D^1 G^\rho(A,x)$ is continuously differentiable on $\mathbb{R}$, by intermediate value theorem $\exists \ x^* \in [x_1, x_N]$ such that $D^1 G^\rho(A,x^*) = 0$. As $G^\rho(A,x)$ is strictly convex in x, x^* is a unique solution to (A.4) and (1.1). This implies, f^ρ is well defined function and $D^1 G^\rho(A, f^\rho(A^i)) = 0$. ■

A.1.2 Examples of norms

Solving for f^ρ when $\rho \to 0$ Consider $d_\alpha(x) = \max_\beta d_\beta(x)$.

$$\Big(\sum_{\beta\in C} d_\beta(x)^{\frac{1}{\rho}}\Big)^\rho = d_\alpha(x)\Big(\sum_{\beta\in C}\frac{d_\beta(x)}{d_\alpha(x)}^{\frac{1}{\rho}}\Big)^\rho \tag{A.6}$$

$$\Big(\lim_{\rho\to 0}\sum_{\beta\in C} d_\beta(x)^{\frac{1}{\rho}}\Big)^\rho = d_\alpha(x)\Big(\sum_{\{\beta: d_\beta(x)=d_\alpha\}} 1\Big)^\rho \tag{A.7}$$

$$\implies \lim_{\rho\to 0} f^\rho(A) = \arg\min_{x\in\mathbb{R}}\max_{\beta\in C} d_\beta(x) \tag{A.8}$$

$$= \frac{\min_{\beta\in C} x_\beta + \max_{\beta\in C} x_\beta}{2} \tag{A.9}$$

Solving for f^ρ when $\rho \to 1$

$$G^1(A,x) = \sum_{\alpha\in C} d_\alpha(x)$$

G^1 is convex but may Guess that solution to A.4 is the median.

Case 1: Suppose N is even and we have two distinct medians, $x^* < x^{**}$. In this case, then $G^\rho(A,x)$ is differentiable at $x \in (x^*, x^**)$, then from our first order condition we get

$$\sum_{\alpha\in C: x_\alpha < x^*} 1 - \sum_{\alpha\in C: x_\alpha > x^{**}} 1 = 0 \tag{A.10}$$

Thus, any $x \in (x^*, x^{**})$ a solution to minimizing $G^1(A,x) = \sum_{\alpha\in C} d_\alpha(x)$.

Case 2: Suppose the median, x^*, is unique.

Let $I_> = \{\beta \in C : x_\beta > x^*\}$, $I_< = \{\beta \in C : x_\beta < x^*\}$

and $I_= = \{\beta \in C : x_\beta = x^*\}$.

For $x < x^*$, our first order condition is :

$$\frac{\partial G^1(A,x)}{\partial x} = |I_<| - |I_=| - |I_>| < 0 \tag{A.11}$$

For $x < x^*$, our first order condition is:

$$\frac{\partial G^1(A,x)}{\partial x} = |I_<| + |I_=| - |I_>| > 0 \tag{A.12}$$

Thus, x^* must be the minimizer of G^1.

Solving for f^ρ when $\rho = 1/2$

At $\rho = 1/2$ we have

$$D^1 G^\rho(A, f^\rho(A)) = - \sum_{\{\alpha \in C: x_\alpha > f^\rho(A)\}} (x_\alpha - f^\rho(A))^{2-1} + \sum_{\{\alpha \in C: x_\alpha < f^\rho(A)\}} \tag{A.13}$$

(A.14)

$$- \sum_{\{\alpha \in C: x_\alpha > f^\rho(A)\}} (x_\alpha - f^\rho(A)) + \sum_{\{\alpha \in C: x_\alpha < f^\rho(A)\}} (f^\rho(A) - x_\alpha) = 0 \tag{A.15}$$

$$m f^\rho(A) = \sum_{\alpha \in C} x_\alpha \tag{A.16}$$

$$f^\rho(A) = \frac{\sum_{\alpha \in C} x_\alpha}{m} \tag{A.17}$$

A.1.3 Properties of norms of compromise

Proof of Lemma 1.3.1

***Lemma* A.1.3.** *Let $\rho \in (0,1)$, and $\alpha \in C \subseteq \mathcal{N}$ with $x_\alpha = x$ then,*

$$\left.\frac{\partial f_\rho(A)}{\partial x_\alpha}\right|_{x_\alpha = x} = \left.\frac{\left(d_\alpha^\rho(A)\right)^{\frac{1-2\rho}{\rho}}}{\sum_{\beta \in C}\left(d_\beta^\rho(A)\right)^{\frac{1-2\rho}{\rho}}}\right|_{x_\alpha = x} \tag{A.18}$$

Proof. Let $x_\alpha > f^\rho(A)$

Differentiating both sides w.r.t x_α we get:

$$\left(\frac{1-\rho}{\rho}\right)\frac{\partial f^\rho(A)}{\partial x_\alpha}\sum_{\left\{x_\beta < f^\rho(A):\ \beta \in C_i\right\}}\left(d_\beta^\rho(A)\right)^{\frac{1}{\rho}-2} = \left(\frac{1-\rho}{\rho}\right)\left[\left(d_\alpha^\rho(A)\right)^{\frac{1}{\rho}-2} - \frac{\partial f^\rho(A)}{\partial x_\alpha}\sum_{\left\{x_\beta > f^\rho(A):\ \beta \in C_i\right\}}\left(d_\beta^\rho(A)\right)^{\frac{1}{\rho}-2}\right] \tag{A.19}$$

$$\implies \frac{\partial f^\rho(A)}{\partial x_\alpha} = \frac{\left(d_\alpha^\rho(A)\right)^{\frac{1}{\rho}-2}}{\sum_{\{\beta \in C_i\}}\left(d_\beta^\rho(A)\right)^{\frac{1}{\rho}-2}} \tag{A.20}$$

Similar proof for $x_\alpha < f^\rho(A)$. ∎

Proof related to example 1.3.1

Proof. Given m number of ideal points with x and $m-l$ number of of ideal points y in any coalition $C_i \subset \mathcal{N}$ from (A.1.2) we get:

$$l.(f^\rho(A)-x)^{\frac{1-\rho}{\rho}} = (m-l)(y-f^\rho(A))^{\frac{1-\rho}{\rho}} \tag{A.21}$$

$$f^\rho(A) = \frac{y\left(\frac{m-l}{l}\right)^{\frac{\rho}{1-\rho}} + x}{1+\left(\frac{m-l}{l}\right)^{\frac{\rho}{1-\rho}}} \tag{A.22}$$

So, for $l > m-l$, f^{ρ} is decreasing in ρ and for $l < m-l$, f^{ρ} is increasing ρ. Also, for $l > m-l$, $\lim_{\rho\to 1} f^{\rho}(A) = x$; and for $l < m-l$, $\lim_{\rho\to 1} f^{\rho}(A) = y$. $\lim_{\rho\to 0} f^{\rho}(A) = 1/2$. ■

A.2 Stability of the Grand Coalition

For a more general proof of stability we allow for a coalition $C_j \in \mathscr{N}$ to block the policy proposed by the grand coalition under norm of compromise ρ. We say C_j is a blocking coalition of $\mathscr{N}$ under ρ if and only if

$$p(m_i,\mathscr{P})u_{\alpha}(f^{\rho}(A_i)) + p(m_j,\mathscr{P})u_{\alpha}(f^{\rho}(A_j)) > u_{\alpha}(f^{\rho}(\mathscr{A})) \qquad \forall \alpha \in C_j$$

where $m_k = |C_k|$ and A_k is the vector of ideal point in coalition C_k for $k \in \{i, j\}$, and $C_i = \mathscr{N} \backslash C_j$.

Stricter Stability condition A grand coalition is strictly stable under ρ if there does not exists a blocking coalition $C \subset \mathscr{N}$ of $\mathscr{N}$ under ρ.

Note, if the grand coalition is strictly stable under ρ then it is also internally stable as defined in the main body of the paper.

For the rest of the proofs in the appendix we use this more general definition of stability to prove our results.

We denote $CE_{\alpha}^{\rho\theta}(\mathscr{P}) = (p(m_i,\mathscr{P})d_{\alpha}(f^{\rho}(A_i))^{\theta} + p(m_j,\mathscr{P})d_{\alpha}(f^{\rho}(A_j))^{\theta})^{\frac{1}{\theta}}$, where $CE_{\alpha}^{\rho\theta}(\mathscr{P})$ is the certainty equilavent distance of α from the gamble of policy proposed generated by the partition $\mathscr{P}$.

For the purposes of these proofs, cost of fragmentation is redefined for coalitional separation.

$$R_{\alpha}^{\rho,\theta}(\mathscr{P}) = CE_{\alpha}^{\rho\theta}(\mathscr{P}) - d_{\alpha}^{\rho}(\mathscr{N}) \tag{A.23}$$

The grand coalition is stable if and only if $R_{\alpha}^{\rho,\theta}(\mathscr{P}) > 0$.

A.2.1 Risk Aversion

***Lemma* A.2.1.** *Fix $\rho \in (0,1)$ and N, then $R_{\alpha}^{\rho\theta}$ increases with θ.*

Proof. Let p_i be the probability that $g = f^{\rho}(A_i)$ under norm ρ.

Since $(x \ln x)$ is a convex function in x and $\left(CE_{\alpha}^{\rho\theta}\right)^{\theta}$ is a convex combination of $\left\{\left(d_{\alpha}^{\rho}(A_i)\right)^{\theta}\right\}_{C_i \in \mathscr{P}}$, we have:

$$\frac{\partial CE_{\alpha}^{\rho\theta}(\mathscr{P})}{\partial \theta} = \frac{\sum_{C_i \in \mathscr{P}} p_i \left(d_{\alpha}^{\rho}(A_i)\right)^{\theta} \ln\left(d_{\alpha}^{\rho}(A_i)\right)^{\theta} - \left(CE_{\alpha}^{\rho\theta}(\mathscr{P})\right)^{\theta} \ln\left(CE_{\alpha}^{\rho\theta}(\mathscr{P})\right)^{\theta}}{\theta(CE_{\alpha}^{\rho\theta}(\mathscr{P}))^{\theta}} > 0 \tag{A.24}$$

By definition

$$R_{\alpha}^{\rho\theta}(\mathscr{P}) = CE_{\alpha}^{\rho\theta}(\mathscr{P}) - d_{\alpha}^{\rho}(A_{\mathscr{N}}) \tag{A.25}$$

$$\implies \frac{\partial R_{\alpha}^{\rho\theta}(\mathscr{P})}{\partial \theta} = \frac{\partial CE_{\alpha}^{\rho\theta}(\mathscr{P})}{\partial \theta} > 0 \tag{A.26}$$

■

Proof of Lemma 1.4.1

***Lemma* A.2.2.** *Given $\mathscr{A}$, let $\mathscr{P} \in \Pi(\mathscr{N})$ such that $\mathbb{E}_{\beta}^{\rho\theta}(\mathscr{P}) \neq \mathbb{E}_{\beta}(\mathscr{N}, \rho\theta)$ for $\beta \in \mathscr{A}$. Then $\exists C_i \in \mathscr{P}$ such that $d_{\beta}^{\rho}(A_i) > d_{\beta}^{\rho}(\mathscr{N})$.*

Proof. We know from (A.1.2)

$$\sum_{\{x\alpha < f^{\rho}(\mathscr{A}):\ \alpha \in \mathscr{N}\}} \left(d_{\alpha}^{\rho}(\mathscr{A})\right)^{\frac{1}{\rho}-1} = \sum_{\{x\alpha > f^{\rho}(\mathscr{A}):\ \alpha \in \mathscr{N}\}} \left(d_{\alpha}^{\rho}(\mathscr{A})\right)^{\frac{1}{\rho}-1} \tag{A.27}$$

Case 1: $x_\alpha < f^\rho(X_{\mathscr{N}})$

If $\mathbb{E}_\alpha^{\rho\theta}(\mathscr{P}) \neq \mathbb{E}_\alpha^{\rho\theta}(\mathscr{N})$ then $\exists C_i \in \mathscr{P}$

$$\sum_{\{x_\alpha < f^\rho(\mathscr{A}):\ \alpha \in C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1} > \sum_{\{x_\alpha > f^\rho(\mathscr{A}):\ \alpha \in C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1} \tag{A.28}$$

$$\implies f^\rho(A_i) > f^\rho(\mathscr{A}) \tag{A.29}$$

$$\implies d_\beta^\rho(A_i) > d_\beta^\rho(\mathscr{A}) \tag{A.30}$$

Case 2: $x_\beta > f^\rho(X_{\mathscr{N}})$

If $\mathbb{E}_\alpha^{\rho\theta}(\mathscr{P}) \neq \mathbb{E}_\alpha^{\rho\theta}(\mathscr{N})$ then $\exists C_i \in \mathscr{P}$

$$\sum_{\{x_\alpha < f^\rho(\mathscr{A}):\ \alpha \in C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1} < \sum_{\{x_\alpha > f^\rho(\mathscr{A}):\ \alpha \in C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1} \tag{A.31}$$

$$\implies f^\rho(A_i) < f^\rho(\mathscr{A}) \tag{A.32}$$

$$\implies d_\beta^\rho(A_i) > d_\beta^\rho(\mathscr{A}) \tag{A.33}$$

■

Lemma 1.4.1 Fix $\rho \in (0,1)$, then $\exists \bar{\theta}_\alpha^\rho \geq 1$ such that $\forall \theta > \bar{\theta}_\alpha^\rho$ and $\forall\ i \in \mathscr{N}$

$$R_{\rho\theta}^i(\mathscr{P}) > 0 \tag{A.34}$$

Proof. Let $C_j = \arg\max_{C_i \in \mathscr{P}} d^{\rho}_{\alpha}(A_i)$ and p_i denote the probability that $g = f^{\rho}(A_i)$.

$$CE^{\rho\theta}_{\alpha}(\mathscr{P}) = \left(\sum_{\alpha \in C_i} p_i \left(d^{\rho}_{\alpha}(A_i)\right)^{\theta}\right)^{\frac{1}{\theta}} \tag{A.35}$$

$$= d^{\rho}_{j}(A_j)(p_j)^{\frac{1}{\theta}} \left(1 + \sum_{\alpha \in C_i} \frac{p_i}{p_j} \left(\frac{d^{\rho}_{\alpha}(A_i)}{d^{\rho}_{\alpha}(A_j)}\right)^{\theta}\right)^{\frac{1}{\theta}} \tag{A.36}$$

Thus, $\lim_{\theta \to \infty} CE^{\rho\theta}_{\alpha}(\mathscr{P}) = d^{\rho}_{j}(A_j)$.
Since $CE^{\rho\theta}_{\alpha}(\mathscr{P})$ is strictly increasing in θ (from inequality (A.24)) and $d^{\rho}_{j}(A_j) > d^{\rho}_{j}(\mathscr{A})$ (from (A.2.2)), by the intermediate value theorem we know $\exists \bar{\theta}^{\rho}_{\alpha}(\mathscr{P}) \geq 1$ such that $CE^{\rho\theta}_{\alpha}(\mathscr{P}) = d^{\rho}_{j}(A_j) > d^{\rho}_{\alpha}(\mathscr{A})$. Since $\mathscr{N}$ is finite, $\theta^{\rho}_{\alpha} = \max_{\mathscr{P} \in \Pi(\mathscr{N})} \theta^{\rho}_{\alpha}$ is well defined. ■

Proof of Proposition 1.4.1

Proof. Using lemma 1.4.1 we know for some $\forall \alpha \in \mathscr{N}$ $\exists \theta^{\rho}_{\alpha} \geq 1$ such that $\forall \theta > \theta^{\rho}_{\alpha}$, $R^{\rho\theta}_{\alpha} > 0$. Since $\mathscr{N}$ is finite, $\theta^{\rho} = \max_{\alpha \in X_{\mathscr{N}}} \theta^{\rho}_{\alpha}$ is well defined. ■

A.2.2 Polarization and Norms

Proof related to example 1.4.2

Proof. Consider cost of fragmentation for the members of the possible blocking coalitions that induces partition $\mathscr{P}$:

Case 1: $C_i \in \mathscr{P}$ such that $\forall \alpha \in C_i, x_{\alpha} = 1$:

$$R^{\rho\theta}_{N}(\mathscr{P}) = CE^{\rho\theta}_{N}(\mathscr{P}) - d^{\rho\theta}_{N}(\mathscr{A}) \tag{A.37}$$

$$= \frac{N - m_i}{N} d_N \left(\left[\frac{1}{1 + \left(\frac{N/2}{N/2 - m_i}\right)^{\frac{\rho}{1-\rho}}}\right]\right)^{\theta} - \left(\frac{1}{2}\right)^{\theta} \tag{A.38}$$

We know from proof (A.1.3) as $\exists\ \underline{\rho} > 1/2$ such that $\forall \rho > \underline{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) < 0$;

and $\exists\ \bar{\rho} < 1/2$ such that $\forall \rho < \bar{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) < 0$.

Case 2: $C_i \in \mathscr{P}$ such that $\forall \alpha \in C_i$, $x_\alpha = 0$:

$$R_1^{\rho\theta}(\mathscr{P}) = CE_1^{\rho\theta}(\mathscr{P}) - d_1^{\rho\theta}(\mathscr{A}) \tag{A.39}$$

$$= \frac{N-m_i}{N} d_1 \left(\left[\frac{1}{1+\left(\frac{N/2-m_i}{N/2}\right)^{\frac{\rho}{1-\rho}}} \right] \right)^{\theta} - \left(\frac{1}{2}\right)^{\theta} \tag{A.40}$$

We know from proof (A.1.3) as $\exists\ \bar{\rho} > 1/2$ such that $\forall \rho > \underline{\rho}$, $R_1^{\rho\theta}(\mathscr{P}) < 0$;

and $\exists\ \underline{\rho} < 1/2$ such that $\forall \rho < \bar{\rho}$, $R_1^{\rho\theta}(\mathscr{P}) > 0$.

∎

Proof related to example 1.4.1

Consider cost of fragmentation for the members of the possible blocking coalitions that induces partition $\mathscr{P}$:

Case 1: $C_i \in \mathscr{P}$ such that $\forall \alpha \in C_i$, $x_\alpha = 1$:

$$R_1^{\rho\theta}(\mathscr{P}) = CE_N^{\rho\theta}(\mathscr{P}) - d_1^{\rho\theta}(\mathscr{A}) \tag{A.41}$$

$$= \frac{N-1}{N} d_1 \left(\left[\frac{1/2+\left(\frac{1}{N-2}\right)^{\frac{\rho}{1-\rho}}}{1+\left(\frac{1}{N-2}\right)^{\frac{\rho}{1-\rho}}} \right] \right)^{\theta} - \left(\frac{1}{2}\right)^{\theta} \tag{A.42}$$

We know from proof (A.1.3) as $\exists\ \underline{\rho} > 1/2$ such that $\forall \rho > \underline{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) < 0$;

and $\exists\ \bar{\rho} < 1/2$ such that $\forall \rho < \bar{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) > 0$.

Case 2: $C_i \in \mathscr{P}$ such that $\forall \alpha \in C_i$, $x_\alpha = 1$:

$$R_N^{\rho\theta}(\mathscr{P}) = CE_N^{\rho\theta}(\mathscr{P}) - d_N^{\rho\theta}(\mathscr{A}) \tag{A.43}$$

$$= \frac{N-1}{N} d_N \left(\left[\frac{0.5}{1 + \left(\frac{1}{N-2}\right)^{\frac{\rho}{1-\rho}}} \right] \right)^{\theta} - \left(\frac{1}{2}\right)^{\theta} \tag{A.44}$$

We know from proof (A.1.3) as $\exists\, \underline{\rho} < 1/2$ such that $\forall \rho > \underline{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) < 0$;
and $\exists\, \bar{\rho} > 1/2$ such that $\forall \rho < \bar{\rho}$, $R_N^{\rho\theta}(\mathscr{P}) > 0$.

Proof of Proposition 1.4.2

This proof uses the weaker definition of stability where we consider deviations by single agents. That is cost of fragmentation considered is

$$R_\beta^{\rho\theta}(\mathscr{A}) = p^{1/\theta} d_\beta^{\rho}(\mathscr{A}_{-\alpha}) - d_\beta^{\rho}(\mathscr{A})$$

Proof.

$$R_\beta^{\rho\theta}(\mathscr{A}) = p^{1/\theta} d_\beta^\rho(\mathscr{A}_{-\alpha}) - d_\beta^\rho(\mathscr{A}) \tag{A.45}$$

Since β is not getting polarized itself (A.46)

$$\implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial x_\alpha} = p^{1/\theta} \frac{\partial d_\beta^\rho(\mathscr{A}_{-\beta})}{\partial f^\rho(\mathscr{A}_{-\beta})} \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_\alpha} - \frac{\partial d_\beta^\rho(\mathscr{A})}{\partial f^\rho(\mathscr{A})} \frac{\partial f^\rho(\mathscr{A})}{\partial x_\alpha} \tag{A.47}$$

$$v.\nabla R_\beta^{\rho\theta}(\mathscr{A}) = \sum_{\gamma \in \mathscr{N}} v_\gamma \left(p^{1/\theta} \frac{\partial d_\beta^\rho(\mathscr{A}_{-\beta})}{\partial f^\rho(\mathscr{A}_{-\beta})} \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_\gamma} - \frac{\partial d_\beta^\rho(\mathscr{A})}{\partial f^\rho(\mathscr{A})} \frac{\partial f^\rho(\mathscr{A})}{\partial x_\gamma} \right) \tag{A.48}$$

$$= p^{1/\theta} \frac{\partial d_\beta^\rho(\mathscr{A}_{-\beta})}{\partial f^\rho(\mathscr{A}_{-\beta})} \sum_{\gamma \in \mathscr{N}} v_\gamma \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_\gamma} - \frac{\partial d_\beta^\rho(\mathscr{A})}{\partial f^\rho(\mathscr{A})} \sum_{\gamma \in \mathscr{N}} v_\gamma \frac{\partial f^\rho(\mathscr{A})}{\partial x_\gamma} \tag{A.49}$$

Since v is symmetric and $\mathscr{A}$ is symmetric (A.50)

$$\sum_{\gamma \in \mathscr{N}} v_\gamma \frac{\partial f^\rho(\mathscr{A})}{\partial x_\gamma} = 0 \tag{A.51}$$

Since v is symmetric, denote the set of polarized individuals as π

and denote $pol-, pol+ \in \pi$, where $v_{pol+} = -v_{pol-} > 0$

$$\sum_{\gamma \in \mathscr{N}} v_\gamma \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_\gamma} = \sum_{pol+ \in \pi} v_\gamma \left(\frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_{pol+}} - \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial x_{pol-}} \right) \tag{A.52}$$

$$= \sum_{pol+ \in \pi} v_\gamma \left(\frac{(d_{pol+}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} - (d_{pol-}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\gamma \in \mathscr{N}_{-\alpha}} (d_\gamma(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}} \right) \tag{A.53}$$

Case 1: $x_\beta > f^\rho(\mathscr{A}) = 1/2$

$$\frac{\partial d_\beta^\rho(\mathscr{A}_{-\beta})}{\partial x_{pol+}} = \frac{\partial d_\beta^\rho(\mathscr{A}_{-\beta})}{\partial x_{pol-}} < 0$$

and

$$d_{pol+}^\rho(\mathscr{A}_{-\beta}) > d_{pol-}^\rho(\mathscr{A}) \qquad \forall\ pol+ \in \pi \text{ and } \forall\ \rho \in (0,1) \tag{A.54}$$

Then using expression A.53, If $\rho > 1/2$ we get $v.\nabla R_{\beta}^{\rho,\theta} > 0$

Then using expression A.53, if $\rho < 1/2$ we get $v.\nabla R_{\beta}^{\rho,\theta} < 0$

Case 2: $x_{\beta} < f^{\rho}(\mathscr{A}) = 1/2$

$$\frac{\partial d_{\beta}^{\rho}(\mathscr{A}_{-\beta})}{\partial x_{pol+}} = \frac{\partial d_{\beta}^{\rho}(\mathscr{A}_{-\beta})}{\partial x_{pol-}} > 0$$

$$d_{pol+}^{\rho}(\mathscr{A}_{-\beta}) < d_{pol-}^{\rho}(\mathscr{A}) \qquad \forall\ pol+ \in \pi \text{ and } \forall\ \rho \in (0,1) \tag{A.55}$$

Then using expression A.53, If $\rho > 1/2$ we get $v.\nabla R_{\beta}^{\rho,\theta} > 0$

Then using expression A.53, if $\rho < 1/2$ we get $v.\nabla R_{\beta}^{\rho,\theta} < 0$ ■

Proof relating to example 1.4.4

Proof. We know from lemma 1.4.2 for **extremists**:

If $\rho < 1/2$ $\frac{\partial R_{\alpha}^{\rho\theta}}{\partial y}(\mathscr{A}) < 0$ and

If $\rho > 1/2$ $\frac{\partial R_{\alpha}^{\rho\theta}}{\partial y}(\mathscr{A}) > 0$. From example(1.4.2) we know that for maximum polarization $(y = 1)$ $\exists\bar{\rho}$ such that $\forall\rho < \bar{\rho}$ $R_{\alpha}^{\rho\theta} < 0$.

Also from example (1.4.1) we know that for minimum polarization $(y = 1/2)$ $\exists\bar{\bar{\rho}}$ such that $\forall\rho < \bar{\bar{\rho}}$, $R_{\alpha}^{\rho\theta} > 0$.

Since R is a continuously decreasing function in y when $\rho < 1/2$, $\exists\ y^{**}$ such that $\forall y < y^{**}$ extremist would prefer to stay in grand coalition.

Similarly,

From example(1.4.2) we know that for maximum polarization $(y = 1)$ $\exists\bar{\rho}$ such that $\forall\rho > \bar{\rho}$ $R_{\alpha}^{\rho\theta} > 0$.

Also from example (**??**) we know that for minimum polarization $(y = 1/2)$, $\exists\bar{\bar{\rho}}$ such that $\forall\rho > \bar{\bar{\rho}}$, $R_{\alpha}^{\rho\theta} < 0$.

Since R is a continuously increasing function in y when $\rho > 1/2$, $\exists\ y^{*}$ such that $\forall y > y^{*}$ would prefer to stay in the grand coalition.

For **moderates**:

Let r be the number of polarized individuals, then using result in lemma 1.3.1

Case 1: For $x_\beta = 1 - y$ we have,

$$\frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial y} = \frac{r(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} - (r-1)(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum_{\gamma \in \mathscr{N}_{-\beta}} (d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}} \tag{A.56}$$

Case 1.1 For $\rho > 1/2$ we have,

$$R_\beta^{\rho\theta}(\mathscr{A}) = p^{1/\theta} d(\mathscr{A}_{-\beta}) - d(\mathscr{A}) \tag{A.57}$$

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} = p^{1/\theta}\left(\frac{\partial f^\rho}{\partial y} + 1\right) - (0+1) \tag{A.58}$$

$$> \frac{N-1}{N}\left(\frac{\partial f^\rho}{\partial y}\right) - \frac{1}{N} > 0 \qquad \text{if } \rho > 1/2 \tag{A.59}$$

So β would not want to leave grand coalition for $\rho > 1/2$ as polarization would increase costs.

Case 1.2 For $\rho < 1/2$:

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} = p^{1/\theta}\left(\frac{r(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} - (r-1)(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\gamma\in\mathscr{N}_{-\beta}}(d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}} + 1\right) - 1 \quad \text{(A.60)}$$

$$= p^{1/\theta}\left(\frac{2r(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_1^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_N^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\gamma\in\mathscr{N}_{-\beta}}(d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}\right) - 1 \quad \text{(A.61)}$$

$$\text{Since } \lim_{\rho\to 0}\left(\frac{2r(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_1^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_N^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\gamma\in\mathscr{N}_{-\beta}}(d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}\right) \to 1 \qquad \text{monotonically} \quad \text{(A.62)}$$

$$\exists\, \bar{\rho} \text{ such that } \forall \rho < r\bar{h}o \quad \text{(A.63)}$$

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} < 0 \quad \text{(A.64)}$$

Therefore, fixing some $\rho < \bar{\rho}$, given examples (1.4.1) and (1.4.2) there must be some y^*_{mod} such that $\forall y < y^*_{mod}$ right leaning moderates will want to leave.

Case 2: For $x_\beta = y$ we have,

$$\frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial y} = \frac{(r-1)(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} - r(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum\limits_{\gamma\in\mathscr{N}_{-\beta}}(d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}} \quad \text{(A.65)}$$

Case 2.1 For $\rho > 1/2$ we have:

$$R_\beta^{\rho\theta}(\mathscr{A}) = p^{1/\theta} d(\mathscr{A}_{-\beta}) - d(\mathscr{A}) \tag{A.66}$$

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} = p^{1/\theta}\left(1 - \frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial y}\right) - (1+0) \tag{A.67}$$

$$> \frac{N-1}{N}\left(-\frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial y}\right) - \frac{1}{N} \tag{A.68}$$

$$\text{if } \rho > 1/2, \text{ then} \qquad \left(-\frac{\partial f^\rho(\mathscr{A}_{-\beta})}{\partial y}\right) > 1/(N-1) \tag{A.69}$$

$$\implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} > 0 \tag{A.70}$$

So β would not want to leave grand coalition for $\rho > 1/2$ as polarization would increase costs.

Case 2.2 For $\rho < 1/2$:

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} = p^{1/\theta}\left(\frac{r(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} - (r-1)(d_y^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum_{\gamma\in\mathscr{N}_{-\beta}} (d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}} + 1\right) - 1 \tag{A.71}$$

$$= p^{1/\theta}\left(\frac{2r(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_1^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_N^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum_{\gamma\in\mathscr{N}_{-\beta}} (d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}\right) - 1 \tag{A.72}$$

$$\text{Since } \lim_{\rho\to 0}\left(\frac{2r(d_{1-y}^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_1^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}} + (d_N^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}{\sum_{\gamma\in\mathscr{N}_{-\beta}} (d_\gamma^\rho(\mathscr{A}_{-\beta}))^{\frac{1-2\rho}{\rho}}}\right) \to 1 \quad \text{monotonically} \tag{A.73}$$

$$\exists\, \bar{\rho} \text{ such that } \forall \rho < \bar{\rho} \tag{A.74}$$

$$\frac{\partial R_\beta^{\rho\theta}(\mathscr{A})}{\partial y} < 0 \tag{A.75}$$

Therefore, fixing some $\rho < \bar{\rho}$, given examples (1.4.1) and (1.4.2) there must be some y^*_{mod} such that $\forall y < y^*_{mod}$ right leaning moderates will want to leave. It would in fact be the same cut off point as the left leaning moderates as the distribution of preferences is symmetric. ■

Proof of Proposition 1.4.3

This proof is split in two parts. The first proof only considers internal stability. The second proof considers blocking coalitions

Proof for internal stability of grand coalition. Consider a vector of preferences given by $\mathscr{A}$ of the set of $\mathscr{N}$ agents. Let $\rho = 1/2$. Then by (A.4) we know for any $C \subset \mathscr{N}$, $f^{\rho}(\mathscr{A}) = \frac{1}{N}\sum_{\beta \in C} x_\beta$.
We also know that $\mathscr{N}$ will be stable if and only if:

$$r^{\rho\theta}_\alpha(\mathscr{A}) = p\left(\frac{d^\rho_\alpha(\mathscr{A}_{-\alpha})}{d^\rho_\alpha(\mathscr{A})}\right)^\theta \geq 1 \tag{A.76}$$

$$\implies r^{\rho,\theta}_\alpha(\mathscr{A}) = p\left(\frac{x_\alpha - \frac{1}{N-1}\sum_{\beta \in C\setminus\alpha} x_\beta}{x_\alpha - \frac{1}{N}\sum_{\beta\in C} x_\beta}\right)^\theta \tag{A.77}$$

$$r^{\rho,\theta}_\alpha(\mathscr{A}) = \frac{p}{((N-1)/N)^\theta} \geq 1 \qquad \text{if } p \geq (N-1)/N \text{ and } \theta \geq 1 \tag{A.78}$$

■

Proof for stability of grand coalition from blocking. Consider any partition, $\mathscr{P}$, of $\mathscr{N}$ with $p(m_i, \mathscr{P}) = \frac{m_i}{N}$, then:

$$\sum_{C_i \in \mathscr{P}} \frac{m_i}{N} f^{\frac{1}{2}}(A_i) = \sum_{C_i \in \mathscr{P}} \frac{m_i}{N} \frac{\sum_{\alpha \in C_i} x_\alpha}{m_i} = f^{\frac{1}{2}}(\mathscr{N})$$

By concavity of u_α,

$$u_\alpha(f^{\frac{1}{2}}(\mathscr{N})) = u_\alpha\left(\sum_{C_i \in \mathscr{P}} (m_i/N) f^{\frac{1}{2}}(C_i)\right) \geq \sum_{C_i \in \mathscr{P}} \frac{m_i}{N} u_\alpha(f^{\frac{1}{2}}(C_i)) = \mathbb{E}^{1/2\theta}_\alpha$$

■

A.2.3 Effect of Moderate preferences

Proof of Proposition 1.4.4

Proof. If $\rho > 1/2$ then $\frac{1}{\rho} - 2 < 0$. Therefore, if $\alpha \in A_i$ such that $\alpha = f^\rho(X_\mathscr{N})$, then for $\beta \neq \alpha$:

$$\frac{\partial f^\rho(A_i)}{\partial x_\alpha} = 1 \qquad \text{(Using lemma ??)} \tag{A.79}$$

$$\implies \frac{\partial R_\beta^{\rho\theta}}{\partial x_\alpha} = \frac{\partial CE_\beta^{\rho\theta}}{\partial x_\alpha} - \frac{\partial d_\alpha^\rho(\mathscr{A})}{\partial f^\rho(\mathscr{A})} \tag{A.80}$$

$$= p_i \left[\frac{d_\beta^\rho(A_i)}{CE_\beta^{\rho\theta}(\mathscr{P})}\right]^{\theta-1} \left[\frac{\partial d_\beta^\rho(A_i)}{\partial f^\rho(A_i)}\right] \left(\frac{\partial f^\rho(A_i)}{\partial x_\alpha}\right) - \frac{\partial d_\alpha^\rho(A_i)}{\partial f^\rho(\mathscr{A})} \tag{A.81}$$

$$= \tag{A.82}$$

If $d_\beta^\rho(A_i) \leq CE_\beta^{\rho\theta}$ then $\left[\frac{d_\beta^\rho(A_i)}{CE_\beta^{\rho\theta}(\mathscr{P})}\right]^{\theta-1} \in [0,1] \implies$

For $\beta = 0$: $\frac{\partial d_\beta^\rho(A_i)}{\partial f^\rho(A_i)}, \frac{\partial d_\alpha^\rho(A_i)}{\partial f^\rho(A_\mathscr{N})} > 0, \implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{P})}{\partial x_\alpha} < 0$.

For $\beta = 1$: $\frac{\partial d_\beta^\rho(A_i)}{\partial f^\rho(A_i)}, \frac{\partial d_\alpha^\rho(A_i)}{\partial f^\rho(A_\mathscr{N})} > 0 \implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{P})}{\partial x_\alpha} > 0$.

If $d_\beta^\rho(A_i) \geq CE_\beta^{\rho\theta}$ then

$$p_i \left[\frac{d_\beta^\rho(A_i)}{CE_\beta^{\rho\theta}}\right]^{\theta-1} = p_i \left[\frac{d_\beta^\rho(A_i)}{d_\beta^\rho(A_i)\left(p_i + \sum_{C_j \in \mathscr{P}} p_j \left(\frac{d_\beta^\rho(A_j)}{d_\beta^\rho(A_i)}\right)^\theta\right)^{\frac{1}{\theta}}}\right]^{\theta-1} \tag{A.83}$$

$$= \frac{p_i\left(p_i + \sum_{C_j \in \mathscr{P}} p_j \left(\frac{d_\beta^\rho(A_j)}{d_\beta^\rho(A_i)}\right)\right)}{p_i + \left(\sum_{C_j \in \mathscr{P}} p_j \left(\frac{d_\beta^\rho(A_j)}{d_\beta^\rho(A_i)}\right)\right)^\theta} \tag{A.84}$$

$$= \frac{p_i}{p_i + \left(\sum_{C_j \in \mathscr{P}} p_j \left(\frac{d_\beta^\rho(A_j)}{d_\beta^\rho(A_i)}\right)\right)^\theta} \left(\frac{\left(\sum_{C_j \in \mathscr{P}} p_j \left(d_\beta^\rho(A_j)\right)^\theta\right)^{\frac{1}{\theta}}}{d_\alpha^\rho(A_i)}\right) \tag{A.85}$$

$$= \frac{p_i}{p_i + \left(\sum_{C_j \in \mathscr{P}} p_j \left(\frac{d_\beta^\rho(A_j)}{d_\beta^\rho(A_i)}\right)\right)^\theta} \left(\frac{CE_\alpha^{\rho\theta}(\mathscr{P})}{d_\alpha^\rho(A_i)}\right) < 1 \tag{A.86}$$

From lemma (1.3.1) we have $\frac{\partial f^\rho(A_i)}{\partial x_\alpha} < 1 \implies p_i \left[\frac{d_\beta^\rho(A_i)}{CE_\beta^{\rho\theta}(\mathscr{P})}\right]^{\theta-1} \left(\frac{\partial f^\rho(X_\mathscr{N})}{\partial \alpha}\right) \in [0,1)$. Thus,

For $\beta = 0$: $\frac{\partial d_\beta^\rho(A_i)}{\partial f^\rho(A_i)}, \frac{\partial d_\alpha^\rho(A_i)}{\partial f^\rho(A_\mathscr{N})} > 0, \implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{P})}{\partial x_\alpha} < 0.$

For $\beta = 1$: $\frac{\partial d_\beta^\rho(A_i)}{\partial f^\rho(A_i)}, \frac{\partial d_\alpha^\rho(A_i)}{\partial f^\rho(A_\mathscr{N})} > 0 \implies \frac{\partial R_\beta^{\rho\theta}(\mathscr{P})}{\partial x_\alpha} > 0.$ ■

Proof Proposition (1.4.5)

Proof. If $\rho < 1/2$ then $\frac{1}{\rho} - 2 > 0$. Therefore, if $\alpha \in C_i$ such that $x_\alpha = f^\rho(X_{\mathscr{N}})$, then for $\beta \neq \alpha$:

$$\frac{\partial f^\rho(A_i)}{\partial x_\alpha} = 0 \qquad \text{(Using lemma 1.3.1)} \tag{A.87}$$

$$\implies \frac{\partial R^{\rho\theta}_\beta}{\partial x_\alpha} = \frac{\partial CE^{\rho\theta}_\beta}{\partial x_\alpha} \tag{A.88}$$

$$= p_i \left[\frac{d^\rho_\beta(A_i)}{CE^{\rho\theta}_\beta}\right]^{\theta-1} \left[\frac{\partial d^\rho_\beta(A_i)}{\partial f^\rho(A_i)}\right] \left(\frac{\partial f^\rho(A_i)}{\partial x_\alpha}\right) \tag{A.89}$$

If $\beta > f^\rho(A_i)$ then $\left[\frac{\partial d^\rho_\beta(A_i)}{\partial f^\rho(A_i)}\right] = -1$ which implies $\frac{\partial R^{\rho\theta}_\beta}{\partial \alpha} < 0$.
If $\beta < f^\rho(A_i)$ then $\left[\frac{\partial d^\rho_\beta(A_i)}{\partial f^\rho(A_i)}\right] = 1$ which implies $\frac{\partial R^{\rho\theta}_\beta}{\partial \alpha} > 0$. ∎

Expression of Influence function in terms of norm of compromise and distribution of preference

Consider a coalition $C_i \subset \mathscr{A}$. Define the Influence function: $I^\rho(\mathscr{A}, \mathscr{A}_{-i}) = |f^\rho(\mathscr{A}) - f^\rho(\mathscr{A}_{-i})|$ where $\mathscr{A}_{-i}$ is the vector of ideal points of $\mathscr{N} \setminus C_i$. This section shows how the influence function, a function that captures the risk of leaving the grand coalition, depends on ρ.

Consider now an arbitrary coalition $C_j \mathscr{N}$ and the distribution of preferences, A_j. From the optimization problem that defines norms in 1.1.

$$\sum_{\{x\alpha < f^\rho(A_j):\ \alpha \in C_j\}} \left(d^\rho_\alpha(A_j)\right)^{\frac{1}{\rho}-1} - \sum_{\{x\alpha > f^\rho(A_j):\ \alpha \in C_j\}} \left(d^\rho_\alpha(A_j)\right)^{\frac{1}{\rho}-1} = 0 \tag{A.90}$$

Define

$$H(A_j, z) \equiv \sum_{\{x_\alpha < z:\ \alpha \in C_i\}} \left(d_\alpha(z)\right)^{\frac{1}{\rho}-1} - \sum_{\{x_\alpha > z:\ \alpha \in C_i\}} \left(d_\alpha(z)\right)^{\frac{1}{\rho}-1} \tag{A.91}$$

Thus,

$$H_z(A_j, z) = \left(\frac{1-\rho}{\rho}\right) \sum_{\{\alpha \in C_i\}} \left(d_\alpha(z)\right)^{\frac{1}{\rho}-2} \tag{A.92}$$

Since $H(A_j, .)$ is differentiable in z we can use Taylor's approximation.

$$H(A_j, z_0 + h) = H(A_j, z_0) + hH_z(A_j, z_0 + \lambda h) \tag{A.93}$$

$$\implies h = \frac{H(A_j, z_0 + h) - H(A_j, z_0)}{H_z(A_j, z_0 + \lambda h)} \tag{A.94}$$

$$\implies |h| \approx \left| \frac{H(A_j, z_0 + h) - H(A_j, z_0)}{H_z(A_j, z_0)} \right| \tag{A.95}$$

$$\text{Let } z_0 = f(\mathscr{A}_{-i}) \text{ with } h = f^\rho(\mathscr{A}) - f^\rho(\mathscr{A}_{-i}) \tag{A.96}$$

$$\text{and } A_j = \mathscr{A}_{-i} \text{ where } \mathscr{A}_{-i} \text{ is the distribution of ideal points of } \mathscr{N} \setminus C_i \tag{A.97}$$

$$|f(\mathscr{A}) - f(\mathscr{A}_{-i})| = \left| \frac{H(\mathscr{A}_{-i}, f^\rho(\mathscr{N}))}{H_z(\mathscr{A}_{-i}, f^\rho(\mathscr{N} \setminus C_i))} \right| \tag{A.98}$$

(A.99)

$$\implies I^\rho(\mathscr{A}, \mathscr{A}_{-i}) \approx \frac{\rho}{1-\rho} \left| \frac{\sum_{\{x_\alpha < f^\rho(\mathscr{A}):\ \alpha \in \mathscr{N} \setminus C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1} - \sum_{\{x_\alpha > f^\rho(\mathscr{A}):\ \alpha \in \mathscr{N} \setminus C_i\}} \left(d_\alpha^\rho(\mathscr{A})\right)^{\frac{1}{\rho}-1}}{\sum_{\{\alpha \in \mathscr{N} \setminus C_i\}} \left(d_\alpha^\rho(\mathscr{A}_{-i})\right)^{\frac{1}{\rho}-2}} \right| \tag{A.100}$$

where $d_\alpha^\rho(A) \equiv d_\alpha(f^\rho(A)) \equiv |x_\alpha - f^\rho(A)|$ for an arbitrary distribution of preferences of agents within a coalition given by A.

Appendix B

Diversity Taxes

B.1 Tables

Table B.1: Variable Description
(Observations are for 1990 unless otherwise noted)

	Ethnicity
Ethnic Fractionalization	Measures the probability that two persons drawn randomly from the population belong to different self-identified ethnic groups; ranges from 0 to 1; ranges from 0 (complete homogeneity) to 1 (complete heterogeneity)
Ethnic Polarization	Captures how far the distribution of the ethnic groups is from the $(1/2, 0, 0, ...0, 1/2)$ distribution (bipolar), which represents the highest level of polarization.
Fractionalization of Minority	Measures the ethnic fractionalization of the population excluding the majority group
Size of Majority Group	Majority group size as a fraction of the population
Size of Biggest Minority Group	Size of second biggest group as a fraction of the population
Black	Fraction of population self-identifying as Black
White	Fraction of population self-identifying as White
Asian	Fraction of population self-identifying as Asian or Pacific Islander
American Indian	Fraction of population self-identifying as American Indian, Eskimo, or Aleut
Other Race	Fraction of population self-identifying as not Black, American Indian, Asian, or White; proxy for Hispanic
	Government
Expenditures per Capita	General local government expenditure per capita, 1990-1991
Health Spending	Fraction of general local government expenditure for health and hospitals (county and metro only)
Education Spending	Fraction of general local government expenditure for education (county and metro only)
Police Spending	Fraction of general local government expenditure for police
Welfare Spending	Fraction of general local government expenditure for public welfare (county and metro only)
Taxes per Capita	Total local government taxes per capita, 1990-1991
Debt per Capita	Per capita local government debt outstanding (county and metro only)
	Income, Education, and Population
Population Size	Log of population size
Percentage BA Graduate	Persons 25 years and over, fraction with Bachelor's degree or higher
Population above 65	Fraction of population that is 65 years or older
Violence per Capita	Violent crimes per capita (murder, forcible rape, robbery, aggravated assault)
Income per Capita	Per capita money income, 1989
Median household income	Median household income, 1989
Mean-to-Median Income	Ratio of mean to median household income

Table B.2: Summary Statistics for City Data (subset of cities with majority)

Variable Name	Mean	St. Dev.	Min	Max	N	Unit
Fractionalization Index	0.283	0.169	0.014	0.737	1,045	Fraction
Polarization Index	0.489	0.271	0.029	0.991	1,045	Fraction
Fractionalization of Minority	0.451	0.205	0.010	0.743	1,045	Fraction
Size of Majority Group	0.818	0.132	0.503	0.993	1,045	Fraction
Size of Biggest Minority Group	0.133	0.117	0.003	0.475	1,045	Fraction
Black	0.112	0.153	0.0004	0.981	1,045	Fraction
White	0.801	0.170	0.016	0.993	1,045	Fraction
Asian	0.037	0.072	0.0003	0.838	1,045	Fraction
American Indian, Eskimo, Aleut	0.006	0.011	0.0003	0.138	1,045	Fraction
Other	0.044	0.074	0.0004	0.669	1,04	Fraction
Expenditures per Capita	872.883	555.759	161.000	7,154.000	991	$ per capita
Taxes per Capita	371.661	276.314	38.487	3,977.627	991	$ per capita
Population Size	10.961	0.759	10.127	15.806	1,045	Log of # people
Percentage BA Graduate	0.230	0.118	0.016	0.712	1,045	Fraction
Population above 65	0.126	0.052	0.020	0.485	1,045	Fraction
Violence per Capita	7.780	6.740	0.023	47.348	923	# crimes per capita
Income Per Capita	14,936.730	5,031.352	5,561	55,463	1,045	$ per capita
Mean-to-Median Income	1.263	0.139	1.030	2.247	1,045	Ratio

Table B.3: Fractionalization, Taxation and Public Spending

	City Level		County Level	
	Tax per Capita	Exp per Capita	Tax per Capita	Exp per Capita
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	166.863***	215.143*	165.145***	616.171***
	(60.108)	(125.338)	(54.653)	(137.133)
Income per Capita	0.020***	0.020***	0.066***	0.041***
	(0.003)	(0.007)	(0.004)	(0.008)
Log Population Size	47.518***	113.099***	−7.101	1.435
	(15.972)	(29.431)	(8.702)	(18.906)
Education	−59.211	64.019	−170.490	−429.975
	(115.207)	(259.619)	(124.581)	(284.288)
Inequality	225.917***	428.224**	131.340*	634.596***
	(80.312)	(183.017)	(75.052)	(201.390)
Violence per Capita	5.270**	14.645***	13.854***	30.861***
	(2.240)	(3.353)	(2.223)	(5.789)
Population Above 65	251.207*	516.633	852.521***	1,061.297**
	(144.967)	(367.303)	(188.155)	(425.416)
Debt per Capita			0.005	0.019*
			(0.003)	(0.011)
State Fixed Effects	Yes	Yes	Yes	Yes
Observations	912	912	1,345	1,345
R^2	0.692	0.629	0.789	0.630
Adjusted R^2	0.673	0.605	0.780	0.614

Note: Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table B.4: Majority Size and Minority Fragmentation

	Dependent Variable: Taxes per Capita			
	City Level		County Level	
	(1)	(2)	(3)	(4)
Size of Majority	−349.450***	−622.170***	−346.184***	−723.692**
	(94.766)	(237.568)	(71.038)	(294.014)
Fractionalization of Minority	182.803***		116.379***	
	(61.521)		(31.264)	
Size of Biggest Minority Group		−491.237*		−522.388*
		(252.308)		(295.083)
Income per Capita	0.020***	0.020***	0.065***	0.065***
	(0.003)	(0.003)	(0.004)	(0.004)
Log Population Size	44.026***	45.795***	−5.835	−8.657
	(16.173)	(16.118)	(8.769)	(8.915)
Education	−76.389	−74.880	−183.286	−160.033
	(113.974)	(117.475)	(124.828)	(124.355)
Inequality	243.188***	234.791***	96.491	90.828
	(79.583)	(81.757)	(72.909)	(73.574)
Violence per Capita	4.934**	4.711**	13.125***	12.908***
	(2.289)	(2.288)	(2.146)	(2.161)
Population Above 65	334.072**	288.670*	952.514***	929.586***
	(154.699)	(151.368)	(186.796)	(190.918)
Debt per Capita			0.004	0.004
			(0.003)	(0.003)
State Fixed Effects	Yes	Yes	Yes	Yes
Observations	885	885	1,341	1,341
R^2	0.698	0.695	0.791	0.790
Adjusted R^2	0.677	0.674	0.782	0.781

Note: Robust standard errors in parentheses. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table B.5: Ethnic Fractionalization and Expenditure Categories (County Level)

	Share of general expenditures for			
	Welfare	Roads	Education	Hospitals
	(1)	(2)	(3)	(4)
Ethnic Fractionalization	0.011	−0.024***	−0.031	−0.003
	(0.007)	(0.007)	(0.031)	(0.037)
Income per Capita	−0.00000***	−0.00000	0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Log Population Size	0.002**	−0.008***	−0.014***	−0.006
	(0.001)	(0.001)	(0.004)	(0.004)
Education	0.035*	0.027*	−0.442***	0.063
	(0.018)	(0.014)	(0.063)	(0.072)
Inequality	0.004	−0.034***	−0.049	0.133***
	(0.012)	(0.008)	(0.043)	(0.041)
Violence per Capita	−0.0001	−0.0001	−0.005***	0.001
	(0.0002)	(0.0003)	(0.001)	(0.001)
Population Above 65	0.080***	0.083***	−0.836***	0.067
	(0.020)	(0.023)	(0.096)	(0.097)
Debt per Capita	0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
State Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,341	1,341	1,341	1,341
R^2	0.750	0.592	0.441	0.236
Adjusted R^2	0.739	0.575	0.417	0.203

Note: Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table B.6: Instrumented Ethnic Fractionalization (County Level)

	Dependent variable:	
	Taxes per Capita	Exp per Capita
	(1)	(2)
Ethnic Fractionalization (instrumented)	142.697**	591.160***
	(55.485)	(136.126)
Income per Capita	0.065***	0.041***
	(0.003)	(0.008)
Log Population Size	−5.596	0.812
	(7.065)	(17.332)
Education	−162.238	−433.594
	(114.169)	(280.099)
Inequality	122.420*	649.493***
	(68.792)	(168.773)
Violence per Capita	844.295***	1,035.676***
	(160.114)	(392.820)
Population Above 65	0.004***	0.019***
	(0.001)	(0.004)
Debt per Capita	13.606***	31.671***
	(1.784)	(4.376)
State Fixed Effects	Yes	Yes
Observations	1,341	1,341
R^2	0.789	0.629
Adjusted R^2	0.780	0.613
F-statistic (first stage)	9994	9166

Note: Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

B.2 Lemmas and Propositions

A.1. Proof of Proposition 2.2.1

The utility function given by:

$$u^i(E_i, c, g, E_{-i}) = \alpha\{\beta(E_i)^\rho + (1-\beta)c^\rho\}^{\frac{\nu}{\rho}} + \gamma g - \delta \sum_{j \neq i} E_j \tag{B.1}$$

The group maximizes $u^i(.)$ subject to the constraint:

$$E_i + \phi_i c = \phi_i(1-t)w \tag{B.2}$$

where t is the tax rate on wages, w.

Second stage results:

Setting up Lagrangian:

$$\begin{aligned}\mathscr{L}(E_i, c, g, E_{-i}, \lambda_i) &= u(E_i, c, g, E_{-i}) - \lambda_i(E_i + \phi_i c - \phi_i(1-t)w) \\ &= \alpha\{\beta(E_i)^\rho + (1-\beta)c^\rho\}^{\frac{\nu}{\rho}} + \gamma g - \delta\sum_{j\neq i} E_j - \lambda_i(E_i + \phi_i c - \phi_i(1-t)w)\end{aligned} \tag{B.3}$$

$$FOCs:$$

$$E_i: \quad \alpha\frac{\nu}{\rho}\{\beta(E_i)^\rho + (1-\beta)c^\rho\}^{(\frac{\nu}{\rho}-1)}\beta(E_i^*)^{(\rho-1)}\rho - \lambda_i = 0 \tag{B.4}$$

$$c: \quad \alpha\frac{\nu}{\rho}\{\beta(E_i)^\rho + (1-\beta)c^\rho\}^{(\frac{\nu}{\rho}-1)}(1-\beta)(c_i^*)^{(\rho-1)}\rho - \lambda_i\phi_i = 0 \tag{B.5}$$

$$\lambda_i: \quad -E_i^* - \phi_i c_i^* + \phi_i(1-t)w = 0 \tag{B.6}$$

Rearranging (B.4) and (B.5) and dividing the two resulting equations we get

$$\begin{aligned}\frac{\beta(E_i^*)^{(\rho-1)}}{(1-\beta)(c_i^*)^{\rho-1}} &= \frac{1}{\phi_i} \\ \frac{\beta\phi_i}{(1-\beta)}(c_i^*)^{1-\rho} &= (E_i^*)^{(1-\rho)} \\ \left(\frac{\beta\phi_i}{1-\beta}\right)^{\frac{1}{1-\rho}} c_i^* &= E_i^*\end{aligned} \tag{B.7}$$

Replacing E_i^* with expression(B.7) in constraint (B.6) we get:

$$\Big(\frac{\beta\phi_i}{1-\beta}\Big)^{\frac{1}{1-\rho}} c_i^* + \phi_i c_i^* = \phi_i(1-t)w \tag{B.8}$$

$$\Big[\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}} + \phi_i^{(1-\frac{1}{1-\rho})}\Big] c_i^* = \phi_i^{(1-\frac{1}{1-\rho})}(1-t)wc_i^* \tag{B.9}$$

$$c_i^* = (1-t)w\left[\frac{\phi_i^{\frac{-\rho}{1-\rho}}}{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}}\right] \equiv (1-t)\kappa_c(\phi_i) \tag{B.10}$$

Using (B.7) we get:

$$E_i^* = (1-t)w\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}}\left[\frac{\phi_i}{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}}\right] \tag{B.11}$$

$$\equiv (1-t)\kappa_E(\phi_i) \tag{B.12}$$

where $\kappa_c(\phi_i)$ and $\kappa_E(\phi_i)$ are the exogenous coefficients of c_i^* and E_i^* respectively that are a function of the distribution of the group.

Thus,

$$E_i^* = \sum_{j\neq i} E_j = (1-t)\sum_{j\neq i}\kappa_E(\phi_j) \tag{B.13}$$

First stage

We normalize δ to 1. Utility maximization solution for the goverment with distribution $\boldsymbol{\phi_m}$ and weights $\boldsymbol{\lambda_m}$ corresponds to critical points of the following function:

$$\mathscr{U}_{\lambda_m\phi_m}(t) = \sum_{i\in M}\lambda_i\phi_i\left(\alpha[\beta\{(1-t)\kappa_E(\phi_i)\}^\rho + (1-\beta)\{(1-t)\kappa_c(\phi_i)\}^\rho]^{\frac{\nu}{\rho}} + \gamma tw - (1-t)\sum_{j\neq i}\kappa_E(\phi_j)\right) \tag{B.14}$$

$$= \sum_{i\in M}\lambda_i\phi_i\left(\alpha(1-t)^\nu\{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}} + \gamma tw - (1-t)\sum_{j\neq i}\kappa_E(\phi_j)\right) \tag{B.15}$$

FOC:

$$t : \sum_{i\in M}\lambda_i\phi_i\left(-\alpha\nu(1-t)^{(\nu-1)}\{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}} + \gamma w + \sum_{j\neq i}\kappa_E(\phi_j)\right) = 0 \tag{B.16}$$

$$\implies (1-t^*)^{(\nu-1)}\nu\alpha\sum_{i\in M}\lambda_i\phi_i\{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}} = \sum_{i\in M}\lambda_i\phi_i\{\gamma w + \sum_{j\neq i}\kappa_E(\phi_j)\} \tag{B.17}$$

$$\implies (1-t^*) = \left[\frac{\sum_{i\in M}\lambda_i\phi_i\{\gamma w + \sum_{j\neq i}\kappa_E(\phi_j)\}}{\nu\alpha\sum_{i\in M}\lambda_i\phi_i\{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}}}\right]^{\frac{-1}{1-\nu}} \tag{B.18}$$

$$\implies t^* = 1 - \left[\frac{\sum_{i\in M}\lambda_i\phi_i\{\gamma w + \sum_{j\neq i}\kappa_E(\phi_j)\}}{\nu\alpha\sum_{i\in M}\lambda_i\phi_i\{\beta(\kappa_E(\phi_i))^\rho + (1-\beta)(\kappa_c(\phi_i))^\rho\}^{\frac{\nu}{\rho}}}\right]^{\frac{-1}{1-\nu}} \tag{B.19}$$

$$t_{\lambda_m\phi_m} = 1 - \left(\Omega_{\lambda_m\phi_m}\right)^{\frac{-1}{1-\nu}} \tag{B.20}$$

where $\Omega_{\lambda_m\phi_m}$ is given by ratio of the weighted gains to increasing tax and the weighted loss of increasing taxes for a given group distribution $\boldsymbol{\phi_m}$ and social weights $\boldsymbol{\lambda_m}$. ■

B.2.1 Proof of Corollary 2.2.1

From Proposition B.20 we have

$$\frac{\partial t_{\lambda_m \phi_m}}{\partial \Omega_{\lambda_m \phi_m}} = \frac{\left(\Omega_{\lambda_m \phi_m}\right)^{\frac{\nu-2}{1-\nu}}}{1-\nu} > 0 \qquad \forall \nu \in (0,1) \tag{B.21}$$

■

A.3. Comparative statics in E_i^* and c_i^* w.r.t ϕ_i

We keep t fixed in this section and only analyze the change in $\kappa_E(\phi_i)$ and $\kappa_c(\phi_i)$ with ϕ_i.

$$\frac{\partial \kappa_E(\phi_i)}{\partial \phi_i} = \frac{\left(\frac{\beta}{1-\beta}\right)^{\frac{1}{1-\rho}} w \left[\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\} + \phi_i \frac{\rho}{1-\rho} \phi_i^{-\frac{\rho}{1-\rho}-1} \right]}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^2} \tag{B.22}$$

$$= \frac{\left(\frac{\beta}{1-\beta}\right)^{\frac{1}{1-\rho}} w \left[(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \frac{1}{1-\rho} \phi_i^{-\frac{\rho}{1-\rho}} \right]}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^2} > 0 \qquad \forall \rho < 0 \tag{B.23}$$

$$\frac{\partial^2 \kappa_E(\phi_i)}{\partial \phi_i^2} = \left(\frac{\beta}{1-\beta}\right)^{\frac{1}{1-\rho}} w \left[\frac{-\frac{\rho}{(1-\rho)^2} \phi_i^{-\frac{\rho}{1-\rho}-1} \{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\} + 2\frac{\rho}{(1-\rho)} \phi_i^{-\frac{\rho}{1-\rho}-1} \{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \frac{1}{1-\rho} \phi_i^{-\frac{\rho}{1-\rho}}\}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^3} \right] \tag{B.24}$$

$$= \left(\frac{\beta}{1-\beta}\right)^{\frac{1}{1-\rho}} w \phi_i^{-\frac{\rho}{1-\rho}-1} \left(\frac{\rho}{1-\rho}\right) \left[\frac{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} (2 - \frac{1}{1-\rho}) + \phi_i^{\frac{-\rho}{1-\rho}} (\frac{1}{1-\rho})}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^3} \right] < 0 \qquad \forall \rho < 0 \tag{B.25}$$

(Since $\frac{1}{1-\rho} \in (0,1)$ for all $\rho < 0$)

Thus, $\kappa_E(\phi_i)$ is increasing and concave in ϕ_i for all $\rho < 0$ which means for a fixed t, total cultural good production is also increasing and concave in ϕ_i.

$$\frac{\partial \kappa_c(\phi_i)}{\partial \phi_i} = w \left[\frac{\frac{-\rho}{1-\rho}\phi_i^{(-\frac{\rho}{1-\rho}-1)}\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\} + \frac{\rho}{1-\rho}\phi_i^{(-\frac{\rho}{1-\rho}-1)}\{\phi_i^{\frac{-\rho}{1-\rho}}\}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^2} \right] \tag{B.26}$$

$$= w\Big(\frac{-\rho}{1-\rho}\Big)\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}} \left[\frac{\phi_i^{(-\frac{\rho}{1-\rho}-1)}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^2} \right] > 0 \qquad \forall \rho < 0 \tag{B.27}$$

$$\frac{\partial^2 \kappa_c(\phi_i)}{\partial \phi_i^2} = w\Big(\frac{-\rho}{1-\rho}\Big)\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}}. \tag{B.28}$$

$$\left[\frac{(-\frac{\rho}{1-\rho}-1)\phi_i^{-\frac{\rho}{1-\rho}-2}\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\} + 2\frac{\rho}{1-\rho}\phi_i^{-\frac{\rho}{1-\rho}-1}\phi_i^{-\frac{\rho}{1-\rho}-1}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^3} \right] < 0 \forall \rho < 0 \tag{B.29}$$

$$= w\Big(\frac{-\rho}{1-\rho}\Big)\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}} \left[\frac{\frac{-1}{1-\rho}\phi_i^{\frac{\rho-2}{1-\rho}}(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \frac{(-1+2\rho)}{1-\rho}\phi_i^{\frac{-2}{1-\rho}}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^3} \right] \tag{B.30}$$

$$= w\Big(\frac{-\rho}{(1-\rho)^2}\Big)\Big(\frac{\beta}{1-\beta}\Big)^{\frac{1}{1-\rho}} \phi_i^{\frac{\rho-2}{1-\rho}} \left[\frac{-(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + (-1+2\rho)\phi_i^{\frac{-\rho}{1-\rho}}}{\{(\frac{\beta}{1-\beta})^{\frac{1}{1-\rho}} + \phi_i^{\frac{-\rho}{1-\rho}}\}^3} \right] \tag{B.31}$$

(Since $\rho < 1/2$ when $\rho < 0$)

Thus, $\kappa_c(\phi_i)$ is increasing and concave in ϕ_i for any $\rho < 0$ which means that for a given t equilibrium private consumption c_i^* is also increasing and concave in ϕ_i for $\rho < 0$. ■

A.4. Proof of Proposition 2.2.1

Define $G_{\lambda_m \phi_m}$ as gains to increasing taxes, and $L_{\lambda_m \phi_m}$ as loss from increasing taxes where $\Omega_{\lambda_m \phi_m} = \frac{G_{\lambda_m \phi_m}}{L_{\lambda_m \phi_m}}$.

Now for government with $\lambda_i = 1$:

$$G_i(\boldsymbol{\phi_m}) = \gamma w + \sum_{j \neq i} \kappa_E(\phi_i) \tag{B.32}$$

$$L_i(\boldsymbol{\phi_m}) = \nu\alpha\left(\beta\kappa_E(\phi_i)^{\rho} + (1-\beta)\kappa_c(\phi_i)^{\rho}\right)^{\frac{\nu}{\rho}} \tag{B.33}$$

Using corollary 2, to show $t^*_{maj}(\boldsymbol{\phi_m}) < t^*_u(\boldsymbol{\phi_m}) < t^*_{min}(\boldsymbol{\phi_m})$ we use $\Omega_{maj}(\boldsymbol{\phi_m}) < \Omega_u(\boldsymbol{\phi_m}) < \Omega_{min}(\boldsymbol{\phi_m})$.

A sufficient condition for this is if for all $i,k \in N$ with $\phi_i < \phi_k$,

$$G_i(\boldsymbol{\phi_m}) > G_k(\boldsymbol{\phi_m}) > 0 \tag{B.34}$$

$$0 < L_i(\boldsymbol{\phi_m}) < L_k(\boldsymbol{\phi_m}) \tag{B.35}$$

$$\implies \frac{G_i(\boldsymbol{\phi_m})}{L_i(\boldsymbol{\phi_m})} > \frac{G_k(\boldsymbol{\phi_m})}{L_k(\boldsymbol{\phi_m})} \tag{B.36}$$

Since $G_u(\phi_m)$ is a convex combination of $\{G_i(\phi_m)\}_{i \in M}$ and $L_u(\phi_m)$ is a convex combination of $\{L_i(\phi_m)\}_{i \in M}$.

Proof of condition B.34

$$G_i(\boldsymbol{\phi_m}) - G_k(\boldsymbol{\phi_m}) = \gamma w + \sum_{j \neq i} \kappa_E(\phi_j) - \gamma w + \sum_{j \neq k} \kappa_E(\phi_j) \tag{B.37}$$

$$= \sum_{j \notin \{i,k\}} \kappa_E(\phi_j) + \kappa_E(\phi_k) - \sum_{j \notin \{i,k\}} \kappa_E(\phi_j) + \kappa_E(\phi_i) \tag{B.38}$$

$$= \kappa_E(\phi_k) - \kappa_E(\phi_i) > 0 \tag{B.39}$$

$$\text{(Since } \kappa_E(\phi_i) \text{ is increasing in } \phi_i\text{)} \tag{B.40}$$

■

Proof of condition B.35: We use

$$\frac{\partial(\beta g(x)^{\rho}+(1-\beta)f(x)^{\rho})^{\frac{\nu}{\rho}}}{\partial x}=\nu\left(\beta g'(x)g(x)^{\rho-1}+(1-\beta)f'(x)f(x)^{\rho-1}\right)^{\frac{\nu-\rho}{\rho}}>0 \tag{B.41}$$

$$\text{(when } g'(x),f'(x),g(x),f(x)>0) \tag{B.42}$$

$$L_k(\boldsymbol{\phi_m})-L_i(\boldsymbol{\phi_m})>0 \tag{B.43}$$

As $\kappa_E'(\phi_i),\kappa_E(\phi_i),\kappa_c'(\phi),\kappa_c(\phi)>0$. ■

A.5. Proof of Proposition 2.2.2

Majority government

Define $G_{\lambda_m\phi_m}$ as gains to increasing taxes, and $L_{\lambda_m\phi_m}$ as loss from increasing taxes where $\Omega_{\lambda_m\phi_m}=\frac{G_{\lambda_m\phi_m}}{L_{\lambda_m\phi_m}}$. Then

$$\frac{\partial\Omega_{maj}(\phi_1)}{\partial\phi_1}<0 \tag{B.44}$$

if $\frac{\partial G_{\lambda_m\phi_m}}{\partial\phi_1}<0$ and $\frac{\partial L_{\lambda_m\phi_m}}{\partial\phi_1}>0$.

Now if $\lambda_1=1$ we have:

$$G_{maj}(\phi_1)=\gamma w+\sum_{j\neq 1}\kappa_E(\phi_j) \tag{B.45}$$

$$L_{maj}(\phi_1)=\nu\alpha(\beta\kappa_E(\phi_1)^{\rho}+\kappa_c(\phi_1)^{\rho})^{\frac{\nu}{\rho}} \tag{B.46}$$

$$\frac{\partial G}{\partial \phi_1} = (m-1)\frac{\partial \kappa_E(\phi_j)}{\partial \phi_j}\frac{\partial(\frac{1-\phi_1}{m-1})}{\partial \phi_1} \tag{B.47}$$

$$= -\frac{\partial \kappa_E(\phi_j)}{\partial \phi_j} < 0 \tag{B.48}$$

$$(\text{Since } \frac{\partial \kappa_E(\phi_j)}{\partial \phi_j} > 0) \tag{B.49}$$

$$\frac{\partial L}{\partial \phi_1} = \frac{\nu^2}{\rho}\alpha\left(\beta\rho\kappa_E(\phi_1)^{\rho-1}\frac{\partial \kappa_E(\phi_1)}{\partial \phi_1} + (1-\beta)\rho\kappa_c(\phi_1)^{\rho-1}\frac{\partial \kappa_c(\phi_1)}{\partial \phi_1}\right)^{\frac{\nu}{\rho}-1} > 0 \tag{B.50}$$

$$(\text{Since } \frac{\partial \kappa_c(\phi_1)}{\partial \phi_1}, \frac{\partial \kappa_E(\phi_1)}{\partial \phi_1} > 0) \tag{B.51}$$

Thus,

$$\frac{\Omega_{maj}(\phi_1)}{\partial \phi_1} < 0$$

Minority Government

For the minority government we have, $\lambda_i = 1$ and ϕ_i' is the size of the minority.

$$\Omega_{min}(\phi_1) = \frac{G_{min}(\phi_1)}{L_{min}(\phi_1)} \tag{B.52}$$

$$\text{where} \tag{B.53}$$

$$G_{min}(\phi_1, \phi_i') = \gamma w + \sum_{j \notin \{i,1\}} \kappa_E(\frac{1-\phi_1-\phi_i'}{m-2}) + \kappa_E(\phi_1) \tag{B.54}$$

$$L_{min}(\phi_1, \phi_i') = \nu\alpha\{\beta\kappa_E(\phi_i')^\rho + (1-\beta)\kappa_c(\phi_i')^\rho\}^{\frac{\nu}{\rho}} \tag{B.55}$$

Clearly,

$$\frac{\partial L_{min}(\phi_1, \phi_i')}{\partial \phi_1} = 0$$

We just need to show that $\frac{\partial G_{min}(\phi_1, \phi_i')}{\partial \phi_1} < 0$

$$G_{min}(\phi_1, \phi_i') = \gamma w + (m-2)\kappa_E(\frac{1-\phi_1-\phi_i'}{m-2}) + \kappa_E(\phi_1) \tag{B.56}$$

$$\frac{\partial G_{min}(\phi_1, \phi_i')}{\partial \phi_1} = -\frac{m-2}{m-2}\kappa_E'(\frac{(1-\phi_1-\phi_i')}{m-2}) + \kappa_E'(\phi_1) < 0 \tag{B.57}$$

$$\iff \quad \kappa_E'(\frac{1-\phi_1-\phi_i'}{m-1}) > \kappa_E'(\phi_1) \tag{B.58}$$

$$\tag{B.59}$$

which is always true since $\kappa_E(.)$ is concave in ϕ_j and by assumption $\phi_1 > \frac{1-\phi_1-\phi_i'}{m-2}$.

■

A.6. Proof of Proposition 2.2.3

Majority Government

Clearly, $L_{maj}(\phi_1)$ is independent of m.

$$\Omega_{maj}(\phi_1, m) = G_{maj}(\phi_1, m)/L_{maj}(\phi_1) \tag{B.60}$$

$$= (m-1)\kappa_E(\frac{1-\phi_1}{m-1})\frac{1}{L_{maj}(\phi_1)} \tag{B.61}$$

$$= \left(\frac{(1-\phi_1)}{(\beta/(1-\beta))^{\frac{1}{1-\rho}} + (\frac{1-\phi_1}{m-1})^{\frac{-\rho}{1-\rho}}} \right) \frac{1}{L_{maj}(\phi_1)} \tag{B.62}$$

$$\text{(Since } \frac{1-\phi_1}{m-1} \text{ decreases with m)} \tag{B.63}$$

$$\implies \frac{\partial \Omega_{maj}(\phi_1, m)}{\partial m} > 0 \tag{B.64}$$

Minority Government

$$\Omega_{min}(\phi_1, m) = G_{min}(\phi_1, m)/L_{min}(\phi_1, m) \tag{B.65}$$

$$= \frac{(m-2)\kappa_E(\frac{1-\phi_1}{m-1}) + \kappa_E(\phi_1)}{\nu\alpha(\beta\kappa_E(\frac{1-\phi_1}{m-1})^\rho + (1-\beta)\kappa_c(\frac{1-\phi_1}{m-1})^\rho)^{\frac{\nu}{\rho}}} \tag{B.66}$$

Clearly, $L_{min}(\phi_1, m)$ decreases with m as $\kappa_E(.)$ and $\kappa_c(.)$ decrease with m. Thus, if we show $G_{min}(\phi_1, m)$ increases with m then we are done with the proof.

$$\frac{\Delta G_{min}(\phi_1, m)}{\Delta m} = (m-1)\kappa_E(\frac{1-\phi_1}{m}) - (m-2)\kappa_E(\frac{1-\phi_1}{m-1}) \tag{B.67}$$

$$= (1-\phi_1)\left(\frac{\frac{m-1}{m}}{(\beta(1-\beta))^{\frac{1}{1-\rho}} + (\frac{(1-\phi_1)}{m})^{\frac{-\rho}{1-\rho}}} - \frac{\frac{m-2}{m-1}}{(\beta(1-\beta))^{\frac{1}{1-\rho}} + (\frac{(1-\phi_1)}{m-1})^{\frac{-\rho}{1-\rho}}}\right) \tag{B.68}$$

$$(\text{Since } \frac{m-1}{m} > \frac{m-2}{m-1} \text{ and } \frac{(1-\phi_1)}{m} < \frac{(1-\phi_1)}{m-1} \text{ and } \rho < 0) \tag{B.69}$$

$$\implies > 0 \tag{B.70}$$

Hence we have $\frac{\Delta\Omega_{min}(\phi_1, m)}{\Delta m} > 0$. ■

A.7. Proof of Proposition 2.2.4

Follows from proof of proposition 2.2.2. ■

A.8. Proof of Proposition 2.2.5

When $\phi_1 > 0.5$ then majority group is the median. This means $\lambda_1 = 1$. We need to show:

$$\frac{\partial G_{maj}(\phi_1, \phi_2, m)}{\partial \phi_2} < 0 \tag{B.71}$$

We have:

$$G_{maj}(\phi_1,\phi_2,m) = \gamma w + (m-2)\kappa_E(\frac{(1-\phi_1-\phi_2)}{m-2}) + \kappa_E(\phi_2) \quad \text{(B.72)}$$

$$\implies \frac{\partial G_{maj}(\phi_1,\phi_2,m)}{\partial \phi_2} = -\kappa_E'(\frac{(1-\phi_1-\phi_2)}{m-2}) + \kappa_E'(\phi_2) < 0 \quad \text{(B.73)}$$

$$\text{(By concavity of } \kappa_E \text{ and by definition } \phi_2 > \frac{(1-\phi_1-\phi_2)}{m-2}) \quad \text{(B.74)}$$

When $\phi_1 < 0.5$ and $\phi_1 + \phi_2 > 0.5$ then group 2 is the median voter. This means that $\lambda_2 = 1$.

We know for any government with $\lambda_i = 1$ then

$$\frac{\partial \Omega_{\lambda_m \phi_m}}{\partial \phi_i} < 0$$

(See condition B.2.1) ∎

A.9. Proof of Lemma 2.3.1

Extensive and intensive margin relationship with fractionalization:

$$FRAC = 1 - \phi_1^2 - (m-1)\left(\frac{1-\phi_1}{m-1}\right)^2 \quad \text{(B.75)}$$

$$= 1 - \phi_1^2 - \frac{(1-\phi_1)^2}{m-1} \quad \text{(B.76)}$$

Extensive margin: Clearly $\frac{\Delta FRAC}{\Delta m} > 0$.

Intensive margin: Differentiating $FRAC$ w.r.t ϕ_1 we get, $\frac{\partial FRAC}{\partial \phi_1} < 0 \; \forall \; \phi_1 > 1/m$.

Extensive and intensive margin relationship with polarization:

$$POL = 1 - \sum_{i \in M} \Big(\frac{0.5 - \phi_i}{0.5}\Big)^2 \phi_i \tag{B.77}$$

$$= 4\{\sum_{i \in M} \phi_i^2 (1 - \phi_i)\} \tag{B.78}$$

$$= 4\{\phi_1^2(1-\phi_1) + (m-1)\Big(\frac{1-\phi_1}{m-1}\Big)^2 (1 - \frac{1-\phi_1}{m-1})\} \tag{B.79}$$

$$= 4\Big\{\phi_1^2(1-\phi_1) + \frac{(1-\phi_1)^2}{m-1} - \frac{(1-\phi_1)^3}{(m-1)^2}\Big\} \tag{B.80}$$

$$= 4(1-\phi_1)\{\phi_1^2 + \frac{1-\phi_1}{m-1} - \frac{(1-\phi_1)^2}{(m-1)^2}\} \tag{B.81}$$

Extensive margin: For $m > 2$ we have $\frac{\Delta POL}{\Delta m} < 0$.

Intensive margin: Differentiating with respect to ϕ_1 we get:

$$\frac{\partial POL}{\partial \phi_1} = DPOL \quad = 4\Big\{ - \big(\phi_1^2 + \frac{1-\phi_1}{m-1} - \frac{(1-\phi_1)^2}{(m-1)^2}\big) + (1-\phi_1)(2\phi_1 - \frac{1}{m-1} + 2\frac{1-\phi_1}{(m-1)^2})\Big\} \tag{B.82}$$

$$= -3\phi_1^2 \frac{m-2}{m-1} + 2\phi_1 \frac{m(m-1)-3}{(m-1)^2} - \frac{2m-5}{(m-1)^2} \tag{B.83}$$

Clearly, the quadratic roots ϕ_1 when setting $DPOL = 0$ are a function of m. Since $DPOL$ is a concave function the maximal root $\bar{\phi_1}(m)$ is such that there exists $\varepsilon > 0$ such that $\forall \phi_1 \in (\bar{\phi_1} - \varepsilon, \bar{\phi_1})$, $DPOL > 0$ and $\forall \phi_1 \in (\bar{\phi_1}, 1)$, $DPOL < 0$. ■

Appendix C

Ideologically Radical, Tactically Conservative

Properties of Euclidean preferences

1. Euclidean preferences are quasi-supermodular

Proof. Let $u = y \vee z$ and $l = z \wedge y$. Define $ZU = \{j \in N | y_j < z_j\}$ and $YU = \{j \in N | y_j \geq z_j\}$ (clearly $ZU \cup YU = N$). Thus, if $j \in ZU$ then $l_j = y_j$ and $u_j = z_j$. If $j \in YU$ then $l_j = z_j$ and $u_j = y_j$.

We need to show that

(a) If $||x^i - y|| < ||x^i - u||$ then $||x^i - l|| < ||x^i - z||$.

(b) If $||x^i - y|| < ||x^i - l||$ then $||x^i - u|| < ||x^i - z||$.

(1) Suppose $||x^i - y|| < ||x^i - u||$ then,

$$\sum_{j \in M} (x^i_j - y_j)^2 < \sum_{j \in M} (x^i_j - u_j)^2 \tag{C.1}$$

Using the definition of u

$$\sum_{j \in M} (x^i_j - y_j)^2 < \sum_{j \in ZU} (x^i_j - z_j)^2 + \sum_{j \in YU} (x^i_j - y_j)^2 \tag{C.2}$$

$$\Leftrightarrow \sum_{j \in ZU} (x^i_j - y_j)^2 < \sum_{j \in ZU} (x^i_j - z_j)^2 \tag{C.3}$$

$$\Leftrightarrow \sum_{j \in ZU} (x^i_j - y_j)^2 + \sum_{j \in YU} (x^i_j - z_j)^2 < \sum_{j \in N} (x^i_j - z_j)^2 \tag{C.4}$$

Using the definition of l

$$\Leftrightarrow \sum_{j \in M} (x^i_j - l)^2 < \sum_{j \in M} (x^i_j - z_j)^2 \tag{C.5}$$

(2) Suppose $||x^i - y|| < ||x^i - l||$ then,

$$\sum_{j \in M} (x^i_j - y_j)^2 < \sum_{j \in M} (x^i - l)^2 \tag{C.6}$$

Using the definition of l

$$\sum_{j \in M} (x^i_j - y_j)^2 < \sum_{j \in YU} (x^i_j - z_j)^2 + \sum_{j \in ZU} (x^i_j - y_j)^2 \tag{C.7}$$

$$\Leftrightarrow \sum_{j \in YU} (x^i_j - y_j)^2 < \sum_{j \in YU} (x^i_j - z_j)^2 \tag{C.8}$$

$$\Leftrightarrow \sum_{j \in YU} (x^i_j - y_j)^2 + \sum_{j \in JU} (x^i_j - z_j)^2 < \sum_{j \in M} (x^i_j - z_j)^2 \tag{C.9}$$

Using the definition of u

$$\Leftrightarrow \sum_{j \in M} (x^i_j - u)^2 < \sum_{j \in M} (x^i_j - z_j)^2 \tag{C.10}$$

■

2. Euclidean preferences with single crossing property must have ordered ideal points.

Proof. Suppose there exists $i, j \in N$ such that $x^i \not> x^j$ and $x^i \not< x^j$. If the preferences satisfy the single crossing property, then there exists an ordering of the types given by θ^i such that if for any $y > z$

$$yR^i z \implies yR^j z \qquad \forall \theta^j > \theta^i$$
$$yR^i z \implies yR^j z \qquad \forall \theta^j \geq \theta^i$$

To show that if $\theta^i < \theta^j$ then $x^i \leq x^j$.

Proof by contradiction:

First we show that $x^i \vee x^j P^i x^j$ if $x^i \not\leq x^j$ and $x^i \not\geq x^j$.

Let $IU = \{k \in N | x^i_k > x^j_k\}$ and $JU = \{k \in N | x^i_k \leq x^j_k\}$. Since $x^i \not\leq x^j$ and $x^i \not\geq x^j$, $JU, IU \neq \phi$.

To show that $x^i \vee x^j P^i x^j$.

$$||x^i - x^j|| = \sum_{j \in M} (x^i_k - x^j_k)^2 \tag{C.11}$$
$$= \sum_{j \in JU} (x^i_k - x^j_k)^2 + \sum_{j \in IU} (x^i_k - x^j_k)^2 \tag{C.12}$$
$$> \sum_{j \in JU} (x^i_k - x^j_k)^2 + \sum_{j \in IU} (x^i_k - x^i_k)^2 = \sum_{j \in M} (x^i_k - \max\{x^i_k, x^j_k\})^2 \tag{C.13}$$

The strict inequality comes from $JU, IU \neq \phi$

$$> ||x^i - x^i \vee x^j|| \tag{C.14}$$
$$\implies x^i \vee x^j P^i x^j \tag{C.15}$$

Say $\theta^j \geq \theta^i$. Suppose $x^j \not\geq x^i$. Then there are two possibilities.

(a) $x^i \not\leq x^j$ and $x^i \not\geq x^j$. Because the policy space is a lattice, we know that $x^i \vee x^j \in \mathbb{R}^m$. We know from (C.15) that $x^i \vee x^j P^i x^j$. Then SCP implies $x^i \vee x^j P^j x^j$. But by definition $x^j R^j x$ for all $x \in \mathbb{R}^m$. Thus, it must be true that $x^j R^j x^i \vee x^j$. This leads to a contradiction.

(b) $x^i \geq x^j$ and $x^i \neq x^j$. We know from optimality and uniqueness that $x^i P^i x^j$. Then SCP implies $x^i P^j x^j$. But by definition $x^j R^j x$ for all $x \in \mathbb{R}^m$. Thus, it must be true that $x^i R^j x^j$. This leads to a contradiction.

■

Proof of Pareto optimality property

Proof. W.l.o.g. let $\mathscr{C} = \{1, ..., m\}$ be the set of types of voters in a coalition. Let $\mathscr{X} = \{x^1, ..., x^m\}$ be the set of ideal points of the types of voters in $\mathscr{C}$. That is, $x^i \in \mathscr{X}$ is the ideal point of voter type $i \in \mathscr{C}$.

Part 1 All Pareto optimal points in a coalition are in the convex hull:

Suppose $y \in \mathbb{R}$ that is a Pareto optimal point. Since preferences of voters is convex and everywhere differentiable we know y is Pareto optimal point iff $\forall i \in \mathscr{C}\ \exists \lambda_i \in \mathbb{R}_+ \backslash \{0\}$ such that $\sum_i \lambda_i \Delta u^i(y) = 0$. In other words for $u^i = ||x^i - y|| = (x^i - y)\dot{(}x^i - y)$ we have,

$$\sum_i \lambda_i 2(x^i - y) = 0 \tag{C.16}$$

$$\sum_i \lambda_i x^i = y \sum_i \lambda_i \tag{C.17}$$

$$\sum_i \left(\frac{\lambda_i}{\sum_j \lambda_j}\right) x^i = y \tag{C.18}$$

Let $\mu_i = \left(\frac{\lambda_i}{\sum_j \lambda_j}\right)$. Since $\sum\limits_{i \in N} \mu_i = 1$, every Pareto optimal point $y \in Conv(\mathscr{X})$.

All points in the convex hull are Pareto optimal:

Let $y \in Conv(\mathscr{X})$ then by definition $\exists \mu_i \in R_+$ with $\sum_{i \in \mathscr{C}} \mu_i = 1$ such that

$$\sum_i \mu_i x^i = y \tag{C.19}$$

$$\sum_i (\mu_i) 2(x^i - y) = 0 \tag{C.20}$$

$$\sum_i \mu_i \nabla u^i(y) = 0 \tag{C.21}$$

Since utilities of the agents are convex and everywhere differentiable y is Pareto optimal. ■ ■

Proof if proposal wins, only ideal points will be proposed in equilibrium

Proof. Let $\boldsymbol{a^j}$ be a winning policy proposed by coalition C_j. Assume $\boldsymbol{a^j}$ is in the interior of X^j. By definition this implies $\forall i \in C^j \ \nabla u^j(\boldsymbol{a^j}) \neq 0$.

Case 1:

If $\boldsymbol{a^j}$ is ordered with respect to ideal points in C^j then there exists a partition of C^j into $S \subset C$ and $S' = C^j \backslash S$ such that $\forall i \in S$, $\nabla u^j < 0$ and $\forall\ i \in S'\ \nabla u^j > 0$. Then, there exists $\boldsymbol{a'}, \boldsymbol{a''}$ in the neighborhood of $\boldsymbol{a^j}$ such that $u^i(\boldsymbol{a'}) > u^i(\boldsymbol{a^j}) > u^i(\boldsymbol{a''})\ \forall\ i \in S$ and $u^i(\boldsymbol{a''}) > u^i(\boldsymbol{a^j}) > u^i(\boldsymbol{a'})\ \forall\ i \in S'$.

Case 2:

Suppose $T = \{i \in C_j | \boldsymbol{a^j} \gneq \boldsymbol{x^i} \text{ and } \boldsymbol{a^j} \lneq \boldsymbol{x^i}\} \neq \emptyset$. Consider $i^* = \arg\min_{i \in T} x^i$. Define $\boldsymbol{a'} = \boldsymbol{x}^{i^*} \wedge \boldsymbol{a^j}$, $\boldsymbol{a''} = \boldsymbol{x}^{i^*} \vee \boldsymbol{a^j}$ and $S \equiv \{i \in C_j | i < i^*\}$. Since $\boldsymbol{a^j}$ is not ordered with respect to $\boldsymbol{x}^{i^*}$, $\boldsymbol{a'}, \boldsymbol{a''} \in X^j$. Now, since by definition $\boldsymbol{a^j}$ is ordered with respect $\boldsymbol{x^i}\ \forall\ i \in S$, $u^i(\boldsymbol{a'}) > u^i(\boldsymbol{a^j}) > u^i(\boldsymbol{a''})\ \forall i \in S$. Also, by single crossing property, $\forall k > i^*\ u^k(\boldsymbol{a''}) > u^k(\boldsymbol{a^j})$ since $\boldsymbol{a''} \geq \boldsymbol{a^j}$ and $u^{i^*}(\boldsymbol{a''}) > u^{i^*}(\boldsymbol{a^j})$. Define $S' \equiv C^j \backslash S$.

Now, consider the permutation $\rho : X^j \to X^j$ where $\rho(\boldsymbol{a'}) = \boldsymbol{a^j}$, $\rho(\boldsymbol{a^j}) = \boldsymbol{a''}$, $\rho(\boldsymbol{a''}) = \boldsymbol{a'}$ and for all $\boldsymbol{x} \in X^j \backslash \{\boldsymbol{a'}, \boldsymbol{a^j}, \boldsymbol{a''}\}$, $\rho(\boldsymbol{x}) = \boldsymbol{x}$. By construction $\forall\ i \in S$, $u^i_\rho(\boldsymbol{a'}) > u^i_\rho(\boldsymbol{a^j})$ and $\forall\ i \in S'$, $u^i_\rho(\boldsymbol{a'}) < u^i_\rho(\boldsymbol{a^j})$. Thus, for all $i \in C^j$ if $u^i_\rho(\boldsymbol{a'}) > u^i_\rho(\boldsymbol{a^j})$ implies $u^i(\boldsymbol{a'}) > u^i(\boldsymbol{a^j})$. But by neutrality $g(u^j_\rho) = \rho(g(u^j)) = \rho(\boldsymbol{a^j}) = \boldsymbol{a'}$. Then by monotonicity $g(u^j) \neq \boldsymbol{a^j}$. This is a contradiction.

■

C.0.1 Lemmas and Propositions

Lemma 3.3.1

Proof. See Dotti (2019) pg. 46 ■

Lemma 3.3.2

Proof. See Dotti (2019) pg. 46 ■

Proof of Proposition 3.3.1

Proof. We know that a necessary condition for legislative gridlock to occur is if an party coalition proposes a policy platform. This will only happen in equilibrium if the core members of the extreme coalition prefer the status quo more than the policy platform proposed by the median. If there does not exist a pair $i, j \in N$ such that $x^0 R^i x^\nu$ and $x^0 R^j x^\nu$ where $i < \nu < j$ then no extreme coalition forms in equilibrium that would create a legislative gridlock where x^0 is passed. Thus, a necessary condition for a gridlock to form is that at least two types of voters prefer x^0 to x^ν. A voter type i prefers x^0 to x^ν if and only if

$$||x^i - x^0|| < ||x^i - x^\nu||$$

$$||x^i - x^\nu - (x^0 - x^\nu)|| < ||x^i - x^\nu||$$

(Using the defintion of z^i and z^ν and squaring both sides)

$$||z^i - z^0||^2 < ||z^i||^2$$

(Using the definition of Eucledian distances)

$$\implies \sum_{j \in M} [(z^i_j)^2 - (z^i_j - z^\nu_j)^2] < 0$$

$$\implies \sum_{j \in M} (z^i_j - z^i_j + z^\nu_j)(z^i_j + z^i_j - z^\nu_j) < 0$$

$$\implies \sum_{j \in M} z^\nu_j (z^\nu_j - 2z^i_j) > 0$$

$$\implies -2 \sum_{j \in M} z^\nu_j z^i_j + \sum_{j \in M} (z^\nu_j)^2 > 0 \tag{C.22}$$

(Using $\boldsymbol{a}.\boldsymbol{b} = |a||b| \cos \phi_{ab}$ where ϕ_{ab} is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.)

$$\implies -2||z^\nu||||z^i|| \cos \theta^i + ||z^\nu||^2 > 0$$

$$\implies \frac{||z^\nu||}{2} > ||z^i|| \cos \theta^i$$

$$\implies \frac{1}{2} < I^i \tag{C.23}$$

Thus, there is a lower bound given by (C.23) of the relative dislike of the status quo that at least two voters $\theta^k < \theta^\nu < \theta^j$ must have in order to form an extreme party coalition that elicits a gridlock. ■

Proof of Corollary 3.3.1

Proof. This amounts to saying that for a given level of extremism E^i there is a minimum level of polarization given by

$$\frac{||x^\nu - x^i||}{||\boldsymbol{z^0}||} = \sqrt{(I^i)^2 + (E^i)^2} > \sqrt{\frac{1}{4} + (E^i)^2} \tag{C.24}$$

Since the minimum level of heterogeneity is zero, $\frac{1}{2}$ is the minimum polarization necessary for legislative gridlock for some $i, j \in \mathcal{N}$ such that $\theta^i < \theta^\nu < \theta^j$. ∎

Proof of Corollary 3.3.2

Proof. Define $x^j = \arg\max_{i>v} \left\|x^v - x^i\right\| x^k = \arg\max_{i<v} \left\|x^v - x^i\right\|$, and we know from Corollary 1 that $\max_{i>v} \left\|x^i - x^v\right\| > 1/2$ and $\max_{i<v} \left\|x^i - x^v\right\| < 1/2$ then the distance between these two extreme points must be at least $\sqrt{(1/2)^2 + (1/2)^2} = 1/\sqrt{2}$). ∎

Proof of Proposition 3.3.2

Proof. We derive necessary conditions for which an arbitrary voter under which prefer the median to any extreme party coalition alternative. This necessary condition gives us the minimum conservative extremism required for an existence of a Condorcet cycle.

Any relevant member of an extreme party must like the status quo more than the median. This means that any party alternative of an extreme party must also lie on or below the plane (C.22).

Thus a necessary condition for a voter i to like any possible coalition is if the shortest distance from plane C.22 is less than the distance between x^i and x^ν. In other words,

$$\frac{|-\boldsymbol{z}^\nu.\boldsymbol{z}^\nu/2 + \boldsymbol{z}^\nu.\boldsymbol{z}^i|}{||\boldsymbol{z}^\nu||} < ||x^i - x^\nu|| = ||z^i - z^\nu|| \tag{C.25}$$

This gives the following condition for a voter which prefers the median to the status quo to vote for proposed policy of extreme parties:

$$\Big|\big(\frac{-||z^0||^2}{2||z^0||} + \frac{||z^\nu||||z^i||\cos\theta^i}{||z^0||}\big)\Big| < ||z^i||$$

$$\implies ||\boldsymbol{z^0}||\Big|\frac{1}{2} - I^i\Big| < ||\boldsymbol{z^0}||\sqrt{(I^i)^2 + (E^i)^2}$$

$$\implies (I^i)^2 + (\frac{1}{2})^2 - 2.I^i.\frac{1}{2} < (I^i)^2 + (E^i)^2 \tag{C.26}$$

$$\implies I^i > \frac{1}{4} - (E^i)^2 \tag{C.27}$$

We get inequality (C.27) because we are looking at the points that prefer the median voter reform proposal to the status quo. We can already see that for a given level of heterogeneity there is a bound on the relative intensity of dislike for the status quo with respect to the median reform that can potentially vote for an extreme party reform. From inequality (C.27) we can see if relative dislike of the status quo is very high then only very high levels of heterogeneity with respect to the status quo will make it possible for the voter to prefer a policy on the half plane. Thus we have:

$$CE \equiv I^i + (E^i)^2 > \frac{1}{4} \tag{C.28}$$

■

Proof of Proposition 3.3.3

Proof. Let $\boldsymbol{a}$ be the policy proposal by the coalition between opposite leaning party members that block the policy of a median legislator. Then define $\boldsymbol{z^a} \equiv \boldsymbol{a} - \boldsymbol{x^\nu}$.

1. If $\mathscr{L}(\succeq)$ (or $\mathscr{R}(\succeq)$) is empty, then no agent from left (right) side of political spectrum has an incentive to sustain a gridlock.

2. Necessary and sufficient condition for agent i to prefer x^g to a:

$$||\boldsymbol{x^i} - \boldsymbol{x^g}|| < ||\boldsymbol{x^i} - \boldsymbol{a}||$$

$$||\boldsymbol{z^0}||\sqrt{(I^i - I^g)^2 + (E^i - E^g)^2} > ||\boldsymbol{z^0}||\sqrt{(I^i - I^a)^2 + (E^i - H^a)^2}$$

$$(I^i)^2 + (I^g)^2 - 2I^iI^g + (E^i)^2 + (E^g)^2 - 2E^iE^g > (I^i)^2 + (I^a)^2 - 2I^iI^a + (E^i)^2 + (E^a)^2 - 2E^iE^a$$

Rearranging terms and dividing by $I^a > 0$.

$$I^i < \frac{I^a - I^g}{2} + \frac{1}{I^a - I^g}\Big[\frac{(E^a)^2 - (E^g)^2}{2} - (E^iE^a - E^iE^g)\Big]$$

3. Necessary and sufficient condition for a to be preferred to x^v.

$$||\boldsymbol{x^i} - \boldsymbol{x^v}|| > ||\boldsymbol{x^i} - \boldsymbol{a}||$$

$$||\boldsymbol{z^i}|| > ||\boldsymbol{z^i} - \boldsymbol{z^a}||$$

$$||\boldsymbol{z^0}||\sqrt{(I^i)^2 + (E^i)^2} > ||\boldsymbol{z^0}||\sqrt{(I^i - I^a)^2 + (E^i - H^a)^2}$$

$$(I^i)^2 + (E^i)^2 > (I^i)^2 + (I^a)^2 - 2I^iI^a + (E^i)^2 + (E^a)^2 - 2E^iE^a$$

Rearranging terms and dividing by $I^a > 0$.

$$I^i > \frac{I^a}{2} + \frac{1}{I^a}\Big[\frac{(E^a)^2}{2} - E^iE^a\Big]$$

4. Necessary and sufficient condition for agent g to prefer x^g to x^0:

$$||\boldsymbol{x^g} - \boldsymbol{x^0}|| < ||\boldsymbol{x^g} - \boldsymbol{a}||$$

$$||\boldsymbol{z^0}||\sqrt{(I^g - I^0)^2 + (E^g - E^0)^2} < ||\boldsymbol{z^0}||\sqrt{(I^g - I^a)^2 + (E^g - H^a)^2}$$

$$1 + (I^g)^2 - 2I^g + (E^g)^2 - 2E^iE^g < (I^g)^2 + (I^a)^2 - 2I^gI^a + (E^g)^2 + (E^a)^2 - 2E^gE^a$$

Rearranging terms and dividing by $I^a > 0$.

$$I^g < \frac{I^a + 1}{2} + \frac{1}{I^a - 1}\Big[\frac{(E^a)^2}{2} - E^gE^a\Big]$$

■