Multidimensional Voting Models

Theory and Applications

Valerio Dotti

September 2016

A Thesis submitted to the Department of Economics in fulfillment

of the requirements for the degree of Doctor of Philosophy (PhD)

University College London
Faculty of Social and Historical Sciences
Department of Economics

Declaration

I, Valerio Dotti confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Valerio Dotti

Abstract

In this thesis I study how electoral competition shapes the public policies implemented by democratic countries. In particular, I analyse the relationship between observable characteristics of the population of voters, such as the distribution of income and age, and relevant public policy outcomes of the political process. I focus on two theoretical issues that have proved difficult to tackle with existing voting models, namely multidimensionality of the policy space and non-convexity of voter preferences. I propose a new theoretical framework to deal with these issues. I employ this new framework to address three popular questions in the Political Economy literature for which a multidimensional policy space is deemed to be a crucial element to capture the underlying economic trade-offs. Specifically, I analyse (i) the relationship between income inequality and size of the government, (ii) the causal link between population ageing and the 'tightness' of immigration policies, and (iii) the role played by the income distribution in shaping public investment in education. I compare the predictions derived under the new theoretical tool with those that prevail in the existing literature. I show that the interaction among multiple endogenous policy dimensions helps to explain why several studies in the literature - in which the analysis in restricted to a unique endogenous policy choice - deliver empirically controversial or inconsistent predictions. For all three questions, the approach proposed in this thesis is shown to be helpful in reconciling the theoretical predictions with empirical evidence, and in identifying the economic channels that underpin the patterns observed in the data.

Contents

List of Tables

List of Figures

Acknowledgments

Firstly I would like to thank my PhD supervisors, Ian Preston and Aureo de Paula for their constant guidance and advice throughout my PhD. Antonio Cabrales and John Roemer have provided insightful and helpful guidance and advice at various points throughout my PhD, for which I would like to state my appreciation. I would also like to thank Marco Bassetto, V. Bhaskar, Christian Dustmann, Antonio Guarino, Johannes Hörner, Tobias König, Suehyun Kwon, Guy Laroque, Gilat Levy, Konrad Mierendorff, Nikita Roketskiy and Larry Samuelson for their helpful comments and suggestions. I gratefully acknowledge financial support from ESRC and UCL. This PhD thesis could not have been accomplished without the support and encouragement of my parents, so I would like to thank them. Lastly, I would like to thank Sanghmitra for her constant support, patience and encouragement.

1 Introduction

This thesis contributes to the analysis of a questions that lies at the very core of the disciplines of Political Economy and Positive Political Theory. Namely, I study how electoral competition shapes the policies implemented by democratic countries, and specifically how individual preferences are aggregated by political processes and institutions into collective choices about public policies. This analysis is particularly important for modern democratic societies, whose political institution are deemed to be representative - at least to some degree - of the preferences of their citizens. The main target of this thesis is to propose a set of theoretical tools that allows one to formulate predictions about the relationship between observable characteristics of the population of voters in the economy, such as the distribution of income, wealth, age, etc., and some relevant policy outcome of such economy. To achieve this goal, I follow a main stream of literature, that identifies the choice of political representation induced by voting in elections as the most important political act in democratic countries (Riker, 1982). Thus, I focus on the role played by electoral and political competition in shaping the relationship between features of the voting population and policy outputs. The importance of such analysis for economists and social scientists in general is vast, both because of the large economic, fiscal and social consequences of the policies implemented by elective bodies, and because of its implications with respect to the design of political institutions and to the well-being of the society. Thus, it is not surprising that the question of how to aggregate heterogeneous individual preferences through electoral competition has been crucial in the development of Social Choice and Positive Political Theory from their very early days. The origins of this stream of literature can be found in the work of the French philosopher and mathematician Nicolas de Condorcet, and in particular in his *"Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix"* (Essay on the application of probability theory to plurality decision making, 1785). This early contribution has identified some of the main theoretical difficulties that would affect the analysis of this problem ever since. Specifically, he describes an issue known as *Condorcet's voting paradox*. Such result states that, for any choice over at least 3 alternatives, for a group of at least 3 individuals, it is possible that none of the alternative is undefeated in a majority voting contest, i.e. there is no *Condorcet winner*. The difficulties in aggregating heterogeneous preferences through a voting process have been shown to be potentially dramatic in the early Social Choice literature. Specifically, Arrow's General Impossibility Theorem (Arrow 1950) implies that one cannot ensure the existence of a stable social preference ordering that satisfies some minimum properties for unrestricted individual preference orderings. Moreover, the so-called Gibbard-Satterthwaite Theorem (Gibbard 1973, Satterthwaite 1975) states that any non-dictatorial deterministic voting rule that allows all potential alternatives to be chosen is prone to *tactical voting*. This means that a voter with full knowledge of how the other voters are to vote and of the rule being used has an incentive to vote in a way that does not reflect his or her preferences. This result implies that such voting rule is *manipulable*, meaning that the incentive compatibility of a voter behavior that reflects the true voter preferences cannot be ensured. The traditional literature - which I extensively survey in the next chapter - has tackled these problems by imposing restrictions on voters' preference orderings. Specifically, the early contributions by Black (1948, 1958) and Downs (1957) have shown that, if voters preferences over electoral outcomes satisfy some ordinal conditions such as *Single-Peakedness,* then a *Condorcet winner* exists and corresponds to the platform that is preferred by the median individual in the population. Thus, in a simple deterministic model of electoral competition - usually referred as *Downsian* model - it is possible to derive a simple and clear characterization of the policy outcome that prevails in equilibrium. This result, known as *Median Voter Theorem,* proved extremely popular in the Political Economy literature. The main reason is that it can be adopted to derive testable implications about the relationship between some characteristics of the voting population and the policy outcome, abstracting from other features of the political process. The shortcoming of this approach is that the conditions under which the Median Voter Theorem holds are extremely restrictive in some cases, because a *Condorcet winner* may not exist. Specifically, two of these cases have been extensively studied in the literature because of their relevance in several economic applications. This thesis focuses primarily on these two cases. The first is the (i) analysis of electoral competition over a multidimensional policy choice domain. A review of the literature about this topic is provided in chapter 2. The second case concerns problems in which (ii) voter preferences exhibit *non-convexities*, and it is extensively described chapter 5. In particular, I study collective choices over policies consisting of the public provision of certain goods and services, for which private alternatives are also available to voters on the market. A typical example of this kind of policy is public intervention in education, which is the object of the analysis in chapter 5. The contribution of this thesis to the literature in Political Economy is twofold. The first contribution is methodological. Specifically, I provide a set of theoretical tools to analyse questions in which either the multidimensionality of the policy space, or the existence of non-convexities in voters' preferences (or both) play a crucial role in shaping the economic trade-offs. Moreover, the proposed theoretical tools preserve - at least to some extent - the strong predictive power that the *Downsian* framework exhibits in presence of *single-peaked* preferences.

The second contribution is the application of such theoretical tools to analyse some of the most important policy decisions that democratic countries face in this decade, which correspond to some of the most popular topic of research in the literature. In detail, the thesis is structured as follows. In chapter 2 I propose a new model of electoral competition that possesses very useful properties to tackle problems of type (i), and I compare strength and weaknesses of this new theoretical framework with the ones that characterize the alternatives in the literature. In the following chapters I present the applications of the theoretical analysis. First, in chapter 3 I study the relationship between income inequality and size of the government. Specifically, I extent the analysis by Meltzer and Richard (1981) to the case in which more than one public spending policy is available to voters. Second, in chapter 4 I analyse the effects of population ageing on collective choices over immigration policies and fiscal policies. A different approach underpins the analysis in chapter 5, in which I study the relationship between income inequality and public investment in education. One has to tackle both issues of type (i) and of type (ii) to analyse such relationship, which implies that the theoretical framework proposed in chapter 2 cannot be successfully employed in this case. Thus, I adopt a more traditional model of electoral competition, namely a *Probabilistic Voting Model*. This approach proves useful to tackle this specific question, but it can only deliver a limited set of analytical results, because of reasons that I describe in section 2.5.5. A common feature of chapters 3, 4 and 5 is to show that the extension of the analysis proposed by existing studies to multiple endogenous policy dimensions can radically change the trade-offs faced by voters. Such exercise allows me to identify the relevant economic channels that drive voters' choices and to shed light on some patterns observed in the data regarding each specific question. Lastly, in chapter 6 I highlight some of the achievements and of the shortcoming of the analysis presented in the previous chapters, and I identify some directions of future research. The theoretical tools proposed in this thesis to address an heterogeneous set of questions are examples of a large and growing literature that explores possible alternatives to the traditional *Downsian* framework. An extensive review of this literature is provided in chapter 2 of this thesis. An important contribution of this thesis is to show that - in order to capture the correct trade-offs that underpin several important questions in the Political Economy literature - it is crucial to relax some of the restriction that are commonly imposed in studies that employ the *Downsian* model, in particular the assumption of a unidimensional of policy space. This is shown to be the case for the three questions that I analyse in chapters 3, 4 and 5 of this thesis. Most papers in the literature - because of the technical issues mentioned above - address these and other questions abstracting from the role played by the interaction among multiple endogenous policy dimensions. As a result, such studies often overlook some crucial trade-offs that underpin the choices made by voters and obtain empirically controversial predictions. In this thesis I show that, if one possesses a tool to tackle the problem of multidimensionality, then the predictions of simple Political Economy models in the literature can be often reconciled with the empirical evidence. I prove that this is the case for some of the most important questions that the recent literature has tried to address with limited success. Moreover, the extension of the analysis to allow for a multidimensional policy space can help in providing more convincing economic channels to explain the patterns observed in the data. Lastly, the new approach proposed in this thesis can be employed to derive new testable implications of existing economic models, and to analyse entirely new questions. Thus, this thesis provides not only a valuable contribution to the literature, but also a set of a promising inputs for future theoretical and empirical research.

2 Generalized Comparative Statics for Voting Models

I investigate the equilibrium properties of a deterministic voting model in which the policy space is multidimensional and politicians have limited ability to commit to platforms. Specifically, a politician running alone can only offer his ideal policy. Voters can form coalitions to increase the commitment ability of politicians. Coalition structures are required to be stable in any equilibrium. This analysis is useful to answer a large class of Political Economy questions in which the multidimensional nature of the policy is crucial to model voters' trade-offs. I show that, under suitable restrictions on voter preferences, a Median Voter Theorem holds. The main result consist of two monotone comparative statics results for the equilibrium policy outcome. Moreover, I characterize the types of coalitions that can be stable in an equilibrium. Lastly, I show how this model relates to popular alternatives in the literature, and that the main result is robust to a variety of different assumptions about the notion of stability.

JEL classification: D72, C71.

Keywords: Multidimensional policy space, Coalitions, Median Voter.

2.1 Introduction

The model of electoral competition proposed by Downs (1957) is a simple and useful tool that has proved to be extremely successful in the Political Economy literature. Under suitable assumptions on voter preferences the model delivers a very sharp result known as the *Median Voter Theorem.* Such result states that in the unique equilibrium the median voter is a *Condorcet Winner*, i.e. he cannot be defeated by any other individual candidate, and that the policy that is chosen in equilibrium is the one that is most preferred by the median voter. The theorem implies that the levels and the comparative statics of the equilibrium policy outcome reduces to the ones of a single pivotal individual. As a consequence, the predictions of the model regarding the comparative statics of the equilibrium policy outcome are typically easy to derive and interpret. These desirable features made this framework very popular in the literature in Political Economy, such that the median voter result has been applied to an incredible variety of questions. Examples are the analysis of the relationship between income inequality and size of governmental intervention in redistributive policies (Meltzer and Richard, 1981), the study of the determinants of immigration policies (Razin and Sadka, 1999), of the extent of taxation on different tyes of income (Bassetto and Benhabib, 2006), and many more. Unfortunately the conditions under which this useful result of the Downsian model holds become extremely restrictive if the policy space is multidimensional. Specifically, the preference restrictions required in order to ensure the existence of a *Condorcet Winner* over a multidimensional choice domain are so demanding that the adoption of such framework to any relevant economic question becomes almost impossible (see section 2). The aim of this paper is to provide a tool that shares the desirable features of the Downsian model regarding the characterization of the equilibrium policy outcome, but that can be easily applied to problems in which the policy space is multidimensional. Following a successful stream of literature (Roemer 1999, Levy 2004, 2005), the approach adopted in this paper in order to achieve the target is to recognize that the political interaction in democratic countries involves a number of actors (voters, politicians) and institutions (e.g. political parties) that interact strategically, and that such interaction is affected by commitment issues. Specifically, I assume that individual politicians have limited ability to commit to specific policy platforms, because they cannot write binding contracts with their electors. As a result, a single citizen-candidate can only credibly propose his own ideal policy. Nevertheless, the ability of political agents to commit to platforms other than their own ideal points is enhanced by the existence of institutions, such as *coalitions* or *political parties*, that can ensure the credibility of commitment through internal self-enforcing agreements. In line with most of the recent literature (e.g. Levy 2004, 2005, Roemer 1999), I assume that a coalition can propose any policy in the Pareto set of its members. I do not explicitly model the process of coalition formation, but I require coalitions to be stable in equilibrium. The notion of stability adopted - that I define formally in section 3 - admits a simple intuitive description. Specifically, a coalition $\mathscr{A}$ is stable if and only if it can credibly commit to a policy platform x such that, given the policy proposed by other coalitions, there is no other platform x' in the policy space that possesses the following features: (i) x' makes each member of a subcoalition $\mathscr{A}' \subseteq \mathscr{A}$ strictly better off with respect to x; (ii) x' is in the Pareto set of the subcoalition $\mathscr{A}'$; (iii) there is no policy x'' in the Pareto set of the complementary subcoalition $\mathscr{A} \backslash \mathscr{A}'$ that can defeat x' by majority voting. This informal description suggests that this is a realtively weak concept of stability. The intuition is that a coalition is not stable if, for any policy platform that this coalition can put forward, there is a subcoalition that can deviate and credibly propose another platform that makes all its member strictly better off, and such that the remaining members of the original coalition do not have access to any alternative that can discourage such deviation. The main results of the paper are robust to stronger stability requirements, such as coalitions being merger-proof, or less extreme assumptions about the ability of the non-deviating members of a coalition to punish the potential deviators.

In this paper I show that under this notion of stability and some specific assumptions on individual preferences - *Supermodularity* and *Strictly Increasing Differences* - a Median Voter Theorem holds in a multidimensional policy space. As a consequence, the equilibrium policy outcome exhibits some monotone comparative statics properties. These results can be used to shed light on the role played by the interaction among mutiple policy dimensions in shaping the equilibrium outcome of some common applications in the literature, such as the analysis of the relationship between income inequality and size of redistribution. Moreover, it can be adopted to answer questions that have been proved difficult to tackle because of the multidimensional nature of the political choice faced by voters. Examples of applications of this framework are provided in the next two chapters. The paper is structured as follows. Section 2 summarizes the existing literature about models of Political Economy in a multidimensional policy space and highlights why none of the existing models is suitable to analyze sufficiently complex problems of comparative statics. In section 3 I describe the model of electoral competition and the notion of stability that I adopt in the rest of the paper. In section 4 I present the main results of this paper: the Generalized Median Voter Theorem and the Monotone Comparative statics of the equilibrium outcome. Moreover, I describe some features of the coalition structures that can be sustained in an equilibrium of the voting game, and some desirable properties of the collective choice rule induced by the model. I also show that the main results are robust to alternative settings. Section 5 describes similarities and differences between the model proposed in this paper and some popular alternatives in the literature, highlighting the advantages and disadvantages of this new framework in the study of comparative statics. Sections 6 concludes providing some comments about the strengths and the weaknesses of the results in the paper and suggesting some promising directions for future research.

2.2 Related Literature

The seminal contributions of Hotelling (1929), Black (1948) and Downs (1957) have introduced in the Political Economy literature the so-called Downsian models of political competition, which has proved to be extremely successful and it is still commonly used in recent applications. The reason for this success relies in the simple and powerful result that this model delivers under suitable restrictions on individual preferences: the *Median Voter Theorem*. The result states that, is voter preferences satisfy some ordinal restrictions such as *single peakedness*, then there exists a unique platform in the policy space that is weakly preferred to any alternative by a majority of voters in a pairwise comparison. This imples that a candidate proposing such policy platform is a *Condorcet*

Winner of the voting game. Moreover, such winning platform is the ideal policy of a specific individual (*Pivotal Voter*), who is the individual that possesses median preferences with respect to the total order induced by the preference restriction. This implies in turn that all predictions about levels and comparative statics of the equilibrium outcomes are very easy to derive and to interpret, because they reduce to the analysis of the preferences of the pivotal individual. For instance, suppose that one wants to analyse the effect on the equilibrium policy outcome of a shock on the distribution of voters, keeping all the other features of the economy unchanged. It is sufficient to identify the median individual before and after the shock, and compare the two ideal policies to obtain the desired comparative statics result. Unfortunately, the class of preference restrictions that are sufficient to ensure that the Median Voter Theorem holds, which are relatively weak and easy to impose whenever preferences are defined over a unidimensional choice domain, tend to be satisfied only in very extreme cases if the policy space is multidimensional. Specifically, the early papers by Plott (1967), Tullock (1967) and Devis *et al.* (1972) have established rather restrictive conditions for the existence of a *Condorcet Winner* in the this case. Grandmont (1978) has elegantly generalized these conditions with the concept of *Intermediate Relations*. The use of Grandmont's result in Political Economy applications is restricted to simple problems of redistribution (e.g. Borge, Rattsø, 2004) because of the extreme constraints that it imposes on preference heterogeneity. These requirements are way too restrictive for applications in which different subgroups of the voting population have sufficiently heterogeneous preferences over the set of available policies[1]. These results suggest that, in the analysis of collective choice over multidimensional choice domains, it is necessary to depart from the traditional framework. Alternatives to Downsian voting models are popular in the literature, and they have been succesfully adopted to answer various questions in the field. The downside of all these approaches is that they do not usually deliver sharp analytical predictions about the comparative statics of the equilibrium policy outcome, except for very specific cases. In the remaining part of this section I review the literature about models of electoral competition over multidimensional choice domains. A more detailed discussion about the similarities and differences between the framework proposed in this paper and the main alternatives in the literature is provided in Section 5. A first and popular alternative is the Citizen-Candidate model, first proposed by Osborne and Slivinski (1996) and Besley and Coate (1997). This class of models is based on the assumption each voter can run for elections but she cannot commit to any platform that is not in the set of her ideal policies. Under this rather restrictive assumption the existence of a political equilibrium is ensured, but multiplicity

[1]In section 5 I provide an example of why the Grandmont conditions usually fail to apply if the policy space is multidimensional, and how the conditions in Grandmont's paper compare with the ones assumed here.

of equilibria is a typical outcome[2]. For instance, for a given set of primitives, there may be equilibria with only one candidate running unopposed, equilibria with two candidates or more, and each of these cases is characterized by a different set of policies that are implemented in equilibrium with positive probability. This implies that such a model - without further restrictions - is not suitable to answer questions about policy outcomes and their comparative statics, because the set of policy vectors that can be sustained in equilibrium is usually too large to deliver any useful prediction. The problem of multiplicity of equilibria is shared with the Party Unanimity Nash Equilibrium (PUNE) proposed by Roemer (1999). In this case the set of equilibrium outcomes is a usually rather large set, with a typical element being a vector containing the platform proposed by each of the two parties that are assumed to compete in the elections. Because the outcome of elections is uncertain, in equilibrium a very large set of policies can be implemented with positive probability. As a consequence, for most applications this framework does not deliver sharp comparative statics result and needs to rely on simulations to formulate predictions about the equilibrium policy outcome. The model developed by Levy (2004, 2005) exploits the role of institutions, such as political parties, in shaping the political process. Her approach is the one that is the most similar to the one proposed in this paper. The crucial assumption is that politicians running for elections as single candidates cannot credibly commit to any platform other then their ideal policy, but they can expand their ability to commit to policies by forming coalitions (called *political parties*). She assumes that each coalition can propose any policy that is in the Pareto set of its members. This assumption is consistent with the idea that within parties members can achieve self-enforcing agreements. Lastly, she adopts a notion of coalition stability borrowed from Ray and Vohra (1998). This set of assumptions ensures the existence of an equilibrium in a multimensional policy space even if the individual preferences are relatively complex. For instance, individuals are allowed to differ in multiple (possibly orthogonal) dimensions. For the purposes of this paper, the downside of Levy's approach is that it allows for a large multiplicity of equilibria, unless voter preferences and policy space are strongly restricted. Lastly, a relative large literature adopts Probabilistic Voting Models (Lindbeck and Weibull 1987, Enelow and Hinich 1989, Banks and Duggan 2005). Such a framework ensure the existence and uniqueness of a voting equilibrium under not very restrictive assumptions. These models deliver a simple analytical characterization of the equilibrium policy outcome under specific preference restriction, for instance if voter preferences are Euclidean. For more general classes of preferences, this framework does not deliver an equally straightforward characterization. The intuition is that in Probabilistic Models the equilibrium policy outcome depends in principle on the preferences of all voters, and not only - as in

[2] There is a particular case, highlighted in Besley and Coate (1997) in which the Citizen-Candidate model admits an equilibium that - under the same preference restrictions assumed in this paper - delivers sharp predictions similar to the one I present in section 4. In section 5.2 I illustrate this case and I show that it can be interpreted as a particular case of the more general setting proposed here.

Downsian models - on the ones of a specific individual. Thus, the direction of any comparative statics exercise that involves a shock on the distribution of voters typically depends on the entire distribution. As a result, analytical predictions can be derived only for specific kinds of shocks on the distribution of preferences. Example of this approach are in the papers by de la Croix and Doepke (2006) and by Dotti (2014). These papers analyse a model of public provision of a private good in which voters differ uniquely in their income. They derive analytical results for comparative statics of the equilibrium outcome induced by a marginal mean preserving spread in the income distribution of voters. On the other hand, for a more general class of shocks on voters' preference distribution, there are no analytical comparative static results for probabilistic voting models. Because of this unappealing feature, in this paper I assume that voting decision are deterministic. I propose a model of electoral competition that shares the characteristics of the one proposed by Levy (2004) in terms of the role played by institutions, and I impose preference restrictions borrowed from the literature of generalized comparative statics (Milgrom and Shannon 1994, Quah 2007). Namely, I assume that preferences satisfy *Supermodularity* and *Strictly Increasing Differences* over a choice set that is a *convex sublattice*. These two modeling choices together are sufficient to deliver a *pivotal voter* result that resembles the one of unidimensional Downsian models. Specifically, a multidimensional version of the *Median Voter Theorem* holds, and this result eases the study of comparative statics.

2.3 The Voting Model

In this section I describe the setup of the voting model. I define the notion of coalition stability that I am going to adopt and I provide a definition for the equilibrium concept. Lastly, I state two crucial restrictions I need to impose on voter preferences in order to achieve the main results of this paper, that are presented in section 4.

2.3.1 Setting

Consider a voting game with N voters[3] such that each voter $i \in \mathscr{N}$ is denoted by a vector of parameters $\theta^i \in \Theta$. Assume $(\Theta, \leqslant)$ is a totally ordered set for some transitive, reflexive, antisymmetric order relation $\leqslant$. This allows one to establish a total order in the set of players $\mathscr{N}$, such that for all $i, j \in \mathscr{N}$, one gets $i \leq j$ if and only if $\theta^i \leqslant \theta^j$. For instance, suppose θ^i is individual i's income, then $\theta^i \in [\underline{\theta}, \overline{\theta}]$ and Θ is a totally ordered

[3]In this chapter I assume that the number of voters is discrete. All the results hold if one assumes a continuum of voters of size N.

set under the order relation $\leq$. Each individual $i \in \mathscr{N}$ is endowed with a reflexive, complete and transitive preference ordering $\succeq^i$ that can be represented by an utility function $V : X \times \Theta \times \Phi \rightarrow R$, which is jointly continuous in x and *θ-concave*[4], where $\varphi \in \Phi$ is a vector of parameters that do not differ across individuals. The policy space X is a subset of the the d-dimensional real space R^d. In order to characterize X it is useful to recall some definitions. Let $(L, \leqslant)$ be a partially ordered set, with the transitive, reflexive, antisymmetric order relation $\leqslant$. For x' and x'' elements of L, let $x' \vee x''$ denote the least upper bound, or join, of x' and x'' in X, if it exists, and let $x' \wedge x''$ denote the greatest lower bound, or meet of x' and x'' in L, if it exists. The set L is a lattice if for every pair of elements x' and x'' in L, the join $x' \vee x''$ and meet $x' \wedge x''$ do exist as elements of L. Similarly, a subset X of L is a sublattice of L if X is closed under the operations meet and join. A sublattice X of a lattice L is a convex sublattice of L, if $x' \leqslant x'' \leqslant x'''$ and x', x''' in X implies that x'' belongs to X, for all elements x', x'', x''' in L. Lastly, a sublattice X of L is complete if for every nonempty subset X' of X, $\inf(X')$ and $\sup(X')$ both exist and are elements of X. Recall the d-dimensional real space R^d is a partially ordered set under the transitive, reflexive, antisymmetric order relation $\leq$[5]. Moreover $(R^d, \leq)$ is a lattice given the definition above. These definitions provide all the elements that are needed in order to characterize the policy space X. Let $X \subseteq R^d$ be a convex sublattice of $(R^d, \leq)$, then $(X, \leq)$ is a partially ordered set with order relation $\leq$. An example of a policy space that satisfies my assumption is given by the family of sets $Y = \{x \in R^d | x_i \in [\underline{x}_j, \bar{x}_j], j = 1, 2, ..., d\}$ where $\underline{x}_i, \bar{x}_i \in R$ for all j. Subset of voters can form coalitions $\mathscr{A} \subseteq \mathscr{N}$. Each voter can be member of only one coalition, i.e. $\mathscr{A}^j \cap \mathscr{A}^k = \varnothing \ \forall \mathscr{A}^j \neq \mathscr{A}^k$. The role of coalitions in this model is to increase the commitment capacity of individuals over policies. Define $p_{X,\mathscr{A}}(x') \equiv \{x \in X : x \succeq^i x' \, \forall i \in \mathscr{A}, x \not\preceq^i x' \, \forall i \in \mathscr{A}\}$ to be the set of allocation in X that are Pareto superior to some vector $x' \in X$ for coalition $\mathscr{A}$. Denote with a^j the platform proposed by coalition $\mathscr{A}^j$. I assume that a coalition can propose any policy in the Pareto set of its members, i.e. $a^j \in P(\mathscr{A}^j)$ where $P(\mathscr{A}^j) \equiv \{x \in X : \ p_{X,\mathscr{A}^j}(x) = \varnothing\}$, or can choose to be inactive. If a coalition is a singleton then the Pareto set reduces to the set of ideal policies of its unique member (as in a citizen-candidate model).

2.3.2 Stability of a Coalition Structure

In order to define stability in this model I need to characterize a coalition structure and the preferences of each coalition. A coalition structure is defined as a partition $\mathbb{P}$ of $\mathscr{N}$,

[4]For any function f defined on the convex subset X of R^d , and any $(d \times 1)$ vector $v \neq 0$, we say that f is concave in direction of v if, for all $x \in X$, the map from the scalar s to $f(x+sv)$ is concave. (The domain of this map is taken to be the largest interval such that $x+sv$ lies in X.) We say that f is *$i-concave$* if it is concave in direction v for any $v > 0$ with the i^{th} element of v equal to 0. See Quah (2007).

[5]For $x', x'' \in R^d$ $x' \leq x''$ if and only if $x'_i \leq x''_i$ for all $i = 1, 2, ..., d$.

i.e. a set of subsets of $\mathscr{N}$ such that $\varnothing \notin \mathbb{P}$, $\cup_{\mathscr{A}\in\mathbb{P}(\mathscr{N})}\mathscr{A} = \mathscr{N}$ and if $\mathscr{A},^j \mathscr{A}^k \in \mathbb{P}(\mathscr{N})$ with $\mathscr{A}^j \neq \mathscr{A}^k$, then $\mathscr{A}^j \cap \mathscr{A}^k = \varnothing$. I define a complete social preference relation $\succ$ and $\succeq$, such that $\succ$ is irreflexive i.e. $x \not\succ x$ and $\succeq$ is reflexive i.e. $x \succeq x$ and the weak and strong relations are dual, i.e. $x' \succeq x'' \Leftrightarrow \neg(x' \succ x'')$ ($\succeq$ is not necessarily transitive)[6]. In particular, I am assuming *Majority Voting,* which is the most common and widely used criterion in order to establish a social preference relation. Formally $x' \succeq x''$ if and only if $\sum_{i=1}^{N} 1[V(x',\theta^i,\varphi) \geq V(x'',\theta^i,\varphi)] \geq N/2$, which implies that $x' \succ x''$ if and only if $\sum_{i=1}^{N} 1[V(x',\theta^i,\varphi) \geq V(x'',\theta^i,\varphi)] \geq N/2$ and $\sum_{i=1}^{N} 1[V(x'',\theta^i,\varphi) \geq V(x',\theta^i,\varphi)] < N/2$[7]. Given this preference relation we can define $SP_A(x') \equiv \{x \in A : x \succ x'\}$ - where $A \subseteq X$ - to be the strictly preferred set of x' in A and $K(A) \equiv \{x \in A : SP_A(x) = \varnothing\}$ to be the set of $SP-maximal$ alternatives in A, or the Core. The crucial aspect of the concept of stability relies on the idea of “deviator” or “credible threat”. Denote with $\nu(\mathbb{P})$ the number of coalitions that are part of $\mathbb{P}$ (including singletons). Recall that a^j represents the policy platform proposed by coalition $\mathscr{A}^j$. If the coalition chooses to be active, such platform must lie in the Pareto set of the coalition $\mathscr{A}^j$. Alternatively, the coalition may choose to be inactive. In such case, I assume that the coalition proposes a default platform (or neutral platform) x^0, such that $a^j \in [P(\mathscr{A}^j) \cup \{x^0\}]$. The default platform satisfies $x \succ^i x^0$ for all $i \in \mathscr{N}$ and all $x \in X$ and thus $x \succeq x^0$ for all $x \in X$, i.e. it is worse than any option available in the policy space for all voters. This assumption corresponds to a strong aversion of voters to the absence of decisions. I define a policy profile as the set containing the platforms proposed by each coalition in a given partition $\mathbb{P}$, i.e. $A_{\mathbb{P}} := \{a^1, a^2, ..., a^j, ..., a^{\nu(\mathbb{P})}\}$ with $a^j \in [P(\mathscr{A}^j) \cup \{x^0\}])$. Moreover, I define a winning policy as follows.

Definition 1. *(Winning policy) A policy vector $a^j \in A_{\mathbb{P}}$ is a winning policy if and only if it is in the Core of $(\mathscr{N}, A_{\mathbb{P}}, V)$, i.e. $a^j \in W(A_{\mathbb{P}})$ if and only if $a^j \in K(A_{\mathbb{P}})$.*

Given that the Social preference relation is given by Majority Voting this is equivalent to saying that, if a^j is a *Condorcet Winner*[8] over the set of alternatives $A_{\mathbb{P}}$, then $a^j \in W(A_{\mathbb{P}})$ implies $a^j \in K(A_{\mathbb{P}})$; if not, then a^j is a “weak” form of *Condorcet winner*, namely a policy vector that is weakly preferred to any alternative in $A_{\mathbb{P}}$ by a majority of voters (I will call it a *weak Condorcet winner*). I assume that, if no *weak Condorcet winner* exists over $A_{\mathbb{P}}$, then the default (neutral) policy x^0 is implemented. Lastly, define the binary relation $T_{\mathscr{A}^j, A_{\mathbb{P}}}$ on $X \times X$ as follows.

[6] Notice that the social preference relation $\succeq$ does not necessarily imply a *tournament* as defined in Dutta (1988), i.e. it is possible that $x' \neq x''$ and $x' \sim x''$.

[7] This definition of the majority rule applies to the case of a discrete number of voters. The analysis can be easily extended to the case in which the set of voters is a continuum.

[8] The relationship between the concepts of (strong) *Core* and *Condorcet Winner* is described in Ordershook, 1986, pp. 347-349.

Definition 2. *$xT_{\mathscr{A}^j,A_{\mathbb{P}}}x'$ if and only if (i) $x \in P(\mathscr{A}')$ for some $\mathscr{A}' \subseteq \mathscr{A}$, $\mathscr{A}' \neq \varnothing$, (ii) $x \succ^i w^*$ for all $i \in \mathscr{A}'$, $,\forall w^* \in W(\{a^1, a^2, ..., x', ..., a^{\nu(\mathbb{P})}\})$ (iii) $x \succeq x''$ for all $x'' \in P(\mathscr{A} \backslash \mathscr{A}') \cup (A_{\mathbb{P}} \backslash \{a^j\})$. Consider the set $T_{\mathscr{A}^j,A_{\mathbb{P}}}(x') \equiv \{x \in X : xT_{\mathscr{A}^j,A_{\mathbb{P}}}x'\}$. The set $T_{\mathscr{A}^j,A_{\mathbb{P}}}(x')$ is the set of deviators to policy x' for a coalition $\mathscr{A}$.*

$T_{\mathscr{A}^j,A_{\mathbb{P}}}(x')$ corresponds to the set of policies in $P(\mathscr{A}')$ that are strictly preferred to x' by each member of any subcoalition $\mathscr{A}' \subseteq \mathscr{A}$ and that are preferred by the society to any policy that can be proposed by the residual coalition $\mathscr{A} \backslash \mathscr{A}'$ and to any platform proposed by other coalitions. Using this concept we can define the T-Core (TK) to be the set of policies that do not face any "credible threat" from any subcoalition of $\mathscr{A}$, or more formally:

Definition 3. *The set $TK(\mathscr{A}, A_{\mathbb{P}}) = \{x \in [P(\mathscr{A}^j) \cup \{x^N\}] : T_{\mathscr{A},A_{\mathbb{P}}}(x) = \varnothing\}$, or T-Core, is the set of $T_{\mathscr{A},A_{\mathbb{P}}}-$maximal alternatives in $P(\mathscr{A})$.*

With this structure I can now define a concept of stability for a coalition structure in this game:

Definition 4. *A coalition $\mathscr{A}$ is stable if and only if $TK(\mathscr{A}, A_{\mathbb{P}})$ is nonempty.*

Example. It is useful to give an example of why a coalition that does not satisfy the definition above is unlikely to survive. Suppose $TK(\mathscr{A}) = \varnothing$. Then for any $x \in [P(\mathscr{A}^j) \cup \{x^0\}]$, $\exists x' \in P(\mathscr{A}')$ and $\mathscr{A}' \subseteq \mathscr{A}$ such that $x' \succ^i x \; \forall i \in \mathscr{A}'$ and $x' \succeq x'' \; \forall x'' \in P(\mathscr{A} \backslash \mathscr{A}') \cup (A_{\mathbb{P}} \backslash \{a^j\})$, i.e. there exists a subset of the coalition $\mathscr{A}$ and a policy $x' \in P(\mathscr{A}')$ such that x' is strictly preferred to x by all members of the subcoalition $\mathscr{A}'$ and x' is also preferred by the society as a whole to any policy x'' that the remaining part of the original coalition $\mathscr{A} \backslash \mathscr{A}'$ can propose and to all the platforms proposed by other coalitions. It is natural to consider this coalition structure unstable because for any policy chosen by this coalition in its Pareto Set (e.g. through some form of bargaining), the choice of this policy would not be self-enforcing. Specifically, a subcoalition $\mathscr{A}'$ can deviate and propose a different policy that makes each member of the subcoalition strictly better off, that can defeat by majority voting any platform proposed by other coalitions, and such that the remaining part of the original coalition $\mathscr{A} \backslash \mathscr{A}'$ cannot prevent this deviation because there is no feasible "punishment" policy that can prevent the deviation.

Definition 5. *A stable coalition structure is a partition $\mathbb{P}$ of $\mathscr{N}$ such that all the coalitions $\mathscr{A} \in \mathbb{P}$ are stable.*

2.3.3 Equilibrium

Define the set of stable coalition structures: $\mathrm{P}(\mathscr{N}) := \{\mathbb{P} | TK(\mathscr{A}, A_{\mathbb{P}}) \neq \varnothing \, \forall \mathscr{A} \in \mathbb{P}\}$ with typical element $\mathbb{P}^*$. I can now define an equilibrium for the voting game as follows.

Definition 6. *A pure strategy equilibrium is a coalition structure* $\mathbb{P}$*, a policy profile* $A_{\mathbb{P}}$ *and a winning policy* $w^* \in W(A_{\mathbb{P}}) \subseteq A_{\mathbb{P}}$ *such that: (i)* $\mathbb{P} \in \mathrm{P}(\mathscr{N})$ *is a stable coalition structure; (ii)* $a^j \in TK(\mathscr{A}^j)$ *for all* $\mathscr{A}^j \in \mathbb{P}$*; (iii) the set of winning policies* $W(A_{\mathbb{P}})$ *is nonempty.*

In other words in an equilibrium each coalition is stable and is represented by one of the policy vectors that makes it stable, and the winning policy is a (weak) *Condorcet Winner* of the reduced games in which the policy space is reduced to $A_{\mathbb{P}} \subseteq X$. As in Levy (2005), I assume that a coalition $\mathscr{A}^j$ in equilibrium is inactive if all its members are just indifferent between proposing a policy platform and not running at all, i.e $(\mathbb{P}, A_{\mathbb{P}}, w')$ is not an equilibrium if $\exists \mathscr{A}^j \in \mathbb{P}$ and $a^j \neq x^0$ such that $w' \in W(A_{\mathbb{P}} \backslash \{a^j\})$[9]. This assumption is consistent with the idea, not explicitly modelled here, of small costs of running for elections, and implies that at any equilibrium only one policy platform is proposed, i.e. $A_{\mathbb{P}} = \{a^j\}$ for some $j \in \{1,2,...,\nu(\mathbb{P})\}$. Notice that this is true because the equilibrium (if it exists) is a (weak) *Condorcet Winner*, and this implies that $W(A_{\mathbb{P}}) \subseteq W(A_{\mathbb{P}} \backslash \{a^j\})$ for all $a^j \notin W(A_{\mathbb{P}})$. Lastly, I define the set that is the object of a the comparative statics exercises of the paper.

Definition 7. *The set of equilibrium policies* $E(\mathscr{N})$ *is the union of the sets of policies that are chosen in any equilibrium of the voting game, i.e.* $E(\mathscr{N}) := \cup_{\mathbb{P} \in \mathrm{P}(\mathscr{N})} W(A_{\mathbb{P}})$.

This last definition completes the description of the voting process and of the equilibrium concept. In the next paragraph I describe some restrictions on voter preferences that ensure that this framework delivers sharp comparative statics results.

2.3.4 Preferences

In order to establish the result I need to restrict individual preferences. The kind of restrictions I impose are very common in the many fields of Economic Theory. Specifically, I assume *Supermodularity (SM)* and *Strictly Increasing Differences (SID).* Recall that individual preferences can be represented by a function $V : X \times \Theta \times \Phi \to R$. A function V satisfies:

1. *SM* in x if and only if $V(x' \vee x'', \theta^i, \varphi) - V(x', \theta^i, \varphi) \geq V(x'', \theta^i, \varphi) - V(x' \wedge x'', \theta^i, \varphi)$ for all $\theta^i \in \Theta$, for all $\varphi \in \Phi$ and for all $x', x'' \in X$.
2. *SID* in (x, θ^i) if and only if $V(x', \overline{\theta}, \varphi) - V(x'', \overline{\theta}, \varphi) > V(x', \underline{\theta}, \varphi) - V(x'', \underline{\theta}, \varphi)$ for all $x' \geq x''$ such that and $x', x'' \in X$, $x' \neq x''$, for all $\varphi \in \Phi$ and for all $\overline{\theta}, \underline{\theta} \in \Theta$ such that $\overline{\theta} > \underline{\theta}$.

[9] Notice that if $W(A_{\mathbb{P}})$ is not a singleton, then this tie-break rule implies an implicit restriction on voters' beliefs off-equilibrium. This turns out to be irrelevant for the main results whenever the voter with median θ possesses a unique ideal policy.

Notice that these conditions on preferences are slightly more restrictive relative to the ones assumed in Migrom and Shannon (1994), namely *Quasisupermodularity* and the *Single Crossing Property*. Thus, if *SM* and *SID* are satisfied, then also the conditions for Monotone Comparative Statics in their paper hold. Moreover, *SM* and *SID* are very general properties, but in the case of a twice differentiable objective function one can simply use the sufficient conditions in Milgrom and Shannon (1994) in order to verify that the function satisfies *SM* and *SID*, namely $\frac{\partial^2 V}{\partial x_i \partial x_j} \geq 0 \ \forall x \in X, \forall i \neq j$ and $\frac{\partial^2 V}{\partial x_i \partial \theta} > 0$ $\forall x \in X, \forall \theta \in \Theta, \forall i$. These sufficient conditions are usually easier to verify in comparison with the one implied by the definitions of *SM* and *SID*.

2.4 Results

Define the set of ideal policies $I(i) \equiv \{x | x \in \arg\max_{y \in X} V(y, \theta^i, \varphi)\}$,[10] and recall that the set of *equilibrium policies* is the union of all the policies that are winning policies in some coalitional equilibrium of the game. Notice that because Θ is a totally ordered set, one can identify a median element θ^v. The individual characterized by this value of the parameter is the *median voter* denoted by the index[11] v.

2.4.1 Main Results

The main results that are relevant for this analysis are stated in the following theorems:

Theorem 1. *(Median Voter Theorem). (i) A coalitional equilibrium of the voting game exists. (ii) In any equilibrium the set of winning policies is a subset of the set of ideal points of the median voter v. (iii) If the median voter has a unique ideal policy, this policy is going to be the one chosen in any equilibrium.*

Proof. See Appendix 2.7.1.

Theorem 2. *(Monotone Comparative Statics). The set of equilibrium policies of the voting game is (i) a sublattice of X which is (ii) monotonic nondecreasing in* θ^v.

Proof. See Appendix 2.7.1.

[10]Notice that the completeness of X implies compactness in the order-interval topology. On bounded sets in R^d, the order-interval topology coincides with the Euclidean topology (Birkhoff 1967). Hence $I(i) \neq \varnothing$ for all i.

[11]In the case of a discrete even number of voters I assume that the ties are broken in favor of the individual with the lower index. Different assumptions would not affect the results in the next paragraphs.

Lastly, consider a totally ordered subset $\Phi' \subseteq \Phi$ and suppose that the objective function $V(x,\theta,\varphi)$ satisfies *Increasing Differences (ID)* in (x,φ), namely $V(x',\theta,\overline{\varphi}) - V(x'',\theta,\overline{\varphi}) \geq V(x',\theta,\underline{\varphi}) - V(x'',\theta,\underline{\varphi})$ for all $x' \geq x''$, and for all $\overline{\varphi},\underline{\varphi} \in \Phi'$ such that $\overline{\varphi} \geq \underline{\varphi}$. Then I can state the following result:

Theorem 3. *(Monotone Comparative Statics 2). The set of equilibrium policies of the voting game is monotonic nondecreasing in φ.*

Proof. See Appendix 2.7.1.

The results in this sections provide a tool to analyze the effects of a shock on the distribution of voters or on a preference parameter on the policy outcome that emerges in a political equilibrium. One only has to verify that an economic model satisfies the conditions stated in this section and then use Theorem 2-3 to formulate the predictions about the comparative statics of the platform that is implemented in equilibrium. An interpretation of this Median Voter Theorem is proposed in Section 4.2. The notion of monotonicity is the same as in Milgrom and Shannon (1994) and it is related to the Strong Set order, namely given two sets Y,Z we say that Y is grater than or equal to Z in the Strong Set order ($Y \geq_s Z$) if for any $y \in Y$ and $z \in Z$ we have $y \vee z \in Y$ and $y \wedge z \in Z$. Notice that the results in Theorems 1-2-3 hold even if the individual objective function V is not differentiable and therefore the First Order Conditions of the maximization problem cannot be used in order to calculate the comparative statics of interest. Moreover, in comparison with Milgrom and Shannon (1994), the only additional restrictions on V I have imposed are joint continuity in x and $\theta - concavity$ (see section 3.1). Thus, the notion of monotonicity adopted is very general.

2.4.2 Stable Coalition Structures

In this section I provide some results about stable and unstable coalition structures in this framework. This characterization is interesting in order to understand the political content of the analytical results in Theorem 1, 2 and 3, and it is robust to the alternative stability concept in section 4.1.1. Assume that individual preferences satisfy *SM* and *SID* and recall that v is the index that denotes the median voter, i.e. the player with median $\theta \in \Theta$. I can state the following results.

Proposition 4. *(Lateral Coalitions). Any coalition $\mathscr{A}^j$ that includes either (a) only individuals with index $(i \leq v)$ or (b) only individuals with index $(j \geq v)$ is always stable. Therefore a coalition structure $\mathbb{P}$ is stable if each coalition $\mathscr{A}^i \in \mathbb{P}$ satisfies either (a) or (b).*

Proof. See Appendix 2.7.2.

Proposition 5. *(Central Coalitions). (i) Any coalition* $\mathscr{A}^j$ *that include both (a) individuals with index* $(i < v)$ *and (b) individuals with index* $(j > v)$ *plus at least one individual with index v is stable if at least one policy* $x^v \in I(v)$ *is in the Core of a game* $(\mathscr{N}, P(\mathscr{A}') \cup \{x^v\}, V)$ *for all* $\mathscr{A}' \subseteq \mathscr{A}$. *(ii) If the Core of the full game* $(\mathscr{N}, X, V)$ *is non-empty, then any "Central Coalition" is stable, including the Grand Coalition of all voters.*

Proof. See Appendix 2.7.2.

Following Levy (2004), I define a "Partisan Equilibrium" as follows.

Definition 8. *A Partisan Equilibrium is an equilibrium in which all party members vote for their party's platform, if it offers one (party members are not restricted in their votes if their party is inactive).*

This definition implies that in a Partisan Equilibrium no voter has a strict incentive to vote for a policy different from the one proposed by the coalition she is part of.

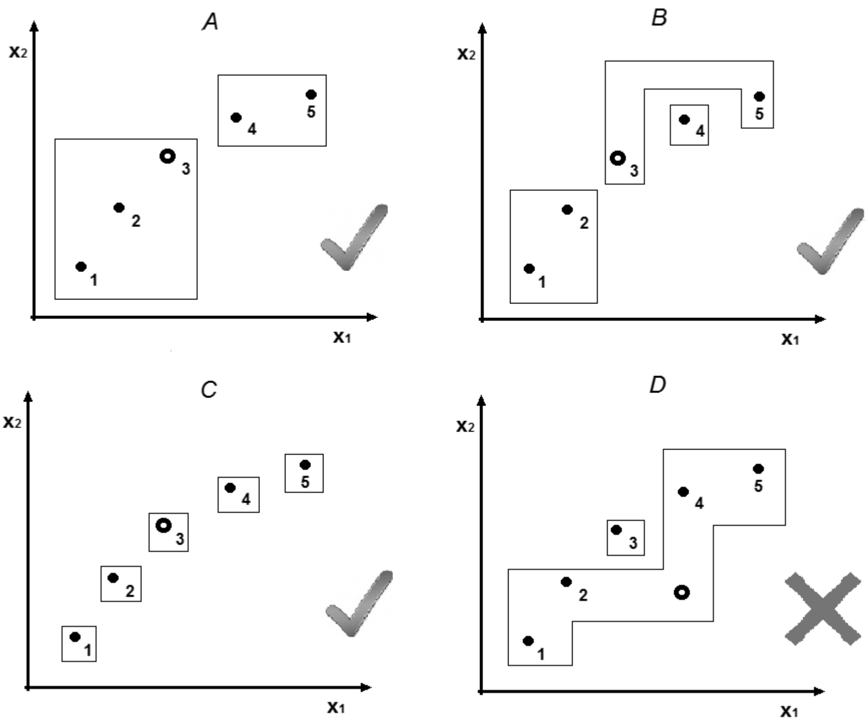
Proposition 6. *(Ends-Against-the-Middle Coalitions). (i) Any coalition* $\mathscr{A}^j$ *that includes both (a) individuals with index* $(i < v)$ *and (b) individuals with index* $(j > v)$ *but that does not include any individual with index v is stable only if either of the following is true: (1)* $a^j \in I(v)$*; (2)* $a^j \geq x^v$ *for all* $x^v \in I(v)$*; (3)* $a^j \leq x^v$ *for all* $x^v \in I(v)$ *(4)* $a^j = x^0$*. Therefore (ii) if* $I(v) \cap P(\mathscr{A}^j) = \varnothing$*, then there is no Partisan Equilibrium in which* $\mathscr{A}^j$ *is stable and* $a^j \neq x^0$.

Proof. See Appendix 2.7.2.

Results in Propositions 4-5-6 provide an intuitive understanding of why I obtain a median voter theorem. If preferences satisfy *SM* and *SID* and if coalitions are constrained to propose credible policies, then the only way to get any policy outcome different from the ideal policy of the median voter is a coalition of the *Ends-Against-the-Middle* type. This kind of coalition is not stable in this framework (unless a losing platform or no platform is offered) and therefore there is no equilibrium that violates the generalized median voter theorem. This is coherent with the idea that coalitions of individuals with opposite political views are less likely to occur, which is further discussed in section 5.

For illustrative purposes it is useful to analyse a simple example that is graphically represented in Fig. 2.1. In this example the policy space is two-dimensional, i.e. $X \subseteq R^2_+$, there are 5 players, i.e. $\mathscr{N} = \{1,2,3,4,5\}$, the median voter is individual $v = 3$, and each individual has a unique ideal point (black dots in Fig. 2.1). The black circle represents the policy platform proposed by a given coalition (if any). *Fig. 2.1 A*, *B*, *C* all represent stable coalition structures because either condition (a) or (b) in Proposition 4 is satisfied by each coalition. On one hand some results of similar models in the literature are shown to hold in this framework, such as the existence of a two-coalitions equilibrium (*Left vs Right* configuration, *Fig. 2.1 A*). On the other hand some less common structures can be stable, such as the one in *Fig. 2.1 B*. Notice that *Fig. 2.1 C* resembles the case of a Citizen-Candidate model (Besley and Coate, 1997), in the sense that each individual can only commit to her own ideal policy. The relationship between the model of electoral competition proposed in this paper and the Citizen-Candidate model is further discussed in section 5.2. Lastly, Fig. *2.1 D* represents a case in which conditions (a) and (b) in Proposition 4 are violated and one of the coalitions is in the *Ends-Against-the-Middle* form. Notice that such coalition does not satisfy either of the conditions 1-2-3 in Proposition 6, thus *Fig. 2.1 D* does not represent a stable coalition structure.

Figure 2.1. Stable and Unstable Coalition Structures



These examples show that, under the assumptions of *SM* and *SID*, even if the median voter is pivotal in all equilibria of the game, the types of coalition structures that can be stable in equilibrium are very heterogeneous. An element that characterizes the three types of structures that are always stable in *Fig. 2.1 A, B, C* is that all are made of *Lateral coalitions* as defined in Proposition 4. The political intuition that underpins the

Median Voter result is that if preferences are ordered by *SM* and *SID,* then in order to defeat the median voter it is necessarily an *Ends-Against-the-Middle* type of coalition, which is intuitively less stable than a "*Lateral Coalition*" or a "*Central Coalition*". In particular, Proposition 6 states that *Ends-Against-the-Middle* coalitions are not stable under the proposed stability concept, except in the four specific cases listed. This represents a major difference with the similar concept proposed by Levy (2004, 2005), which is discussed in detail in section 5.3. This feature of the coalition structures in equilibrium is a consequence of employing a stability concept in which agents take other coalitions' strategies as given. Specifically, this implies that if a coalition optimally chooses to be inactive given a certain partition, then it does not represent a threat for potential deviators in other coalitions. The main consequence of this feature of the model is that an *Ends-Against-the-Middle* coalition is unlikely to be sustained in an equilibrium. The modeling choices are motivated by the empirical evidence provided by the literature in Political Science. First of all, an outcome in the *Ends-Against-the-Middle* form - sometimes referred as *Coalition of Extremes* (COFEX) in such literature - is not commonly observed in the voting behaviour of members of elective bodies. For instance, Poole and Rosenthal (1997) analyse the history of voting behaviour of the members of the Congress and provide evidence that COFEX voting "does not to appear very often", "except in very unusual circumstances". Their results show little COFEX activity throughout Congress, and further reduction after 1950. Moreover, the same literature supports the claim that the party structure of a political system does not change very quickly. This suggests that assuming that politicians take other parties' strategies as given may be appropriate. In particular, there is evidence that politically inactive factions cannot easily become active and obtain representation. For instance, the birth of "genuine" new parties is a relatively rare event. Travits (2006) shows that, on average, only 1.4 new parties gained some representation in elections in 22 OECD democracies surveyed in 1960-2002. Moreover, among those new parties that have obtained some political representation, the "genuinely" new ones - i.e. the ones that are not the outcome of the division or the merger of other pre-exisiting parties - are a very small fraction (Krouwel, 2012). In the light of this evidence, one can conclude that the behaviour of political agents in the model seems to match fairly well some features of the political process in democratic countries. It also suggests that, if voters preferences are ordered along a single dimensions, then the structure of the political competition tends to be organized in some generalized "*Left vs Right*" form, even if the policy space is multidimensional.

2.4.3 Properties of the Voting Rule

It is useful to analyse the characteristics of the Social Choice rule generated by this voting model. The theory of voting typically attempts to evaluate voting rules systematically by examining which fundamental properties or axioms they satisfy. Even if this paper does not propose a voting rule, but a way of modeling the political process, such properties may still be desirable. Maskin and Dasgupta (2008) have listed as desirable the following five properties. The first is the so-called *(Weak) Pareto Property (PP),* the principle that if all voters prefer option x to option y, then y should not be chosen over x. The second property is *Anonymity (A),* the notion that no voter should have more influence on the outcome of an election than any other (anonymity is sometimes called the "one person–one vote" principle). The third is *Neutrality (N),* which requires that no candidate should get special treatment. The fourth is *Independence of Irrelevant Alternatives* (*IIA*)*,* which has attracted considerable attention since its emphasis by Nash (1950) and Arrow (1951). *IIA* dictates that if option x is chosen from the feasible set, and now some other option y is removed from that set, then x is still chosen. Lastly, *Decisiveness (D)*, which requires that a voting rule always picks a (possibly unique) winner. It is well-known that, because of Arrow's Impossibility Theorem (Arrow 1950), there is no Social Preference Ordering that satisfies at the same time all the properties listed above plus *Unrestricted domain (UD)*, the notion that all possible individual preference orderings should be allowed. Strictly related to Arrow's famous result is the Gibbard-Satterthwaite Theorem (Gibbard 1973; Satterthwaite 1975). Their result shows that for any deterministic voting system and for three or more candidates, one of the following three things must hold for every voting rule: *(i)* the rule is *dictatorial* (i.e., there is a single individual who can choose the winner), or *(ii)* there is some candidate who can never win under the rule, or *(iii)* the rule is susceptible to *tactical voting*, in the sense that there are conditions under which a voter would have an incentive to misrepresent their true preference orderings. The approach of this paper is the one of restricting the domain of voter preferences. Specifically, I assume that voter preferences satisfy *Supermodularity* and *Strictly Increasing Differences*. This implies that *UD* is not satisfied. Moreover, such preference restrictions are important to ensure that the voting rule satisfies other desirable properties. For instance, if *SM* and/or *SID* are not satisfied, then the voting process presented in this paper may imply violations of *D* and *IIA*. A violation of *D* can be shown with a simple example. Suppose there are three citizens, 1, 2, 3, with the following orderings for three policies, x^1, x^2 and x^3: $x^1 \succ_1 x^2 \succ_1 x^3$; $x^2 \succ_2 x^1 \succ_2 x^3$ and $x^3 \succ_3 x^1 \succ_3 x^2$. The tie-break rule implies that, if an equilibrium exists, it must involve only one active coalition. Because of such result, it is easy to show that, given other coalitions' strategies, there is always a profitable deviation for at least one player. It is interesting to notice that under the preference restrictions of *SM* and *SID*, no violation of *D* occurs. As previously stated, an equilibrium always exists in this case, such that a certain party wins and a certain policy $x \in I(v)$ is implemented. About possible violations of *IIA*, these cannot be ruled out and the reason is analogous to the one proposed by Bergson (1976) for the class of Citizen-Candidate models. Lastly, it is easy to verify that *A,N* and *PP* are satisfied. *A* and *N* are satisfied because all voters have same weight in determining the outcome of elections, and because the rule to win elections is the same for all parties, specifically the simple majority voting rule. The weak *PP* is always satisfied because the outcome is the ideal policy of one individual (Notice that the *Strong PP* may not be always satisfied if preferences are not strictly convex and mergers of parties are not allowed). Summarizing, the voting process assumed in this paper does not satisfies all desired properties in an unrestricted domain of preferences, but it performs fairly well under the preference restrictions under which the main results are derived.

2.4.4 Robustness

In this section I show that the main results of this paper in Theorems 1-2-3 are robust to slightly different assumptions about coalition stability. Namely, I explore the case in which voters and party members have preferences driven by ideological motives rather than by a selfish interest over the policy that is implemented in equilibrium, and the case in which deviations in the form of mergers between two or more coalitions are allowed.

2.4.4.1 Ideological Voters

If voters are purely policy motivated and sufficiently sophisticated, they evaluate the profitability of a deviation by comparing the policies that are implemented if no deviation occurs with the one that would be chosen by the society if the deviation occurs. In other words, deviating players calculate their payoffs anticipating the final outcome in terms of policy that is implemented after the elections. This may not be the case if voter preferences over policies are driven by ideological motives. In such case, members of a coalition may want their party to propose a policy that is close to their ideological preferences, independently of what other coalitions are going to propose. Nevertheless, stability of a coalition may still be an issue because a subcoalition may want to deviate in order to be able to propose a platform that is closer to their preferred policy. In this section I present a stability concept alternative to the one proposed in section 3.2, in which voters are driven by ideological motives. In this alternative concept stability is fully "internal" to coalitions, in the sense that the profitability of a deviation for a subgroup of members of a given coalition is independent of the strategies of other coalitions. This assumption captures the idea that ideologically motivated party members only care about how close a platform is to their ideology, and are not concerned about obtaining a victory in elections. The main results in Theorem 1-2-3 and Propositions 4-5-6 are unaffected. An important simplification with respect to the stability concept described in the previous section is that one does not need to assume that coalitions choose to be inactive whenever it is weakly better to do so for any of its members. Given that party members are driven by ideological motives, they optimally propose a platform $a^j \neq x^0$ even if they expect to lose elections with probability equal to 1. In this case, I need to propose alternative formulations of definitions 2-3-4-5-6 that describe the stability concept. In detail, define the relation $S_{\mathscr{A}}$, as follow:

Definition 2b. *$xS_{\mathscr{A}}x'$ if and only if (i) $x \in P(\mathscr{A}')$ for some $\mathscr{A}' \subseteq \mathscr{A}, \mathscr{A}' \neq \varnothing$, (ii) $x \succ^i x'$ for all $i \in \mathscr{A}'$, (iii) $x \succeq x''$ for all $x'' \in P(\mathscr{A} \backslash \mathscr{A}')$. The set $S_{\mathscr{A}}(x') \equiv \{x \in X : xS_{\mathscr{A}}x'\}$ is the set of deviators to policy x' for a coalition $\mathscr{A}$.*

This means that a deviation occurs if x cannot be defeated not only by any policy x'' that the remaining part of the coalition can credibly commit to propose, but also by all the policy proposed by other coalitions. Notice that under the relation $S_{\mathscr{A}}$, x is not required to make all the deviators better off in comparison with the platform that wins if no deviation occur (which was the case under the relation $T_{\mathscr{A},A_{\mathbb{P}}}$). This second stability concept uses a slightly different set of supporting policies:

Definition 3b. *The set $SK(\mathscr{A}) = \{x \in P(\mathscr{A}) : S_{\mathscr{A}}(x) = \varnothing\}$, or S-Core, is the set of $S_{\mathscr{A}}-$maximal alternatives in $P(\mathscr{A})$.*

Similarly to the first stability concept, a coalition $\mathscr{A}$ is stable if and only if $SK(\mathscr{A})$ is nonempty.

Definition 4b. *A coalition $\mathscr{A}$ is stable if and only if $SK(\mathscr{A})$ is nonempty.*

The set of stable coalition structures is now defined as $\mathrm{P}(\mathscr{N}) := \{\mathbb{P} | SK(\mathscr{A}, A_{\mathbb{P}})) \neq \varnothing \forall \mathscr{A} \in \mathbb{P}\}$.

Definition 5.b *A stable coalition structure is a partition $\mathbb{P}$ of $\mathscr{N}$ such that all the coalitions $\mathscr{A} \in \mathbb{P}$ are stable.*

The definition of a winning policy in Definition 1 is unchanged. A pure strategy equilibrium is defined as follows.

Definition 6b. *A pure strategy equilibrium is a coalition structure $\mathbb{P}$, a policy profile $A_{\mathbb{P}}$ and a winning policy $w^* \in W(A_{\mathbb{P}}) \subseteq A_{\mathbb{P}}$ such that: (i) $\mathbb{P} \in \mathrm{P}(\mathscr{N})$ is a stable coalition structure; (ii) $a^j \in SK(\mathscr{A}^j)$ for all $\mathscr{A}^j \in \mathbb{P}$; (iii) the set of winning policies $W(A_{\mathbb{P}})$ is nonempty.*

These definitions clarify why I do not need to assume that a coalition $\mathscr{A}^j$ in equilibrium chooses to be inactive if all its members are weakly better off with such choice. The intuition is that the additional condition for a policy to be a deviator in the baseline setting, namely $x \succeq a^k \; \forall a^k \in A_{\mathbb{P}} \backslash \{a^j\}$, only matters at an equilibrium for partitions in which $\mathscr{A}$ is a not winning coalition because in all other cases (in equilibrium) $a^k = x^0$ for all $k \neq j$. The latter cases do not matter on the equilibrium policy outcome, unless the deviation generates a winning coalition. Hence, even if the set of coalition structures that are stable using $TK(\mathscr{A}, A_{\mathbb{P}})$ may differ from the corresponding set under the notion of stability $SK(\mathscr{A})$, the implications about the equilibrium policy described in Theorems 1-2-3 are unaffected (see Appendix 2.7.1).

2.4.4.2 Mergers of Coalitions

Now suppose that deviations in the form of a merger of two (or more) existing coalitions are allowed under the same conditions under which a deviation of a subcoalition is assumed to be successful in section 3.2. Namely, such deviation is profitable if there is a policy vector x such that (i) $x \in P(\cup_{\mathscr{A}^J \in D} \mathscr{A}^j)$ for some $D \subseteq \mathbb{P}$, (ii) $x \succ^i w^*$ for all $i \in \mathscr{A}'$, $\forall w^* \in W(\{a^1, a^2, ..., x, ..., a^{\nu(\mathbb{P})}\})$, (iii) $x \succeq x''$ for all $x'' \in (A_{\mathbb{P}} \backslash \{a^j\})$. In words, a group of coalitions in $\mathbb{P}$ can merge in a new coalition $\cup_{\mathscr{A}^J \in D} \mathscr{A}^j$ and propose a platform x that makes all the members of the new coalition strictly better off relative to the policy that would be implemented in absence of such deviation, and the new proposed platform is a *Condorcet Winner* of the voting game given the platform proposed by other coalitions. The equilibrium concept induced by this new definition of deviation is a *refinement* of the original one, because it implies that new types of deviations are successful relative to the baseline. Thus, in order to show that the results in Theorems 1-2-3 holds in this new setting it is sufficient to show that an equilibrium still exists and that the set of equilibrium policies E preserves some simple ordinal properties. About the first part of the statement, consider any equilibrium under the original stability concept in which there is a winning (lateral) coalition $\mathscr{A}^v$ such that either (i) $i \in \mathscr{A}^v$ if and only if $i \geq v$ and $a^v = \sup\{I(v)\}$ or (ii) $i \in \mathscr{A}^v$ if and only if $i \leq v$ and $a^v = \inf\{I(v)\}$. It is easy to show that this is also an equilibrium in the case in which mergers can occur. The intuition is that all coalitions other than $\mathscr{A}^v$ cannot propose any platform capable of defeating a^v (even if they merge), and any merger involving coalition $\mathscr{A}^v$ cannot make all its member strictly better off, because the median voter v is already achieving her ideal policy under the current partition, thus she cannot get something that is strictly better. About the second part, the set of equilibrium policies E as previously defined may not be a convex sublattice of X, but it still contains $\sup\{I(v)\}$ and $\inf\{I(v)\}$. Moreover, all the other equilibrium platforms w must be such that $\inf\{I(v)\} \leq w \leq \sup\{I(v)\}$, because $w \in I(v)$ which is a complete sublattice of X. Thus, a weaker notion of monotonicity over the set of equilibria holds[12]. Moreover, if the preferences of voter v are strictly convex, then the set $I(v)$ is a singleton and therefore the main results in Theorems 1-2-3 hold in this different setting under the previously employed notion of Strong Set Order. Lastly notice that, if on one hand the set of partitions that can be stable in equilibrium is reduced relative to the baseline setup, on the other hand Propositions 4-5-6 hold even if mergers are allowed.

2.4.5 Extension: Constrained Problems

Several questions in Political Economy concern the analysis of voters' choice over a set of alternatives conditional on a costraint being satisfied. A typical example is the choice of spending policies and tax policies, which is made under the condition that the government has to break even. In such cases, the monotonicity results in Milgrom-Shannon (1994) may not apply because the choice set may not be a lattice. One can adopt two strategies to study the outcome of the voting process using an approach like the one shown in the previous sections of this chapter. The first possibility is to assume that the constraint is always binding at any policy that can be offered by some coalition. In most examples, policies in which the government budget constraint is not binding are typically never in the Pareto set of a coalition. Thus, such assumption is innocuous for the analysis because the option ruled out cannot be proposed by any coalition. Examples of this approach are proposed in the next two chapters of this work. Neverthless, after substituting the constraint into the objective function of a voter, it may not be always easy to show that *SM* and the *SID* are satisfied. In such cases, one can adopt a second approach, which involves the use of the Comparative Statics results for constrained maximization problems in Quah (2004, 2007). This involves a choice set X which is a convex sublattice of R^d and voters maximizing a concave objective function $V : X \times \Theta \to R$ subject to a constraint $B(x) \leq 0$ where $\theta^i \in \Theta$ is the parameter the identifies the voter's type and $B : X \to R$ is a convex function of x. The downside of this second approach is that the results of the coalitional equilibrium stated in Theorem 1, 2 and 3 may not apply for all the possible stable coalition structures. Neverthless, it is easy to show that the corresponding results hold for a Citizen-Candidate version of the voting model, i.e. a model in which only coalitions that are singletons are allowed. Thus, under specific preference restrictions, the result in Theorems 1 holds and some comparative statics result can be derived. In detail, assume that *(i)* $V(x,\theta) = \mathscr{V}(x) + \upsilon(x_j,\theta)$, in which *(ii)* $\upsilon(x_j,\theta)$ satisfies *ID* in (x_j,θ) on X, and *(iii)* $B(x)$ is a convex function of X. Notice that in this case there is no need to assume that the set $\{x|x \in X, B(x) \leq 0\}$ is itself a lattice. The intuition that underpins this

[12] Specifically, one needs to adopt the following weaker concept of set order. For any two sets S,S', S is higher than S' ($S \geq_W S'$) if for any $x \in S$, $\exists x' \in S'$ such that $x \geq x'$, and for any $y' \in S'$, $\exists y \in S$ such that $y' \leq y$.

result is simple. Proposition 8 in Quah (2004) implies that under the stated conditions the set of optimal choices $I(i)$ of individual with parameter θ^i has the (weakly) increasing property in θ^i for policy dimension j. This means that, for any two policies $x^k \in I(k)$ and $x^l \in I(\theta^l)$ for $\theta^k > \theta^k$, either $x^k \geq_j x^l$ of there exist $y' \in I(k)$ and $y'' \in I(l)$ such that $y^k \geq_j x^l$ and $y^l \leq_j x^k$. Thus, the sets of ideal points are ordered under the relation $\leq_j$ and it is intuitive that a median voter result will hold. Then one can state the following.

Theorem 7. *(Monotone Comparative Statics - Constraint Problems). If (i)$V(x,\theta) = \mathscr{V}(x) + \upsilon(x_j,\theta)$, in which (ii) $\upsilon(x_j,\theta)$ satisfies ID in (x_j,θ) on X, and (iii) $B(x)$ is a convex function of X, then the set of equilibrium policies of the voting game is nondecreasing in θ^v for policy dimension j.*

Proof. See Appendix 2.7.1.

The result in Theorem 7 represents an alternative tool to study problems of comparative statics of voting models. Unfortunately, the rather restrictive assumptions on voters' preferences that must be satisfied in order for such result to hold make the range of its possible applications very narrow. Thus, in all the examples proposed in the next chapters, I am going to adopt the first approach. Namely, I will transform the voter problem into an unconstrained maximization problem, rather than dealing directly with a constrained one.

2.5 Discussion

In this section I discuss similarities and differences between the voting model proposed in this paper and some alternatives that are popular in the theoretical literature. I show that the framework proposed here is usually more flexible, in the sense that the main comparative static results apply to a larger class of voter preferences relative to the alternatives in the literature.

2.5.1 Grandmont Conditions for Downsian Models

In this subsection I compare the sufficient conditions for existence of a *Condorcet Winner* in a multimensional choice domain with the ones that I have assumed in this paper to derive Theorems 1-2-3. The aim is to show that the class of preferences that admits a coalitional equilibrium is much larger than the one that ensures the existence of a *Condorcet Winner,* and that the latter are excessively restrictive to be succesfully adopted for most application in the literature. In a famous paper Grandmont (1978) has established sufficient conditions for the existence of a *Condorcet Winner*. Such conditions apply to a specific class of individual preferences that he named *Intermediate Preferences,* and that generalize previous results by Plott (1969) and by Davis *et al*. (1972). A formal definition for this class of preferences and details on how one can construct them are provided in Appendix 2.7.3. Unfortunately, the conditions for the existence of a *Condorcet Winner* remain extremely restrictive. Specifically, in order to have a transitive majority voting relation one needs that

> [...] the shape of the distribution of preferences has nice symmetry properties.

This sentence in Grandmont's paper means that the sufficient conditions for the existence of a Majority Voting Equilibrium are strictly related to specific restrictions on the distribution of voter preferences. In particular, his result can be interpreted as a generalization of the notion of *"unique median in all directions*" proposed by Davis *et al*. (1972, see also p. 326 in Grandmont 1978) for spatial voting models in which preferences are generalized Euclidean. The concept of unique median in all directions requires the existence, in a subset of the $d-$dimensional Euclidean Space, of a platform $x^* \in X$, such that all the hyperplanes passing through x^* divide the voters' probability distribution with respect to the preference for any two policies $y,z \in X$ into two sets of equal size. For instance, in the two-dimensional Euclidean space, if the ideal points are unformly distributed over a rectangle, it can be shown that all the hyperplanes passing through the centre of the rectangle divide it in two parts of equal size. Thus, in such example, the Lebuesgue measure corresponding to the size of the voting population with ideal point on each side of the hyperplane is always 1/2. How restrictive is this condition? In order to answer this questions it is easy to analyze the example of Euclidean preferences in a very simple scenario. Say there are 3 voters, $1,2,3$, with ideal points x^1, x^2, x^3 respectively. It is easy to show that the conditions for the existence of a *Condorcet Winner* are satisfied if and only if the three ideal points lie on a straight line (a proof is provided in Appendix 2.7.3). This example also clarifies why such conditions become severely restrictive if the choice domain has dimensionality greater than 1, and in particular much more restrictive than the conditions assumed in this paper. For instance, it is easy to show that, in a sublattice of the Euclidean space, Euclidean preferences satisfy *SM* and *SID* if the set of voters' ideal points is a *chain*. This is the case if, for any two platforms $x^i \in I(i)$ and $x^j \in I(j)$ for $i,j \in \mathscr{N}$, either $x^i \geq x^j$ or $x^i \leq x^j$ (or both). Thus, it is sufficient that the ideal points are *totally ordered*, but they do not have to lie on a straight line. Another way to compare the two set of conditions is the following. A consequence of Grandmont's result is that a *Condorcet Winner* exists if preferences admit a representation in the form (i) $V(x,\theta) = a(\theta^i)u(x)$, in which θ^i is a (possibly multidimensional) parameter capturing voter i's intensity of preferences. Notice that V in such form always satisfies *SM* and *SID* over a covex sublattice X[13]. The opposite is not true. Namely, not all functions that satisfy *SM* and *SID* over X admit a representation like the one above. Thus, for preferences defined over sublattices, the condition (i) is more restrictive than the one adopted in this paper. More importantly, for most applications in Political Economy the voter preferences are unlikely to meet condition (i), if one departs from very simple problems such as the one proposed in Borge and Rattsø (2004). If voter preferences do not satisfy these conditions, then a *Condorcet Winner* may not exist and this implies that one cannot use a Downsian model to characterize the policy outcome of the voting process. In such cases, the framework proposed in this paper can prove useful.

2.5.2 Citizen-Candidate Model

A class of models that allow for the existence of a political equilibrium even if the policy space is multidimensional is the one of Citizen-Candidate models (Besley and Coate 1997; Osborne and Slivinski, 1996). The crucial assumption of this class of models is that each voter can run for elections as a candidate, and that each candidate i can credibly commit only to a policy that is in the set of her ideal points $x^i \in M(i)$. In their model a citizen faces a cost δ to run for elections, and has preferences defined over the policy that is implemented. The main shortcoming of this class of models if one aims to get predictions about the policy choice of a certain group of individuals is the multiplicity of equilibria. But notice that, if one restricts the analysis to coalition structures in which all coalitions contain a single individual, then the model of electoral competition proposed in this paper resembles a Citizen-Candidate model in which the cost of running for election is zero. It is therefore intuitive that, if one imposes the same preference restrictions, then there must be an equilibrium in which the predictions of the two models are the same. In Corollary 2 (ii) of Besley and Coates (1997), one can find a hint of whether this is the case. They state that:

> (ii) if x^i is a strict Condorcet winner in the set of alternatives $\{x^j : j \in \mathscr{N}\}$ and if $x^i \neq x^0$, then a political equilibrium exists in which citizen i runs unopposed for sufficiently small δ[14].

[13] One needs to define a transformation of the taste parameter to ensure that the *ID* are strict.

[14] They denote with x^0 the deafult policy that is implemented if no candidate runs for election. This is similar to the definition of x^0 in this paper, except that they do not assume that $x \succ_i x^0$ by all i for all $x \in X$.

Consider one particular stable coalition structure in the model presented in this paper, namely the one in which each coalition is a singleton. In this case the Pareto set of each coalition coincides with the set of ideal points of its single member. In this setting, the assumptions about individual preferences (*SM* and *SID*) are sufficient to ensure that there is at least one $x^v \in I(v)$ that is a *Condorcet Winner* over the set $A_{\mathbb{P}}$, as a consequence of Theorem 1 and Definition 1. This implies that x^v is within the set of equilibrium policies under the notion of coalitional equilibrium. The fact that x^v is a *Condorcet Winner* also implies that the conditions of Corollary 2 (ii) in Besley and Coate's paper are satisfied, thus there is an equilibrium of the Citizen-Candidate model in which the same candidate is elected and the same policy is implemented. Lastly, notice that in the Citizen-Candidate model, even if preferences satisfy *SM* and *SID* and $\delta \to 0$, there may be other equilibria in which the policy implemented is not an ideal policy of the median voter. The reason is that in the Citizen-Candidate model, even if there is a *Condorcet Winner*, there are equilibria in which two (or more) candidates run for elections and each of them wins with positive probability. One can conclude that, if preferences satisfy *SM* and *SID*, then (i) there is an equilibrium in Besley and Coate's model for $\delta \to 0$ that resembles the coalitional equilibrium proposed in this paper; (ii) if there is a *Condorcet Winner* in the set of voters' ideal policies, then the Citizen-Candidate model allows for equilibria in which platforms different from the *Condorcet Winner* are implemented, while the coalitional equilibrium does not.

2.5.3 Levy's Coalitional Equilibrium

The theoretical framework described in this paper possesses several similarities with the one in Levy (2004, 2005). She proposes a model of political parties based on the notion of coalition stability in Ray and Vohra (1997). The main difference relies in one aspect of the notion of stability. That is, in Levy's paper the equilibrium - and consequently the political outcomes - is defined given any party structure. Given the equilibrium outcome of each possible party structure, she analyzes which party structures are stable. A stable party structure is an array of political parties in which no group of politicians wishes to quit its party and form a smaller one, thus inducing a different equilibrium outcome. Conversely, in the model I propose, the stability of the coalition structure and the equilibrium strategies of political agents are jointly determined, because potential deviators take strategies of other coalitions as given. The main consequence of this different assumption when voter preferences satisfy *Supermodularity* and *Strictly Increasing Differences* is that there is a class of coalition structures and corresponding equilibrium outcomes that are stable if the notion of equilibrium in Levy is adopted, and that are not stable if the equilibrium concept is the one proposed here. Specifically, in the former case *Ends-Against-the-Middle* coalitions proposing a policy platform that is not an ideal point of the median voter may be winning coalitions in some equilibrium, while in the latter case there is no equilibrium of this kind (see Proposition 6). The reason is that, in both models, if there is a winning coalition in the *Ends-Against-the-Middle* form, this implies that the platform proposed is a compromise between members of the two sides of the coalition. In such case moderate voters, including the median, optimally decide to be inactive. This implies that, given such choice, there is at least one side of the winning coalition that has incentive to deviate, rejecting the compromise policy and achieving an outcome that is closer to the preferences of the deviators. In the case of Levy anyway, such kind of deviations are not optimal because the strategy of moderate voters is not taken as given. Thus, the deviation is prevented by the threat that inactive players may become active and compete for elections. As highlighted in section 4.2, there is empirical evidence suggesting that the threat of inactive political agents may not be so relevant for the choices of politicians, and that *Ends-Against-the-Middle* coalitions are a very unlikely outcome on the political competition in democratic countries. One can conclude that, if voters preferences satisfy *Supermodularity* and *Strictly Increasing Differences*, then the equilibrium concept proposed in this paper delivers predictions about the set of equilibrium policy outcomes that are similar to the one in Levy. Specifically, there is a subset of equilibria under the notion of Levy that shares the same properties described in Theorems 1-2-3. The main difference is that the coalitional equilibrium rules out some outcomes that are stable in Levy - the ones in which an *Ends-Against-the-Middle* coalition is a winning coalition[15]- that are deemed to be less likely to occur by the recent empirical evidence.

2.5.4 Party Unanimity Nash Equilibrium (PUNE)

The way of modeling the political process proposed by Roemer (1999, 2001) and the corresponding equilibrium concept represent a useful tool to analyse voting models in which the policy space is multidimensional. The political process is modeled as a competition between two parties. Each party proposes a platform and there is uncertainty about the outcome of elections. The membership to each party is endogenously determined by the choice of policy platforms made by the parties in equilibrium. The equilibrium concept, named Party Unanimity Nash Equilibrium (PUNE), is based on the idea that - within each party - members can agree to propose a platform that is in the Pareto Set of its members[16]. The PUNE is a flexible concept and allows for voters to differ in two or more (possibly orthogonal) characteristics (Lee and Roemer, 2006).

[15] Notice that, even if one restricts the analysis to the particular class of voter preferences adopted in this paper, the equilibrium is not a refinement of the one in Levy, because the stable coalition structures that can support a certain equilibrium outcomes may not be the same under the two notions of stability.

[16] Roemer classifies party members in three categories, office-motivated, policy-motivated and an intermediate class of members that cares about both aspects. Conversely, I assume that all individuals are solely policy-motivated. Thus the Pareto set of a coalition in Roemer's paper does not coincide with the one proposed in this paper, except for the case in which all party members are in the first category.

Neverthless, such a model possesses features that makes it usually unsuitable to derive analytical results about the comparative statics of the policy outcome. The reason is that multiplicity of equilibria is a typical outcome of this model and the set of equilibria is a multidimensional manifold that does not usually admit a straightforward analytical characterization. Specifically, each equilibrium consist of a pair of policy platforms. The consequence it that in most cases one has to rely on simulation in order to understand how the set of equilibria move as a consequence of a shock affecting some characteristic of the voting population, such as a a change in the distribution of voter preferences. Because of these features this framework is not suitable to answer the specific question of this paper. Nevertheless, Roemer's model can prove useful to analyse voting environments characterized by strong restrictions on the policy space and on voter preferences (e.g. Roemer, 1999).

2.5.5 Probabilistic Voting Models

Lastly, the class of probabilistic voting models is often employed in the literature to analyse voting behaviour over a multidimensional policy space (Lindbeck and Weibull 1987, Enelow and Hinich 1989, Banks and Duggan 2005). The reason is that this kind of model of electoral competition dramatically eases the problems of existence of a voting equilibrium, that are typical of the deterministic voting framework. Unfortunately, the characterization of the equilibrium policy outcome in such class of models may not always be suitable to answer questions regarding comparative statics. A simple and useful characterization can be derived in the case of (i) the spatial models of elections (in which voter preferences are generalized Euclidean), and of (ii) applications in which voter preferences have specific features, such as symmetry conditions (Hinich, Ledyard, and Ordeshook 1973, Banks and Duggan 2005). Notice that in these cases, if one imposes the preference restrictions adopted in this paper, a monotone comparative static result similar to the one in Theorems 2-3 can be shown to hold, thus the main result of this paper is robust to this alternative setting. The reason is that under the specific restrictions assumed in these papers, the equilibrium outcome of probabilistic voting models reduces to a pivotal voter result (sometimes referred as *Mean Voter Theorem*, in contrast with median voter result that is typical of Downsian models). Thus, in those specific cases, the comparative statics of the equilibrium outcome is the same as the one of the ideal policy of a pivotal voter. In all the other cases, the derivation of analytical comparative statics results may not be straightforward. The reason is that the choice of policy in equilibrium depends on the entire distribution of voter preferences, and not only on a pivotal individual. In chapter 5 of this work I will show cases in which some comparative statics results can be derived for this class of models even in cases different from (i) and (ii) mentioned above. Nevertheless, the kind of comparative static exercises that can be performed is limited and the derivation requires specific preference restrictions. For instance, some papers (de la Croix and Doepke, 2009, Dotti, 2014) study the effects on the policy outcome of a marginal mean preserving spread in the income distribution of voters for specific classes of preferences. Conversely, the deterministic voting model proposed in this chapter allows for a large class of voter preferences and delivers fairly general comparative statics predictions as stated in Theorems 2-3. This will become clear in the next two chapter, in which I show how this framework can be adopted to provide an answer to some important questions in the Political Economy literature. Lastly, the Probabilistic Voting framework can prove useful to address theoretical issues other than the multidimensionality of the choice domain. For instance, in chapter 5 I show that it can tackle questions that cannot be easily addressed using the Downsian framework because voter preferences exhibit non-convexities.

2.6 Concluding remarks

This paper proposes a model of electoral competition in which agents cannot credibly commit to policy platforms before the elections take place. As a result, an individual running for election as a single candidate can only credibly commit to implement her own ideal policy. Moreover, in line with the literature, I assume that voters can form coalitions in order to increase the set of policy platforms that can be proposed in a credible way. Lastly, I require coalitions to be stable in an equilibrium, meaning that, given a certain coalition structure and the strategies of other coalitions, there is no subcoalition that can profitably deviate by creating a subcoalition. I show that the assumptions of *Supermodularity* and *Strictly Increasing Differences* of voters' objective functions are sufficient for the existence of a political equilibrium in a multidimensional policy space. Moreover, I show that under the same assumptions a version of the Median Voter Theorem holds. As a consequence, a monotone comparative statics result of the equilibrium outcomes is derived. I show that this result holds in various alternative settings, specifically (i) a version of the model in which mergers of coalitions are allowed, (ii) a version in which individual agents are idealogically motivated, and (iii) a version in which coalitions are not allowed and the political process resembles the one of a Citizen-Candidate model.

The main results stated in Theorems 2 and 3 of this paper represent a tool that can be applied to a large number of questions in Political Economy. For instance, it can be adopted to revisit many traditional questions in the theoretical literature. In several of those the multidimensionality of the policy space is crucial to shape the direction of the predictions. In most of the traditional studies, however, the policy space is restricted to a unique dimension in order to exploit the useful properties of Downsian voting models.

Such restriction may have profound effects on the predictions of the model. In some cases (see for instance Haupt and Peters, 1998), the sign of the comparative statics is entirely driven by the restrictions on the policy space. Examples are the traditional analysis of (i) the relationship between income inequality and size of redistributive policies (Meltzer and Richard, 1981), (ii) the study of the determinants of the degree of restrictiveness of immigration policies (Razin and Sadka, 1999), and (iii) the question of how the wealth distribution shapes the tax rates on labour and capital income. For instance, in the latter example, the intrinsic bi-dimensional nature of the problem has often constrained the analysis to a very restricted policy space (e.g. Benhabib and Bassetto, 2006). In the next sections of this work I show how the voting model presented in this chapter can be useful to analyse questions (i) and (ii) in a new way, and to deliver predictions that are more empirically sound relative to the ones derived using the traditional framework. Question (iii) represents a promising topic for future research. Lastly, the model is characterized by a sufficient degree of flexibility and its predictions, because of the pivotal voter result, are relatively easy to interpret. These desirable features make this framework potentially suitable to analyse several other questions in Political Economy. Moreover, it allows for various extensions, some of which are presented in this paper. Many other extensions are possible, so there is large scope for future research. Regarding this aspect, it is important to highlight that the model proposed in this paper relies on assumptions borrowed from the literature of generalized comparative statics (Milgrom and Shannon 1994, Quah, 2007). This literature has been largely exploited to study the equilibrium properties of games characterized by *strategic complementarities*. A typical result of this literature is that in this class of games, even if there may be multiplicity of equilibria, some monotone comparative statics properties are preserved. Thus, this framework represents a promising starting point for extensions aiming to tackle a larger set of Political Economy questions. The current analysis focuses on questions in which voters only choose a vector of policies and the outcome for each individual only depends upon the policy that is implemented in equilibrium. A natural extension of this framework would be the analysis of more complex voting games, for instance the ones in which competing elective institutions decide simultaneously their policies, and in which voters' strategies across different institutions are strategic complements. This may be useful, for instance, to study how policy changes in one country propagate to the neighbors, if such policy changes affect voters' trade-offs in the other countries. Such extension could be applied to study various kinds of public policies in which such interdependency may arise, such as immigration policies, corporate tax policies, etc. In conclusion, the framework proposed in this paper represents a useful tool, that allows one to analyse a potentially large range of applications without imposing excessively strong restrictions on the policy space. Its main downside is that its range of possible applications is restricted to a class of problem in which voter preferences possess some rather restrictive ordinal properties.

2.7 Appendix

2.7.1 Proof of Main Result

Theorem 1. *(Median Voter Theorem). (i) A coalitional equilibrium of the voting game exists. (ii) In any equilibrium the set of winning policies is a subset of the set of ideal points of the median voter v. (iii) If the median voter has a unique ideal policy, this policy is going to be the one chosen in any equilibrium.*

The proof to Theorem 1 procedes as follows. First, I am going to prove that the result holds under the alternative stability concept ($SK-Core$) presented in section 4.4.1. Then I am going to show that the result is unaffected if the baseline stability concept ($TK-Core$) is adopted. In order to prove this result we need to introduce some additional notation. Define $\theta_{\mathscr{A}} := \{\theta^i | i \in \mathscr{A}\}$ and $\theta_{\mathscr{A}} \in \Theta_{PS}$ where Θ_{PS} is the power set of Θ. Suppose the coalition $\mathscr{A}$ has k members. Consider a set of $k \times 1$ weighting vectors $\Lambda_{\mathscr{A}} \equiv \{\lambda^1_{\mathscr{A}}, \lambda^2_{\mathscr{A}}\lambda^3_{\mathscr{A}}, ..., \lambda^k_{\mathscr{A}}\}$ such that $\sum_{i\in\mathscr{A}} \lambda^i_{\mathscr{A}} = 1$ for each coalition $\mathscr{A}$ and a function $G: X \times \Lambda_{\mathscr{A}} \times \Theta_{PS} \times \Phi \to R$ defined as follows: $G(x, \lambda_{\mathscr{A}}, \theta_{\mathscr{A}}, \varphi) = \sum_{i\in\mathscr{A}} \lambda^i_{\mathscr{A}} V(x, \theta^i, \varphi)$. Lemmas 8, 9, 10 and 11 are results that will be used as part of the proofs of Lemma 12 and 13, which constitute the main part of the proof of Theorem 1.

Lemma 8. *If V is a continuous function of x and X is a convex set then any point $\tilde{x}$ in the Pareto set of $\mathscr{A}$ is a solution to* $\max_{x\in X} G(x, \lambda, \theta_{\mathscr{A}}, \varphi)$ *for some vector $\lambda_{\mathscr{A}}(\tilde{x}) \in \Lambda_{\mathscr{A}}$.*

Proof. M.W.G., Proposition 16.E.2.

I need to define four additional objects. For each $\tilde{x} \in P(\mathscr{A})$ define:

(i) a vector $\underline{\lambda}_{\mathscr{A}}(\tilde{x}, j)$ such that $\underline{\lambda}^i_{\mathscr{A}}(\tilde{x}, j) = \lambda^i_{\mathscr{A}}(\tilde{x})$ for all $i \in \mathscr{A}$ s.t. $\theta^i < \theta^j$, $\underline{\lambda}^i_{\mathscr{A}}(\tilde{x}, j) = 0$ for all $i \in \mathscr{A}$ s.t. $\theta^i > \theta^j$, $\underline{\lambda}^j_{\mathscr{A}}(\tilde{x}, j) = \sum_{\substack{i\in\mathscr{A} \\ i \geq j}} \lambda^i$;

(ii) a vector $\bar{\lambda}_{\mathscr{A}}(\tilde{x}, j)$ such that $\bar{\lambda}^i_{\mathscr{A}}(\tilde{x}, j) = \lambda^i_{\mathscr{A}}$ for all $i \in \mathscr{A}$ s.t. $\theta^i > \theta^j$, $\bar{\lambda}^i_{\mathscr{A}}(\tilde{x}, j) = 0$ for all $i \in \mathscr{A}$ s.t. $\theta^i < \theta^j$, $\bar{\lambda}^j_{\mathscr{A}}(\tilde{x}, j) = \sum_{\substack{i\in\mathscr{A} \\ i \leq j}} \lambda^i$;

(iii) the sets $\Lambda_{\mathscr{A}}(\tilde{x}, j) = \{\underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \lambda_{\mathscr{A}}(\tilde{x})\}$, $\Lambda'_{\mathscr{A}}(\tilde{x}, j) = \{\lambda_{\mathscr{A}}(\tilde{x}), \bar{\lambda}_{\mathscr{A}}(\tilde{x}, j)\}$;

(iv) an order relation $\leq^{\lambda}$ given by: $\lambda_1 \leq^{\lambda} \lambda_2$ iff $\lambda^i_1 \geq \lambda^i_2 \; \forall i \leq j$ and $\lambda^i_1 \leq \lambda^i_2 \; \forall i > j$. It follows that $(\Lambda_{\mathscr{A}}(\tilde{x}, j), \leq^{\lambda})$ and $(\Lambda'_{\mathscr{A}}(\tilde{x}, j), \leq^{\lambda})$ are totally ordered sets.

Lemma 9. *If V satisfies SM and SID then the Pareto Set $P(\mathscr{A})$ of a coalition of players $\mathscr{A} \subseteq \mathscr{N}$ is such that $y \in P(\mathscr{A})$ only if $y \geq \sup\{I(l)\}$ and $y \leq \inf\{I(h)\}$ where* $\min(\mathscr{A})$ *and $h = \max(\mathscr{A})$.*

Proof. Denote $\bar{x}^l = \sup\{I(l)\}$ and $\underline{x}^h = \inf\{I(h)\}$. Suppose $y \ngeq \bar{x}^l$ but $y \in P(\mathscr{A})$. Because of the optimality of $\bar{x}^k$ and because X is a lattice, it must be true that $V(\bar{x}^l, \theta^l) \geq V(y \wedge \bar{x}^l, \theta^l)$. Supermodularity implies $V(y \vee \bar{x}^l, \theta^l) \geq V(y, \theta^l)$. Notice that $y \ngeq \bar{x}^l$ implies $y \vee \bar{x}^l \neq y$. Hence Strictly Increasing Differences imply $V(y \vee \bar{x}^l, \theta^i, \varphi) > V(y, \theta^i, \varphi)\ \forall \theta^i > \theta^l$. Given that $\theta^i > \theta^l$ is true for all $(\theta \in \mathscr{A}) \cap (\theta \neq \theta^l)$ we have that $\exists x \in X$ such that $V(x, \theta, \varphi) \geq V(y, \theta, \varphi)\ \forall \theta \in \mathscr{A}$ and $V(x, \theta, \varphi) > V(y, \theta, \varphi)$ for at least one $\theta \in \mathscr{A}$, i.e $p_{X,\mathscr{A}}(y) \neq \varnothing$. Hence $y \notin P(\mathscr{A})$. Similarly one can show that $y \in P(\mathscr{A})$ only if $y \leq \underline{x}^h$. Q.E.D.

Lemma 10. *The function $G(x, \lambda, \theta_{\mathscr{A}}, \varphi)$ satisfies (i) SM in x and (ii) ID in (x, λ) over $\Lambda_{\mathscr{A}}(\tilde{x}, j)$ ($\Lambda'_{\mathscr{A}}(\tilde{x}, j)$) for all $x \in X$.*

Proof. (i) *SM*. G is the weighted sum of *SM* functions so it is supermodular (proof in Milgrom, Shannon, 1994). (ii) *ID*. Using the definition of *ID*, G satisfies the *ID* if and only if: $G(\bar{x}, \lambda, \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \lambda, \theta_{\mathscr{A}}, \varphi) \geq G(\bar{x}, \lambda, \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \lambda, \theta_{\mathscr{A}}, \varphi)\ \forall \bar{x} \geq \underline{x}, \lambda \in \Lambda_{\mathscr{A}}(\tilde{x}, j)$. Use the definitions of G and $\lambda_{\mathscr{A}}(\tilde{x})$and $\underline{\lambda}_{\mathscr{A}}(\tilde{x}, j)$:

$$[G(\bar{x}, \lambda_{\mathscr{A}}(\tilde{x}), \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \lambda_{\mathscr{A}}(\tilde{x}), \theta_{\mathscr{A}}, \varphi)] - [G(\bar{x}, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi)] =$$

$$= \Big(\sum_{\substack{i \in \mathscr{A} \\ i \geq j}} \lambda^i [V(\bar{x}, \theta^i, \varphi) - V(\underline{x}, \theta^i, \varphi)]\Big) - \Big(\sum_{\substack{i \in \mathscr{A} \\ i \geq j}} \lambda^i \Big)[V(\bar{x}, \theta^j, \varphi) - V(\underline{x}, \theta^j, \varphi)] =$$

$$= \sum_{\substack{i \in \mathscr{A} \\ i \geq j}} \lambda^i \left([V(\bar{x}, \theta^i, \varphi) - V(\underline{x}, \theta^i, \varphi)] - [V(\bar{x}, \theta^j, \varphi) - V(\underline{x}, \theta^j, \varphi)] \right)$$

Notice that $[V(\bar{x}, \theta^i, \varphi) - V(\underline{x}, \theta^i, \varphi)] - [V(\bar{x}, \theta^j, \varphi) - V(\underline{x}, \theta^j, \varphi)] \geq 0$ for $\forall i \geq j$ and $\lambda^i \geq 0\ \forall i$ hence the sum above is also weakly positive, which implies $[G(\bar{x}, \lambda_{\mathscr{A}}(\tilde{x}), \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \lambda_{\mathscr{A}}(\tilde{x}), \theta_{\mathscr{A}}, \varphi)] - [G(\bar{x}, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi) - G(\underline{x}, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi)] \geq 0$. Similarly one can show that this is also true for $\Lambda'_{\mathscr{A}}(\tilde{x}, j)$. Q.E.D.

Define $\widetilde{M}(\mathscr{A}, \lambda) := \arg\max_{x \in X, \lambda \in \Lambda} G(x, \lambda, \theta_{\mathscr{A}}, \varphi)$ and the sets $\mathscr{A}^{\leq j} := \{i | (i \in \mathscr{A}) \cap (i \leq j)\}$ and $\mathscr{A}^{\geq j} := \{i | (i \in \mathscr{A}) \cap (i \geq j)\}$ for some $j \in \mathscr{A}$.

Lemma 11. *(i) If $x' \in \widetilde{M}(\mathscr{A}, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j))$ and $x \in \widetilde{M}(\mathscr{A}, \lambda_{\mathscr{A}}(\tilde{x}))$, then either $x' = x$, or if $x' \neq x$, then $V(x', \theta^j, \varphi) \geq V(x, \theta^j, \varphi)$ and $V(x', \theta^i, \varphi) > V(x, \theta^i, \varphi)\ \forall i < j$ and $x' \leq x$.* Moreover, (ii) if $x \in \widetilde{M}(\mathscr{A}, \lambda_{\mathscr{A}}(\tilde{x}))$ and $x'' \in \widetilde{M}(\mathscr{A}, \bar{\lambda}_{\mathscr{A}}(\tilde{x}, j))$, then either $x'' = x$, or if $x'' \neq x$, then $V(x'', \theta^j, \varphi) \geq V(x, \theta^j, \varphi)$ and $V(x'', \theta^i, \varphi) > V(x, \theta^i, \varphi)\ \forall i > j$ *and* $x'' \geq x$.

Proof. The results $x' \leq x$ and $G(x', \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi) \geq G(x, \underline{\lambda}_{\mathscr{A}}(\tilde{x}, j), \theta_{\mathscr{A}}, \varphi)$ follow from Milgrom and Shannon 1994. Suppose $V(x', \theta^j, \varphi) < V(x, \theta^j, \varphi)$ and $x' \neq x$. Then it must be true that $\sum_{i\in\mathscr{A}} \underline{\lambda}^i_{\mathscr{A}}(\tilde{x},j)\left[V(x',\theta^i,\varphi)-V(x',\theta^i,\varphi)\right] > V(x',\theta^j,\varphi)-V(x,\theta^j,\varphi)$. Using $\sum_{i\in\mathscr{A}} \underline{\lambda}^i_{\mathscr{A}}(\tilde{x},j)=1$ the above can be rearranged as follows:

$$\sum_{i\in\mathscr{A}} \underline{\lambda}^i_{\mathscr{A}}(a,j)\left(\left[V(x',\theta^i,\varphi)-V(x',\theta^i,\varphi)\right]-\left[V(x',\theta^j,\varphi)-V(x,\theta^j,\varphi)\right]\right)>0$$

Notice that $x'\leq x$ and $i\leq j$ $\forall i\in\mathscr{A}$, hence *SID* implies $\left[V(x',\theta^i,\varphi)-V(x',\theta^i,\varphi)\right]-\left[V(x',\theta^j,\varphi)-V(x,\theta^j,\varphi)\right]\leq 0$ $\forall i\in\mathscr{A}$ and hence

$$\sum_{i\in\mathscr{A}} \underline{\lambda}^i_{\mathscr{A}}(a,j)\left(\left[V(x',\theta^i,\varphi)-V(x',\theta^i,\varphi)\right]-\left[V(x',\theta^j,\varphi)-V(x,\theta^j,\varphi)\right]\right)\leq 0$$

which leads to a contradiction. Hence it must be true that $V(x',\theta^j,\varphi)\geq V(x,\theta^j,\varphi)$. Given that $x'\leq x$, $x'\neq x$, *SID* implies $V(x',\theta^i,\varphi)>V(x,\theta^i,\varphi)$ $\forall i<j$. This statement implies that x is not part of the Pareto set of $\mathscr{A}^{\leq j}$ if $x'\neq x$. In the same way it is easy to show that the statement (ii) is also true. Q.E.D.

Lemma 12. *The coalition $\mathscr{A}^v$ (could be a singleton) that includes the median voter v is stable only if $a^v\in I(v)$.*

Proof. Suppose $a^v\notin I(v)$. This cannot be the case if $\theta^i=\theta^v$ for all members of $\mathscr{A}^v$. Hence consider the case in which there is at least one $i\in\mathscr{A}^v$ such that either $i>v$ or $i<v$ (or both). This situation is illustrated in Fig. SM.8.1 in the additional material.

(i) If $a^v\geq \underline{x}^v(\leq)$ for any $\underline{x}^v=\inf\{I(v)\}$ and $a^v\wedge x^v\in P(\mathscr{A}^v)$. $a^v\notin I(v)$ implies $V(x^v,\theta^m,\varphi)>V(a^v,\theta^v,\varphi)$. *SID* implies $V(x^v,\theta^i,\varphi)>V(a^v,\theta^i,\varphi)$ and $x^v\succ^i a^v$ $\forall i\in\mathscr{N}:\theta^i\leq\theta^v(\geq)$. Recall that any $c\in P(\mathscr{A}^v\backslash\mathscr{A}^{\leq v})$ is $c\geq\underline{x}^v$ (because of Lemma 9). Hence either $c\in I(v)$ or $\sum_{i=1}^n 1[V(\underline{x}^v,\theta^i,\varphi)>V(c,\theta^i,\varphi)]>n/2$ $\forall c\in P(\mathscr{A}^v\backslash\mathscr{A}^{\leq v})$ which implies $x^v\succ c\rightarrow a^v\notin SK(\mathscr{A}^v)$.

(ii) If $a^v\ngeq x^v, a^v\nleq x^v$. Consider $a^v\vee x^v(a^v\wedge x^v)$. Revealed preferences and $a^v\notin I(v)$ imply $V(x^v,\theta^v,\varphi)>V(a^v\vee x^v,\theta^v,\varphi)$. *SM* implies $V(a^v\wedge x^v,\theta^v,\varphi)>V(a^v,\theta^v,\varphi)$. *SID* implies $V(a^v\wedge x^v,\theta^i,\varphi)>V(a^v,\theta^i,\varphi)$ $\forall i\in\mathscr{N}:\theta^i\leq\theta^v$. Recall that any $c\in P(\mathscr{A}^v\backslash\mathscr{A}^{\leq v})$ is $c\geq\bar{x}^v\geq a^v\wedge x^v$ because of Lemma 9. Hence either $c\in I(v)$ or $\sum_{i=1}^n 1[V(a^v\wedge x^v,\theta^i_i,\varphi)\geq V(c,\theta^i,\varphi)]>n/2$ which implies $a^v\wedge x^v\succ c$.

If $a^v\wedge x^v$ is part of the Pareto set of $\mathscr{A}^{\leq v}$ this constitute a feasible and profitable deviation.

If and $a^v\wedge x^v\notin P(\mathscr{A}^{\leq v})$ (case described in Fig. SM.8.2 in the additional material), recall that X is a convex set and $V(x,\theta,\varphi)$ is $\theta-concave$, hence as $a^v\in A^v$ it has to be the solution to a problem in the form $a^v\in\arg\max_{x\in X} G(x,\lambda_{\mathscr{A}^v}(a^v),\theta_{\mathscr{A}^v},\varphi)$. Consider the following alternative: $\tilde{x}\in\widetilde{M}(\mathscr{A},\underline{\lambda}_{\mathscr{A}^v}(a^v,v))$ (see Lemma 11). We know from Lemma 11 that $\tilde{x}\leq a^v$. First of all notice that $\widetilde{M}(\mathscr{A},\underline{\lambda}_{\mathscr{A}^v}(a^v,v))=\widetilde{M}(\mathscr{A}',\lambda')$ for some λ', which implies that $\tilde{x}\in P(\mathscr{A}')$, i.e. it is in the Pareto set of $\mathscr{A}'$. One needs to show that $\tilde{x}\neq a^v$ and that $\tilde{x}\succeq^i a^v$ $\forall i\in\mathscr{A}'$. Suppose $\tilde{x}=a^v\rightarrow a^v\in P(\mathscr{A}')$. But from point (ii) we know that

$V(a^v \wedge x^v, \theta_i, \varphi) > V(a^v, \theta^i, \varphi) \forall i \in \mathscr{N} : \theta^i \leq \theta^v \rightarrow a^v \notin P(\mathscr{A}')$ $\rightarrow$Contradiction. Hence $\tilde{x} \neq a^v$ and $\tilde{x} \leq a^v$. Moreover, Lemma 9 implies $\tilde{x} \succeq^v a^v$. This means that $V(\tilde{x}, \theta^v, \varphi) \geq V(a^v, \theta^v, \varphi)$ and because $\tilde{x} \neq a^v$, the *SID* implies $V(\tilde{x}, \theta^i, \varphi) > V(a^v, \theta^i, \varphi)$ $\forall i \in \mathscr{N}$: $\theta^i \leq \theta^v$. Recall that any $c \in P(\mathscr{A}^v \backslash \mathscr{A}^{\leq v})$ is $c \geq \bar{x}^v \geq \tilde{x}$. Hence either $c \in M(v)$ or $\sum_{i=1}^n 1[V(\tilde{x}, \theta_i, \varphi) \geq V(c, \theta_i, \varphi)] > n/2$ which implies $\tilde{x} \succ c \rightarrow a^v \notin SK(\mathscr{A}^v)$.

One can also show that some kind of coalitions containing v are stable. Suppose $a^v \in I(v)$, and in particular say $a^v = \bar{x}^v = \sup\{I(v)\}$ $(a^v = \underline{x}^v = \inf\{I(v)\})$. Consider any coalition $\mathscr{A}^{\leq v}$ $(\mathscr{A}^{\geq v})$ such that $\theta^i \leq \theta^v(\geq)$ $\forall i \in \mathscr{A}^{\leq v}$. From Lemma 9 we know that any $b \in P(\mathscr{A}^{\leq v})$ it must be true that $b \leq x^v(\geq)$. Moreover, a deviation implies Optimality implies $V(x^v, \theta^v, \varphi) > V(b, \theta^v, \varphi)$. *SID* implies $V(x^v, \theta^i, \varphi) > V(b, \theta^i, \varphi)$ and $x^v \succeq^i b$ $\forall i \in \mathscr{N}$: $\theta^i \geq \theta^v(\leq)$. Hence $\sum_{i=1}^n 1[V(x^v, \theta^i, \varphi) > V(b, \theta^i, \varphi)] \geq n/2$ $\forall b \in P(\mathscr{A}^{\leq v})$ which implies $x^v \succ b$ $\forall b \in P(\mathscr{A}^v) \rightarrow a^v \in SK(\mathscr{A}^v)$.

Finally Consider any coalition $\mathscr{A}^v$ such that $\theta^i \leq \theta^v(\geq)$ $\forall i \in \mathscr{A}^v$ and $a^v = \bar{x}^v(\underline{x}^v)$. From Lemma 9 we know that any $b \in P(\mathscr{A}^v)$ it must be true that $b \leq \bar{x}^v(\geq \underline{x}^v)$. This implies $V(\bar{x}^v, \theta^v, \varphi) > V(b, \theta^v, \varphi)$. *SID* implies $V(\bar{x}^v, \theta^i, \varphi) > V(b, \theta^i, \varphi)$ and $\bar{x}^v \succ^i b$ $\forall i \in \mathscr{N}$: $\theta^i \geq \theta^v(\leq)$. Hence $\sum_{i=1}^n 1[V(\bar{x}^v, \theta^i, \varphi) > V(b, \theta^i, \varphi)] > n/2$ $\forall b \in P(\mathscr{A}^v)$ which implies $\bar{x}^v \succ a$ $\forall a \in P(\mathscr{A}^v) \rightarrow SP_{P(\mathscr{A}^v)}(a^v) = \varnothing \leftrightarrow a^v \in K(A^v)$. Q.E.D.

Lemma 13. *Any coalition $\mathscr{A}^j$ that does not contain the median voter v is stable only if $\exists a^j$ such that either of the following is true: (i) $a^j \in I(v)$; (ii) $a^j \geq x^v$ for all $x^v \in I(v)$; (iii) $a^j \leq x^v$ for all $x^v \in I(v)$.*

Proof. Suppose $a^j \notin I(v)$ and $a^j \ngeq x^v, a^j \nleq x^v$ for some $x^v \in I(v)$. There are two possible cases.

(i) say $x^k \in I(k)$ and $\forall k \in \mathscr{A}^j$ it is true either $x^k \geq a^j$ or $x^k \leq a^j$. This case is illustrated in Fig. SM.8.3 in the additional material. Consider x_j such that $x_j \in \arg\max_{x \in \{\bar{x}, \underline{x}^{v-1}\}} V(x, \theta^v, \varphi)$. Consider $\underline{x}^w$ $(\bar{x}^z)$ where $w = \max_{i<v, i \in \mathscr{A}^j} i$ $(z = \min_{i>m, i \in \mathscr{A}^j} i)$. Suppose $\underline{x}^w \succeq^v \bar{x}^z$ $(\underline{x}^w \prec^v \bar{x}^z)$ Notice that $\underline{x}^w \neq a^j$, optimality implies $V(\underline{x}^w, \theta^w, \varphi) > V(a^j, \theta^w, \varphi)$ (strict because $a^j \ngeq x^v, a^j \nleq x^v$ *SM* and *SID* imply $x^v \wedge a^j \succ^w a^j$ and optimality implies $\underline{x}^w \succeq^w x^v \wedge a^j$). Notice that because $\underline{x}^w \neq a^j$ *SID* implies $V(\underline{x}^w, \theta^i, \varphi) > V(a^j, \theta^i, \varphi)$ $\forall i \in \mathscr{N} : \theta^i < \theta^v(>)$. Also notice that $\underline{x}^w$ $(\bar{x}^z)$ is in the Pareto set $P(\mathscr{A}^{<v}) = P(\{i \in \mathscr{A} : i < v\})$ $(P(\mathscr{A}^{>v}) = P(\{i \in \mathscr{A} : i > v\}))$ because it is the lowest (highest) ideal point of the highest (lower) member of the subcoalition $\mathscr{A}^{<v}$ (see Lemma 9). Finally notice that any policy $b \in P(\mathscr{A}^j \backslash \mathscr{A}^{<v})$ must be $b \geq \bar{x}^z$ $(\leq \underline{x}^w)$ (because of Lemma 9). Hence given that $x^w \neq a^j$ (see above), then $\sum_{i=1}^n 1[V(x^w, \theta^i, \varphi) \geq V(b, \theta^i, \varphi)] > n/2$ $\forall b \in P(\mathscr{A}^j \backslash \mathscr{A}^{<v})$ which implies $x^w \succeq b \forall b \in P(\mathscr{A}^j \backslash \mathscr{A}^{<v}) \rightarrow a^j \notin SK(\mathscr{A}^j)$.

(ii) $\exists x^k \in I(k)$, $k \in \mathscr{A}^j$, $\theta^k > \theta^m(<)$ such that $x^k \ngeq a^j$ and $x^k \nleq a^j$. Consider $x^k \wedge a^j$. Notice that $x^k \ngeq a^j$ and $x^k \nleq a^j$ imply $x^k \wedge a^j \neq a^j$. Optimality implies $V(x^k, \theta^k) \geq V(x^k \vee a^j, \theta^k)$. *SM* implies $V(x^k \wedge a^j, \theta^k, \varphi) \geq V(a^j, \theta^k, \varphi)$. *SID* implies

$V(x^k \wedge a^j, \theta^i, \varphi) > V(a^j, \theta^i, \varphi) \quad \forall i \in \mathscr{N} : \theta^i < \theta^k$. Hence $\sum_{i=1}^{n} 1[V(x^k \wedge a^j, \theta^i, \varphi) > V(a^j, \theta^i, \varphi)] > n/2$. which implies $x^k \wedge a^j \succ a^j$.

If $x^k \wedge a^j$ is part of the Pareto set of $\mathscr{A}^{<k}$ this constitute a feasible and profitable deviation. If not, recall that X is a convex set and $V(x, \theta, \varphi)$ is $\theta - concave$, hence as $a^j \in P(\mathscr{A}^j)$ it has to be the solution to a problem in the form $a^j \in \arg\max_{x \in X} G(x, \lambda_{\mathscr{A}^j}(a^j), \theta_{\mathscr{A}^j}, \varphi)$.This case is illustrated in Fig. SM.8.4 in the additional material. If $x^k \wedge a^j$ is not part of the Pareto set of $\mathscr{A}^j$, consider the following alternative: $\tilde{x} \in \widetilde{M}(\mathscr{A}, \underline{\lambda}_{\mathscr{A}^j}(a^j, k))$ (see Lemma 11). We know from Lemma 11 that $\tilde{x} \leq a^j$. First of all notice that $\widetilde{M}(\mathscr{A}, \underline{\lambda}_{\mathscr{A}^j}(a^j, k)) = \widetilde{M}(\mathscr{A}', \lambda')$ for some λ', which implies that $\tilde{x} \in P(\mathscr{A}')$, i.e. it is in the Pareto set of $\mathscr{A}^{\leq k}$.We need to show that $\tilde{x} \neq a^j$ and that $\tilde{x} \succ^i a^j \ \forall i \in \mathscr{A}^{\leq k}$. Suppose $\tilde{x} = a^j \rightarrow a^j \in P(\mathscr{A}^{\leq k})$. From point (ii) we know that $V(x^k \wedge a^j, \theta^i, \varphi) > V(a^j, \theta^i, \varphi) \ \forall i \in \mathscr{N} : \theta_i \leq \theta^v \rightarrow a^j \notin P(\mathscr{A}^{\leq k}) \rightarrow$ Contradiction. Hence $\tilde{x} \neq a^j$ and $\tilde{x} \leq a^j$. Moreover, Lemma 11 implies $\tilde{x} \succ^k a^j$. This means that $V(\tilde{x}, \theta^k, \varphi) > V(a^j, \theta^k, \varphi)$ and because $\tilde{x} \neq a^j$ *SID* implies $V(\tilde{x}, \theta^i, \varphi) > V(a^j, \theta^i, \varphi) \ \forall i \in \mathscr{N} : \theta_i \leq \theta_j$. Recall that any $c \in P(\mathscr{A} \backslash \mathscr{A}^{\leq k})$ is $c \geq \bar{x}_k \geq \tilde{x}$. Hence either $c \in I(v)$ or $\sum_{i=1}^{n} 1[V(\tilde{x}, \theta^i, \varphi) \geq V(c, \theta^i, \varphi)] > n/2$ which implies $\tilde{x} \succ c \rightarrow a^j \notin SK(\mathscr{A}^j)$. Q.E.D.

Now consider the same analysis but in the case in which the stability concept is given by the TK instead of the *SK*. The crucial intuition of this case is that at any equilibrium only one coalition propose a policy.

Lemma 30. *In any TK-stable equilibrium only one coalition proposes a policy.*

Proof. Suppose that more than one coalition propose a policy platform. Then either there is a weak *Condorcet Winner*, in which case each of the non-winning active coalitions have an incentive to withdraw without a change in their payoff, as the same policy will still be the *Condorcet Winner* (because of the tie-break rule, and the fact that other coalitions' strategies are taken as given by potential deviators). If there is no *Condorcet Winner*, given that the default policy x^0 is a platform that is strictly worse than any other, then a withdraw implies a weakly better outcome for everybody. Thus, a deviation occurs thanks to the tie-break rule. Lastly, if no coalition propose any platform, x^0 is implemented, thus each coalition can deviate, propose a feasible platform and being strictly better off. Q.E.D.

I can now state two Lemmas that are equivalent to Lemma 12-13 for the case of TK stability. They are given by:

Lemma 12b. *The coalition $\mathscr{A}^v$ (could be a singleton) that includes the median voter v is stable at an equilibrium only if $a^v \in I(v)$ or $a^v = x^0$.*

Proof. At an equilibrium either $A_{\mathbb{P}(\mathscr{N})} \backslash \{a^v\}$ only contains $a^j = x^0$ for all j, in which case the analysis of stability is totally equivalent to the one in lemma 13, or $a^j \in A_{\mathbb{P}(\mathscr{N})} \backslash \{a^v\}$, such that $a^j \neq x^0$ and $A_{\mathbb{P}(\mathscr{N})}$ can be an equilibrium policy profile only if $a^v = x^0$ (because it is optimal that only the winning coalition runs with a platform given the tie-break rule). Q.E.D.

Lemma 13b. *Any coalition $\mathscr{A}^j$ that does not contain the median voter v is stable at an equilibrium only if $\exists a^j$ such that either of the following is true: (i) $a^j \in I(v)$, (ii) $a^j = x^0$.*

Proof. At an equilibrium either $A_{\mathbb{P}(\mathscr{N})} \backslash \{a^j\}$ only contains $a^j = x^0$ for all j, in which case the analysis of stability is totally equivalent to the one in lemma 13, or or $a^k \in A_{\mathbb{P}(\mathscr{N})} \backslash \{a^j\}$, such that $a^k \neq x^0$, such that $A_{\mathbb{P}(\mathscr{N})}$ can be an equilibrium policy profile only if $a^j = x^0$ (because it is optimal that only the winning coalition runs with a platform given the tie-break rule). Q.E.D.

I can now use Lemmas 7 to 13 to prove Theorem 1. Recall that the theorem states the following.

Theorem 1. *(Median Voter Theorem). (i) A coalitional equilibrium of the voting game exists. (ii) In any equilibrium the set of winning policies is a subset of the set of ideal points of the median voter v. (iii) If the median voter has a unique ideal policy, this policy is going to be the one chosen in any equilibrium.*

Proof. The results in Lemma 12 (12b) and Lemma 13 (13b) imply that the only policies that can be proposed by stable coalitions in equilibrium are either $a^v = x^v \in I(v)$ or $a^l \leq x^v$ or $a^h \geq x^v$ for all $x^v \in I(v)$. Recall optimality implies $V(a^v, \theta^v, \varphi) > V(a^l, \theta^v, \varphi)$ and *SID* implies $V(a^v, \theta^i, \varphi) > V(a^l, \theta^i, \varphi)$ $\forall i \in \mathscr{N} : \theta^i \geq \theta^v$. Similarly $V(a^v, \theta^v, \varphi) > V(a^h, \theta^v, \varphi)$ and *SID* implies $V(a^v, \theta^i, \varphi) > V(a^h, \theta^i, \varphi)$ $\forall i \in \mathscr{N} : \theta^i \leq \theta^v$. Also, the coalition structure in which every coalition is a singleton is always stable. Hence a (weak) Condorcet winner among the proposed policies exists, which is also the policy chosen in an equilibrium of the coalitional game (i). The total order in the policy space effectively available in all reduced games generated by a stable coalition structure implies the weak Condorcet winner must be always some $a^v \in I(v)$ (ii). The proof of (iii) is straightforward from (i) and (ii). In the case of the TK stability concept, there is no equilibrium in which a coalition other than $\mathscr{A}^v$ wins proposing a policy $a^j \notin I(v)$, because in such case the fact that either $a^j \leq x^v$ or $a^j \geq x^v$ for all $x^v \in I(v)$ implies that the median voter can change his strategy (e.g. leaving his coalition) and being strictly better off. For instance suppose $a^j \geq \bar{x}^v$ and $a^j \notin I(v)$. Then optimality implies $V(x^v, \theta^v) \geq V(x^v \vee a^j, \theta^v)$ and by *SM* one gets $V(x^v \wedge a^j, \theta^v, \varphi) > V(a^j, \theta^v, \varphi)$. *SID* implies $V(x^v \wedge a^j, \theta^i, \varphi) > V(a^j, \theta^i, \varphi)$ for all $i \leq v$. Hence either $x^v \wedge a^j$ is in $P(\mathscr{A}^{\leq v})$ or one can use part (ii) of the proof of Lemma 12 (12b) to show that there is a policy $\tilde{x} \in P(\mathscr{A}^{\leq v})$ such that $V(\tilde{x}, \theta^i, \varphi) > V(a^j, \theta^i, \varphi)$ for all $i \leq v$, which means that the subcoalition $\mathscr{A}^{\leq v} \subseteq \mathscr{A}^v$) possess a profitable deviation. Q.E.D.

Corollary 31. *(i) The equilibrium policy is in the Core of a winning coalition, i.e.* $x \in W(A_{\mathbb{P}(\mathscr{N})^*}) \rightarrow x \in K_{\mathscr{A}^v}(X)$ *for some winning coalition* $\mathscr{A}^v$*. Moreover, (ii) the equilibrium policy is in the Core of the reduced game, i.e.* $x \in W(A_{\mathbb{P}(\mathscr{N})^*}) \rightarrow x \in K(A_{\mathbb{P}(\mathscr{N})^*})$ *for any equilibrium policy profile* $A_{\mathbb{P}^*}$*.*

Proof. Straightforward from Theorem 1.

Theorem 2. *(Monotone Comparative Statics). The set of equilibrium policies of the voting game is (i) a sublattice of X which is (ii) monotonic nondecreasing in* θ^v*.*

Proof. Notice that the coalition structure in which the winning coalition includes all the individuals $\theta^i = \theta^v$ and all other coalitions are singleton is always stable if *SM* and *SID* are satisfied, because no individual in $\mathscr{A}^v$ has strict incentive to deviate and all the other coalition do not admit any deviation. Given the definition of $E(\mathscr{N})$ and that $I(v)$ is a superset of the unions of subsets of $I(v)$, all I need to show is that all elements of $I(v)$ are equilibria in that particular coalition structure, and this implies $E(\mathscr{N}) = I(v)$. Suppose this is not true. Then $\exists i \in \mathscr{N}$ with $\theta^i \neq \theta^v$ and $x \in I(i)$, such that $V(x, \theta^i, \varphi) > V(a^v, \theta^i, \varphi)$ for a majority on individuals. If (i) $x \geq a^v$ ($\leq$) then optimality implies $V(x, \theta^v, \varphi) \leq V(a^v, \theta^v, \varphi)$ and by the *SID* $V(x, \theta^v, \varphi) < V(a^v, \theta^v, \varphi)$ for all $i < j$ ($>$) which means that there is no strict majority that supports x against a^v. Contradiction. If (ii) $x \ngeq a^v$ and $x \nleq a^v$ then optimality implies $V(a^v, \theta^v, \varphi) \geq V(x \wedge a^v, \theta^v, \varphi)$ and $V(a^v, \theta^v, \varphi) \geq V(x \vee a^v, \theta^v, \varphi)$. Using *SM* one gets $V(x \vee a^v, \theta^v, \varphi) \geq V(x, \theta^v, \varphi)$ and $V(x \wedge a^v, \theta^v, \varphi) \geq V(x, \theta^v, \varphi)$. Finally the *SID* implies $V(x \vee a^v, \theta^i, \varphi) > V(x, \theta^i, \varphi)$ for all $i > v$ and $V(x \wedge a^v, \theta^j, \varphi) > V(x, \theta^j, \varphi)$ for all $j < v$, which means that $x \notin I(i)$ for all $i \in \mathscr{N}$ such that $i \notin \mathscr{A}^v$, and this means that x cannot be proposed by any other coalition, and therefore it cannot defeat a^v. Contradiction. This means that $E(\mathscr{N}) = I(v)$, and $I(v)$ is monotone nondecreasing in θ^v by Theorem 4 in Milgrom and Shannon, 1994. Q.E.D.

Theorem 3. *(Monotone Comparative Statics 2). The set of equilibrium policies of the voting game is monotonic nondecreasing in* φ*.*

Proof. (i) In the proof of Theorem 2 I have shown that $E(\mathscr{N}) = I(v)$. *SM* and *ID* imply that $I(v)$ is monotone nondecreasing in φ (Theorem 4 in Milgrom and Shannon, 1994). Q.E.D.

Theorem 7. *(Monotone Comparative Statics - Constraint Problems). If (i)*$V(x, \theta) = \mathscr{V}(x) + \upsilon(x_j, \theta)$*, in which (ii)* $\upsilon(x_j, \theta)$ *satisfies the ID in* (x_j, θ) *on X, and (iii)* $B(x)$ *is a convex function of X, then the set of equilibrium policies of the voting game is nondecreasing in* θ^v *for policy dimension* j*.*

Proof. Suppose any $x \notin I(v)$ is implemented in equilibrium. Because of the citizen-candidate assumption, $x \in I(i)$ for some $\theta^i > \theta^v$ ($<$) and the candidate i must run for election with no other candidates. Then the median voter v possess in $M(v)$ a policy $x^v \neq x$ such that $x^v \geq_j x$ ($\leq_j$). Optimality implies $\mathscr{V}(x^v) + \upsilon(x_j^v, \theta^v) > \mathscr{V}(x) + \upsilon(x_j, \theta^v)$. The *ID* of υ implies $\mathscr{V}(x^v) + \upsilon(x_j^v, \theta^i) > \mathscr{V}(x) + \upsilon(x_j, \theta^i)$ for all $i \geq v$ ($\leq$), which means that x^v defeats x by majority voting. Thus, being inactive is not a best response to other voters' strategies for player v. She can be strictly better off by running for election with policy platform x^v and winning the elections. Now one has to show that there is an equilibrium in which the median voter runs unopposed. Suppose the median voter run for election proposing a platform $\underline{x}^v = \min_{x_j} I(v)$ and that there is a platform $x^k \in I(k)$ for some k that can strictly defeat $\underline{x}^v$. Because the candidate proposing x^i must have a strict incentive to do so, it must be $k \neq v$ and $\mathscr{V}(x) + \upsilon(x_j, \theta^i) > \mathscr{V}(\underline{x}^v) + \upsilon(\underline{x}_j^v, \theta^i)$. Say $k > v$. Quah's result implies $\underline{x}^v \leq_j x^k$, thus optimality implies $\mathscr{V}(\underline{x}^v) + \upsilon(\underline{x}_j^v, \theta^v) > \mathscr{V}(x) + \upsilon(x_j, \theta^v)$, and because of the *ID* $\mathscr{V}(\underline{x}^v) + \upsilon(\underline{x}_j^v, \theta^i) \geq \mathscr{V}(x) + \upsilon(x_j, \theta^i)$ for all $i \leq v$, thus there is a majority of voters that weakly prefer $\underline{x}^v$ to x. Say $k < v$. Either $\underline{x}^v \geq_j x^k$, in which case there is a majority of voters that weakly prefer $\underline{x}^v$ to x for the same reason as before, or $\underline{x}^v <_j x^k$. In the latter case, optimality implies $\mathscr{V}(\underline{x}^v) + \upsilon(\underline{x}_j^v, \theta^v) \geq \mathscr{V}(x) + \upsilon(x_j, \theta^v)$ and thus $\mathscr{V}(\underline{x}^v) + \upsilon(\underline{x}_j^v, \theta^k) \geq \mathscr{V}(x) + \upsilon(x_j, \theta^k)$, which means that voter k does not have strict incentive to deviate. Similarly, one can show the same for the case in which $\bar{x}^v = \max_{x_j} I(v)$. Because there are equilibria in which both $\underline{x}^v$ and $\bar{x}^v$ are implemented, then the set of equilibria has the $i-$*increasing* property in θ^v. Q.E.D.

2.7.2 Stable Coalition Structures

Proposition 4. *(Lateral Coalitions). Any coalition $\mathscr{A}^j$ that include either (a) individuals with index $(i \leq v)$ or (b) individuals with index $(j \geq v)$ is always stable. Therefore a coalition structure $\mathbb{P}(\mathscr{N})$ is stable if each coalition $\mathscr{A}^i \in \mathbb{P}(\mathscr{N})$ satisfies either (a) or (b).*

Proof. Straightforward from Lemma 12 (12b) and Lemma 13 (13b) and the definition of a stable coalition structure.

Proposition 5. *(Central Coalitions). (i) Any coalition $\mathscr{A}^j$ that include both (a) individuals with index $(i < v)$ and (b) individuals with index $(j > v)$ plus individual v is stable if at least one policy $x^v \in I(v)$ is in the Core of a game $(\mathscr{N}, P(\mathscr{A}') \cup \{x^v\}, V)$ for all $\mathscr{A}' \subseteq \mathscr{A}$. (ii) If the Core of the full game $(\mathscr{N}, X, V)$ is non-empty, then any "Central Coalition" is stable, including the Grand Coalition of all voters.*

Proof. (i) $x^v \in I(v)$ is in the Core of a game $(\mathscr{N}, P(\mathscr{A}') \cup \{x^v\}, V)$ for all $\mathscr{A}' \subseteq \mathscr{A}^j$ implies that for any deviation of a subcoalition $\mathscr{A}' \subseteq \mathscr{A}$ any policy a' in the Pareto set

$P(\mathscr{A}')$ of this coalition is defeated by x^v by majority voting. Moreover $x^v \in P(\mathscr{A} \backslash \mathscr{A}')$, hence a' it is not a "credible threat" to x^v for coalition $\mathscr{A}$. As the statement implies that this is true for all possible $\mathscr{A}' \subseteq \mathscr{A}$, it implies that $x^v \in SK(\mathscr{A})$ $(TK(\mathscr{A}))$ and therefore $\mathscr{A}$ is stable. (ii) Notice that given that $P(\mathscr{A}') \cup \{x^v\} \subseteq X$ this implies that if the Core of the full game $(\mathscr{N}, X, V)$ is non-empty, then any "Central Coalition" is stable, including the Grand Coalition of all voters. Q.E.D.

Following Levy (2004), I define a ''partisan'' equilibrium as an equilibrium in which all party members vote for their party's platform, if it offers one (party members are not restricted in their votes if their party is not offering a platform).

Proposition 6. *(Ends-Against-the-Middle Coalitions). (i) Any coalition $\mathscr{A}^j$ that include both (a) individuals with index $(i < v)$ and (b) individuals with index $(j > v)$ but it does not include any individual with index v is stable only if either of the following is true: (1) $a^j \in I(v)$; (2) $a^j \geq x^v$ for all $x^v \in I(v)$; (3) $a^j \leq x^v$ for all $x^v \in I(v)$ (4) $a^j = x^0$. Therefore (ii) if $I(v) \cap P(\mathscr{A}^j) = \varnothing$, then there is no Partisan Equilibrium in which $\mathscr{A}^j$ is stable and $a^j \neq x^0$.*

Proof. (i) is straightforward from Lemma 12 (12b) and Lemma 13 (13b) and the definition of a stable coalition structure. For (ii) notice that condition $I(v) \cap P(\mathscr{A}^j) = \varnothing$ implies $x^v \neq a^j$ for all $x^v \in I(v)$, therefore, under the *SK* stability concept, lemma 12 implies that such coalition is stable only if either $a^j \geq x^v$ or if $a^j \leq x^v$ for all $x^v \in I(v)$. Then, there is at least one voter $i \in \mathscr{A}^j$ such that prefers x^v to a^j. Notice that this is relevant only under the *SK* stability concept. In the case of *TK*, an Ends-Against-the-Middle kind of coalition is stable *only if* either of the following is true: (1) $a^j \in I(v)$; (2) $a^j = x^0$ (Lemma 13b), thus part (i) is proved. Part (ii) of the statement does hold in this case because in equilibrium only one platform is proposed, and therefore all equilibria are of the "partisan" type and $a^j = x^0$ for all $\mathscr{A}^j \neq \mathscr{A}^v$. Q.E.D.

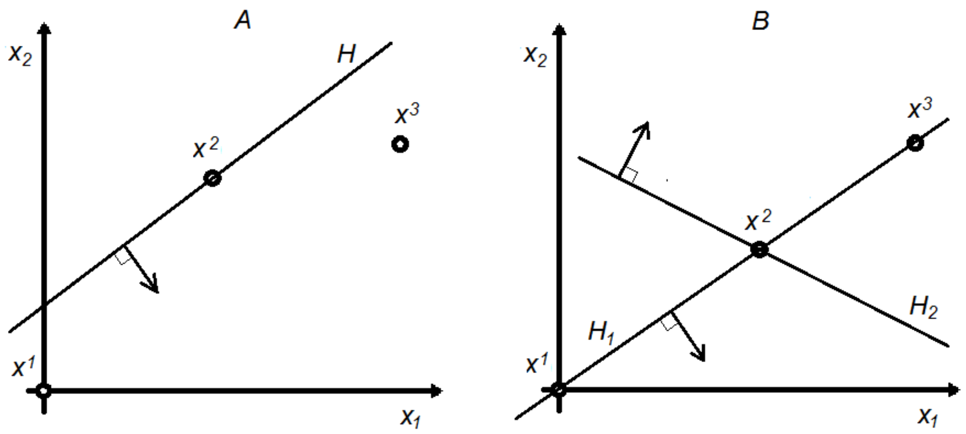
2.7.3 Grandmont Conditions

Consider a weak preference relation R^i and the corrisponding strict relation P^i and indifference relation I^i. Recall the definition of *betweeness*.

Definition. *(Grandmont, 1978).R is between R_1 and R_2 (noted $R \in (R_1, R_2)$) if for all x and y in X, (1) xR_1y and xR_2y imply xRy; (2) xP_1y and xP_2y imply xPy; (3) (xI_1y and xP_2y) or (xP_1y and xI_2y) imply xPy.*

Gandmont has shown that if preferences are Euclidean, this condtion corresponds to the concept of *unique median in all directions* in Davis *et al.* (1972). Such condition requires the existence of a point $x^* \in X$ such that any hyperplane passing through x^* divides the probability distribution over the set of ideal policies in two the set of ideal points into two parts with equal Lebesgue measure. Consider an example with 3 voters in which the policy space X is a convex subset of R^2. The set of ideal points IP is made by 3 points: $IP := \{x^1, x^2, x^3\}$. Given a hyperplane H, the probability assigned to an ideal point to lie on one side of the hyperplane is either 0 or 1 if the point does not lie on H, and it is assumed to be 0.5 otherwise. Such assumption corresponds to the case in which a voter that is just indifferent between two alternatives, support each of them with equal probability. Fig. 2.2 A shows that for any point x in X there is always a hyperplane H that divides the set IP into two sets of unequal size. The only exception is the case in which the three points lie on a straight line, as shown in Fig 2.2 B.

Figure 2.2. Grandmont Conditions in the Two-Dimensional Euclidean Space



The example in Fig. 2.2 A implies a violation of betweeness. A *Condorcet Winner* does not exist in this example[17]. On the other hand, it is easy to show that *SM* and *SID* are satisfied in this example. Thus one could use the concept of coalitional equilibrium to study the equilibrium policy outcome, but not the one of simple majority voting equilibrium.

[17] Notice that betweeness is a *sufficient* condition for the existence of a *Condorcet Winner*, thus a violation of such condition does not necessarily imply that existence fails. Neverthless, if the policy space is sufficiently, rich as in this case, a violation of betweeness usually corresponds to a fail in existence. See Grandmont, 1978.

3 A Multidimensional Theory of the Size of Government

I investigate the relationship between income inequality and size of the public sector in a theoretical framework. I extend the Meltzer-Richard model of public intervention in redistribution allowing for two different kinds of public spending. Specifically, voters choose - through the political process - a two-dimensional policy consisting of a linear tax on labour income and of the provision of a Public Good. Under the assumption of balanced governmental budget, these two choices determine the amount of a uniform lump-sum grant. The multidimensionality of the policy space implies that the traditional Median Voter Theorem does not hold. I adopt the model of electoral competition proposed by Dotti (2015) to tackle such problem. I show that, if in the proximity of a political equilibrium the endogenous progressivity of the tax system is sufficiently low, then a rise in the median-to-mean income ratio increases the size of the government. This prediction has opposite sign relative to the one in the traditional analysis. Moreover, I show that the progressivity of the tax system is increasing in the median-to-mean income ratio. Such results are consistent with most findings in the empirical literature.

JEL classification: D72, C71, H30, H41.

Keywords: Redistribution, Government, Public Goods, Policy, Voting.

3.1 Introduction

Does the shape of the income distribution affect the degree of public intervention in public spending in democratic countries? If so, how? Do highly unequal countries redistribute larger shares of their wealth? These questions have been at the very core of the literature in Political Economy for decades. In one influential paper Meltzer and Richard (1981) analyse the relationship between income distribution of a society and the extent of redistributive policies adopting a unidimensional Downsian model of electoral competition. A famous result in in their paper is that

> [..] An increase in mean income relative to the income of the decisive voter increases the size of the government.

This result imply a positive relationship between the size of the governmental sector and a measure of income skewness, the mean-to-median income ratio. Such prediction seems to be in sharp contrast with most anecdotal evidence. Examples of such evidence are provided by Bénabou (2000) and are mostly based on cross-country correlations. He emphasizes that, among industrial economies, the more unequal ones tend to redistribute less, not more. For instance, one can notice the differences in the extent of redistribution between Europe and the U.S.A., or - within Europe - the case of Scandinavian countries, which are are not strongly unequal but have adopted highly redistributive policies. These examples seem to suggest that the countries that are more equal also redistribute more. Regarding the developing world, Bénabou observes a similar contrast in the size of public spending in policies with redistributive effects, such as public education and health care. Namely, the size of this kind of public intervention is much larger in those areas where income inequality is lower - such as East Asia - than in the ones where income inequality is higher - such as Latin America. Lastly, he concludes this anecdotal evidence by noticing that the cuts in welfare spending experienced in most industrial democracies during the last two decades occurred at the same time as an unprecedented rise in income inequality. Such concerns are reinforced by several findings in the empirical literature. The predictions of Meltzer and Richard's model have been tested empirically in several studies - which I survey in the next section - and found very little support. Specifically, there is limited evidence of a causal link between the degree of inequality or of skewness of the income distribution and the size of public intervention in redistributive policies. Moreover, if any significant relationship is found at all in such studies, it often exhibits opposite direction relative to the one implied by Meltzer and Richard's paper. There are several theoretical and empirical reasons that could potentially explain this puzzle, and some of them - which I briefly mention in the next section - have been extensively investigated in the literature. In this paper I focus on one specific theoretical aspect that I claim to be crucial in order to understand why higher income inequality is likely to be associated with a smaller size of the government, rather than with a larger one. Specifically, I study the role played by the interaction among different kinds of policies with redistributive effects in shaping the relationship of interest. The consequences of such interaction have been mostly overlooked in the literature because of technical reasons. Specifically, the analysis of the choice of voters over different forms of public intervention in redistribution requires a multidimensional policy space. Unfortunately, the simple tools usually adopted to formulate predictions about the equilibrium policy outcome in traditional models, such as the *Median Voter Theorem*, do not generally apply when the choice set faced by voters is multidimensional. The reason is that, in such case, a *Condorcet winner* - i.e. a policy platform that is preferred to any alternative in the policy space by a majority of voters - does not usually exist. Moreover, the theoretical alternatives to Downsian framework in the literature did not prove useful to study this theoretical problem. The reason is that - as extensively described in chapter 2 - such models of electoral competition do not always deliver sharp analytical predictions about the comparative statics of interest. Thus, the papers in the theoretical literature that study the relationship between income inequality and size of the government usually focus on a single endogenous redistributive policy in isolation from the others. In this paper I tackle the problem of multidimensionality by adopting the model of electoral competition proposed by Dotti (2015), and the notion of *coalitional equilibrium* introduced in such paper. I apply this tool to study a labour economy in which redistribution can be achieved through two different channels: the tax system and the provision of a public good. Specifically, I assume a tax system characterized by two variables: a linear tax on labour income and a uniform lump-sum grant. Both are assumed to be endogenous outcomes of the political process. The main contributions of this paper are the following. First, I show that the result in Meltzer and Richard does not generally survive if their model is augmented by allowing for forms of public spending other than uniform transfers, such as the provision of a public good. Second, I show that the opposite prediction prevails in such augmented model if the endogenous degree of progressivity of the tax system is sufficiently low in the neighborhood of a coalitional equilibrium. Lastly, I show that the progressivity of the tax system is weakly decreasing in median-to-mean income ratio. These predictions are consistent with most recent findings in the empirical literature. The conclusion of this analysis is that the interaction between different kinds of redistributive policies is crucial to assess the relationship between income inequality and size of the government, and that the empirical puzzle generated by Meltzer and Richard's theoretical result does not survive in a model in which a richer policy space is assumed. The paper is organized as follows. In section two I survey the related literature and I describe the contrast between theoretical prediction and empirical findings, which represents a puzzle in the Political Economy literature. In section 3 I describe a simple model of labour economy that extends the ones in the traditional literature. Section 4 includes all the main results, which are about the comparative statics of the equilibrium policy outcome, of the size of the government and of the progressivity of the tax system. Section 5 concludes by comparing the results in section 4 with some recent empirical findings and suggesting directions for future research.

3.2 Related Literature

In their seminal paper Meltzer and Richard (1981) analyse in a simple general equilibrium model of a labour economy the relationship between a specific feature of the income distribution - the mean-to-median ratio - and the size of public intervention in redistribution. They adopt the Downsian framework of electoral competition and assume a unidimensional policy space, in which the collective choice only concerns the amount of a uniform lamp-sum grant financed through a linear tax on labour income. They prove the famous result that the size of the government - defined as the share of income redistributed by the government - is increasing in the mean-to-median income ratio. Such result implies a sharp relationship between the degree of income inequality - or more precisely, a measure of skewness of the income distribution - and the extent of redistribution in democratic countries. The analysis of such relationship has been a major topic of research in the Political Economy literature ever since. On one hand, several theoretical papers in this literature support similar predictions. For instance, Persson and Tabellini (1994) and Bénabou (1996) find that higher income inequality leads to more redistribution through higher taxation. On the other hand, some studies find opposite implications. The channels that underpin the non-standard predictions of this second group of papers are very diverse. For instance, Soares (1998) and de la Croix and Doepke (2009) focus on different forms of redistribution, and find that more unequal societies spend less on public goods. Saint-Paul (1994) and Bénabou (2000) show that if there are capital market imperfections, then the relationship between inequality and redistribution may not be monotonic increasing. Piketty (1995) proves that a negative relationship between income inequality and size of the government prevails in a dynamic model with imperfect learning of the social mobility process. Glomm (2004) shows that the relationship between income inequality and the amount of redistribution through public education services depends on the elasticity of substitution between consumption and the quality of publicly provided education in the parents' utility. He finds that for empirically relevant value of this parameter, higher inequality generates less redistribution. Lastly Dotti (2014) shows that, in a model of parental investment in education, income inequality is positively related to public spending if the returns to education are decreasing in parental income. An attempt to analyse a theoretical model of redistribution that adopts a multidimensional policy space is provided by Borge and Rattsø (2004) in a paper that is primarily empirical. They propose a model in which voters choose a two-dimensional policy consisting of a property tax and a poll tax, and in which tax revenues are used to provide local public services. Their model is extremely simple and it does not include income tax and labour supply decisions. This ensures that voter preferences satisfy the rather restrictive conditions stated by Grandmont (1978) and extensively described in chapter 2 of this work. Unfortunately, their approach is unlikely to prove successful in more complex theoretical environments, and in particular it is not suitable to analyse the one that is the object of my analysis. An example of why this is the case is provided in the appendix of chapter 2.

On the empirical side the evidence of a link between income inequality and size of the government is extremely mixed. Only a limited number of studies supports the predictions of the traditional theoretical framework. Some of the findings are summarized in Table 3.1. For instance, Meltzer and Richard (1983) test the hypothesis formulated in their earlier paper using US time series data of government spending. Their analysis support a negative relation between public spending and the median-to-mean income ratio. Panizza (1999) studies the relationship between the third quintile of the income distribution and various measures of redistribution using state averages of 46 U.S. in the period 1970-1980. He finds a positive and significant relationship between income inequality and various measures of redistribution, such as progressivity indexes and total governmental spending, and a non-significant relationship for other ones, such as total tax revenues and spending in welfare policies. Milanovic (2000) adopts the net income gains from tax and transfers of different income groups as a measure of redistribution. His estimates support the hypothesis that countries with greater inequality redistribute more. Notice that such analysis does not account for any form of redistribution other than direct in-cash policies. A more recent paper by Boustan *et al.* (2010) finds that rising inequality in cities and districts is associated with higher local revenue collection and expenditures. A majority of early papers based mostly on cross-country data did not find any statistically significant relationship between various features of the income distribution and some measure of the size of the government or of redistribution (for instance Perotti 1992, 1994; Persson and Tabellini 1994, Rodriguez, 1999). Perotti (1996) analyzes a cross-section of 67 countries and estimates the direction of the relationship of interest, separating democracies from non-democracies. He adopts various measures of redistribution, such as the marginal tax rate and different expenditure components, and finds no relationship between inequality and redistribution. Bassett *et al.* (1999) adopt a different set of measures of redistribution and spending and find a negative but poorly significant relationship of such measures with the degree of income inequality. A subsequent stream of more recent studies based on panel data has often found evidence of a significant negative relationship between the two variables. For instance, Gouveia and Masia (1998) use panel data for U.S. states during the period 1979-1991 and find a significant negative relationship between the mean-to-median income ratio and various measure of the size of the government based on public expenditures. Razin *et al.* (2002) reach similar conclusions using a panel of 13 countries over the period 1965-1992 for a total of 330 observations and including country fixed effects in their specification. Their measure of income skewness is the is the ratio of the income share of the top quintile to the combined share of the middle three quartiles and their measure of the size of the government is the average tax rate on labour income.

Their findings are summarized in Table 3.2. They find a negative and statistically significant relationship between income skewness and average tax rate. Moreover, they show that if they perform the same analysis using a different dependent variable, namely the total amount of social transfers, the relationship becomes positive and statistically significant. The latter analysis supports the idea that different redistributive policies may have different relationship with income inequality. Specifically, it suggests that higher income inequality may have a positive effect on redistribution through in-cash policies and progressive taxation, but its effect on the overall size of the government - which include spending in publicly provided goods and services - may have opposite sign. Lastly, some papers explore closely related questions. For instance, a number of papers have found that support for redistribution and public goods provision is weaker in more unequal or more heterogeneous societies. In such studies the notions of inequality and heterogeneity may not relate exclusively to income (Goldin and Katz 1997, Alesina *et al.* 1999, 2001, Luttmer 2001). Within this group, some papers find mixed evidence. For instance, Lind (2007), using data from the U.S. General Social Survey, finds that inequality between different groups reduces redistribution, while within group inequality increases it. Although the evidence is highly mixed, this review of the literature suggest that there is no convincing evidence in support of Meltzer and Richard's result. On one hand some limited evidence in favor of their hypothesis is provided by empirical studies that focus on measures of redistribution such as the progressivity of the tax system or the amount of transfers. On the other hand the results tend to be insignificant or show opposite sign in analyses that include in-kind policies, such as public provisions of good and services, and other policies with redistributive effects.

3.3 The Model

The setup is similar to the one proposed by Meltzer and Richard. The main difference is that the budget of the government is spent not only in in-cash redistribution, but also in Public Goods. For simplicity I am going to assume that the voters' population is a continuum with Lebesgue measure 1. This assumption is meant to represent an economy with an arbitrarily large number of voters.

3.3.1 Policy Space and Parameter Set

The policy space denoted by $X \subseteq R^2$ is two-dimensional with typical element (x,Y), in which x is 1 minus the linear tax rate on labour income and Y is the amount of a public good provided by the government. Thus X is defined as follows:

$X := \{(x,Y)|x \in [\underline{x},\bar{x}], Y \in [0,\bar{Y}]\}$, with $0 < \underline{x} < \bar{x} \leq 1$. Notice that the partially ordered set $(X, \leq)$ is a complete and convex sublattice of R^2. Individuals differ only in their productivity $\omega^i \in [\underline{\omega}, \bar{\omega}]$, which is perfectly observed by all agents, and such that $\underline{\omega} > 0$. The distribution of productivity is right-skewed with c.d.f. $R(\omega, \theta)$ and p.d.f. $r(\omega, \theta)$, where r is continuous and such that $r(\omega, \theta) \geq 0$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$ and equal to zero otherwise. Lastly, $\theta \in [0,1]$ is a parameter capturing the degree of inequality in the distribution of productivity.

3.3.2 Preferences

Each voter i has preferences represented by a concave utility function in the form:

$$U(c^i, Y, l^i) = u(c^i) + a(Y) + \gamma l^i \tag{3.1}$$

where c^i is i's consumption of private goods, Y is the the quantity of public goods that is provided by the government and l^i is leisure. The functions u and a are continuous, twice differentiable, strictly increasing and strictly concave in c, Y. I assume weakly positive consumption, i.e. $c^i \geq 0$ for all i. Individuals allocate their time between consumption and leisure such that $l^i + h^i = T$, where $h^i \in [0, T]$ is i's hours of work and T is the total endowment of time. Denote with $y^i = \tilde{w}(\omega^i)h^i$ the pre-tax income of an individual i facing hourly wage $\tilde{w}(\omega^i)$ and that supplies an amount h^i of hours of labour.

3.3.3 Public Finances

The tax system is the same as in Meltzer and Richard's model, namely individual post-tax income is determined by a linear tax rate t and by a lump-sum grant g. Recall $x = 1 - t$ and assume that the private good is the numéraire with price normalized to 1. Thus, the after-tax income - that is equivalent to the amount of private good consumed by individual i - is given by:

$$c^i = xy^i + g \tag{3.2}$$

The government has to break even, such that the governmental budget constraint is given by:

$$(1-x)\bar{y}(x,Y) - Y - g \geq 0 \tag{3.3}$$

where $\bar{y}(x,Y)$ is the average income. The inequality 3.3 simply states that the total governmental spending cannot exceed the total tax revenues. I restrict the analysis to the cases in which the budget constraint above is satisfied with strict equality. Notice that if $(1-x)\bar{y}(x,Y)-Y-g>0$, then a triple (x,Y,g) that makes all voters strictly better off is feasible, thus the condition above must be satisfied in any coalitional equilibrium. This implies that the restriction of balanced budget does not affect the results. Under this assumption, one can solve the budget constraint for g and define the function $g(x,Y)$ as follows:

$$g(x,Y) = (1-x)\bar{y}(x,Y) - Y \tag{3.4}$$

and then substitute g into equation (2) to obtain:

$$c^i = xy^i + (1-x)\bar{y} - Y \tag{3.5}$$

In order to express individual income and consumption as functions solely of the policy (x,Y) and of the individual wage I need to analyse how labour demand and supply decisions are made.

3.3.4 Production

There is a continuum of identical firms of size 1. Firms are characterized by a production function that is linear in effective labour in the form $Q(\theta) = \int_{\underline{\omega}}^{\bar{\omega}} \omega H(\omega) dR(\omega,\theta)$, where Q is output, the right hand side is the total demand for effective labour, and $H(\omega)$ is demand for labour of type ω. Firms supply a consumption good C^S and a public good Y^S, which are perfect substitutes in consumption, i.e. $Q = C^S + Y^S$. These goods are sold on the market at prices p_C and p_Y, respectively. Perfect substitutability in production implies that only the most expensive good is produced in a positive quantity, thus one can define a unique price for the output $p = \max\{p_C, p_Y\}$. Assuming that the consumption good is produced in positive quantity, this price is equal to the normalized price of such good, i.e. $p = 1$. The price of a unit of effective labour is w, such that the hourly wage of an worker of type ω is $\tilde{w}(\omega) = w\omega$. Firms are profit maximizers. The production of Q is positive as long as profits are non-negative. This occurs if and only if $Q(\theta) - w\int_{\underline{\omega}}^{\bar{\omega}} \omega H(\omega) dR(\omega,\theta) \geq 0$, and for each type ω they demand a positive amount of labour of type ω if $\omega \geq w\omega$, and no labour of that type otherwise. Lastly, firms are perfectly competitive, so they make zero profits, i.e. $\int_{\underline{\omega}}^{\bar{\omega}} [\omega - w\omega] H(\omega) dR(\omega,\theta) = 0$. This implies $w = 1$, i.e. each worker i is paid her marginal productivity ω^i per hour of work supplied to the firm. Notice that the total output of the firms is also equal to $Q(\theta)$.

3.3.5 Labour supply

The labour supply of each individual is endogenous in (ω^i, x, Y). Specifically, at each policy vector $(x,Y) \in X$ an individual i solves a problem of choice over consumption and leisure (C-L problem). Recall that hourly wage is $\tilde{w}(\omega) = \omega$. After substituting the time endowment constraint into the utility function, I define the optimal labour supply function of individual i as a continuous function h as follows.

$$h(\omega^i, x, Y) = \arg\max_{h\in[0,T]} u\left(x\omega^i h + g(x,Y)\right) + a(Y) + \gamma(T-h) \tag{3.6}$$

One can define a continuous function representing optimal earned income (pre-tax income) as follows.

$$y(\omega^i, x, Y) = \omega^i h(\omega^i, x, Y) \tag{3.7}$$

Thus, the pre-tax income of an individual i defined above lies in the range $[0, \omega^i T]$ and the formula for the average income can be written as follows.

$$\bar{y} = \bar{y}(x,Y) = \int_{\underline{\omega}}^{\bar{\omega}} y(\omega, x, Y) dR(\omega, \theta) \tag{3.8}$$

Notice that $\bar{y}$ is endogenous in (x,Y) and depends on individual labour supply decisions. Moreover, g is itself a function of $\bar{y}$ and hence of h^i for each i. Nevertheless, because h^i is finite, the effect of the individual choice of h^i on g tends to zero as the number of voters grows large. Therefore, the assumption of an arbitrarily large number of voters implies that g is constant in h^i for each i in this model. This simply means each individual's labour supply decisions are made treating g as constant given the policy (x,Y).

3.3.6 Market Equilibrium

Recall that the private consumption good is the numéraire. Notice that perfect substitutability in production implies that, if both the private and the public good are produced in positive quantities, then they must be sold at same price, which is equal to 1 in this case. The total demand of public good at fixed policy (x,Y) is simply Y, thus $Y^s = Y$ clears the market for such good at price equal 1. The hourly wage is equal to the worker's productivity ω because of the zero-profit condition (see section 3.4), and given this wage schedule the labour market clears at $H(\omega) = h(\omega^i, x, Y)$ for all $\omega \in [\underline{\omega}, \bar{\omega}]$ given policy (x,Y). Lastly, the market for the private good must clear by Walras' Law. To verify this result, notice that the total demand of such good is given by the total disposable income, i.e. $C^D(x,Y,\theta) = \int_{\underline{\omega}}^{\bar{\omega}} [xy(\omega,x,Y)+g(x,Y)]\,dR(\omega,\theta) = x\bar{y}(x,Y) + g(x,Y)$, while the total supply is simply the part of output that is not a public good, i.e. $C^S(x,Y,\theta) = \int_{\underline{\omega}}^{\bar{\omega}} \omega h(\omega,x,Y)dR(\omega,\theta) - Y$. The government budget constraint implies $Y = (1-x)\bar{y}(x,Y) - g(x,Y)$ at any market equilibrium, thus one can rewrite $C^S(x,Y,\theta) = \bar{y}(x,Y) - (1-x)\bar{y}(x,Y) + g(x,Y) = x\bar{y}(x,Y) + g(x,Y)$ which implies $C^S(x,Y,\theta) = C^D(x,Y,\theta)$ at $p = w = 1$ as expected. Lastly, I need to rule out the possibility that $c^i = xy^i + g$ is negative for some individuals at some $(x,Y) \in X$. Thus, I assume that $(1-x)\bar{y} - Y + xy(\omega,x,Y) \geq 0$ for all $\omega \in [\underline{\omega},\bar{\omega}]$ and for all $(x,Y) \in X$. This condition ensure that $g(x,Y)$ cannot get too negative and induce negative consumption for some individual. Also notice that if $\bar{y} > 0$ for all $x \in [\underline{x},\bar{x}]$ at $Y = 0$ the continuity of $y(\omega^i,x,Y)$ and $\bar{y}(x,Y)$ in (x,Y) ensures that there is always a range $[0,\bar{Y}]$ with $\bar{Y} > 0$ such that the condition above is satisfied within such range for all $x \in [\underline{x},\bar{x}]$.

3.3.7 Voter Objective Function

The objective function of an individual i in the voting game is a strictly increasing transformation of i's indirect utility function. It represents the utility achieved by an individual at the optimal level of labour supply as a function of the policy vector (x,Y) and of the individual wage ω^i multiplied by ω^i. Notice that such monotone transformation does not imply any change in voter preferences. In order to obtain the objective function, I substitute the equations for $g, c^i, y^i, \bar{y}$ from the previous paragraphs into U and I get the following formula.

$$V^i = V(x,Y;\omega^i) = \omega^i \left[u\left(xy(\omega^i,x,Y) + (1-x)\bar{y} - Y\right) + a(Y) - \gamma(T-h^i)\right] \qquad (3.9)$$

Using formula (3.9) I can derive the main results, which are stated in the next section.

3.4 Results

In this section I present the main results of the paper, namely the conditions for existence and uniqueness of a coalitional equilibrium and the comparative statics results. The latter consist of the analysis of the effects of a shock on the wage distribution on the equilibrium policy variables, and of the implied relationship between various measures of size of the government and the median-to-mean income ratio. Then I compare the predictions of this augmented model with the ones implied by Meltzer and Richard's famous result (1981).

3.4.1 Existence of a Coalitional Equilibrium

I derive sufficient conditions for the existence of a voting equilibrium of the economic model of redistribution proposed in the previous sections of this paper. The equilibrium concept adopted is the one of coalitional equilibrium proposed by Dotti (2015) and described in chapter 2. In order to prove the existence of a coalitional equilibrium I need to show that the following conditions are satisfied: (i) The policy space X in a convex and complete sublattice of R^2 (ii) $V(x,Y;\omega)$ is jointly continuous and concave in (x,Y), (iii) $V(x,Y;\omega)$ satisfies *Supermodularity (SM)* in (x,Y) and *Strictly Increasing Differences (SID)* in $(x,Y;\omega)$ for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$. Notice that condition (i) is satisfied by the policy space defined in section 3.1. Continuity of V is ensured by the continuity of the utility function U and by the following Lemma.

Lemma 1. *Individual income $y(\omega^i,x,Y)$ is weakly increasing in productivity ω^i. If $T > h^i > 0$, then $y(\omega^i,x,Y)$ is strictly increasing in $\omega^i$and twice differentiable with respect to x,Y and ω^i for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$,* and *$h(\omega^i,x,Y)$ is twice differentiable with respect to x,Y and ω^i for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$.*

Proof. see Appendix 3.6.1.1.

Condition (iii) is satisfied thanks to the following Lemma.

Lemma 2. *If the individual labour supply is such that $T > h^i > 0$ for all i and all $(x,Y) \in X$, then the function $V(x,Y;\omega)$ satisfies Supermodularity in (x,Y) and the Strictly Increasing Differences in $(x,Y;\omega)$ for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$.*

Proof. See Appendix 3.6.1.1.

Denote with ω^m the median wage under the distribution with c.d.f $R(\omega,\theta)$ at $\theta = 0$. Lemmas 1 and 2 imply that all the conditions for the existence of a coalitional equilibrium are satisfied if the objective function is concave and if the solution of the consumption/leisure problem is interior for all voters. This, I can state the following result.

Theorem 3. *A Coalitional Equilibrium exists if (i) $V(x,Y;\omega^i)$ is concave in (x,Y) for all i and if (ii) the individual labour supply is such that $T > h^i > 0$ for all i and all $(x,Y) \in X$. Moreover, if $V(x,Y;\omega^m)$ is strictly concave in (x,Y), then the set of equilibrium policies is a singleton.*

Proof. See Appendix 3.6.1.1.

Theorem 3 simply states that if the solution to the consumption/leisure problem is always interior, i.e. if all individuals allocate a positive amount of time both to consumption and to labour for any possible policy platform in the policy space, then all that is needed for the existence of a coalitional equilibrium is the concavity of the objective function with respect to (x,Y). Notice that such requirement is not trivially satisfied if $u(xy^i+(1-x)\bar{y}-Y)+a(Y)+\gamma(T-h^i)$ is concave in (x,Y) for given $y^i,\bar{y},h^i$ for all i and all $(x,Y)\in X$. One has to recognize that $y^i,\bar{y},h^i$ are endogenous functions of (x,Y) and must account for such endogeneity in order to asses the concavity of $V(x,Y;\omega^i)$. In the rest of the paper I am going to assume that V is strictly concave in (x,Y). In Appendix 3.6.2 I provide an example with a parametric utility function in which such condition is satisfied.

3.4.2 Comparative statics

In this section I describe a comparative statics exercise and use the result in Theorem 2 in Dotti (2015) to derive the sign of the policy change induced in equilibrium by such exercise (see chapter 2). The aim is to show that the predictions of the augmented model proposed in this paper are, under certain conditions, more consistent with the empirical evidence relative to the ones implied by the Meltzer and Richard's paper. Suppose that the sufficient conditions for the existence of a coalitional equilibrium and for uniqueness of the policy outcome described in the previous section are satisfied. Define a continuous distribution for ω with p.d.f. $f(\omega)$, such that $f(\omega)>0$ for all $\omega\in[\underline{\omega},\overline{\omega}]$ and $f(\omega)=0$ otherwise. Denote with (x^*,Y^*) the equilibrium policy vector given such wage distribution. The comparative statics exercise is the following. First, define another continuous distribution with p.d.f. $\tilde{f}(\omega)$, such that (i) the mean income is the same under both distribution f and $\tilde{f}$ at constant policy (x^*,Y^*), i.e.

$$\int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x^*,Y^*)f(\omega)d\omega=\int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x^*,Y^*)\tilde{f}(\omega)d\omega=\bar{y}(x,Y) \tag{3.10}$$

and (ii) the median income is strictly higher under the second distribution conditional on policy (x^*,Y^*), i.e.

$$\tilde{F}(\omega^m)d\omega<F(\omega^m)d\omega \tag{3.11}$$

in which ω^m is the median under distribution f. Notice that such a distribution with p.d.f. $\tilde{f}$ exists as long as there exists a corresponding income distribution conditional on policy (x^*,Y^*) with continuous p.d.f. that is a mean-preserving spread of the original income distribution given policy (x^*,Y^*), and that has higher median relative to the the one that prevails under the original distribution. This is always the case for income distributions with a continuum p.d.f. that are not symmetric about the mean. See Appendix 3.6.1.3. Second, define a distribution with p.d.f. $r(\omega,\theta) = (1-\theta)f(\omega) + \theta\tilde{f}(\omega)$ for $\theta \in [0,1]$. Notice that the expected value of income under r for such distribution is given by

$$(1-\theta)\int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x^*,Y^*)f(\omega)d\omega + \theta\int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x^*,Y^*)\tilde{f}(\omega)d\omega = \bar{y}(x^*,Y^*) \tag{3.12}$$

for any $\theta \in [0,1]$. The comparative statics exercise consists of a marginal increase in θ evaluated at $\theta = 0$. This corresponds to marginally increasing the median income keeping the mean income constant, for a policy vector fixed at $(x,Y) = (x^*,Y^*)$. Notice that $(x^*,Y^*) = (x^*(0),Y^*(0))$, where the latter is the equilibrium policy under the distribution with p.d.f. $r(\omega,0)$. I define the comparative statics exercise and the notion of monotonicity of such comparative statics as follows.

Definition 1. A *(marginal) increase in median income* is a (marginal) increase in parameter θ at constant policy (x,Y).

Definition 2. The equilibrium policy vector is *weakly increasing in the median income* if the vector $(x^*(\theta),Y^*(\theta))$ is weakly increasing in θ.

In the previous section I have shown that - under the assumption stated - the conditions for the existence of a coalitional equilibrium are satisfied. Thus, the corresponding monotone comparative static results must also apply. Specifically, I can use Theorem 2 in Dotti (2015) to state the following results.

Theorem 4. *(i) The spending in Public Goods is weakly increasing in the median income, and (ii) the tax rate on labour income is weakly decreasing in the median income.*

Proof. See Appendix 3.6.1.2.

Notice that the comparative statics result in Theorem 4 does not state that the mean income is unaffected by changes in θ, because the endogenous adjustment in the equilibrium policy $(x^*(\theta),Y^*(\theta))$ implies a change in $\bar{y}(x,Y)$, and possibly in $y(\omega^m,x,Y)$, too. Thus, in order to address the endogenous relationship between the median-to-mean income ratio and the policy outcome one has to analyse how $y(\omega^m,x,Y)/\bar{y}(x,Y)$ moves as the equilibrium policy changes. The direction of such effect is given by the following Lemma.

Lemma 5. *If the income distribution at the equilibrium policy is such that $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$, then (i) the median-to mean income ratio is weakly increasing* *in θ in a neighborhood of (x^*,Y^*). Thus, if $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$ in the neighborhood of a coalitional equilibrium, then (ii) a marginal increase in median income results in both a weakly higher equilibrium policy and a weakly larger median-to-mean income ratio.*

Proof. See Appendix 3.6.1.2.

Theorem 4 and Lemma 5 deliver the sign of the comparative statics exercise on the equilibrium policy outcome and on the median-to-mean income ratio. Thus, I can now use these results to assess the relationship between this ratio and some measures of the size of the government implied by such exercise.

3.4.3 Size of the Government and Progressivity

In order to answer the main questions of this paper I need to define measures of the size of the government and of the progressivity of the tax system. In Meltzer and Richard's paper the size of the government is simply the marginal tax rate $t = 1-x$. In their unidimensional framework this is correct because the total government spending and the total tax revenues are weakly increasing functions of t. In the two-dimensional framework proposed here this is not necessarily true. Thus, I define the size of the government as either (i) the total government spending, or as (ii) the total tax revenue, or as (iii) the average tax rate. Notice that the predictions of the Meltzer and Richard's model would be qualitatively the same with respect of these three alternative measures. Thus, the results of the two papers are directly comparable. The formula for the total revenue is:

$$TR(x,Y,\theta) \equiv \int_{\underline{\omega}}^{\bar{\omega}} \left(1-\mathbb{I}\left[g(x,Y)-(1-x)y(\omega^i,x,Y) \geq 0\right]\right) \left[(1-x)y(\omega^i,x,Y)-g(x,Y)\right] r(\omega,\theta)d\omega \tag{3.13}$$

where $\mathbb{I}(\cdot)$ is an indicator function that has value equal to 0 if individual i is a net tax payer and value equal to 1 if individual i is a net receiver of subsidies. So the integral defining TR in (3.13) represents the sum of net tax paid. Notice that at an equilibrium policy (x,Y) two possible situations can occur. The first possibility is that (i) the net tax payments if individual i are such that $g(x,Y)-(1-x)y(\omega^i,x,Y) < 0$ for all $\omega^i \in [\underline{\omega},\bar{\omega}]$. Because $y(\omega^i,x,Y)$ is strictly increasing in ω^i, then this case occurs if $g(x,Y)-(1-x)y(\underline{\omega},x,Y) < 0$. The second possibility is that $g(x,Y)-(1-x)y(\underline{\omega},x,Y) \geq 0$. In such case, because $y(\omega^i,x,Y)$ is continuous and strictly increasing in ω^i, there exists a unique $\hat{\omega}(x,Y) \in [\underline{\omega},\bar{\omega}]$ such that $(1-x)y(\omega^i,x,Y)-g(x,Y) \geq 0$ for all $\omega^i \in [\hat{\omega}(x,Y),\bar{\omega}]$ and

$(1-x)y(\omega^i,x,Y)-g(x,Y)<0$ for all $\omega^i \in [\underline{\omega},\hat{\omega}(x,Y))$. Specifically, $\hat{\omega}(x,Y)$ - if it exists - solves the following:

$$g(x,Y)-(1-x)y(\hat{\omega}(x,Y),x,Y)=0 \tag{3.14}$$

Notice that it cannot be the case that $g(x,Y)-(1-x)y(\omega^i,x,Y)>0$ for all i because this would imply that total public spending exceed total tax revenues, which implies in turn a violation of the balanced budget rule. Thus, if $g(x,Y)-(1-x)y(\underline{\omega},x,Y)\geq 0$, one can rewrite the formula above as follows:

$$TR^1(x,Y,\theta)=\int_{\hat{\omega}(x,Y)}^{\bar{\omega}}[(1-x)y(\omega,x,Y)-g(x,Y)]\,r(\omega,\theta)d\omega \tag{3.15}$$

Similarly, if $g(x,Y)-(1-x)y(\underline{\omega},x,Y)<0$, then the formula for the total revenue is simply the following:

$$TR^2(x,Y,\theta)=\int_{\underline{\omega}}^{\bar{\omega}}[(1-x)y(\omega,x,Y)-g(x,Y)]\,r(\omega,\theta)d\omega=Y \tag{3.16}$$

The second measure of the size of the government proposed is the total public spending. Given that the net tax revenue is used to finance the provision of a public good Y and the transfers to the individuals who pay no net taxes, one can define the total public spending as the sum of net transfers plus the expenditure in public goods. Under the same assumptions imposed for the previous index, this has the following form. If $g(x,Y)-(1-x)y(\underline{\omega},x,Y)\geq 0$, then the formula is:

$$TS^1(x,Y,\theta)=\int_{\underline{\omega}}^{\hat{\omega}(x,Y)}[g(x,Y)-(1-x)y(\omega,x,Y)]\,r(\omega,\theta)d\omega+Y \tag{3.17}$$

while if $g(x,Y)-(1-x)y(\underline{\omega},x,Y)<0$, the formula for total spending is simply $TS^2(x,Y)=TR^2(x,Y)=Y$. Lastly, another possible way to measure the size of the government is to use the average tax rate, which is defined as the total net tax revenues divided by total income:

$$AT(x,Y,\theta)=\left\{\int_{\underline{\omega}}^{\bar{\omega}}[(1-x)y(\omega,x,Y)-g(x,Y)]\,r(\omega,\theta)d\omega\right\}/\bar{y}=Y/\bar{y} \tag{3.18}$$

Theorem 6. *There exists a threshold $\hat{g}\geq 0$ such that, if $g(x,Y)\leq\hat{g}$, then (i) the total government spending, (ii) the total tax revenue, and (iii) the average tax rate are all* *weakly increasing in the median income in a neighborhood of* (x^*, Y^*). *Thus, if* $y(\omega^m, x^*, Y^*) \leq \bar{y}(x^*, Y^*)$ *and the degree of tax progression is sufficiently low for all income levels in the neighborhood of the coalitional equilibrium, then (iv) a marginal increase in median income results both in a weakly larger size of the government and in a weakly larger median-to-mean income ratio.*

Proof. See Appendix 3.6.1.2.

This Theorem suggests that, differently from what is implied by the result in Meltzer and Richard's paper, a society with lower income skewness may exhibit - *ceteris paribus* - a larger size of the government. Specifically, this is the case for economies that are characterized by a tax system that is not strongly progressive, which occurs if transfers $g(x,Y)$ are not too large. The link between progressivity and transfers will become clear in the next few lines, in which I analyse the comparative statics of the degree of progression of the tax system. The tax schedule implied by the balanced budget in equation (3.4) is simply given by the formula:

$$T(x,Y;y) = (1-x)y - g(x,Y) \tag{3.19}$$

Following Lambert (1989), I focus on measures of the degree of *tax progression*, which is a property of an income tax schedule alone, rather then the degree of *tax progressivity*, which is instead a property of the interaction between the tax schedule and the pre-tax income distribution to which it is applied. Specifically, I am going to employ two indexes proposed by Musgrave and Thin (1948). The first measure of tax progression is the *average rate progression (ARP)*:

$$ARP(x,Y;y,\theta) \equiv \left.\frac{\partial[T(x,Y;y)/y]}{\partial y}\right|_{x=x^*(\theta),Y=Y^*(\theta)} = \frac{g(x^*(\theta),Y^*(\theta))}{y^2} \tag{3.20}$$

The *ARP* simply records the rate at which the average tax rate increases with income. If it is increased at an income y (or at all income level) the tax has become more average rate progressive at y (or everywhere). Lastly, $ARP = 0$ for a proportional tax system and $ARP > 0$ for progressive tax systems. A shortcoming of the *ARP* is that it is not unit free. A unit free alternative to the *ARP* is proposed by Lambert (1989). The *inverse of residual progression (IRP)*[1] is defined as the inverse of the elasticity of post-tax income to pre-tax

[1] Notice that the *IRP* is the inverse of the *residual progression* measure proposed by Musgrave and Thin (1948). The former index is chosen because it increases when the tax becomes more progressive.

income:

$$IRP(x,Y;y,\theta) \equiv \left(\frac{\partial[y-T(x,Y;y)]}{\partial y} \frac{y}{[y-T(x,Y;y)]} \Big|_{x=x^*(\theta),Y=Y^*(\theta)} \right)^{-1} = \\ = \frac{yx^*(\theta)+g(x^*(\theta),Y^*(\theta))}{yx^*(\theta)} \tag{3.21}$$

The *IRP* can be interpreted as the elasticity of pre-tax income to post-tax income. An increase in the *IRP* makes the tax more residual progressive. The *IRP* has value equal to 1 if the tax system is proportional and greater than 1 if it is progressive. These two measures of progression are income dependent. This is not going to matter for the sign of the comparative statics exercise that I present in the next lines of this section, because the direction of the effect is the same for all income levels. Thus, I can establish a notion of monotonicity for the degree of progression of the tax system[2]. Specifically, I define monotonicity for the degree of progression of the tax system as follows.

Definition 3. The degree of progression of the tax system is *weakly increasing in the median income* if *ARP* (*IRP*) is weakly increasing in θ for all income levels $y \geq 0$.

Notice Using this definition, I can state the following result.

Theorem 7. *If the tax system is weakly progressive at* (x^*,Y^*)*, then (i) the degree of progression of the tax system is weakly decreasing in the median income in a neighborhood of* (x^*,Y^*)*. Thus, if* $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$ *in the neighborhood of the coalitional equilibrium, then (ii) a marginal increase in median income results both in a weakly lower progression of the tax system and in a weakly larger median-to-mean income ratio.*

Proof. See Appendix 3.6.1.2.

Theorem 7 states an additional prediction of the augmented model. Namely, an increase in the median-to-mean ratio due to a change in the wage distribution tends to be associated with a rise in the degree of progression of the tax system. Such prediction is consistent with the findings in Razin *et al.* (2002). Specifically, they find that the effect of an increase in income skewness is a fall in the average tax rate on labour income and a weak increase in the total amount of social transfers. The link with the predictions of the model presented in this section is given by the fact that a tax system that exhibits a lower average tax rate and (weakly) larger social transfers is actually a more progressive tax system, which is in line with the prediction of Theorem 7. The main consequence of

[2]Notice that this statement does not necessarily imply that the *progressivity* faced by a specific individual *i* moves in the same direction for all *i*.

Theorems 6 and 7 is that, if one allows for different kinds of public spending to be chosen by the population of voters, then the size of the government is not necessarily decreasing in the median-to-mean income ratio, as postulated in Meltzer and Richard (1981). On the contrary, the size of the government is weakly increasing in such ratio if - in the proximity of an equilibrium - the tax system is not strongly progressive. Moreover, the progressivity of the tax system is weakly decreasing in the median-to-mean income ratio. These results are relevant for empirical purposes. They suggest that Meltzer and Richard's hypothesis may be more empirically sound if formulated as a theory of in-cash redistribution and of progressivity of the tax system rather than a theory of the size of the government. This alternative interpretation is more consistent with the empirical evidence presented in section 2. Nevertheless, additional research is needed in order to claim that the empirical puzzle generated by Meltzer and Richard's paper has found a convincing explanation.

3.5 Concluding Remarks

In this paper I study the effects of a marginal increase in income inequality on the size of the government. To achieve this goal, I adopt a positive model of labour supply and redistribution in the spirit of the one proposed by Meltzer and Richard (1981). The policy outcomes are determined by the behavior of voters through electoral competition. The main novelty of this analysis is that voters can decide endogenously both the tax rate on labour income and the amount of a public good provided by the government. The budget of the public sector is assumed to be balanced. Thus, the choice of the two policy variables determine the amount of in-cash redistribution, which is carried out through a uniform lump-sum grant. The policy space is two-dimensional, thus a *Condorcet winner* does not usually exist. As a consequence, the traditional Downsian framework of electoral competition cannot be adopted to study the comparative statics of the equilibrium policy outcome. I adopt the model of electoral competition introduced by Dotti (2015), which ensures sharp comparative statics predictions even over multidimensional choice domains. I study the effects on the equilibrium policy outcome of a specific change in the wage distribution that translates - given a certain policy - into a higher median income with no changes in the mean income (*marginal increase in median income*). I find that a marginal increase in median income implies an increase in the provision of the public good and a fall in the marginal tax rate on labour income. Moreover, at the new equilibrium, the median-to-mean income ratio is weakly larger. I use this results to study the relationship between such ratio and the size of the government, defined either as total public spending, or total tax revenues, or as average tax rate. I show that, if in the neighborhood of an equilibrium the progressivity of the tax system is sufficiently low, then a rise in the median-to-mean income ratio is associated with an increase in the size of the government. This result contrasts the prediction of the traditional unidimensional analysis by Meltzer and Richard, and is more consistent with recent empirical evidence. Lastly, I show that an increase in the median-to-mean income ratio in associated with a more progressive tax system, at all income levels. The latter prediction is also consistent with the empirical evidence in Panizza (1999), Milanovic (2000) and Razin *et al.* (2002). Two main conclusions can be derived from this analysis. The first is that the positive relationship between income inequality and size of the government implied by traditional positive models of redistribution is mainly an outcome of an excessively restricted policy space, and that such result will not survive if forms of redistribution other than in-cash transfers - such as the provision of public goods - are included in the analysis. The second is that this analysis does not rule out the possibility that income inequality may help explaining the degree of public intervention in redistribution in democratic countries, but only for those policies that imply direct in-cash effects, such as transfers, changes in marginal tax rates and in the progressivity of the tax system. In the light of the encouraging but sometimes inconsistent evidence in the literature, this new prediction of the model should be tested empirically as an alternative to the traditional one. Lastly, the theoretical framework adopted in this paper to tackle the problems induced by the multidimensionality of the policy space represents a promising tool to address other important questions regarding the political economy of fiscal policies. For instance, a relatively recent stream of literature has been studying the choice of voters over different kinds of taxes. Within this group, some papers analyse the role of wealth inequality in shaping tax rates on labour and capital income (Bassetto and Benhabib, 2006), while other papers are interested in the determinants of the relative size of a property tax and a poll tax (Borge and Rattsø, 2004). Such questions are characterized by an intrinsic multidimensionality of the policy space. As a consequence, these analyses typically rely on very strong restrictions on voter preferences and/or on the policy space in order to ensure the existence of a *Condorcet winner*. Thus, an approach like the one adopted in this paper may prove helpful in order to verify the robustness of their findings in a more flexible theoretical environment.

3.6 Appendix

This section is structured as follows. Appendix 3.6.1 contains all the proofs not included in the main body of the paper. Appendix 3.6.2 proposes a parametric example of the framework, and shows that the assumption of strict concavity is satisfied for a simple utility representation of voter preferences.

3.6.1 Proofs

3.6.1.1 Existence of a Coalitional Equilibrium

Lemma 1. *Individual income $y(\omega^i,x,Y)$ is weakly increasing in productivity ω^i. If $T > h^i > 0$, then $y(\omega^i,x,Y)$ is strictly increasing in $\omega^i$and twice differentiable with respect to x,Y and ω^i for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$,* and *$h(\omega^i,x,Y)$ is twice differentiable with respect to x,Y and ω^i for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$.*

Proof. Using the F.O.C. of the C/L problem one gets $u'(xy^i + (1-x)\bar{y} - Y)x\omega^i - \gamma \leq 0$. If the solution to this problem is a corner for agent i given policy (x,Y), then $y(\omega^i,x,Y) = 0$ and $\frac{\partial y(\omega^i,x,Y)}{\partial \omega^i} = 0$. If the solution is interior, given that $\bar{y}(x,Y)$ is constant in ω^i, differentiability of y and h follows directly from the twice differentiability of function u. Then:

$$\frac{\partial y(\omega^i,x,Y)}{\partial \omega^i} = -\frac{u'(xy^i + (1-x)\bar{y} - Y)}{u''(xy^i + (1-x)\bar{y} - Y)x\omega^i} > 0 \tag{3.22}$$

because of the strict concavity of u. Notice that $y(\omega,x,Y)$ is continuous in ω but may be not differentiable at all $\omega \in [\underline{\omega},\bar{\omega}]$ if the solution is a corner one for some ω at some policy (x,Y). Lastly, continuity and differentiability of y and h with respect to ω for interior solutions of the C/L problem are ensured by the twice differentiability of u with respect to ω. Regarding continuity and twice differentiability with respect to x and Y, one also needs to show that $\bar{y}(x,Y)$ is continuous and twice differentiable. Recall that $\bar{y}(x,Y) = \int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x,Y)dR(\omega,\theta)$. Denote the inverse of function u with $u^{-1}(\cdot)$. In an interior solution the F.O.C. of the C/L problem implies $y(\omega,x,Y) = [u^{-1}(\gamma/x\omega^i) - (1-x)\bar{y}(x,Y) + Y]/x$. Integrating both sides under the wage distribution with c.d.f. $R(\omega,\theta)$ one gets $\bar{y}(x,Y) = \int_{\underline{\omega}}^{\bar{\omega}} u^{-1}(\gamma/x\omega^i)dR(\omega,\theta) + Y$, which is continuous and twice differentiable in x,Y. Q.E.D.

Lemma 2. *If the individual labour supply is such that $T > h^i > 0$ for all i and all $(x,Y) \in X$, then the function $V(x,Y;\omega)$ satisfies Supermodularity in (x,Y) and the Strictly Increasing Differences in $(x,Y;\omega)$ for all $(x,Y) \in X$ and all $\omega \in [\underline{\omega},\bar{\omega}]$.*

Proof. Denote $V^i_{ab} = \frac{\partial V(x,Y;\omega^i)}{\partial a \partial b}$. Recall that - given that V^i is a twice differentiable function - sufficient conditions for *SM* and *SID* are $V^i_{xY} \geq 0$, $V^i_{x\omega} > 0$ and $V^i_{Y\omega} > 0$ for all i and all $(x,Y) \in X$. First of all we need to calculate the marginal effects of x and Y on V_i, denoted with $\frac{\partial V_i}{\partial x}$ and $\frac{\partial V_i}{\partial Y}$ respectively. The objective function is:

$$V(x,Y;\omega^i) = \omega^i \left[u'\left(c(\omega^i,x,Y)\right) + a(Y) + \gamma\left(1 - h(\omega^i,x,Y)\right)\right] \tag{3.23}$$

where $c^i = c(\omega^i,x,Y) = xy(\omega^i,x,Y) + \hat{g}(x,Y)$ and $\hat{g}(x,Y) = (1-x)\bar{y}(x,Y) - Y$. Recall assumption *(1)* implies: $y^T > y^*_i > 0$ for all i and all $x,Y \in X$, which is equivalent to say that all individuals are in an internal maximum of their problem of utility maximization over consumption and leisure for any policy (x,Y). This assumption allows me to use an Envelope theorem when calculating $\frac{\partial V_i}{\partial x}$ and $\frac{\partial V_i}{\partial Y}$, for instance:

$$V^i_x = \omega^i \left[u'(c^i)\left(y^i_i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) + u'(c^i)x\omega\frac{dh^i}{dx} - \gamma\frac{dh^i}{dx}\right] \tag{3.24}$$

Because we have assumed to be in an interior solution of of the consumption/leisure problem, then the F.O.C. is: $u^i_c x\omega_i - \gamma = 0$. Using this result into (3.24) one gets:

$$V^i_x = \omega^i u'(c^i)\left(y^i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) \tag{3.25}$$

In the same way one can show that:

$$V^i_Y = \omega^i u'(c^i)\left((1-x)\frac{d\bar{y}}{dx} - 1\right) + \omega^i a'(Y) \tag{3.26}$$

SID. Calculate the derivative of V^i_x and V^i_Y w.r.t. ω using (3.25) and (3.26). One gets:

$$V^i_{x\omega} = \left[\omega^i u''(c^i)\cdot\frac{\partial y^i}{\partial \omega^i} + u'(c^i)\right]\left(y_i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) + \omega^i u'(c^i)\frac{\partial y^i}{\partial \omega^i} \tag{3.27}$$

$$V^i_{Y\omega} = -\omega^i u''(c^i)x\frac{\partial y^i}{\partial \omega^i}\left(1-(1-x)\frac{d\bar{y}}{dY}\right) + u'(c^i)\left((1-x)\frac{d\bar{y}}{dY} - 1\right) + a'(Y) \tag{3.28}$$

SM. Calculate the cross derivative V^i_{xY} using (3.25). One gets:

$$\begin{aligned} V^i_{xY} = \quad & \omega^i\left[u''(c^i)\left(x\tfrac{dy^i}{dY} + (1-x)\tfrac{d\bar{y}}{dY} - 1\right)\right]\left(y_i - \bar{y} + (1-x)\tfrac{d\bar{y}}{dx}\right) + \\ & +\omega^i u'(c^i)\left(\tfrac{dy^i}{dY} - \tfrac{d\bar{y}}{dY} + (1-x)\tfrac{d^2\bar{y}}{dxdY}\right) \end{aligned} \tag{3.29}$$

Therefore a coalitional political equilibrium exists if the three inequalities above are satisfied for all i at all $(x,Y) \in X$. I need to show that the cross derivatives above satisfy the inequalities (a) $V^i_{x\omega} > 0$ and (b) $V^i_{Y\omega} > 0$, and (c) $V^i_{xY} \geq 0$. (a) I know from Lemma 1 that for interior solutions of the C/L problem, i.e. $T > h^i > 0$ one gets $\frac{dy^i}{d\omega^i} = \frac{-u'(c^i)}{u''(c^i)\omega^i x} > 0$. Under the same assumption it is easy to calculate:

$$\frac{\partial y^i}{\partial x} = \frac{-u'(c^i)}{u''(c^i)x^2} - \frac{1}{x}\left(y^i - \bar{y} + (1-x)\frac{\partial \bar{y}}{\partial x}\right) \tag{3.30}$$

Notice that:

$$\frac{\partial \bar{y}}{\partial x} = \frac{\partial}{\partial x}\int_{\underline{\omega}}^{\bar{\omega}} y^i dR(\omega,\theta) = \int_{\underline{\omega}}^{\bar{\omega}} \frac{\partial y^i}{\partial x} dR(\omega,\theta) \tag{3.31}$$

Thus, integrating both sides of (3.30), one can write:

$$\frac{\partial \bar{y}}{\partial x} = \int_{\underline{\omega}}^{\bar{\omega}} \left[\frac{-u'(c(\omega^i,x,Y))}{u''(c(\omega^i,x,Y))x^2} - \frac{1}{x}\left(y^i - \bar{y} + (1-x)\frac{\partial \bar{y}}{\partial x}\right)\right] dR(\omega,\theta) \tag{3.32}$$

Solving equation (3.32) for $\frac{\partial \bar{y}}{\partial x}$ one gets the following:

$$\frac{\partial \bar{y}}{\partial x} = \int_{\underline{\omega}}^{\bar{\omega}} \frac{-u'(c(\omega^i,x,Y))}{u''(c(\omega^i,x,Y))x} dR(\omega,\theta) \tag{3.33}$$

and substituing (3.33) into (3.30) one gets:

$$\frac{\partial y^i}{\partial x} = \frac{-u'(c^i)}{u''(c^i)x^2} - \frac{1}{x}\left(y^i - \bar{y} + \frac{(1-x)}{x}\int_{\underline{\omega}}^{\bar{\omega}} \frac{-u'(c^i)}{u''(c^i)} dR(\omega,\theta)\right) \tag{3.34}$$

Substitute (3.33) into (3.25) to get:

$$V^i_x = \omega^i u'(c^i)\left(y^i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) \tag{3.35}$$

and (3.33) and (3.34) into (3.27) to get:

$$V^i_{x\omega} = \left[\omega^i u''(c^i)x\frac{\partial y^i}{\partial \omega^i} + u'(c^i)\right]\left(y_i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) + \omega^i u'(c^i)\frac{\partial y^i}{\partial \omega^i} \tag{3.36}$$

Notice that for interior solution of the C/L problem one gets:

$$\frac{\partial y^i}{\partial \omega^i} = \frac{-u'(c^i)}{u''(c^i)x\omega^i} \tag{3.37}$$

Substitute (3.37) into (3.36):

$$V^i_{x\omega} = \left[u''(c^i)\left(\frac{-u'(c^i)}{u''(c^i)}\right) + u'(c^i)\right]\left(y_i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) + u'(c^i)\left(\frac{-u'(c^i)}{u''(c^i)x}\right) =$$

$$= u'(c^i)\left(y_i - \bar{y} + (1-x)\frac{\partial\bar{y}}{\partial x}\right)(1-1) - \frac{u'(c^i)^2}{u''(c^i)x} = \tag{3.38}$$

$$= -\frac{u'(c^i)^2}{u''(c^i)x} > 0$$

Hence $V^i_{x\omega} > 0$ for all $\omega^i \in [\underline{\omega}, \bar{\omega}]$ and for all $(x,Y) \in X$. About condition (b), use the definition of $g(x,Y) = (1-x)\bar{y} - Y$ to derive:

$$\frac{dg(x,Y)}{dY} = (1-x)\frac{d\bar{y}}{dY} - 1 \tag{3.39}$$

Now totally differentiate the F.O.C. of the C/L problem, that is $\omega^i[u\left(xy^i + g(x,Y)\right) - \gamma] = 0$, $\omega^i \neq 0$:

$$\frac{\partial y^i}{\partial Y} = \frac{1-(1-x)\frac{d\bar{y}}{dY}}{x} \tag{3.40}$$

Recall that:

$$\frac{\partial \bar{y}}{\partial Y} = \frac{\partial}{\partial Y}\int_{\underline{\omega}}^{\bar{\omega}} y^i dR(\omega,\theta) = \int_{\underline{\omega}}^{\bar{\omega}} \frac{\partial y^i}{\partial Y} dR(\omega,\theta) \tag{3.41}$$

Thus, integrating both sides of (3.40) and solving for $\frac{\partial\bar{y}}{\partial Y}$ one gets:

$$\frac{\partial \bar{y}}{\partial Y} = \frac{\partial y^i}{\partial Y} = 1 \tag{3.42}$$

Substitute (3.42) into (3.28) to get:

$$\begin{aligned} V^i_{Y\omega} = & -\omega^i u''(c^i)x^2\left(\frac{-u'(c^i)}{u''(c^i)\omega^i x}\right) - xu'(c^i) + a'(Y) = \\ & = xu'(c^i) - xu'(c^i) + a'(Y) = a'(Y) > 0 \end{aligned} \tag{3.43}$$

Hence $V^i_{Y\omega} > 0$ for all $\omega^i \in [\underline{\omega}, \bar{\omega}]$ and for all $(x,Y) \in X$. About condition (c), use the result (3.42) into (3.29) to get:

$$\begin{aligned} V^i_{xY} = & \ \omega^i\left[u''(c^i)(x+1-x-1)\right]\left(y^i - \bar{y} + (1-x)\frac{d\bar{y}}{dx}\right) + \\ & +\omega^i u'(c^i)(1-1+(1-x)0) = 0 \end{aligned} \tag{3.44}$$

Hence $V^i_{xY} = 0$ for all $\omega^i \in [\underline{\omega}, \bar{\omega}]$ and for all $(x,Y) \in X$. Thus, we got that (a) $V^i_{x\omega} > 0$ and (b) $V^i_{Y\omega} > 0$, and (c) $V^i_{xY} \geq 0$ for all $\omega^i \in [\underline{\omega}, \bar{\omega}]$ and for all $(x,Y) \in X$. Q.E.D.

Theorem 3. *A Coalitional Equilibrium exists if (i) $V(x,Y;\omega^i)$ is concave in (x,Y) for all i and if (ii) the individual labour supply is such that $T > h^i > 0$, for all i and all $(x,Y) \in X$. Moreover, if $V(x,Y;\omega^m)$ is strictly concave in (x,Y), then the set of equilibrium policies is a singleton.*

Proof. Recall that the sufficient conditions for the existence of a Coalitional Equilibrium are: (i) The policy space X is a convex and complete sublattice of R^d; (ii) the objective function V^i is jointly continuous and concave in (x,Y), (iii) the objective function V^i satisfies *SM* and *SID* for all *i* and all $(x,Y) \in X$. Condition (i) is always satisfied as the policy space assumed in this example is a convex and complete sublattice of R^2. Regarding condition (ii), notice that V^i is jointly continuous in (x,Y) because $\omega^i \left[u'\left(c(\omega^i,x,Y)\right) + a(Y) + \gamma\left(1 - h(\omega^i,x,Y)\right)\right]$ is a continuous function of c,Y,h and the functions $c(\omega^i,x,Y)$ and $h(\omega^i,x,Y)$ are also jointly continuous and differentiable (see Lemma 1). Concavity in (x,Y) is assumed to hold for all $(x,Y) \in X$ and all $\omega^i \in [\underline{\omega}, \bar{\omega}]$. Condition (iii) is satisfied under the conditions stated in the Theorem because of Lemma 2. Q.E.D.

3.6.1.2 Comparative Statics.

Theorem 4. *(i) The spending in Public Goods is weakly increasing in the median income, and (ii) the tax rate on labour income is weakly decreasing in the median income.*

Proof. The proof to Theorem 3 implies that all the conditions for a coalitional equilibrium are satisfied. Thus, the comparative static results in Dotti (2015) apply to the equilibrium policy outcome. Specifically, this result is an immediate consequence of Theorem 2 in Dotti (2015). See chapter 2.

Lemma 5. *If the income distribution at the equilibrium policy is such that $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$, then (i) the median-to mean income ratio is weakly increasing in θ in a neighborhood of (x^*,Y^*). Thus, if $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$ in the neighborhood of a coalitional equilibrium, then (ii) a marginal increase in median income results in both a weakly higher equilibrium policy and a weakly larger median-to-mean income ratio.*

Proof. It is sufficient to show that:

$$\frac{dx^*}{d\theta}\frac{\partial}{\partial x}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\bigg|_{x=x^*,Y=Y^*}\geq 0 \tag{3.45}$$

and

$$\frac{dY^*}{d\theta}\frac{\partial}{\partial Y}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\bigg|_{x=x^*,Y=Y^*}\geq 0 \tag{3.46}$$

Notice that:

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\bigg|_{x=x^*,Y=Y^*} &= \frac{1}{\bar{y}(x,Y)^2}\left[\underbrace{\frac{-u'(c^m)\bar{y}(x,Y)^2}{u''(c^m)x^2}}_{A}+\right.\\ &\left.+\underbrace{\frac{-\bar{y}(x,Y)^2}{x}\left(y(\omega^m,x,Y)-\bar{y}+(1-x)\frac{\partial\bar{y}}{\partial x}\right)}_{B}+\underbrace{y(\omega^m,x,Y)\frac{\partial\bar{y}}{\partial x}}_{C}\right]\end{aligned} \tag{3.47}$$

Recall that $\frac{dx^*}{d\theta}\geq 0$ and $\frac{dY^*}{d\theta}\geq 0$ because of Theorem 4. First, part A in the formula above is strictly positive because of the strict concavity of u. Second, notice that in a coalitional equilibrium the median voter achieve her ideal policy. Thus, if $x^*<\bar{x}$, the F.O.C. of the maximization problem with respect to x for voter m implies $V_x^m=\omega^m u'(c^i)\left(y^m-\bar{y}+(1-x)\frac{d\bar{y}}{dx}\right)\leq 0$. Because $\omega^m u'(c^i)>0$, this implies $y^m-\bar{y}+(1-x)\frac{d\bar{y}}{dx}\leq 0$ which means that part B is also positive. Lastly, $\frac{\partial\bar{y}}{\partial x}\geq 0$ as shown in the proof to theorem 2. Thus, $\frac{dx^*}{d\theta}\frac{\partial}{\partial x}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\Big|_{x=x^*,Y=Y^*}\geq 0$. About the case in which $x^*=\bar{x}$, Theorem 3 implies $\frac{dx^*}{d\theta}=0$, and the strict concavity of $V(x,Y;\omega)$ in (x,Y) implies that $\frac{\partial}{\partial x}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\Big|_{x=x^*,Y=Y^*}$ is finite. Thus the inequality $\frac{dx^*}{d\theta}\frac{\partial}{\partial x}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\Big|_{x=x^*,Y=Y^*}\geq 0$ is also trivially satisfied. Regarding condition in (3.46), this can be rewritten as follows:

$$\begin{aligned}&\frac{dY^*}{d\theta}\frac{\partial}{\partial Y}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\bigg|_{x=x^*,Y=Y^*}=\\ &=\frac{dY^*}{d\theta}\frac{1}{\bar{y}(x,Y)^2}\left(\frac{\partial y(\omega^m,x,Y)}{\partial Y}\bar{y}(x,Y)-\frac{\partial\bar{y}(x,Y)}{\partial Y}y(\omega^m,x,Y)\right)\end{aligned} \tag{3.48}$$

Using the results in (3.42), formula (3.48), formula (3.48) becomes:

$$\frac{dY^*}{d\theta}\frac{\partial}{\partial Y}\left(\frac{y(\omega^m,x,Y)}{\bar{y}(x,Y)}\right)\bigg|_{x=x^*,Y=Y^*}=\frac{dY^*}{d\theta}\frac{\bar{y}(x,Y)-y(\omega^m,x,Y)}{\bar{y}(x,Y)^2} \tag{3.49}$$

which is weakly positive as long as the income distribution at $(x,Y)=(x^*,Y^*)$ is such that

$y^m \leq \bar{y}$. Q.E.D.

Theorem 6. *There exists a threshold $\hat{g} \geq 0$ such that, if $g(x,Y) \leq \hat{g}$, then (i) the total government spending, (ii) the total tax revenue, and (iii) the average tax rate are all weakly increasing in the median income in a neighborhood of (x^*,Y^*). Thus, if $y(\omega^m,x^*,Y^*) \leq \bar{y}(x^*,Y^*)$ and the degree of progressivity is sufficiently low for all income levels in the neighborhood of the coalitional equilibrium, then (iv) a marginal increase in median income results both in a weakly larger size of the government and in a weakly larger median-to-mean income ratio.*

Proof. Part (i) I use the definition of the size of the government as the total government spending corresponding to (3.15):

$$\begin{gathered} TS(x,Y,\theta) \equiv Y + \\ + \int_{\underline{\omega}}^{\bar{\omega}} \mathbb{I}\left[(1-x)y(\omega,x,Y) - g(x,Y) \leq 0\right]\left[g(x,Y) - (1-x)y(\omega,x,Y)\right] r(\omega,\theta) d\omega \end{gathered} \tag{3.50}$$

(1) if $g(x,Y) \geq 0$ and $g(x,Y) - (1-x)y(\underline{\omega},x,Y) \geq 0$ (this implies $g(x,Y) \geq 0$), recall formula (3.17) states:

$$TS^1(x,Y,\theta) = \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} \left[g(x,Y) - (1-x)y(\omega,x,Y)\right] r(\omega,\theta) d\omega + Y$$

differentiate (3.17) w.r.t. θ and evaluate at $\theta = 0$:

$$\begin{aligned} \left.\frac{dTS(x,Y,\theta)}{d\theta}\right|_{\theta=0} = {} & \frac{d\hat{\omega}(x,Y)}{d\theta}\left[g(x,Y) - (1-x)y(\hat{\omega}(x,Y),x,Y)\right] f(\hat{\omega}(x,Y)) + \\ & + \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} \left[\frac{dx^*}{d\theta}\left(y^i - (1-x)\frac{\partial y^i}{\partial x} + \frac{\partial g(x,Y)}{\partial x}\right) + \right. \\ & \left. + \frac{dY}{d\theta}\left((1-x)\frac{\partial y^i}{\partial Y} - \frac{\partial g(x,Y)}{\partial Y}\right)\right] f(\omega) d\omega + \\ & + \frac{dY^*}{d\theta} + \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} (1-x)y(\omega,x,Y)\left[f(\omega) - \tilde{f}(\omega)\right] d\omega \end{aligned} \tag{3.51}$$

Notice that the definition of $g(x,Y)$ implies that $g(x,Y) - (1-x)y(\hat{\omega}(x,Y),x,Y) = 0$. Thus (3.51) reduces to:

$$\begin{aligned} = {} & -\frac{dx^*}{d\theta}\int_{\underline{\omega}}^{\hat{\omega}(x,Y)} \frac{\partial}{\partial x}\left\{(1-x)\left[y(\omega,x,Y) - \bar{y}(\omega,x,Y)\right]\right\} f(\omega) d\omega + \\ & + \left[1 + F(\hat{\omega}(x,Y))\right]\frac{dY^*}{d\theta} + \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} (1-x)y(\omega,x,Y)\left[f(\omega) - \tilde{f}(\omega)\right] d\omega \end{aligned} \tag{3.52}$$

Theorem 4 implies that $\frac{dx^*}{d\theta} \geq 0$. If at the equilibrium policy $\hat{\omega}(x,Y) - \underline{\omega}$ is small enough, then $\left.\frac{dTS(x,Y,\theta)}{d\theta}\right|_{\theta=0} \geq 0$. Notice that if $g(x,Y)-(1-x)y(\underline{\omega},x,Y) \leq 0$ then $\hat{\omega}(x,Y) \notin [\underline{\omega},\bar{\omega}]$ and the formula above is weakly positive. Because of the continuity of g and y, there exists a threshold $\hat{g}$ such that if at the equilibrium $g(x,Y) \leq \hat{g}$, then $\hat{\omega}(x,Y)-\underline{\omega}$ is small enough, to ensure that the derivative of interest is weakly positive.

(2) if $g(x,Y)-(1-x)y(\underline{\omega},x,Y) < 0$, then $TS(x,Y,\theta) = Y$ and therefore:

$$\left.\frac{dTS(x,Y,\theta)}{d\theta}\right|_{\theta=0} = \frac{dY^*}{d\theta} > 0 \tag{3.53}$$

Part (ii) total tax revenues, it is easy to show that if (1) $g(x,Y) \geq 0$ and $g(x,Y)-(1-x)y(\underline{\omega},x,Y) \geq 0$, then

$$TR(x,Y,\theta) = \int_{\hat{\omega}(x,Y)}^{\bar{\omega}} [(1-x)y(\omega,x,Y)-g(x,Y)]\, r(\omega,\theta)d\omega \tag{3.54}$$

Rewrite (3.54) as follows:

$$\begin{aligned} &= \int_{\underline{\omega}}^{\bar{\omega}} [(1-x)y(\omega,x,Y)-g(x,Y)]\, r(\omega,\theta)d\omega + \\ &- \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} [(1-x)y(\omega,x,Y)-g(x,Y)]\, r(\omega,\theta)d\omega \end{aligned} \tag{3.55}$$

Thus, (3.55) reduces to:

$$= \int_{\underline{\omega}}^{\hat{\omega}(x,Y)} [g(x,Y)-(1-x)y(\omega,x,Y)]\, r(\omega,\theta)d\omega + Y = TS(x,Y,\theta) \tag{3.56}$$

As expected, the total tax revenue has same formula as the total spending in equilibrium. Thus, it must be the case that

$$\left.\frac{dTR(x,Y,\theta)}{d\theta}\right|_{\theta=0} = \left.\frac{dTS(x,Y,\theta)}{d\theta}\right|_{\theta=0} \tag{3.57}$$

For the same reasons explained in part (i), if at the equilibrium policy $\hat{\omega}(x,Y)-\underline{\omega}$ is small enough, which is the case if at the equilibrium $g(x,Y) \leq \hat{g}$ for a threshold $\hat{g} \geq 0$, then $\left.\frac{dTR(x,Y,\theta)}{d\theta}\right|_{\theta=0} \geq 0$. Lastly, (iii) consider the average tax rate defined in (3.18):

$$AT(x,Y) = \int_{\underline{\omega}}^{\bar{\omega}} [(1-x)y(\omega,x,Y)-g(x,Y)]\, r(\omega,\theta)d\omega/\bar{y}(x,Y) = Y/\bar{y}(x,Y)$$

Differentiate w.r.t. θ and evaluate at $\theta = 0$:

$$\left.\frac{dAT(x,Y,\theta)}{d\theta}\right|_{\theta=0} = \frac{dY^*}{d\theta}\frac{1}{\bar{y}(x,Y)}\left(\bar{y}(x,Y) - Y\frac{d\bar{y}}{dY}\right) = \frac{dY^*}{d\theta}\left(\frac{\bar{y}(x,Y)-Y}{\bar{y}(x,Y)}\right) \geq 0 \quad (3.58)$$

Because $\bar{y}(x,Y) - Y \geq 0$ the total income in the society must (weakly) exceed the total public spending in the public good, and $\frac{dY^*}{d\theta} \geq 0$ because of Theorem 4. Regarding part (iv), it is straightforward from points (i), (ii), (iii) and Lemma 5. Q.E.D.

Theorem 7. *If the tax system is weakly progressive at (x^*, Y^*), then (i) the degree of progression of the tax system is weakly decreasing in the in the median income in a neighborhood of (x^*, Y^*). Thus, if $y(\omega^m, x^*, Y^*) \leq \bar{y}(x^*, Y^*)$ in the neighborhood of the coalitional equilibrium, then (ii) a marginal increase in median income results both in a weakly lower progression of the tax system and in a weakly larger median-to-mean income ratio.*

Proof. (i) Recall avarage rate progression of the tax system has formula (3.20):

$$ARP(x,Y;y) = \frac{g(x^*(\theta), Y^*(\theta))}{y^2}$$

I check how this index moves, for each income level y, if a shock on the wage distribution that causes a marginal increase in median income at constant mean in the proximity of the equilibrium policy (x^*, Y^*).

$$\left.\frac{dARP(y,\theta)}{d\theta}\right|_{y} = \frac{1}{y^2}\left\{\frac{dx^*}{d\theta}\left(-\bar{y}(x^*,Y^*) + (1-x^*)\frac{\partial \bar{y}(x,Y)}{\partial x}\right) - \frac{dY^*}{d\theta}x^*\right\} \quad (3.59)$$

Notice that the above must be (weakly) negative. The reason is that either $(x^*(0), Y^*(0))$ is an interior solution for the median voter with respect to x, and therefore $y(\omega^m, x, Y) - \bar{y}(x,Y) + (1-x)\frac{\partial \bar{y}(x,Y)}{\partial x} = 0$, or it is a corner solution with respect to x, and therefore $\frac{dx^*}{d\theta} = 0$. In the former case, $y(\omega^m, x, Y) > 0$ implies $-\bar{y}(x,Y) + (1-x)\frac{\partial \bar{y}(x,Y)}{\partial x} \leq 0$. Theorem 4 implies $\frac{d^*x}{d\theta} \geq 0$ and $\frac{dY^*}{d\theta} \geq 0$, thus the overall sign of the formula above is weakly negative. In the latter case, $\frac{dx^*}{d\theta} = 0$, $\frac{dY^*}{d\theta} \geq 0$ and the fact that $-\bar{y}(x,Y) + (1-x)\frac{\partial \bar{y}(x,Y)}{\partial x}$ is finite ensure that the sign is also weakly negative. Similarly, using the definition of *IRP* in (3.22) one gets:

$$\left.\frac{dIRP(y,\theta)}{d\theta}\right|_{y} = \frac{1}{yx}\left\{\frac{dx^*}{d\theta}\left(-\bar{y}(x^*,Y^*) + (1-x^*)\frac{\partial \bar{y}(x,Y)}{\partial x} - g(x,Y)\right) - \frac{dY^*}{d\theta}x^*\right\} \quad (3.60)$$

which is weakly negative if $g(x,Y) \geq 0$. Part (ii) is straightforward from Lemma 5 and Theorem 7 (i). Q.E.D.

3.6.1.3 Change in the the wage distribution

Lemma 8. *If there exists a continuous income distribution that is (i) a mean-preserving spread of the income distribution given policy (x,Y) and that (ii) has different median relative to the the original distribution, then there exists a continuous wage distribution with p.d.f. $\tilde{f}(\omega)$, such that (i) the mean income is the same under both distribution at constant policy (x,Y), i.e. $\int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x,Y)f(\omega)d\omega = \int_{\underline{\omega}}^{\bar{\omega}} y(\omega,x,Y)\tilde{f}(\omega)d\omega = \bar{y}(x,Y)$ and (ii) the median income is strictly higher under the new distribution conditional on policy (x,Y), i.e.$\tilde{F}(\omega^m)d\omega < F(\omega^m)d\omega$.*

Proof. Define c.d.f. $S(y)$ such that $S(y(\omega,x,Y)) = F(\omega)$ for all ω at given policy (x,Y). Suppose there exists a distribution with c.d.f. $\tilde{S}(y)$ such that:

$$\int_{y(\underline{\omega},x,Y)}^{y(\bar{\omega},x,Y)} \tilde{S}(y)dy = \int_{y(\underline{\omega},x,Y)}^{y(\bar{\omega},x,Y)} S(y)dy \tag{3.61}$$

i.e. $\tilde{S}$ is a mean-preserving transformation of S, and

$$\tilde{S}(y(\omega^m,x,Y)) < S(y(\omega^m,x,Y)) \tag{3.62}$$

which implies that the median under distribution $\tilde{S}$ is strictly higher relative to the median under distribution S. Denote with $y^{-1}_{x,Y}(y)$ the inverse of function $y(\omega,x,Y)$ at fixed (x,Y). Such function exists because $y(\omega,x,Y)$ is strictly increasing in ω. Lastly, define a wage distribution with c.d.f. $\tilde{F}$ such that $\tilde{F}(y^{-1}_{x,Y}(y)) = \tilde{S}(y)$ for all y. It is easy to show that such function satisfies the requirement stated above. Q.E.D.

3.6.2 Example

Consider the following example: there is a continuum of voters of Lebesgue measure 1, individuals have wages $\omega \in [\underline{\omega},\bar{\omega}]$ with $E(\omega) = \mu$ and c.d.f. $F(\omega)$. Preferences are represented by the utility function:

$$U(c^i,Y,l^i) = \alpha\ln(c_i) + (1-\alpha)\ln(Y) + \gamma l_i \tag{3.63}$$

The policy space is $X := \{(x,Y)|x \in [\underline{x},1], Y \in [0,\overline{Y}]\}$ with $\underline{x} > 0$ and with $\underline{x}$ and $\bar{Y}$ chosen such that $\underline{x} \geq \frac{\mu-\underline{\omega}}{\mu}$ and $\bar{Y} \leq (T-\alpha/\gamma)\underline{\omega}$ with $T-\alpha/\gamma > 0$. These restrictions on the policy space and the parameter set ensure that the solution to the Consumption/Leisure problem is interior for all voters and for all policies in X. Solve for the optimal labor supply conditional on the policy (x,Y) :

$$h^i = h(\omega^i,x,Y) = \frac{\alpha}{\gamma} - \frac{(1-x)\bar{y}}{x\omega^i} + \frac{Y}{x\omega^i} \tag{3.64}$$

This implies that:

$$\bar{y}(x,Y) = \frac{\alpha x\mu}{\gamma} + Y \tag{3.65}$$

Thus, it is easy to derive the formula for individual income:

$$y^i = y(\omega^i,x,Y) = \frac{\alpha}{\gamma}x[\omega^i - (1-x)\mu] + Y \tag{3.66}$$

and for individual consumption of private goods:

$$c(\omega^i,x,Y) = y(\omega^i x,Y) + g(x,Y) = \frac{\alpha x\omega^i}{\gamma} \tag{3.67}$$

which implies that consumption is positive for all ω^i. Substitute (3.64) and (3.67) into (3.63) to get, for each voter i with wage ω^i:

$$V^i = \alpha\ln\left(\alpha x\omega^i\right) - \alpha\ln(\gamma) + (1-\alpha)\ln(Y) - \alpha + (1-x)\frac{\alpha\mu}{\omega_i} - \frac{\gamma Y}{\omega_i} \tag{3.68}$$

The partial derivative with respect to x and Y are $V^i_x = \frac{\alpha}{x} - \frac{\gamma\mu}{\omega^i}$ and $V^i_Y = \frac{1-\alpha}{Y} - \frac{\gamma}{\omega^i}$ respectively. The cross derivatives are:

$$V^i_{x\omega} = \frac{\gamma\mu}{(\omega^i)^2} > 0 \tag{3.69}$$

$$V^i_{Y\omega} = \frac{\gamma}{\omega^i} > 0 \tag{3.70}$$

$$V^i_{xY} = 0 \tag{3.71}$$

Hence *SM* in (x,Y) and *SID* in $(x,Y;\omega)$ hold for all $(x,Y) \in X$ and for all $\omega \in [\underline{\omega},\bar{\omega}]$. Notice that the indirect utility function is strictly concave in x,Y for all i if the Hessian Matrix of the objective function is negative definite for all $\omega \in [\underline{\omega},\bar{\omega}]$. In this case one gets: $V^i_{xx} = -\frac{\alpha}{x^2} < 0$, $V^i_{YY} = -\frac{1-\alpha}{Y^2} < 0$ and $V^i_{xY} = 0$ for all $\omega \in [\underline{\omega},\bar{\omega}]$. Thus the condition is satisfied. Notice that:

$$x^i = \begin{cases} \underline{x} & if \quad \omega^i \leq \gamma\mu\underline{x}/\alpha \\ \frac{\alpha}{\gamma}\frac{\omega^i}{\mu} & if \quad \gamma\mu\underline{x}/\alpha \leq \omega^i \leq \gamma\mu/\alpha \\ 1 & if \quad \omega^i \geq \gamma\mu/\alpha \end{cases} \tag{3.72}$$

and that:

$$Y^i = \max\left[(1-\alpha)\omega^i/\gamma, \overline{Y}\right] \tag{3.73}$$

Formulas (3.72) and (3.73) confirm the expected result that the optimum is monotonic nondecreasing in ω^i. Now recall that $g(x,Y) = (1-x)\bar{y}(x,Y) - Y$. In this example for $Y < \overline{Y}$ this is equivalent to:

$$g(x^*,Y^*) = \begin{cases} \frac{\alpha\omega^m}{\gamma^2}\left(\alpha - \frac{\alpha^2\omega^m}{\gamma\mu} - \frac{(1-\alpha)\omega^m}{\mu}\right) & if \quad \underline{x} < x < 1 \\ (\alpha\mu - (1-\alpha)\omega^m)x/\gamma & if \quad x = \underline{x} \\ -\frac{(1-\alpha)\omega^m}{\gamma} & if \quad x = 1 \end{cases} \tag{3.74}$$

Notice that $g \leq 0$ if $x = 1$ or if $x < 1$ and $\frac{\omega^m}{\mu} \geq \frac{\alpha\gamma}{\alpha^2+\gamma(1-\alpha)}$, but interior solutions occur only if $\frac{\omega_m}{\mu} < \frac{\gamma}{\alpha}$. This implies that for $\frac{\gamma}{\alpha} > \frac{\omega_m}{\mu} \geq \frac{\alpha\gamma}{\alpha^2+\gamma(1-\alpha)}$ there is an interior solution with $g \leq 0$. Moreover we have internal solutions with $g > 0$ for $\frac{\gamma}{\alpha} > \frac{\omega_m}{\mu} \geq \frac{\alpha\gamma}{2[\alpha^2+\gamma(1-\alpha)]}$ which is the threshold that ensures that the total tax revenue does not exceed the total income. Also one can prove that $\mu \geq \frac{\alpha}{\alpha^2-\gamma(1-\alpha)}$ and a suitable choice of $\underline{Y}$ is sufficient to ensure that the objective function is$\omega_i - concave$ for all i. Finally notice that for interior solutions:

$$\frac{\partial g}{\partial \omega^m} = \frac{\alpha}{\gamma^2}\left(\alpha - \frac{2\alpha^2}{\gamma\mu} - \frac{2(1-\alpha)}{\mu}\right) < 0 \tag{3.75}$$

within the range of parameters for which the solution is interior. Q.E.D.

Tables

Table 3.1. Summary of Recent Studies: Inequality and Redistribution

Source	Sample Size	Period	Data Structure	Measures: Redistribution[a]	Measures: Inequality	Correlation: Sign[b]	Correlation: Significance
Bassett, Burkett, and Putterman (1999)	Up to 54 countries	1970-1985	Cross-country average	Social security and welfare	Mostly Q3 in 1960s[c]	Generally negative	Inconsistent
Easterly and Rebelo (1993)	Not available	1970-1988	Cross-country average	Spending variables	Gini, various income shares	Positive	Significant
Figini (1998)	Up to 63 countries	1970-1990	Cross-country average	Tax rates, total revenue, and total spending	Gini coefficient in 1970	Nonlinear	Significant
Gouveia and Masia (1998)	50 U.S. states	1979-1991	Panel	Spending variables	Ratio of mean to median income	Generally negative	Generally significant
Lindert (1996)	14 OECD countries	1962-1981	Panel	Spending variables	Income gap index	Negative	Generally insignificant
Meltzer and Richard (1983)	U.S. average	1937-1977	Time series	Spending variables	Ratio of mean to median income	Positive	Significant
Milanovic (2000)	24 mostly OECD countries	1967-1997[d]	Panel	Gain by poorest quintile or poorest half	Pretransfer Gini coefficient	Positive	Significant
Panizza (1999)	46 U.S. states	1970-1980	Cross-state average	Tax, tax progressivity, and spending variables	Q3 in 1970[c]	Generally positive	Inconsistent
Partridge (1997)	48 U.S. states	1960-1990[e]	Panel	Tax, employment, and spending variables	Pretax Gini, Q3	Inconsistent	Generally significant
Perotti (1992)	40 democracies	1970-1985	Cross-country average	Spending variables	Q3 in 1970[c]	Negative	Insignificant
Perotti (1994)	52 countries	1970-1985	Cross-country average	Transfers	Q3 in 1960[c]	Negative	Insignificant
Perotti (1996)	49 countries	1970-1985	Cross-country average	Tax rates and spending variables	Q3 and Q4 in 1960[c]	Generally positive	Generally insignificant
Perrson and Tabellini (1994)	13 OECD countries	1960-1981	Cross-country average	Transfers	Q3 in 1965[c]	Positive	Insignificant
Rodriguez (1999a)	50 U.S. states	1984-1994	Time series and cross-state average[f]	Spending variables	Distribution skewness	Inconsistent	Insignificant
Tanninen (1999)	Up to 45 countries	1970-1988	Cross-country average	Spending variables	Adjusted Gini coefficient in 1970s[g]	Inconsistent	Generally insignificant

Note: See text for further discussion of these results.
a. In percentage of GDP unless otherwise indicated.
b. Negative means greater inequality is associated with *less* spending.
c. A higher Q3 means greater income equality. For consistency with other studies in the reporting of main results, Q3 and Q4 are taken to mean –Q3 and –Q4.
d. Total number of observations is 79.
e. Total number of observations is 144.
f. Refers to national time-series.
g. Adjusted for variations in the Gini definition.

Source: De Mello and Tiongson, 2006

Table 3.2. Determinants of Labor Tax Rate and Social Transfers

	Labor Tax Rate (1)	Labor Tax Rate (2)	Social Transfers (3)	Social Transfers (4)
Dependency ratio	−.382	−.383	−7.493	−7.492
	(−4.02)	(−4.40)	(−8.81)	(−8.80)
Government jobs/employment	.915	.729	4.467	4.611
	(12.17)	(10.01)	(6.64)	(6.47)
Trade openness	.198	.131	.740	.792
	(8.09)	(5.45)	(3.73)	(3.37)
Per capita GDP growth	−.187	−.127	−2.716	−2.762
	(−2.83)	(−2.09)	(−4.59)	(−4.63)
Rich/middle income share	−.055	−.049	.276	.271
	(−2.77)	(−2.66)	(1.55)	(1.52)
Unemployment rate		.480		−.370
		(7.82)		(−.62)
R^2	.753	.793	.617	.618

Note.—All specifications include country fixed effects (coefficients not shown).

Source: Razin et al., 2002

4 The Political Economy of Immigration and Population Ageing

I investigate the effects of population ageing on immigration policies. Voters' attitude towards immigrants depends on how the net gains from immigration are divided up in the society by the fiscal policy. In the theoretical literature this aspect is treated as exogenous to the political process because of technical constraints. This generates inconsistent predictions about the policy outcome. I adopt a new equilibrium concept for voting models to analyse the endogenous relationship between immigration and fiscal policies and solve this apparent inconsistency. I show that the elderly and the poor have a common interest in limiting immigration and in increasing public spending. This exacerbates the effects of population ageing on public finances and results in a high tax burden on working age individuals and further worsens the age profile of the population. Moreover, I show that if the share of elderly population is sufficiently large, then a society is unambiguously harmed by the tightening in the immigration policy caused by the demographic change. The implications of the model are consistent with the patterns observed in UK attitudinal data and in line with the findings of the empirical literature about migration.

JEL classification: D72, C71, J610, H550.

Keywords: Immigration, Ageing, Policy, Voting.

4.1 Introduction

What are the effects of population ageing on immigration policies? Do ageing societies tend to impose excessive restrictions on the inflow of foreign workers and if so, why? Should we expect an adjustment in immigration and spending policies to mitigate the impact of population ageing on public finances? This paper attempts to answer these questions using a theoretical model. In particular, I investigate why rapidly ageing countries - that arguably need more legal immigration - are imposing increasing restrictions on the inflow of immigrant workers and how this choice affects the tax burden faced by the working population. I also analyse the effects of these policy changes on the welfare of current and future generations. The importance of these questions is related to the vast fiscal effects of population ageing and immigration. The increase in longevity implies rising costs for the public sector, in particular the ones of public pensions and health care. The fall in the fertility rates causes an insufficient growth in the tax base. Both result in a pressure on public finances and tax rates. Several scholars and policy makers suggest that legal immigration can help in mitigating the effects of this problem, but this can happen only if there is political support for an increasingly open immigration policy. This analysis is therefore crucial to assess the fiscal soundness of ageing societies in the long run. Immigration also have demographic, social and cultural implications. Hence the study of immigration policies is also important to understand the evolution of the structure of our society in a broader sense.

4.1.1 Methods

In keeping with previous literature (Razin and Sadka, 1999), I analyse a political economy model with overlapping generations, in which voters differ in their income and in their age. In contrast with previous literature, however, I depart from a unidimensional policy space. Specifically, in each period the society chooses a two-dimensional policy consisiting of an immigration quota and of the provision of an imperfect public good. The elderly receive an exogenous public pension that is financed by the tax revenues. The government budget is balanced, hence the political choice determines the tax rate on labour income. The bi-dimensionality of the policy allows one to model endogenously both the immigration policy and how the net fiscal benefits from immigration are divided up in the society. In detail, if immigrants generate a fiscal surplus, voters can employ it to increase public spending and/or to reduce taxes. The first choice mostly benefits the elderly and the low-income individuals, while the second favours the high earners. This implies that the way in which the net gains are divided up by the fiscal system is crucial to correctly assess the attitude towards immigration of different groups of voters. An endogenous analysis of both the immigration and the fiscal policy requires a bi-dimensional policy. Thus the standard tools in the Political Economy literature - based on unidimensionality - cannot be used to answer this question. In order to address this problem, I adopt a dynamic version of the model of electoral competition and of the concept of coalitional equilibrium proposed in Dotti (2015). In such theoretical framework simple ordinal preference restrictions are sufficient to deliver existence of equilibrium and sharp comparative static results on the policy outcome. This is a consequence of a key restriction on the political process. Specifically, single politicians cannot commit to any platform other than their ideal policies, but they can form coalitions to enhance their ability to commit through internal agreements. Coalitions must be stable in equilibrium, in the sense that no subcoalition has a strict incentive to deviate and propose a different policy platform. I adopt this notion of equilibrium to study an overlapping generations model of immigration and public spending. This allows me to analyse how shocks on the longevity and on the fertility of the population affect immigration policy, public spending and the tax rate faced by the working population.

4.1.2 Summary of Results

I show that the elderly and the low income individuals have a common interest in reducing immigration and increasing public spending. Population ageing causes both an increase in the political power of these groups and a pressure on the government budget due to the rising cost of pensions. These two channels underpin the main results of this paper, which are as follows.

First, I show that, if the share of elderly is sufficiently large, a rise in the longevity and/or a fall in the natural growth rate of the population cause a tightening in the immigration policy and an increase in public spending. The reduced inflow of immigrant workers implies a reduction in the tax base. This, together with the rise in public spending in public goods and pensions, causes a sharp rise in the tax rate. Hence the political process tends to exacerbate the effects of population ageing on public finances.

Second, the effects of demographic shocks tend to worsen with time. In detail, a reduction in the immigration quota in the current period implies a change in the future age profile of the population because immigrants are mostly young and have weakly higher fertility rates relative to the natives. This causes further population ageing in the following periods and reinforces the effects.

Third, if the share of retired population is sufficiently large, then the tightening in the immigration policy generates a welfare loss for the society as a whole and harms the future generations.

These results suggest that ageing countries, that arguably need more immigration, tend to reduce it instead. This causes vast and persistent welfare and demographic effects and can affect the fiscal sustainability of the public sector in these countries.

4.1.3 Related Literature

Population ageing has been significant since the mid-twentieth century and it is expected to have dramatic demographic consequences in the next decades (see Figure 1). On one hand there are strong theoretical and empirical arguments in support of legal immigration as an instrument to ensure the financial soundness of a rapidly ageing society (Razin and Sadka, 1999 and Dustmann and Frattini, 2014). On the other hand the recent political debate in many countries is dominated by the discussion about how to limit the inflow of foreigners by introducing increasingly restrictive immigration policies. Immigration policies have been a dominant topic of recent political campaigns, such as the one for the referendum on Britain's European Union membership held in June 2016 and the one for the primary elections of the Republican Party in the USA in the same year. In many European countries this political agenda has led to a substantial tightening of immigration restrictions from 1994 (Boeri and Brucker, 2005) as shown in Figure 2. For the USA, Ortega and Peri (2009) provide evidence of an increase in the restrictiveness of immigration policy in the period 1994-2005. These trends in the implemented policies are consistent with a widespread and increasing aversion to immigration in those countries. Attitudinal data show that in the UK the share of citizens that would like immigration into their country to decrease has risen from 72.8% to a staggering 79.1% during the last 10 years (British Social Attitude Survey, 2003-2013). Moreover, the elderly are consistently more averse to immigration relative to the young. In the UK 85.7% of the individuals aged 60 or over would like less immigration while 71.2% of the individuals under 40 years old share the same opinion (British Social Attitude Survey, 2013). In the USA, the corresponding values are 47.3% and 39.2% (General Social Survey, 2014). These statistics suggest that population ageing may play an important role in the collective choice about immigration policies.

The empirical studies of the determinants of immigration policy are mostly based on attitudinal data and provide two main consistent facts that are relevant for this paper. The first fact is that age, education and income have a significant impact on the disapproval of further immigration and that in particular the elderly tend to have stronger preferences against further immigration in comparison with the young. Dustmann and Preston (2007), Facchini and Mayda (2007) and Card *et al.* (2012), using respectively data from the British Social Attitude Survey, the International Social Survey Programme and the European Social Survey, all support this finding. The latter paper also provides evidence that this result is mainly due to the perceived effect of immigration on the composition of the community in which the respendents live (or "compositional amenities") and to its economic effects. The second important fact is that economic hostility to immigration is driven by concern about effects on public finances at least as much as by effects on labour market outcomes (Dustmann and Preston, 2006, 2007; Boeri, 2010). Consistently with this finding, Milner and Tingley (2009) show that public finance aspects play a major role in shaping the immigration policy in the US. This is somewhat surprising given that there is not convincing empirical evidence about negative net effects of legal immigration on public finances (Preston, 2014), and that on the contrary some studies suggest that legal immigrants may be net contributors to the fiscal system in several countries (Dustmann *et al.* 2010, Dustmann and Frattini 2014).

Lastly, the empirical literature about public spending provides an important result for this analysis. That is, population ageing affects fiscal policies in two key ways. On one hand there are direct effects - largely exogenous to the political process - due to changes in the cost of pensions, health care and education (Banks and Emmerson, 2003). On the other hand there is evidence that indirect political effects play an important role in shaping spending policies (Persson and Tabellini, 1999; Galasso and Profeta, 2004). Accounting for these two aspects is crucial to understand how demographic shocks affect the tax rates.

These three empirical findings justify some of the modelling choices of this paper. In particular: (i) the choice of an overlapping generation model with a crucial role for the elderly in shaping the equilibrium policy, (ii) the main role played by the political determination of tax rates and public spending in shaping the attitudes of different individuals towards immigration, (iii) the explicit account for the "compositional amenities" in the preferences of native individuals, (iv) the inclusion of both exogenous and endogenous effects of ageing on the size of public spending.

The theoretical literature has analyzed the effects of population ageing on three political outcomes that are crucial for this paper, namely: (i) the immigration policies, (ii) the public spending policies, and how these two affect (iii) the tax policy (Razin and Sadka 1999, 2000, Razin *et al.* 2002). The use of unidimensional models to study this problem (that is largely prevalent in the literature) has constrained the analysis to a unique endogenous outcome variable. The implication is that fiscal and immigration policies have been studied separately. This resulted in two complementary streams of literature whose key trade-offs are going to be relevant in the model proposed in this paper.

The first analyzes the political effects of ageing on public spending and intergenerational redistribution. Persson and Tabellini (1999) show that in a simple overlapping generation model the extent of intergenerational redistribution towards the elderly is increasing in the share of elderly population, and Tabellini (1990), Lindert (1996) and Perotti (1996) provide a partial empirical support to this hypothesis. Razin *et al.* (2002) propose a second channel: a larger share of elderly implies a higher tax burden on the median voter, because it corresponds to a lower share of taxpayers relative to the share of net benefit receivers. These two channels imply opposite effects of ageing on the level of public spending in equilibrium: the pro-tax coalition becomes larger but each taxpayer is relative less supportive of public spending.

The second stream of literature analyzes the determinants of immigration policy. If on one hand some papers focus on immigration policies related to the *quality* of immigrants, such as skill requirements (Benhabib, 1996 and Ortega, 2005), on the other hand the prevalent approach - of which this paper is an example - analyses policies that restrict the *number* of immigrants such as immigration quotas (see Preston, 2014 for a survey). These papers (Kemnitz, 2003; Krieger, 2003; Ben-Gad, 2012) emphasize the importance of intergenerational aspects such as the pension system and the investment in education in explaining the determinants of the political choice about immigration policies.

A crucial finding in this literature is that the unidimensionality assumption has important consequences on the predictive power of these models. In particular, it generates inconsistent predictions about the comparative statics of the outcome variable depending on the specific restrictions that are imposed in order to satisfy the required condition. An example of these paradoxical effects is described in Facchini and Mayda (2008, 2009) and Haupt and Peters (1998). They study a simple economy characterized by a linear income tax and assume that revenues are lump-sum rebated to all citizen. In this setting one may choose to meet the requirement of unidimensionality by imposing the exogeneity of either (i) the level of public spending in benefits or of (ii) the income tax rate. These two assumptions corresponds respectively to the classes of "Tax adjustment models" (TAM, e.g. Scholten and Thum, 1996) and "Benefit adjustment models" (BAM, e.g Razin and Sadka, 1999, 2000) and imply opposite predictions about the relationship between pre-tax income, age and attitude towards immigration (Figure 1-2-3-4). Specifically, the first model implies that the elderly and the low income individuals are more hostile to immigration than the young and high income, while the opposite is true in the second model. The intuition that underpins these two apparently contradictory results lies in the consequence of an increase in the legal inflow of immigrants. Consider for instance the case in which immigrants are net contributors to the fiscal system. If publicly provided benefits are set exogenously, then the effect of an increase in immigration is a fall in the tax rate. Conversely, if the exogenous variable is the tax rate, then the effect is a rise in public spending per capita. As a result, in the former case immigration benefits mostly the young and high income voters, while in the latter the elderly and the low income individuals enjoy the largest share of the gains. In a recent paper Preston (2014) clarifies that the source of this inconsistency lies in how the social gains generated by immigration are divided up among different groups. This division is an output of the political process, but existing models treat it as an input. The issue is even more relevant for the purposes of this paper because I aim not only to understand the patterns of immigration policy, but more generally to address how a democratic society responds to population ageing in terms of immigration and fiscal policy, and the overall consequences on the public finances. These questions can be addressed only in a framework that allows immigration, spending and tax policy to be endogenously determined.

The theoretical literature has recognized the crucial importance of multidimensionality of the policy space in order to study the determinants of immigration policies, but all the existing studies are based on unidimensional models because of technical reasons. The early papers by Plott (1967), Tullock (1967) and Devis *et al.* (1972) have established rather restrictive conditions for the existence of a *Condorcet Winner* - a platform that is preferred to any alternative by a majority of voters - if the policy space is multidimensional. Grandmont (1978) has elegantly generalized these conditions with the concept of *Intermediate Relations*. The use of Grandmont's result in Political Economy applications is restricted to simple problems of redistribution (e.g. Borge and Rattsø, 2004) because of the extreme constraints that it imposes on preferences' heterogeneity. These requirements are way too restrictive for applications in which different subgroups of the voting population (such as the working age and the retired individuals in this paper) have sufficiently heterogeneous preferences over the set of available policies[1].

Alternatives to unidimensional voting models are popular in the literature, but they are not generally useful to answer questions about the comparative statics of the equilibrium policy outcomes because they do not deliver sharp analytical predictions about the policy response to a shock to the voters' distribution. This can be due either to a large multiplicity of equilibria, like in the Citizen-Candidate models (Besley and Coate, 1997) and in the Party Unanimity Nash Equilibrium (Roemer, 1999), or to the lack of sufficiently robust analytical comparative statics results, like in Probabilistic Voting models (Lindbeck et al. 1987, Banks et al. 2003). A more detailed analysis of the advantages and disadvantages of different theoretical framework in the study of comparative statics in models of electoral competition is provided in chapter 2. An attempt to model collective choices over immigration policies and welfare spending allowing for a multidimensional policy space is in Razin *et al.* (2011, 2014). They characterize the type of political coalitions that may prevail among skilled, unskilled and elderly voters in an overlapping generation models that shares several features with the one proposed in this paper. Neverthless, their approach is unsuitable to answer the questions of this paper, because of two reasons. First, they assume exogenous tax rates. Thus, the implications in terms of preferences for immigration are the same as the ones of Benefit adjustment models. Secondly, the assumptions they impose to tackle the multidimensionality of the policy space severely limit the possibility of deriving comparative statics results about the equilibrium policy outcome.

This paper is based on another stream of literature (Levy 2004, 2005) which exploits the role of coalitions and political parties in ensuring stability in a multidimensional deterministic voting model. I adopt a dynamic version of the model of electoral competition proposed by Dotti (2015). Such framework, under appropriate preferences restrictions, delivers sharp predictions about the equilibrium policy outcome, and it is therefore suitable to answer the questions of the paper.

[1] In the Appendix to chapter 2 I provide an example of why the Grandmont conditions usually fail to apply in this framework, and in particular to the model that I present in section 3 of this chapter.

4.1.4 Organization of the Paper

The paper is organized as follows. In the next section I introduce the main model and an equilibrium concept that allows me to answer the questions. Section 3 presents the main results of the paper, which are stated in Theorems 7-8. In section 4 I propose four extensions of the basic framework. In section 5 I analyze the welfare implications of the main predictions of the paper. Section 6 provides an analysis of the determinants of the attitude towards immigration in the UK based on the British Social Attitudes Survey and show that they are consistent with the one implied by the model proposed in this paper. Lastly, in section 7 I discuss some limitations of this work and future directions of research.

4.2 A Political Model of Immigration and Spending Policy

This section is constituted by two parts. In the first I describe the features of the political process. In the second I present the economic model of immigration and public spending and I formally define the notion of equilibrium. These two theoretical tools are then used to derive the main results of this paper, which are stated in section 3.

4.2.1 The Political Process

I define a political process that translates individual preferences into a policy outcome x_t in each period t. The elements of the vector x_t represent the relevant policy outcomes, namely the immigration quota (M_t) and the uniform provision of an imperfect public good (Y_t). I adopt a dynamic version of the political model of electoral competition introduced in a companion working paper (Dotti 2015). It is a general tool with a potentially large range of applicability, some of which are mentioned in the concluding section of chapter 2. The closest example in the literature is in Levy (2004, 2005). A formal definition of the equilibrium concept is provided in section 2.3 (Definition 1), while a detailed description of the political process and its properties in the static case is available in chapter 2.

The political process is based on the assumption that voters can form coalitions in order to enhance their capacity to influence the policy outcome. Each individual can be the member of only one coalition, thus a *coalition structure* is defined as a partition of the set of voters. As in Levy (2005), a coalition can only offer credible policies, that is, policies in the Pareto set of its members. Thus, when a voter runs as an individual candidate, he can only offer his ideal policy, as in the "citizen-candidate" model. On the other hand, when heterogeneous individuals join together in a coalition, their Pareto set is larger than the set of their ideal policies. This assumption captures the idea that within a coalition individuals can commit to policies that represents a compromise among the members, and that these internal agreements are credible for the voting population provided that not all the members have an incentive to renegotiate the terms of the deal. Individuals play a two stage game: in the first stage they form coalitions in support of a certain proposed policy platform (or no policy) and in the second stage a voting game is played over the set of policies that are proposed by at least one coalition in the previous stage. Coalitions are required to be stable in equilibrium, in the sense that each coalition must possess at least one policy vector in its Pareto set such that - if the policy is proposed - there is no subcoalition that have a strict incentive to deviate and propose a different platform (named a *deviator* in this case)[2]. If the deviation occurs the policy initially proposed by the coalition may become unfeasible. Therefore the profitability of a deviation depends on the behavior of the remaining part of the coalition that did not participate in the deviation. I assume that this subgroup responds to the deviation by proposing a policy (if any) that is capable of reducing the final payoff of some (or of all) the deviating players and therefore to prevent the deviation, and no policy if such platform does not exist. It can be shown that the main results of this section are robust to different assumptions about such behaviour (see chapter 2). Moreover, I assume that the profitability of a deviation is determined by the final outcome of the voting process[3]. Specifically, voters fully anticipate not only the effects of their strategies in the current period, but also the effects on the equilibrium in the following periods. The latter effects are derived assuming rational expectations that satisfy the Markov property. This means that expectations about future equilibrium outcomes depend uniquely on the state of the economy in the current period. Details are provided in section 2.3. I assume a tie-breaking rule for the case in which, given the other platforms that are offered in equilibrium, all members of a given coalition are indifferent between offering a platform and running at all. Specifically, I impose that in equilibrium a coalition facing such a situation does not propose any platform. The same restriction is assumed in Levy (2005) and it is justified if one considers some small costs of running for elections which are not explicitly assumed in the model. If there is at least one policy in the Pareto set of a certain coalition that does not face any deviator, then this policy is *feasible* and the coalition is *stable*. A *stable coalition structure* is a partition of the set of voters in which all coalitions that are part of such partition are stable in the sense described above. Before observing the coalition structure each coalition (including one-member

[2] One can also allow for mergers between coalitions with no effects on the results in Theorems 3-4-5.

[3] Alternatively one can assume that the equilibrium choices of other coalitions do not affect the behavior of potential deviators (in such case the stability is purely *internal* to the coalition), with no effects on the comparative statics results, see Ch. 2.

coalitions) proposes either a feasible policy platform or no policy. Then the coalition structure and the proposed platforms are observed by all the players. Voters (the whole population) vote for one of the available policy platforms and the election's outcome is a weak *Condorcet Winner*, which I name a *winning policy*. If no policy is offered or no weak *Condorcet Winner* exists, a default policy is implemented which is worse for all players than any other outcome[4]. A set of platforms (named a *policy profile*), a stable coalition structure and a winning policy given expectations about future policy outcomes constitute a *Markov-Perfect coalitional equilibrium* of the game if one of the coalition is a (weak) *Condorcet Winner* of the voting game at the second stage (see chapter 2 for a formal definition). Notice that, differently from Levy (2005), I do not assume sincere voting: the existence of a *Condorcet Winner* at the second stage of the voting game implies a result that is robust to a fully sophisticated voting behavior and to a number of different voting protocols.

The main difficulty in applying the concept of coalitional equilibrium to the analysis in this paper is related to the dynamic nature of the problem. Specifically, voters' expectations about the effects of current policy choices on future outcomes may affect the equilibrium behaviour. Moreover, because of this dynamic aspect, multiple equilibria are, in principle, possible. However, under appropriate restrictions on voters' expectations, the analysis in each period t becomes equivalent to the one of a static problem, as I am going to clarify in section 4.2.3. In the next section I present the economic model of immigration that I adopt in this paper, and in section 4.2.3 I will provide sufficient conditions on voters' expectations such model satisfy, in each period t, the condition for a Markov-Perfect coalitional equilibrium.

4.2.2 The Economic Environment

In this section I introduce an economic model of immigration and public spending in the spirit of the ones in the literature, in particular of Razin and Sadka (1999). Differently from the latter, I allow for the endogeneity of both the spending variable (an imperfect Public Good) and the immigration policy (in the form of a quota in each period t).

4.2.2.1 Demographic Structure

Consider an overlapping generation model with three generations in each period t: the children (ch), the working age population (y) and the elderly (o). In each period only

[4]The comparative statics results apply even if the default policy is the platform implemented in the previous period, see Appendix 4.8.2.4.

the native individuals of working age and the elderly (which include both the native and immigrants of the previous period) have voting rights (highlighted in capital letters in Fig. 4.5). Each period has length normalized to 1 and it is characterized by a native working age population of size n_t and a number of immigrants m_t in their working age. Natives and immigrants have potentially different exogenous expected fertility rates denoted by σ_t^n and σ_t^m respectively. An elderly individual at time t has life expectancy $l_{t-1} \leq 1$. At the end of each period immigrants and their children are fully assimilated to the native population in terms of costs and fertility behavior. The size of each part of the population is summarized in Fig. 4.6. Denote with o_t the size of the elderly population, i.e. $o_t = l_{t-1}(n_{t-1} + m_{t-1})$. Notice that o_t is an increasing function of longevity. This assumption captures in a simple way the implications of a more realistic continuous time model[5]. Thus, the total number of individuals that possess voting rights at time t is $N_t = n_t + o_t$. Also notice that the way in which I define the size of different groups in the population implies a number of voters that is not necessarily a natural number, while in reality that must be the case. Given that the object of this study are policies that are typically decided at country level, and that the effects of this approximation tend to disappear as the number of individuals grows large, these assumptions are reasonable and commonly used in the literature (e.g. Razin and Sadka, 1999).

4.2.2.2 Individual Preferences

An individual i of working age (y) at time t has preferences that are represented by a utility function whose arguments are consumption of private goods C_s and the imperfect Public Good Y_s, and the share of immigrants in the total population of working age M_s in the form:

$$\begin{aligned} U_t^{i,y}\left(C_t^{i,y}, C_{t+1}^{i,o}, M_t, M_{t+1}, Y_t, Y_{t+1}\right) = \quad & C_t^{i,y} + b(Y_t) - c(M_t) + \\ & + \beta l_t \left[C_{t+1}^{i,o} + d(Y_{t+1}) - c(M_{t+1})\right] \end{aligned} \tag{4.1}$$

where β is a parameter capturing how an individual discounts future utility. The function c captures the perceived effect of immigration on the composition of the community in which the voter lives (or *compositional amenities*, see Card *et al.*, 2012). It is a function of the ratio of immigrants to total working age population, because the effects of immigration on compositional amenities are likely to be larger in communities that face a relative large number of immigrants. The function c take positive values if immigration has a positive effect on the way in which the native population enjoys such compositional amenities,

[5] In a continuous time model the number of elderly in each moment in time t is given by $\int_{s=1}^{1+l} n_{t-s}(s) + m_{t-s}(s) ds$ which is also linearly increasing in the longevity l and in the size of the oldest generation of elderly $n_{t-1-l}(1+l) + m_{t-1-l}(1+l)$.

and negative values otherwise. The latter case is going to be the most interesting one for this analysis because it implies an equilibrium policy that restricts immigration even if immigrants are, on average, net contributors to the fiscal system. The functions b and d are restricted to take only weakly positive values. Moreover, b and d are strictly concave while c is strictly convex. For retired individuals $U_t^{i,o}$ is constructed in a similar way, except that it only includes consumption and share of immigrants in the current period of life:

$$U_t^{i,o}\left(C_t^{i,o}, M_t, Y_t\right) = l_{t-1}\left[C_t^{i,o} + d(Y_t) - c(M_t)\right] \tag{4.2}$$

One can allow for different effects of immigration on compositional amenities at different ages, with no changes in the results of this analysis. Lastly, the function c is assumed to be the same across different cohorts of voters (at given age). This assumption ensures that, if the features of the society - in terms of longevity, birth rates, etc., - are constant over time, then the choice of voters tends to converge to a stationary policy vector.

4.2.2.3 Production

Individual productivity is given by ε_t^i and has average $\bar{\varepsilon}_t$. The distribution of ε_t^i is perfectly observed by all agents and it does not change over time. I denote its continuous c.d.f. with Q, its p.d.f. with q and I assume $q(0) > 0$. Immigrants have the same expected productivity as the natives. Individuals are endowed with 1 unit of time and their labour supply is perfectly inelastic. I assume a linear production function $F_t(L_t) = \xi_t L_t$ in which the total supply of effective labour is given by $L_t = (m_t + n_t)\bar{\varepsilon}_t$. Perfect competition on the labour market implies a wage rate per unit of effective labour $w_t = \xi_t$. Therefore individual pre-tax income is given by:

$$y_t^i = w_t \varepsilon_t^i \tag{4.3}$$

and has average $\bar{y}_t$. The assumption of inelastic labour supply simplifies the results and it is not crucial for driving the pay-offs of the model (in the supplementary online material[6] I show that the results are identical if all individuals have the same tax elasticity of labour supply). The assumption of a linear production function rules out the effects of changes in the aggregate labour supply on wages and it is common in the literature (e.g. Razin and Sadka, 2000). It is justified if one considers that in a more complex economy these effects tend to be offset by the adjustment in the stock of capital of the economy - not explicitly assumed in this analysis - that occurs in the relatively long time framework of a generation. This adjustment is particularly strong if firms have access to international

[6]Available for download at http://valeriodotti.github.io/research.html.

capital markets (see Ben-Gad, 2012). In the additional online material I show that the main results of this paper are mostly unaffected in the case of a strictly concave production function.

4.2.2.4 Public Finances

The public sector raises revenues through a linear tax τ_t on labour income and spends them on the publicly provided good Y_t and on pensions for the elderly. In section 4.3 I introduce an extension of the model in which the government also provides public education. The government faces an exogenous amount of forgone tax revenues $\lambda_t = \lambda(w_t)$ per immigrant. This assumption captures the idea that certain skills may be country-specific and therefore the immigrants may earn less than native individuals with similar productivity levels. Alternatively one can assume that immigrants and natives have different average productivities $\bar{\varepsilon}_t^m$, $\bar{\varepsilon}_t^n$, and assume λ_t to be a function $\lambda(w_t, \bar{\varepsilon}_t^m, \bar{\varepsilon}_t^n)$ that captures the net forgone government revenue due to the difference in income[7].

I assume a Pay-As-You-Go pension system (in section 4.1 I present an extension in which I allow for a partially funded system). The state pension paid to an individual i at time t is denoted by p_{t-1}^i and has average $\bar{p}_{t-1}$. It is promised to a working age individual at time $t-1$ and it is predetermined at time t. It is a constant flow, such that the total transfer is $l_{t-1}p_{t-1}^i$ (the flow amount times the time the pension is going to be paid for). It is a function of the relative income of the pensioner in the previous period $y_{t-1}^i/\bar{y}_{t-1}$ and of the growth rate of working age population. At time $t-1$, when the promise is made, m_t is not yet determined, because it is a function of the immigration policy at time t. Thus the promised pension is a function of an exogenously fixed amount of immigrants $\hat{m}_t$ (which can be equal to zero). This assumption allows voters to ease the burden of pension on the working age population by choosing an immigration quota larger than $\hat{m}_t$. The assumptions on the pension system ensure that a certain positive amount of pensions is provided even if the pivotal voter typically prefers no pensions at all. Although not explicitly modeled in this paper, the assumption of an exogenous positive provision of public pensions in an overlapping generation model is justified in a game theoretical framework like the one in Rangel and Zeckhauser (2001). The state pension p_{t-1}^i is given by the formula:

$$p_{t-1}^i = \left(\alpha + \gamma \frac{y_{t-1}^i}{\bar{y}_{t-1}}\right) \frac{n_t + \hat{m}_t}{n_{t-1} + m_{t-1}} = \left(\alpha + \gamma \frac{y_{t-1}^i}{\bar{y}_{t-1}}\right) \frac{\bar{\sigma}_{t-1}}{(1 - \widehat{M}_t)} \tag{4.4}$$

where $\bar{\sigma}_{t-1} = \frac{n_t}{(m_{t-1}+n_{t-1})}$ is the natural growth factor of the working population between period $t-1$ and t and $\widehat{M}_t = \frac{\hat{m}_t}{n_t+\hat{m}_t}$ is the share of immigrants implied by the default level

[7]Notice that these two assumptions have consequences on the post-tax income of the immigrants.

of immigration $\widehat{m}_t$. The parameters $\alpha \geq 0$, $\gamma \geq 0$ determine the features of the benefits provided to the elderly and, as a consequence, the type of public pension system. Pay-As-You-Go pension systems are often classified into two categories. Specifically, a system is (i) Beveridgean if it provides flat-rate benefits, and (ii) Bismarkian if it provides earnings-related (or contribution-related) benefits. Formula (4.3) implies that the pension system can be either Beveridgean (if $\gamma = 0$), Bismarckian (if $\alpha = 0$) or a combination of the two. Notice that if native and immigrants have different birth rates, i.e. $\sigma_t^n \neq \sigma_t^m$, then the natural growth rate of the population $\bar{\sigma}_t$ is itself endogenous in the immigration policy, and in particular: $\bar{\sigma}_t = \frac{\sigma_t^n n_t + \sigma_t^m m_t}{n_t + m_t} = \sigma_t^m M_t + \sigma_t^n (1 - M_t)$. Lastly, notice that the total cost of the pension system per taxpayer is decreasing in the number of immigrant workers that are allowed to enter the country in period t, while p_t^i is increasing in M_{t-1} if $\sigma_{t-1}^m > \sigma_{t-1}^n$. I assume that the government budget is balanced in every period. The choice of not allowing for public debt simplifies the analysis and does not affect the trade-offs of the model. The government budget constraint ensure that the total public spending in public goods, pensions and the costs of immigration do not exceed the total tax revenue, and has form:

$$Y_t(m_t + n_t) + l_{t-1}\bar{p}_{t-1}(m_{t-1} + n_{t-1}) + \lambda_t m_t \leq \tau_t(m_t + n_t)\bar{y}_t \tag{4.5}$$

Assume that the governmental budget constraint is satisfied with equality (it must be true at any equilibrium of the voting game[8]). Using formula for the pensions (4.3) the governmental budget constraint (4.4) can be rewritten as follows:

$$\tau_t = \tau(M_t, Y_t, \bar{y}_t) = \bar{y}_t^{-1}\left(\lambda_t M_t + (\alpha + \gamma) l_{t-1} \frac{(1 - M_t)}{(1 - \widehat{M}_t)} + Y_t\right) \tag{4.6}$$

Notice that this formula implies that working age voters can ease the tax burden on their income by voting for a more open immigration policy. The intuitition is that, if the number of immigrants increases, then the expenditure in pensions is going to be shared among a larger number of taxpayers. This results in lower income taxes. I can use this formula to state the feasibility condition of the policy space:

$$0 \leq \tau_t(M_t, Y_t, \bar{y}_t) \leq k \tag{4.7}$$

for some $k < 1$. This restriction ensures that the implied tax rate on income will not exceed 1 or become negative. Notice that this restriction is crucial for the results in the next section to apply: if the tax rate hits the upper bound then the model and its predictions become similar to the ones of a standard Benefit Adjustment Model (See Appendix 4.8.2.5). It is easy to show that the consumption of private goods of a young

[8]In the case in which the pivotal voter is retired or has zero income one has to rule out Pareto inferior outcomes to ensure this result.

individual is given by her post-tax income such that:

$$C_t^{i,y} = (1-\tau_t)y_t^i \tag{4.8}$$

Lastly, the consumption of old people at time t depends only on the amount of pensions provided by the government, i.e.

$$C_t^{i,o} = p_{t-1}^i \tag{4.9}$$

Formulas (4.8) and (4.9) rule out the possibility of savings. I also abstract from bequest motives.

4.2.2.5 Policy Space

I assume that voters face a two-dimensional policy space in each period t. Namely, a policy platform consist of an immigration policy M_t, and of a level of public spending in the imperfect public good Y_t. Moreover, I assume that both the immigration policy M_t and the spending policy Y_t lie between zero and an upper bound, i.e. $0 \leq M_t \leq \overline{M}$ and $0 \leq Y_t \leq \overline{Y}$. A typical platform is given by a two dimensional vector $x_t = (x_{1t}, x_{2t})$ with $x_{1t} = M_t$ and $x_{2t} = -Y_t$.

4.2.2.6 Voters' Objective Function

Substituting the formulas for C_t^y (4.8) and C_{t+1}^o (4.9) into the utility function of a young voter given by formula (4.1), one gets the indirect utility function $v_t^{i,y} = v^y(M_t, Y_t, M_{t+1}, Y_{t+1}; y_t^i)$:

$$\begin{aligned} v_t^{i,y} = \quad & (1-\tau_t)y_t^i + b(Y_t) - c(M_t) + \\ & +\beta l_t \left[\left(\alpha + \gamma \frac{y_t^i}{\bar{y}_t}\right) \frac{\bar{\sigma}_t}{(1-\widehat{M}_{t+1})} + d(Y_{t+1}) - c(M_{t+1}) \right] \end{aligned} \tag{4.10}$$

The next step is to state the objective function of the elderly. Using the formula for $C_t^{i,o}$ (4.9) into the utility function of an elderly voter (4.2) I get $v_t^{i,o} = v^o(M_t, Y_t; y_{t-1}^i)$:

$$v_t^{i,o} = l_{t-1} \left[\left(\alpha + \gamma \frac{y_{t-1}^i}{\bar{y}_{t-1}}\right) \frac{\bar{\sigma}_{t-1}}{(1-\widehat{M}_t)} + d(Y_t) - \hat{c}(M_t) \right] \tag{4.11}$$

The formula (4.11) delivers the main intuition that underpins the results in this paper. Notice that retired individuals internalize (indirectly) the positive effects of immigration through the level of public spending in the imperfect Public Good. The key difference with traditional models is that the tax rate on income is also an endogenous variable. Thus, the elderly always prefer, given a certain level of public spending, a policy that finances it with high taxes on the income of native workers rather than with a larger number of immigrants. This result follows from the fact that in this model the elderly face the same costs of immigration as the young but, differently from the latter, they do not internalize the negative effects of high taxes on the working age population. Moreover, notice that the same preferences represented by $v_t^{i,o}$ are also represented by the function $v_t^o = d(Y_t) - c(M_t)$ for all the elderly at time t. This objective function implies that the attitude of the elderly towards immigration is always more hostile than the one of any working age individual. This is true even if immigrants are net contributors in financing the public spending of which the elderly are net beneficiaries. This implication of the model is consistent with the empirical findings outlined in section 1 and it is crucial in order to understand the comparative statics of the equilibrium outcomes of the model that I will present in the next sections of this paper. Define $\theta_t^{i,y}$ as the ratio of i's income to mean income at time t:

$$\theta_t^{i,y} = \frac{\varepsilon_t^i}{\bar{\varepsilon}_t} = \frac{y_t^i}{\bar{y}_t} \tag{4.12}$$

The preferences of each young native individual i are uniquely identified by the parameter $\theta_t^{i,y} \in \Theta_t^y$ with $\Theta_t^y = [\underline{\theta}_t^y, \bar{\theta}_t^y]$. Notice that the function $v_t^{i,y}$ can be written as a function of one exogenous parameter $\theta_t^{i,y}$ and of the choice variables $(M_t, Y_t, M_{t+1}, Y_{t+1})$ at time t and $t+1$, plus the parameters $\varphi_t = (\{\alpha, \beta, \gamma, \sigma_{t+s-1}^n, \sigma_{t+s-1}^m, l_{t+s-1}\}_{s=0}^{\infty})$. Moreover, the definition of $\theta_t^{i,y}$ implies that the cumulative distribution of $\theta_t^{i,y}$ is the same as the one of ε_t^i. The value of $y_{t-1}^i / \bar{y}_{t-1}$ does not affect the preferences of an elderly individual j over x_t, therefore all the elderly have the same preferences. This means that we can set a unique parameter $\theta_t^{j,o} = \theta_t^o \in \Theta_t^o$ which identifies the preferences of each elderly individual j at time t such that $\Theta_t^o = \{\theta_t^o\}$. I assign to all the elderly a parameter $\theta_t^o = -1$. I can now define the parameter set:

$$\Theta_t = \left\{\Theta_t^y \cup \Theta_t^o\right\} \tag{4.13}$$

which is a totally ordered set. In order to show that the preferences described in this section satisfy the conditions for the existence of a coalitional equilibrium I define a new objective function that includes both $v_t^{i,y}$ and $v_t^{i,o}$and has the following form:

$$v_t^i = v(x_t, x_{t+1}; \theta_t^i, \varphi_t) = \begin{cases} v_t^{i,y} & if \quad age = y \\ \kappa v_t^o & if \quad age = o \end{cases} \tag{4.14}$$

with $x_{1t} = M_t$ and $x_{2t} = -Y_t$ and for large enough $\kappa > 0$. Notice that multiplying by κ represents a strictly increasing transformation of the original objective function of the elderly therefore κv_t^o implies the same preferences as v_t^o. The role of κ is the one of translating an ordinal property of voter preferences into a cardinal property of the function v_t^i. Specifically, consider two policy vectors x_t' and x_t'' such that $x_t' \geq x_t''$, i.e. policy x_t' implies a weakly larger share of immigrants and a weakly larger provision of the public good relative to x_t''. Then a large enough κ ensures that, if an elderly voter is more averse to immigration and more favourable to public spending than a young voter, then the difference in utility between x_t' and x_t'' (positive or negative) should be larger for the former relative to the latter type of voter. This transformation will prove useful to show that the voters' objective function satisfies *Strictly Increasing Differences*, as I state in Lemma 2.

4.2.3 Markov-Perfect Coalitional Equilibrium

The equilibrium concept is a dynamic version of the coalitional equilibrium in Dotti (2015), that is described in chapter 2. I assume rational expectations on and off equilibrium. This implies that, given the history up to the the current period, the expectations are the same for all the voters. I also assume that such expectations only depend on the state of the economy in that period. Notice that, under this assumption, the state of the economy at the beginning of period $t+1$ is fully summarized by the ratio of elderly to young natives $g_{t+1} = o_{t+1}/n_{t+1}$, which is therefore the unique endogenous state in the dynamic process. Denote with $h_t = \{x_s, g_s\}_{s=0}^t$ the full history of policy choices and states observed by all agents up to time t, with $h_t \in H_t$. Denote with $x_{t+s}^{**}(g_t, \varphi_t, h_{t-1})$ the expectation at time t about the equilibrium policy at time $t+s$ given the state of the economy at time t and the history up to time $t-1$[9], and with $x_t^*(g_t, \varphi_t, h_{t-1})$ the policy actually implemented at time t. I assume that:

$$x_{t+s}^{**}(g_t, \varphi_t, h_{t-1}) = x_{t+s}^{**}(g_t, \varphi_t, h_{t-1}') \tag{4.15}$$

for all histories $h_{t-1}, h_{t-1}' \in H_{t-1}$ and all $s \geq 0$. Moreover, the assumption of rational expectations implies that $x_{t+s}^{**}(g_{t+s}, \varphi_{t+s}, h_{t-1+s}) = x_{t+s}^{*}(g_{t+s}, \varphi_{t+s}, h_{t-1+s})$ for all $s \geq 0$. These two assumptions imply that the dynamic system satisfies the Markov property. That is, there is no equilibrium in which different histories correspond to different equilibrium choices given an identical economic environment. I also assume that $x_{t+s}^{**}(g_{t+s}, \varphi_{t+s}, h_{t-1+s})$ is twice differentiable with respect to g_{t+s}. This condition will prove to be satisfied in any Markov-Perfect coalitional equilibrium under appropriate restrictions (see Lemma 6). Notice that expectations are assumed to be a function of the state of the economy at time t, which is fully summarized by the state of the economy g_t. Also notice that g_{t+1} is perfectly known at the end of time t because there is no

[9]The value of x_t^{**} is also a function of the distribution of productivity Q, but this is omitted in the formula.

uncertainty about the distribution of future productivity[10]. Given these assumption, I define a Markov-Perfect coalitional equilibrium as follows.

Definition 1. *A Markov-Perfect coalitional equilibrium at time t is:*

(i) a partition $\mathbb{P}_t$ of the set of voters at time t,

(ii) a policy profile A_t,

(iii) a winning policy x_t^*, and

(iv) a set of expectations about future policies $\{x_{t+s}^{**}(g_t, \varphi_t, h_{t-1})\}_{s=0}^{\infty}$,

such that:

(a) $(\mathbb{P}_t, A_t, x_t^*)$ is a coalitional equilibrium of the voting game given state g_t and given expectations about current and future policies $\{x_{t+s}^{**}(g_t, \varphi_t, h_{t-1})\}_{s=0}^{\infty}$,

(b) expectations are rational, i.e. $x_{t+s}^{**}(g_{t+s}, \varphi_{t+s}, h_{t-1+s}) = x_{t+s}^{*}(g_{t+s}, \varphi_{t+s}, h_{t-1+s})$ for all i and for all $s \geq 0$, *and*

(c) satisfy the Markov Property, i.e. $x_{t+s}^{**}(g_t, \varphi_t, h_{t-1}) = x_{t+s}^{**}(g_t, \varphi_t, h'_{t-1})$, for all $s \geq 0$, all i and all $\varphi_t \in \Phi_t$.

For ease of notation, I am going to denote a Markov-Perfect coalitional equilibrium with $(\mathbb{P}_t, A_t, x_t^* \{x_{t+s}^{**}\}_{s=0}^{\infty}; g_t)$, in which I have suppressed the arguments of x_t^* and of each x_{t+s}^{**}. Using this notion of equilibrium, I can state the following Lemma:

Lemma 1. *In a Markov-Perfect coalitional equilibrium - if it exists - (i) each individual's ideal policy x_t^i and (ii) the equilibrium policy x_t^* at time t are invariant - conditional on g_t - to the history up to time $t-1$, i.e. $x_t^i(g_t, \varphi_t, h_{t-1}) = x_t^i(g_t, \varphi, h'_{t-1})$ and $x_t^*(g_t, \varphi_t, h_{t-1}) = x_t^*(g_t, \varphi_t, h'_{t-1})$ $\forall t$ and $\forall h_{t-1}, h'_{t-1} \in H_{t-1}$.*

Proof. See Appendix 4.8.1.1.1.

These results imply that the history up to time $t-1$ is irrelevant for all aspects of the model conditional on g_t. Thus, from now on I am going to suppress the argument h_{t-1+s} from the formulas of x_{t+s}^*, x_{t+s}^{**} and x_{t+s}^i. Summarizing, g_t is the unique endogenous state variable of this dynamic system and a coalitional equilibrium in this model (if it exists) is a temporary equilibrium that depends only on the value of the state variable g_t at time t and is independent of the previous history conditional on g_t. Notice that g_t is the ratio of elderly relative to native individuals of working age, and therefore it represents the crucial variable in order to determine the identity of the pivotal voter. Lemma 1 allows one to

[10] The result is the same if one allows for uncertainty and the size of the population is very large, because the law of large numbers implies that the identity of the median voter in the next period is known with probability equal to 1.

disregard the effects of current policy choices (other than the effects on g_{t+1}) on future equilibrium outcomes when calculating the optimality conditions for each voter. This implies that future equilibrium policy outcomes affect the individual objective functions at time t only through their effects on g_{t+1}. The consequence is that, given expectations $x^{**}_{t+1}(g_t, \varphi_t)$, I can write a working age voter's objective function $V^i_t = V(x_t; \theta^i_t, \varphi_t, g_t)$ as follows:

$$V^i_t = V(x_t; \theta^i_t, \varphi_t, g_t) = v(x_t, x^{**}_{t+1}(g_t, \varphi_t); \theta^i_t, \varphi_t) \tag{4.16}$$

where $x^{**}_{t+1}(g_t, \varphi_t)$ represents the expected equilibrium policies at time $t+1$ which are a function solely of g_t and φ_t. Similarly, one can define the corresponding objective functions of young and old voters, $V^{i,y}_t = V^y(M_t, Y_t; \theta^i_t, \varphi_t, g_t)$ and $V^{i,o}_t = V^o(M_t, Y_t; \theta^i_t, \varphi_t, g_t)$ respectively. Notice that these two objective function implies that an interior solution for the optimal policy of individual i with a partially open immigration policy $x_{2t} = M_t > 0$ may exist even if immigrants "contribute less than what they take out" in the current period, or more precisely if - at a given policy $x_t = (M_t, -Y_t)$ - a marginal increase in the number of immigrants at constant Y_t implies, *ceteris paribus*, a rise in the income tax rate τ_t. This is true because if immigrants have higher fertility rates in comparison with the natives ($\sigma^m_t > \sigma^n_t$), then a native individual of working age will have a future benefit from immigration. Specifically, higher immigration today implies a lower dependency ratio tomorrow and, as a consequence, a more generous state pension system. This implies that this model is not affected by the dichotomy between "skilled migration" and "unskilled migration" in the patterns of attitude towards immigration and income that is typical of traditional models such as Facchini and Mayda (2008). In the model proposed in this paper the attitude towards immigration may improve with income even if the immigrants are a net burden for the society in the short run, because preferences accounts for the future positive effect of immigration. Moreover, these future benefits are increasing with income if the Bismarckian component of the pension system is positive ($\gamma > 0$). Using the previously defined V^i_t function I can state the following result:

Lemma 2. *The function $V(x_t; \theta^i_t, \varphi_t, g_t)$ satisfies SM and SID in $(x_t; \theta^i_t)$ for all $\theta^i_t \in \Theta_t$ and all $\varphi_t \in \Phi_t$ for any given state g_t.*

Proof. See Appendix 4.8.1.1.2.

This Lemma is crucial in order to establish existence of a Markov-Perfect coalitional equilibrium, and therefore to derive all the results in the next section of this paper. The intuition about how this result can be proved relies on the effect of the Markov assumption. Recall that the expectations about the policy outcome in any future period

$t+s$ are assumed to depend uniquely on the state of the economy at the beginning of such period (g_{t+s}), and that $g_{t+s} = \frac{o_{t+s}}{n_{t+s}} = \frac{l_{t+s}}{M_{t+s}(\sigma^m_{t+s} - \sigma^n_{t+s}) + \sigma^n_{t+s}}$. Thus, the only choice at time t that can affect the value of g_{t+1} - and therefore all future expectations - is the one about the immigration policy M_t. As a consequence, conditional on M_t and given parameters φ_t, the expectations about the policy implemented in the future periods are unaffected by changes in Y_t or θ^i_t. This makes the cross-partial derivatives of V with respect to each policy dimension $x_{k,t}, x_{j,t}$, $k \neq j$ and with respect to $x_{k,t}, \theta^i_t$ for all k realtively easy to calculate. Thus, the sufficient conditions for *SM* and *SID* - that are based on the sign of such cross derivatives - can be shown to hold.

4.2.3.1 Conditions for a Markov-Perfect Coalitional Equilibrium

Following the static analysis in Dotti (2015), denote with $\wedge$ and $\vee$ the meet and joint operators over a lattice (see chapter 2). Recall that (i) $\theta^i_t \in \Theta_t$ is the parameter that identifies the preference of a voter i , that (ii) the parameter space Θ_t is a totally ordered set, and that (iii) $\varphi_t \in \Phi_t$ is a vector of parameters that do not differ across voters. I state the conditions for a Markov-Perfect coalitional equilibrium to exist and satisfy some desirable properties.

1. The policy space X_t must be a subset of the the d-dimensional real space R^d with typical element x_t, such that the partially ordered set $(X_t, \leq)$ is a convex and complete sublattice of R^d.
2. Each individual i must be endowed with a reflexive, complete and transitive preference ordering $\succeq^i$ represented by an objective function $V : X_t \times \Theta_t \times \Phi_t \to R$ that is jointly continuous in x_t and $\theta_t - concave$ [11].
3. Individual preferences are such that the function V satisfies, given the state g_t:

   a) *Supermodularity (SM)* in x_t: $V(x'_t \vee x''_t; \theta_t, \varphi_t, g_t) - V(x'_t; \theta_t, \varphi_t, g_t) \geq V(x''_t; \theta_t, \varphi_t, g_t) - V(x'_t \wedge x''_t; \theta_t, \varphi_t, g_t)$ for all $\theta_t \in \Theta_t$, for all $\varphi_t \in \Phi_t$ and for all $x'_t, x''_t \in X_t$.

   b) *Strictly Increasing Differences (SID)* in (x_t, θ_t): $V(x'_t; \overline{\theta}_t, \varphi_t, g_t) - V(x''_t; \overline{\theta}_t, \varphi_t, g_t) > V(x'_t; \underline{\theta}_t, \varphi_t, g_t) - V(x''_t; \underline{\theta}_t, \varphi_t, g_t)$ for all $x'_t, x''_t \in X_t$ such that $x'_t \geq x''_t$ and $x'_t \neq x''_t$, for all $\varphi_t \in \Phi_t$ and for all $\overline{\theta}_t, \underline{\theta}_t \in \Theta_t$ such that $\overline{\theta}_t > \underline{\theta}_t$.

[11] For any function f defined on the convex subset X_t of R^d , we say that f is concave in direction $v \neq 0$ if, for all x, the map from the scalar s to $f(x+sv)$ is concave. (The domain of this map is taken to be the largest interval such that $x+sv$ lies in X_t.) We say that f is $i-concave$ if it is concave in direction v for any $v > 0$ with $v_i = 0$. See Quah (2007).

Regarding condition 1, I assume that X_t is the same in all periods. The assumptions on the policy space stated in section 2.2.5 ensure that this condition is satisfied. Condition 2 simply requires that the objective function satisfies some basic properties. Lastly, condition 3 is equivalent to stating that all voters can be ordered along a single preference dimension over a multidimensional choice set. These assumptions on individual preferences are common in many fields of Economic Theory. Notice that condition 3 is stated in a very general form, but in the case of a twice differentiable objective function one can simply adopt the sufficient conditions in Milgrom and Shannon (1994) in order to verify that the function satisfies *SM* and *SID*. Namely, one needs to check that the following conditions hold. (i) $\frac{\partial^2 V}{\partial x_{i,t} \partial x_{j,t}} \geq 0 \; \forall x_t \in X_t, \; \forall i \neq j$, and (ii) $\frac{\partial^2 V}{\partial x_{i,t} \partial \theta_t} > 0 \; \forall x_t \in X_t, \; \forall \theta_t \in \Theta_t, \; \forall i$. These sufficient conditions are usually easier to verify in comparison with the one implied by the definitions of *SM* and *SID*. Because of that, in the next sections I am going to make frequent use of these sufficient conditions.

4.2.3.2 Monotone Comparative Statics

Denote the set of ideal policies of voter i in period t given state g_t (and for a given expectations $\{x^{**}_{t+s}(g_t, \varphi_t)\}_{s=0}^{\infty}$) with $I_t(i_t) \equiv \{x_t | x_t \in \arg\max_{y \in X_t} V(y; \theta^i_t, \varphi_t, g_t)\}$,[12], and define the set of *equilibrium policies* as the union of all the policies that are winning policies in some coalitional equilibrium of the game for given expectations $\{x^{**}_{t+s}(g_t, \varphi_t)\}_{s=0}^{\infty}$. Notice that because Θ_t is a totally ordered set, one can identify a median element θ^v_t. The individual characterized by this value of the parameter is the *median voter* denoted by the index v_t[13]. In this setting, conditional on g_t and given expectations, the political process at time t is identical to the one of the static model described in chapter 2. Thus, the results stated in Dotti (2015) hold in the framework proposed here with minor modifications. Specifically, if the three conditions stated in the previous section are satisfied, then the following theorems hold for any value of the state g_t.

Theorem 3. *(Median Voter Theorem). If conditions 1-2-3 are satisfied, then (i) A Markov-Perfect coalitional equilibrium of the voting game exists; (ii) in any equilibrium the set of winning policies is a subset of the set of ideal points of the median voter v_t; (iii) if the median voter has a unique ideal policy, then the set of equilibrium policies is a singleton.*

Proof. See Appendix 4.8.1.1.3.

[12]Notice that the completeness of X_t implies compactness in the order-interval topology. On bounded sets in R^d, the order-interval topology coincides with the Eucilidean topology (Birkhoff 1967). Hence $I_t(i_t) \neq \varnothing$ for all i.

[13]In the case of a discrete even number of voters I assume that the ties are broken in favor of the individual with the lower index. Different assumptions would not affect the results in the next paragraphs.

Theorem 4. *(Monotone Comparative Statics). If conditions 1-2-3 are satisfied, then the set of equilibrium policies of the voting game is (i) a sublattice of X_t which is (ii) monotonic nondecreasing in θ_t^v.*

Proof. See Appendix 4.8.1.1.3.

Lastly, consider a totally ordered subset $\Phi' \subseteq \Phi$ and suppose that the objective function $V(x_t, \theta_t, \varphi_t)$ satisfies *Increasing Differences (ID)* in (x_t, φ_t), namely $V(x_t', \theta_t, \overline{\varphi}_t) - V(x_t'', \theta_t, \overline{\varphi}_t) \geq V(x_t', \theta_t, \underline{\varphi_t}) - V(x_t'', \theta_t, \underline{\varphi_t})$ for all $x_t' \geq x_t''$, and for all $\overline{\varphi}_t, \underline{\varphi_t} \in \Phi'$ such that $\overline{\varphi}_t \geq \underline{\varphi_t}$. Then I can state the following result:

Theorem 5. *(Monotone Comparative Statics 2). If conditions 1-2-3 are satisfied, then the set of equilibrium policies of the voting game is monotonic nondecreasing in φ_t.*

Proof. See Appendix 4.8.1.1.3.

The interpretation of this generalized Median Voter Theorem is identical to the one provided for the static case in chapter 2. Notice that, while $I_t(v_t)$ depends on voters' expectations, the identity of the median voter v_t is independent of expectations, thus the median voter is the same in all coalitional equilibria of the voting game. The results in this sections provide a tool to analyze the effects of a shock on the distribution of voters or on a preference parameter on the policy outcome that emerges in a political equilibrium. One only has to verify that an economic model satisfies the conditions stated in this section and then use Theorems 4-5 to formulate the predictions about the comparative statics of the platform that is implemented in equilibrium. Following this approach, I derive the main results of this paper, which are stated in the next section.

4.3 Results

In this section I present the main results of the paper, namely the existence and characterization of the voting equilibrium, the analytical comparative statics results, the dynamics of the equilibrium outcome and the simulation of the other long-run implications of the model. Notice that all the results described in this section - except for the cases in which the opposite is explicitly stated - are also valid for the extended version of the model with endogenous public education presented in section 4.3. The proofs in Appendix A include both the basic model and the extended one (the objects that refer to the extended model are denoted with a tilda in the proofs).

4.3.1 Equilibrium Existence and Characterization

Using the results in the previous sections we get that (i) the Policy Space $(X_t, \leq)$ is a convex and complete sublattice of R^2; (ii) the parameter set Θ_t is a totally ordered set; (iii) the objective function $V(x_t;\theta_t^i,\varphi_t,g_t)$ satisfies *Supermodularity* in x_t and the *Strictly Increasing Differences* in $(x_t;\theta_t^i)$. Therefore all the conditions for the existence of a Markov-Perfect coalitional equilibrium are satisfied provided that $V(x_t;\theta_t^i,\varphi_t,g_t) = v(x_t,x_{t+1}^{**}(g_t,\varphi_t);\theta_t^i,\varphi_t)$ is such that $x_{t+1}^{**}(g_t,\varphi_t)$ are rational expectations and V is concave in x. Moreover, if the objective function of each working age individual is strictly concave, then - given expectations - the ideal policy of the median voter is unique. Notice that, because the indirect utility function v is continuous, twice differentiable and strictly concave in x_t in each period t, and because of the assumptions on Q and on expectations previously stated, there exists a threshold on $\hat{\sigma}$ such that if $|\sigma_t^m - \sigma_t^n| \leq \hat{\sigma}$, then V is also continuous and strictly concave in x_t. Thus, I can state the following result.

Lemma 6. *If $|\sigma_{t+s}^m - \sigma_{t+s}^n| \leq \hat{\sigma}$ for some $\hat{\sigma} > 0$ and all $s \geq 0$, then (i) a Markov-Perfect coalitional equilibrium for the voting game exists. Moreover, (ii) in any Markov-Perfect coalitional equilibrium at time t the equilibrium policy is the unique ideal point of the median voter $x_t^v = x_t^* \in I_t(v_t)$. (iii) The parameter θ_t^v that identifies the median voter is weakly decreasing in g_t. If $\sigma_t^m - \sigma_t^n$ is arbitrarily small, then (iv) there is a unique equilibrium policy that is chosen in any Markov-Perfect Coalitional Equilibrium in period t.*

Proof. See Appendix 4.8.1.1.4.

Notice that the condition on $\sigma_{t+s}^m - \sigma_{t+s}^n$ is sufficient but not necessary for results (i), (ii) and (iii) in Lemma 6. In the rest of the paper, I am going to assume that v and φ_t are such that continuity and strict concavity are satisfied for any x_{t+1}^{**} that implies rational expectations[14]. Notice that Lemma 6 does not postulate the *uniqueness* of the Markov-Perfect coalitional equilibrium in points (i), (ii), (iii). The reason is that, even if the equilibrium is unique conditional on g_t and on expectations $\{x_{t+s}^{**}(g_t,\varphi_t)\}_{s=0}^{\infty}$, there may be different rational expectations $\{x_{t+s}^{**}(g_t,\varphi_t)\}_{s=0}^{\infty}$ that may support different policies in equilibrium. Neverthless, the comparative statics results in the next sections are valid in any equilibrium, thus the analysis is not affected by arbitrary equilibrium selection rules[15]. Having established existence of an equilibrium and (conditional)

[14] Notice that the requirement of joint continuity and concavity in x_t are necessary for a coalitional equilibrium, but not for the Citizen-Candidate version of the equilibrium in which individuals only run as single candidates. Thus, one does not have to impose these two restrictions if such simpler model of electoral competition is adopted.

[15] The Markov assumption implies that each function $x_{t+s}^{**}(g_t,\varphi_t)$ is uniquely affected by g_t and φ_t, thus

uniqueness of the policy outcome, I can use the result of the Monotone Comparative Statics of the equilibrium outcome in order to study the effects of shocks on the voters' distribution on the equilibrium policy outcome.

4.3.2 Main Result: Comparative Statics

In this section I analyse the short-run effects of shocks on the parameters that are related to population ageing on the equilibrium policy outcome. That is, how the equilibrium policy vector changes as a consequence of a shock - in the period in which the shock is observed - relative to the equilibrium level in absence of any shock. One has to account for four aspects: (i) how the direct preferences over policies of each native individual of working age are affected by the shock ("preference effect"), (ii) how the indirect preferences change because of the effects of the shocks on the governmental budget constraint ("budget effect") and (iii) how the identity of the pivotal voter changes as a consequence of the changes in the demographic composition of the population induced by the shock ("political effect"). Lastly, one has to account for the ability of a fully rational agent to anticipate that if $\sigma_t^m \neq \sigma_t^n$, then the choice of the immigration policy at time t affects the demographic structure of the voting population in the following periods and can therefore change the political equilibrium in the future. One may think that voters are unlikely to really anticipate this (iv) "sophisticated effect", therefore whenever this aspect is relevant in this section I will distinguish between the predictions that emerge with "naive" agents - i.e. if voters expectations do not account for future political effects of current policies - and the ones implied by fully "sophisticated" agents. The approach used is the following. First I verify if there is any effects of type (i), (ii) and (iv). In detail, if V_t^i satisfies the condition of Theorem 5 for a given value of g_t, then the theorem can be used to establish the sign of these effects. Then I study the effects of type (iii). If g_t is affected by the shock, then Lemma 6 (iii) implies a change in the parameter that identifies the pivotal voter and therefore Theorem 4 can be used to formulate the predictions. The results about the tax rate τ_t stated in this section refer to the case in which immigrants provide, on average, a contribution to public finances sufficient to ensure that τ_t is weakly decreasing in M_t. This is true whenever the average cost per pensioner is sufficiently large, namely if $l_{t-1}\bar{p}_{t-1} \geq \lambda_t$. The results about M_t and Y_t are valid even if the latter condition does not hold.

the comparative statics results are valid in *any* equilibrium of the game, provided that no changes in the functions $x_{t+s}^{**}(g_t, \varphi_t)$ occur.

4.3.2.1 Unanticipated Rise in the Longevity of the Retired Population

I analyse the effects of a marginal increase in l_{t-1} keeping other parameters constant. That is, the longevity of the current elderly increases, keeping the longevity of other generations and birth rates unchanged. Recall that $x_t = (M_t, -Y_t)$. For effects of type (i)-(ii)-(iv) one can verify that V_t^i satisfies *Increasing Differences (ID)* by studying the cross derivatives of $V_t^{i,y}$ with respect to each policy dimension and l_{t-1}. Denote with $V_{M_t}^{i,y}$ ($V_{Y_t}^{i,y}$) the partial derivative of $V_t^{i,y}$ with respect to the policy dimension M_t (Y_t) and with $V_{M_t l_{t-1}}^{i,y}$ ($V_{Y_t l_{t-1}}^{i,y}$) the cross derivative of $V_t^{i,y}$ with respect to M_t (Y_t) and a parameter l_{t-1}. In this case one gets:

$$V_{M_t l_{t-1}}^{i,y} = \frac{\theta_t^i(\alpha+\gamma)}{(1-\widehat{M}_t)} \geq 0 \tag{4.17}$$

$$V_{Y_t l_{t-1}}^{i,y} = 0 \tag{4.18}$$

Consider a vector of parameters $\tilde{\varphi}_t \in \Phi_t$. Define a subset $\Phi_{j,t} \subseteq \Phi_t$ as follows: $\Phi_{j,t} := \{\varphi_t \in \Phi_t | \varphi_{t,k} = \tilde{\varphi}_{t,k} \forall k \neq j\}$, where j is the position of the longevity of the elderly l_{t-1} at time t in vector φ_t. Notice that Φ_j is a totally ordered set. Moreover, the signs of the cross derivatives imply that $V(x_t; \theta_t^v, \varphi_t, g_t)$ satisfies *SM* and *ID* in $(x_t; \varphi_t)$, it also satisfies *SM* and *SID* in $(z_t; \varphi_t)$ where $z_t = (x_{1t}, -x_{2t})$. The conditions of Theorem 5 are satisfied, therefore at constant g_t the effect is a weak rise in M_t and no changes in Y_t. Moreover, $V_{M_t}^{i,y}$ is solely affected through the budget constraint hence the effect is of type (ii) ("budget effect"). For effects of type (iii) notice that $g_t = \frac{l_{t-1}}{\bar{\sigma}_{t-1}}$ is increasing in l_{t-1}. Lemma 6 (iii) implies that θ_t^v is decreasing in g_t. Hence Theorem 4 implies a weak increase in the public spending variable Y_t and a weakly more restrictive immigration policy M_t. The total effect of an increase in l_{t-1} is therefore weakly positive on the public spending variables Y_t and ambiguous on the immigration policy M_t. There are cases in which one effects dominates and therefore the comparative statics result for the immigration policy is also sharp. In particular, I can state the following results:

Theorem 7. *The effect of an increase in the life expectancy l_{t-1} is weakly positive on the spending policy and ambiguous on the immigration policy. Moreover, there exists a threshold $\hat{g}_t \in [0,1]$ such that if $g_t \geq \hat{g}_t$ then the effect on immigration policy is unambiguously (weakly) negative and the effect on the tax rate is strictly positive.*

Proof. See Appendix 4.8.1.2.1.

In order to get an intuition of what drives this results, consider the following cases. If $g_t = 1$ (i.e. there are as many working age individuals as elderly), then the pivotal voter has $\theta_t^v = 0$, which implies that $V^{v,y}_{M_t l_{t-1}} = 0$. Thus, there is no "budget effect" and the "political effect" weakly dominates. On the other hand, consider the case in which the variance of the income distribution is arbitrarily close to zero (e.g. $y_t^i = y_t$ for all i). In this case, as long as $g_t \neq 1$, the θ_t^v of the pivotal voter is unaffected by changes in the share of elderly, which implies that the "political effect" is zero and that the "budget effect" weakly dominates. Theorem 7 is a consequence of the negative relationship between age and attitude towards immigration and the positive one between age and attitude towards public spending implied by the the model. This result suggests the existence of a link between the size of the two effects and two characteristics on the voting population: the share of elderly and the degree of income inequality. Moreover, it implies that the sign of the effect of an increase in longevity on the equilibrium level of the immigration quota is the one implied by the Tax Adjustment Model if the share of elderly is large enough and there is sufficient income inequality, and the one implied by the Benefit Adjustment Model in societies characterized by opposite features (see section 1.3).

4.3.2.2 Unanticipated Fall in the Natural Growth Rate of the Woking Age Population

The natural growth rate of the native population is $\frac{n_t}{n_{t-1}+m_{t-1}} - 1 = \bar{\sigma}_{t-1} - 1$. In this model the effect of an unanticipated fall in such rate has same sign as the one of a decrease in the lagged birth rate of the natives σ^n_{t-1}. This is true because one can show that $\bar{\sigma}_{t-1} = \sigma^m_{t-1} M_{t-1} + \sigma^n_{t-1}(1 - M_{t-1})$, which implies that $\bar{\sigma}_{t-1}$ is predetermined at time t. This kind of shock corresponds for instance to the case in which the birth rate actually experienced during the period $t-1$ is smaller than the one expected at the beginning of that period. Therefore I analyse the effects of a shock on σ^n_{t-1}. I can state the following:

Theorem 8. *The effect of a decrease in the growth rate of the working age population is a weak decrease in the openness of the immigration policy and a weak increase in spending in the imperfect Public Good and in the tax rate.*

Proof. See Appendix 4.8.1.2.2.

Notice that conditional on g_t the shock has no effect on the equilibrium policy outcome (i.e. there are no effects of type (i), (ii), (iv)). The reason is that the pension system adjusts its size to changes in the birth rate for the reasons described in section 2. Nevertheless a fall in σ^n_{t-1} implies a rise in g_t, which corresponds to a "political effect". Using Theorem 4 one gets the result stated above.

4.3.2.3 Rise in Life Expectancy of the Working Age Population

I analyse the effects of a shock on the life expectancy of the current working age population l_t, keeping all the other elements of vector φ_t unchanged. First of all notice that g_t is unaffected by changes in l_t, which means that there is no "political effect". The results of this paragraph are summarized in Theorem 9.

Theorem 9. *The effect of an increase in the life expectancy l_t is ambiguous on the immigration policy. If voters are "naive" then the effect is weakly positive. If the birth rate of the native is the same as the one of the immigrants, then there is no effect.*

Proof. See Appendix 4.8.1.2.3.

In order to understand this result it is useful to analyze the cross derivative of $V_t^{i,y}$ with respect to the immigration policy M_t and the parameter l_t.

$$V^{i,y}_{M_t l_t} = \underbrace{\frac{\beta(\alpha+\gamma\theta^i_t)}{(1-\widehat{M}_{t+1})}(\sigma^m_t-\sigma^n_t)}_{preferences \quad effect} + \underbrace{-\frac{d}{dl_t}\left\{\frac{\beta l_t^2}{\bar{\sigma}_t{}^2}\left[d'(Y^{**}_{t+1})\frac{dY^{**}_{t+1}}{d\theta^v_{t+1}} - \hat{c}'(M^{**}_{t+1})\frac{dM^{**}_{t+1}}{d\theta^v_{t+1}}\right]\frac{d\theta^v_{t+1}}{dg_{t+1}}(\sigma^m_t-\sigma^n_t)\right\}}_{sophisticated \quad effect} \tag{4.19}$$

First of all, notice that if $\sigma_t^m = \sigma_t^n$, then the cross derivatives are equal to zero and g_{t+1} is unaffected by changes in l_t, therefore a shock on l_t has no effects on the equilibrium outcome. If $\sigma_t^m \geq \sigma_t^n$ the sign of $V^{i,y}_{M_t l_t}$ is ambiguous. The reason is that two different effects enter the formula. On one hand a rise in the life expectancy makes consumption after retirement more attractive. This increases the desirability of better future pensions and therefore implies a more favorable attitude towards immigration ("preferences effect"). On the other hand more immigration today reduces the value of g_{t+1}. This changes the expected equilibrium policy in the next period in a way that harms a retired individual ("sophisticated effect"). In particular, a decrease in g_{t+1} causes a weak rise in M^{**}_{t+1} and a weak fall in Y^{**}_{t+1}, because of the future political effect. Which of the two effects dominates depends on many aspects, including the income distribution at time $t+1$ and the values of g_t and g_{t+1}. In particular notice that if the variance of the income distribution of the working age population tends to zero, then $\frac{d\theta^v_{t+1}}{dg_{t+1}} = 0$ and therefore the "preferences effect" dominates. Finally, if agents are "naive" then there is no "sophisticated effect" and therefore an increase in l_t has a weakly positive effect on the openness of the immigration policy.

4.3.2.4 Fall in the Birth Rate of the Native Population

I analyse the effects of a fall in σ_t^n keeping all other parameters constant. The results are summarized in the following theorem.

Theorem 10. *The effect of a decrease in the birth rate of the native population σ_t^n is ambiguous on the immigration policy and on the tax rate. If voters are "naive", then the effect is weakly positive on the immigration policy and weakly negative on the tax rate.*

Proof. See Appendix 4.8.1.2.4.

Similarly to the previous case, the presence of a "sophisticated effect" and of a "preferences effect" that can have opposite sign implies that the sign of the comparative statics is ambiguous. If voters are "naive", then the preferences effect implies a weakly less restrictive immigration policy M_t. Moreover, if the immigrants are net contributors to the fiscal system, this also implies a weak fall in the tax rate τ_t. The intuition is that a fall in the birth rate of the natives implies a stronger positive impact of immigration of future pensions and no fiscal effects in the short run. Theorem 10 implies that a fall in the birth rate can have positive effects on public finances and cause a fall in the tax rate because of an increasingly liberal immigration policy in the short run. If this result may seem paradoxical, section 3.4 clarifies that this effect is true only in the current period, while in the long run a fall in the birth rate may have strong negative effects on public finances and tax rates.

4.3.2.5 Shocks to the Income Distribution of the Working Age Population

Given the state g_t, a shock on the income distribution of the working age population affects the equilibrium outcome if and only if it implies a change in the pivotal voter θ_t^v. If this is the case, it represents a shock of type (iii), if it is not, it has no effects. For instance, a shock that results in a median preserving spread of the distribution of θ_t does not imply any change in the identity of the median voter and therefore it does not affect the policy outcome. Thus, I can state the following result.

Theorem 11. *An increase in the median to mean income ratio implies in equilibrium (i) a weak increase in the openness of the immigration policy M_t and (ii) a weak decrease in* *the public spending in the imperfect Public Good Y_t. Moreover, (iii) if the immigrants are net contributors to the fiscal system then it also implies a weak fall in the tax rate τ_t.*

Proof. Results (i), (ii), follows directly from Theorem 4. Result (iii) follows directly from the governmental budget constraint and results (i), (ii). Q.E.D.

Corollary 12. *The equilibrium levels of Y_t and M_t respond in opposite directions to shocks to the voters' distribution.*

Proof. Straightforward from Theorem 11.

Notice that this result implies a positive correlation between the tightness of the immigration policy and the spending on the imperfect public good. This suggests that the concerns about the relationship between an open immigration policy and cuts to public benefits, which are documented in all attitudinal studies, may have some ground in the observed policy outcomes even if immigrants are net contributors to the tax system.

4.3.3 Steady-State Equilibrium

I define a long-run equilibrium of the overlapping generation model as a sequence of Markov-Perfect coalitional equilibria from time t onwards. Within this class, I define a steady-state as follows.

Definition 2. *A steady-state at time t is a sequence of Markov-Perfect coalitional equilibria $\left\{(\mathbb{P}_{t+s}, A_{t+s}, x^*_{t+r}, \{x^{**}_{t+s+r}\}_{r=0}^{\infty}; g_{t+s})\right\}_{s=0}^{\infty}$ such that, in each time $t+s$, and in absence of shocks on the parameters φ_{t+s}, (i) the policy platform implemented in equilibrium is the same in each period $t+s$, i.e. $x^*_{t+s} = x^{**}_{t+s} = x^{ss}$ for all $s \geq 0$, and (ii) the state of the economy is constant $g_{t+s} = g^{ss}$ for all $s \geq 0$.*

In the definition above, the superscript ss denote the steady-state value of a state or a control variable. In other words, in a steady state the equilibrium policy and the natural growth rate of the population are constant over time. Recall that g is the only state that evolves endogenously in the dynamic system and that at a Markov-Perfect coalitional equilibrium in each period $t+s$ - conditional on g_{t+s} and on expectations $\{x^{**}_{t+s+r}\}_{r=0}^{\infty}$ - the equilibrium policy x^*_t may not be unique. Neverthless, if the conditions of Lemma 6 are satisfied, then the set of equilibrium policies is a singleton, and in order to show that the economy is at a steady state one has to show that $g_s = g^{ss}$ for all $s > t$. Conditional uniqueness also implies that if $g_{t+s} = g_{t+s+1}$ in period $t+s$ and if the parameters are such that $(l_{t+s}, \sigma^m_{t+s}, \sigma^m_{t+s}) = (l, \sigma^m, \sigma^m)$ for all $s > 1$, i.e. $\varphi_{t+s} = \varphi$ for all $s > 1$, then

$g_{t+s} = g_t = g^{ss}$ for all $s > t$. In such case, if $g_{t+1} = g_{t+2}$, then the economy is at a steady state.

Lemma 13. *If there exists a Markov-Perfect Coalitional Equilibrium in each period $t+s$, for all $s \geq 0$, then (i) an equilibrium for the OLG model at time t exists . Moreover, if $\varphi_{t+s} = \varphi$ for all $s > 0$, then (ii) there is an equilibrium that always converges to a steady-state. Lastly, if $\sigma_t^m = \sigma_t^n = \sigma_t$, then (iii) the political equilibrium at time t is independent of the previous political choices and the economy converges immediately to the steady state after a shock.*

Proof. See Appendix 4.8.1.2.5.

Notice that this statement does not necessarily imply that the steady state is unique, except for case (iii).

4.3.4 Dynamics

The analysis of the dynamics of the OLG model is a complex exercise because of the number of different short-run effects described in the previous sections. There are anyways interesting results that can be stated about the long-run effects of shocks in this framework. In particular, I present two analytical results: (i) the long-run effects of an unanticipated permanent shock on the longevity of the elderly l_{t-1} and/or on the natural growth rate of the native population $\bar{\sigma}_{t-1}$ on the sequence of political equilibria from the period after the shock until the economy converges to a new steady state (keeping other parameters constant); (ii) the long-run effects of an unanticipated permanent shock in the life expectancy (l_t) and/or on the expected birth rate of the native population (σ_t^n) in the case in which immigration does not cause changes in the age profile of the society (i.e. $\sigma_t^m - \sigma_t^n \leq \eta$ for arbitrarily small η). For the other cases which I cannot address analytically I propose a simulation in section 3.5 which show that the results are not qualitatively different from the one presented in the following paragraphs.

4.3.4.1 Long-Run Effects of a Permanent Shock on the Longevity of the Retired Population and on the Natural Growth Rate of the Working Age Population

The sign of the long-run effects of a positive shock on the longevity l_{t-1} or on the natural growth rate of working age population $\bar{\sigma}_{t-1} - 1$ at time t depend on the ambiguous short-run effects on the immigration policy stated in Theorems 7-8. In order to address the effects at period $t+1$ and the following ones it is sufficient to notice that, given that the shock is permanent (i.e. $l_{t+s} = l_t$ or $\sigma^n_{t+s} = \sigma^n_t$ for all $s \geq 0$), the collective choice problem at time $t+1$ is identical to the one at time t except for the value of g. Thus, there is an equilibrium in which expectations are time-independent. In such equilibrium, all the changes in the policy choices at time $t+1$ must be due to the evolution of the endogenous state g_{t+1}. The results in Lemmas 7b-8b apply to this class of equilibrium. In particular notice that g_{t+1} is strictly decreasing in M_t, and therefore if $M_t \geq M_{t-1}$ $(M_t \leq M_{t-1})$ then $g_{t+1} \leq g_t$ $(g_{t+1} \geq g_t)$. Lemma 6 (iii) ensures that the parameter that identifies the pivotal voter changes accordingly $\theta^v_{t+1} \geq \theta^v_t$ $(\theta^v_{t+1} \leq \theta^v_t)$. Therefore I can state the following results.

Theorem 7b. *The long-run effect of an increase in l_{t-1} on the immigration policy has same sign as the short-run effect and a weakly larger magnitude. If $g_t \geq \hat{g}_t$ then the effect on immigration policy is (weakly) negative and the effect on the public spending and the tax rate is strictly positive.*

Proof. See Appendix 4.8.1.2.5.

Similarly one can analyse the long-run effects of a fall in the natural growth rate of the native population of working age (or equivalently of σ^n_{t-1}, see section 3.2.2). The result is the following.

Theorem 8b. *The long-run effect of a decrease in the natural growth rate of the native population is a weak decrease in the openness of the immigration policy and a weak increase in spending in the imperfect Public Good and in the tax rate. All the effects have weakly larger magnitude relative to the short-run effects.*

Proof. See Appendix 4.8.1.2.5.

These results imply that the effects of population ageing are persistent and tend to increase in magnitude in the periods after the shock. The reason is that - if immigrants have higher fertility rate relative to the natives - then a change in the size of the immigration flow affects the distribution of voters in the following periods. In particular, a more restrictive immigration policy in the current period implies further population ageing in the future and therefore an increase in magnitude of the initial effects.

4.3.4.2 Long-Run Effects of a Permanent Rise in Life Expectancy

I analyse the long run effects of changes in l_t. This shock generates a number of effects that affect the temporary equilibrium as described in the previous sections. Moreover, the specific path of policies depends on the timing of the different shocks (for instance shocks on l_t and l_{t-1} may occur simultaneously). I study the case in which σ_t^m is arbitrarily close to σ_t^n and the shock is permanent, i.e. l_{t+s} is equal to the new value of l_t for all $s > t$. This case is simple to analyze because the long run effects at time t and $t+1$ after the shock correspond, respectively, to the temporary effects of a rise in l_t and l_{t-1} described in the previous sections. Moreover, given that $\sigma_{t+s}^m - \sigma_{t+s}^n$ is arbitrarily small, then the "preference effect" and the "sophisticated effect" can be disregarded and the economy converges to the new steady state one period after the shock[16]. Under the proposed restrictions I can state a sharper result:

Theorem 14. *The long-run effects of an increase in the life expectancy is a weak rise in public spending and, if $g_t \geq \hat{g}_t$, a weak decrease in the openness of the immigration policy.*

Proof. In the case of $\sigma_t^m - \sigma_t^n \leq \eta$ the sign of the long-run effect corresponds to the short-run effect of an increase in l_{t-1}. Q.E.D.

4.3.4.3 Long-Run Effects of a Permanent Fall in the Birth Rate of the Natives

I study the long-run effects of a marginal fall in σ_t^n, in the case in which $\sigma_t^m - \sigma_t^n$ is arbitrarily small and the shock is permanent, i.e. the rise of σ_t^n implies that σ_{t+s}^n will be equal to the new value of σ_t^n for all $s > t$.

Theorem 15. *The long-run effects of a marginal decrease in the birth rate of the native population is a weak rise in public spending. The effect on the openness of the immigration policy is ambiguous at time t and weakly negative in the following periods. If voters are "naive" the effect on the openness of the immigration policy is weakly positive at time t and weakly negative in the following periods.*

Proof. In the case of $\sigma_t^m - \sigma_t^n \leq \eta$ the long-run effect corresponds to the short-run effect of an decrease in σ_t followed by a decrease in σ_{t-1}. Q.E.D.

The results in this section suggest that if immigrants are not too different from the natives in terms of fertility rates, then the long run effects of population ageing follow the patterns of the corresponding short-run effects.

[16] Notice that under such restriction, the Markov-Perfect coalitional equilibrium is unique in every period (see Lemma 6), hence - in contrast with the previous paragraph - there is no need to select a class of equilibria.

4.3.5 Simulation

Some interesting cases cannot be fully described analytically, in particular the long-run effects of permanent shocks to the parameters in the case in which the birth rate of immigrant is different from the one of the natives. In order to study these cases I run a simulation of the model whose results are extensively presented in the supplementary material of this paper. This exercise shows that the effects due to the sophistication of voters may be substantial in terms of levels of the equilibrium policy, but they do not generally imply qualitatively different predictions about the shape of the curves describing the policy response to shocks on the parameters. I find that for several different parametrizations - if the difference in the birth rate of immigrants and natives $(\sigma_t^m - \sigma_t^n)$ is not too large[17] - then the predictions of Theorems 14 and 15 are valid even if voters are "sophisticated" (Figures 4.7-4.8) . Figures 4.9 A and 4.9 B show the response of the immigration policy M_t and of the spending policy Y_t to a permanent rise in the life expectancy of the retired population, both for the case of "naive" voters (dashed lines) and of sophisticated voters (solid lines). Although the shape of the two lines is very similar, the equilibrium level is different. Sophisticated individuals fully internalize the effect of current immigration on the composition of the society in the following period. In particular, they anticipate that more immigration in the current period would imply a higher share of young individuals in the next period, and therefore an equilibrium policy that is less favorable to them when they will be retired. Therefore the equilibrium with "sophisticated" voters features a more restrictive immigration policy and a higher public spending in comparison with the case of "naive" voters.

The simulation exercise can also help to understand the factors that determine the speed of convergence to the steady state after a shock. The crucial aspect is that the speed is decreasing in the size of the "sophisticated effect", specifically in the value of $\sigma_t^m - \sigma_t^n$. Figures 4.10 A and 4.10 B show the path of convergence of the immigration policy M_t after a positive (solid line) and a negative (dashed line) shock on the endogenous state g_t, in the case of high difference (4.10 A, $\sigma_t^m - \sigma_t^n = 1$) and low difference (4.10 B, $\sigma_t^m - \sigma_t^n = 0.2$) in the birth rates of immigrants and natives. This exercise suggests that the key results in the previous sections still apply even to the cases in which the long run effects of shocks in the model cannot be characterized analytically. Thus, I can conclude that - in a society characterized by a very large share of retired individuals - population ageing leads to a policy that is closer to the needs of the elderly. In particular, high public spending and increasingly restrictive immigration policies are going to be implemented. These policy changes imply an increasing tax burden on the individuals of working age and may affect the fiscal sustainability of public spending in the long run.

[17] For large value of $\sigma_t^m - \sigma_t^n$ the steady-state may not be unique and a shock may cause a transition to a different equilibrium path. Moreover, the conditions in Lemma 6 may not be satisfied.

4.4 Extensions

In this section I propose three extensions in which I introduce alternative forms of public intervention in social spending and a different legal status of the immigrants. I describe if and how the equilibrium political choices differ from the one presented in section 3. Specifically, I analyse the implication of the model if (i) the pension system is partially funded, if (ii) immigrants do not acquire voting rights and if (iii) the government provides public education. On one hand the main comparative statics results of this paper remain generally valid. On the other hand these exercises deliver some understanding of how different rules in the public sector may affect the attitude towards immigration of the voting population. Lastly, in section 4.4, I describe (iv) an extension of the model in which the labour market is segmented. In particular, I study the case in which the elderly demand specific services, such as home care, and only immigrants possess the skills to provide such services. In this case the results may differ substantially from the ones of the baseline model.

4.4.1 Partially Funded Pension System

The assumption of a pure Pay-As-You-Go pension system is a very stylized description of how the social security for the elderly is organized in most developed countries. In particular partially funded pension schemes are becoming increasingly common. There is empirical evidence of an increasing size of the funded part of the pension relative to the "state pension" in European countries (Galasso and Profeta, 2004). The theoretical analysis proposed by Rangel and Zeckhauser (2001) suggests that this phenomenon may also be related to the increase in the number of elderly relative to the working age population. In the model proposed in this paper I did not explicitly account for savings. One simple possibility is to model the funded part of the pension system as a form of compulsory savings. Under this assumption each individual has to save an amount $\psi(\rho)s(y_t^i)$ when young and she will receive $(1+r)\psi(\alpha+\gamma)s(y_t^i)$ when retired, where r is the exogenous interest rate and ψ is a strictly decreasing function. The total pension received by i at time $t+1$ becomes:

$$p_t^i = \left(\alpha + \gamma \frac{y_t^i}{\bar{y}_t}\right) \frac{\bar{\sigma}_t}{(1-\widehat{M}_{t+1})} + (1+r)\psi(\alpha+\gamma)s(y_t^i) \quad (4.20)$$

This formulation implies that if the state-pension component falls (e.g. if γ decreases) then the funded pension part rises. Notice that - because the utility function is linear in consumption - the size of the compulsory saving does not affect voter preferences over policies. Thus, the effect of a marginal transition towards a fully funded pension system simply corresponds to the effect of a fall in the Beveridgean part of the state pension α or of a fall in the Bismarckian part γ (or both). Hence I can state the following result:

Theorem 16. *The effect of a marginal decrease in the size of the public pension system in the short run is an increase in the restrictiveness of the immigration policy. In the long run, the effect is an increase in restrictions to immigration and an increase in public spending in the imperfect Public Good. The total effect on the tax rate is ambiguous.*

Proof. See Appendix 4.8.2.1.

The intuition that underpins this result is simple. If the share of the Pay-As-You-Go component of the pension system decreases in favor of a fully funded scheme, then the fiscal gains from immigration for a worker decrease because the total size of public pension expenditures to be shared among the working age population is smaller. Moreover, the future gains from immigration also decrease, because public finance aspects have a lower impact on the overall pension enjoyed by a retired individual. Therefore all voters become more averse to immigration and ask for a more restrictive policy. In the long run, if the immigrants have higher fertility rates relative to the natives, this political choice causes an increase in the share of elderly individuals, with consequences that are similar to the ones described in section 3.4.1 for the case of increasing life expectancy. Namely, a further tightening in the immigration policy and an increase in the endogenous part of public spending prevail in equilibrium. An important aspect of this analysis is that if the size of the state pension system becomes too small (e.g. small $\alpha + \gamma$) then the total gains from immigration for a working age individuals may become negative, which implies an equilibrium in which the most restrictive immigration policy is implemented.

4.4.2 Voting Rights: *Ius Soli* vs. *Ius Sanguinis*

In the previous sections I have assumed that the children of immigrants that are born in the guesting country are awarded the voting right when they become adults (*Ius Soli*). Moreover, in the model voting rights can be also acquired after a sufficiently long period of legal residency. These assumptions are consistent with the legal procedures to obtain citizenship - and consequently voting rights - in several countries such as the US, Canada and France. In many other countries - such as the UK, Japan, Germany and Italy - the legal requirements are often quite different and they do not typically imply an automatic award of the citizenship based of the place of birth only. The most common case is that at least one of the parents must possess the citizenship in order for the children to obtain the same status (*Ius Sanguinis*). It is out of the scope of this paper to formulate assumptions that precisely describe the law of different countries. Nevertheless, in order to understand the possible effects of different legal requirements, it is useful to analyze the consequences of the opposite assumption in comparison with the one in section 3.2 of this paper. Namely, in this section I assume a pure form of *Ius Sanguinis,* in which neither the immigrants nor their children ever obtain the nationality. This assumption is clearly extreme and only serves as a term of comparison.

The main implications of the model stated in Theorems 7-8 are unaffected by this modification, except for one aspect. Specifically, immigrants and their children do not become members of the voting population at any point in time. Therefore the choice of the immigration policy does not affect the future composition of the voting population. This implies that there is no "sophisticated effect" in this case, and therefore some of the results in section 4 are sharper. Namely for any σ_t^n, σ_t^m such that $\sigma_t^m \geq \sigma_t^n$ one gets:

Theorem 17. *(i) the short-run effect of a rise in l_t in unambiguously (weakly) positive on the immigration policy M_t and weakly negative on the tax rate τ_t; the long-run effects of (ii) an increase in the life expectancy and of (iii) a decrease in the birth rate of the native population is a weak rise in public spending and, if $g_t \geq \hat{g}$, a weak increase in the openness of the immigration policy at time t followed by a weak fall in the following periods.*

Proof. The relevant variable for determining the pivotal voter is in this case $\tilde{g}_t = \frac{l_{t-1}\tilde{n}_{t-1}}{\tilde{n}_t} = \frac{l_{t-1}}{\sigma_{t-1}^n}$ where $\tilde{n}_{t-1} \leq n_{t-1}$ is the number of young individuals that possess voting rights at time $t-1$ and it is smaller than or equal to the number of individuals that are born in the country. Notice that $\tilde{g}_t$ is independent of M_{t-1}. The rest of the analysis is unaffected. All the proofs are identical to the ones for Theorems 9-10 except that no "sophisticated effect" occurs.

Theorem 17 suggests that the analytical predictions of the model are not strongly affected by the cross-country differences in the law that regulates the acquisition of the citizenship, and that - on the contrary - some results tend to become sharper and less sensitive to changes in parameter values if an extreme version of the *Ius Sanguinis* is assumed.

4.4.3 Endogenous Public Education

I analyse an extension of the model in which the income of an individual depends not only on the wage rate and on her productivity, but also on the amount of education she received when she was a child. I assume that education is uniformly provided by the government and has decreasing returns given by the strictly concave function f. Individual fertility of natives is given in this alternative setting by the random variable k_t^i, that is i.i.d. and with $E[k_t^i] = \sigma_t^n$. I get that the income of an individual i at time t can be written as follows:

$$y_t^i = f(e_{t-1}) w_t \varepsilon_t^i \tag{4.21}$$

and the total supply of effective labour at time t becomes $L_t = f(e_{t-1})\bar{\varepsilon}_t(n_t + m_t)$. The budget constraint accounts for the public spending in education, such that the formula for the tax rate on labour income becomes:

$$\tau_t = \tau(e_t, M_t, Y_t, \bar{y}_t) = \bar{y}_t^{\;-1}\left[\bar{\sigma}_t e_t + \lambda_t M_t + (\alpha+\gamma) l_{t-1}\frac{(1-M_t)}{(1-\widehat{M}_t)} + Y_t\right] \tag{4.22}$$

Moreover, I assume that working age individuals (retired individuals) care about the utility of their children (grandchildren) such that the utility function of an individual of generation a can be written as follows:

$$\widetilde{U}_t^{i,a} = U_t^{i,a}\left(\left\{C_s^i, M_s, Y_s\right\}_{s=t}^{t+1}\right) + \delta^a E\left[k_t^i U_t^{j,y}\left(\left\{C_s^j, M_s, Y_s\right\}_{s=t+1}^{t+2}\right)\right] \tag{4.23}$$

Lastly, I assume that the number of voters is large, such that the uncertainty about the size of the future generation does not affect the result. Notice that given the assumptions about k_t^i the preferences shown above can also represent individuals that care about the *next generation* rather than about their children and grandchildren. The structure of the overlapping generations model in the same as in the baseline model, except for the presence of an additional endogenous state e_{t-1} which affects the average income at time t. A coalitional equilibrium exists under the assumptions stated in Lemma 6 and most results of this augmented model about the comparative statics of shock on life expectancy are the same as the ones described in the previous section. (See Appendix A.2). The interesting aspect of this analysis is the counterintuitive effect of population ageing on public investment in education (per child). Such effects are stated in the following theorem.

Theorem 18. *The effects of an increase in the longevity of the retired population l_{t-1} and /or of a decrease in the growth rate of native population σ_{t-1}^n is a weak increase in the public spending in education per child e_t.*

Proof. See Appendix 4.8.2.1.

The intuition that underpins this result is that if an elderly individual cares about her grandchildren (i.e. $\delta^o > 0$), then she will always support any policy that increases the spending in education through a rise in the taxes on the working age population, because she is not affected by this rise in the tax rate. The consequence of Theorem 18 is that the next generation may enjoy a better education and a higher pre-tax income as a consequence of population ageing. Notice that the overall welfare effect of the policy adjustment is not necessarily positive for these individuals. The negative side for future generations may come from the results in Theorems 7 and 8, which hold also in the augmented model (see Appendix A). In particular, in period t a more restrictive immigration policy is implemented. Thus, the future generations may have to face a society with a larger share of elderly which implies, *ceteris paribus*, higher tax rates on labour income and more public spending. Such policy can be harmful for the most productive individuals of the next generation. The second result is the following.

Theorem 19. *If voters are "naive" and* $\frac{l_t p_{t+1}^v}{e_t} \geq \frac{\theta_t^v}{\beta}$*, then the effects of a decrease in the birth rate of the native population* σ_t^n *is a weak fall in the public spending in education per child* e_t *and a weak increase in the openness of the immigration policy. Otherwise both effects have an ambiguous sign.*

Proof. See Appendix 4.8.1.2.4.

Theorem 19 suggests that the cost of public education may play a role in shaping the effects of shocks on fertility rates on the immigration policy. On one hand immigration tends to reduce the pressure of the pensions system on public finances, on the other hand it causes an increase in the total costs of public education. If the latter effect is sufficiently strong, the predictions are going to be different from the one implied by the baseline model. This is particularly relevant if one considers that several countries are implementing reforms in order to reduce the Pay-As-You-Go share of the pensions received by the elderly in favor of a fully funded system (see section 4.1). Nevertheless, the public expenditures for the elderly represents a large share of the governmental budget in most western countries and, more importantly, they consistently exceed the ones on education and childcare (OECD 2015, 2015b). Notice that the assumption $\frac{l_t p_{t+1}^v}{e_t} \geq \frac{\theta_t^v}{\beta}$ is satisfied if β is close to 1 and the median cost of a pensioner is weakly larger than the cost of educating a child. Thus, OECD data suggest that such an assumption is consistent with the facts about public spending in most OECD countries.

4.4.4 Services for the Elderly ("Elderly Goods")

In this section I present the results of an extension of the model in which the labour market is segmented. In particular, I study the case in which immigrants possess the skills to provide those services that are needed only by the elderly, such as home care, while the natives workers do not. This may be the case if immigrants are selected by the firms in the receiving country on the basis of their qualifications and previous work experience. In the next line I describe informally the characteristics that differentiate this setting from the baseline model. A detailed description of the economic environment is provided in Appendix B.2. Suppose that the elderly consume a different private good denoted by O_t. This good is produced with the same technology as the consumption good C_t and the imperfect public good Y_t, but only the immigrant workers are capable of producing it. Immigrants can also be employed in the production of the other goods. For simplicity I assume that there is no difference in the average tax payments of immigrants and natives, i.e. $\lambda_t = 0$, that the default immigration is $\widehat{M}_t = 0$ and I analyse the case in which $\sigma_t^m - \sigma_t^n$ is arbitrarily small. There are two possibilities. If at the equilibrium there are enough immigrant workers to satisfy the demand for "elderly goods" at a sufficiently low price, then the segmentation of the labour market is irrelevant and the results are identical to the baseline model. The perfect substitutability in production and the perfect competition ensure that all prices are are unaffected by immigration choices. The implications change dramatically if in the proximity of an equilibrium there are not enough immigrant workers to satisfy the demand for the "elderly good" at the constant price[18]. I can state the following result.

Theorem 20. *If $g_t \leq 1$ then at the equilibrium, if it exists, the immigration policy is $M_t = 0$, else a positive level of immigration is possible.*

Proof. Appendix 4.8.2.2.

This result implies that, as long as the majority of voters is of working age, the society always chooses the most restrictive immigration policy. Moreover, a shock on the longevity or the fertility of the native population does not affect the immigration policy in equilibrium. The channel that underpins this result is the effect of immigration on equilibrium prices. Specifically, immigrants are endogenously hired in the sector that produces the "elderly good" O_t, but they consume only the other two goods C_t and Y_t. As a result, immigration in equilibrium implies a rise in the relative prices faced by the young natives, offsetting the fiscal benefits generated by immigrants and making working age voters extremely hostile to immigration. The conclusion one can derive from this section is that some implications of the analysis presented in section 3 of this paper are true for this extended case only if in the proximity of the equilibrium the immigration policy is not too restrictive. If the number of immigrants is too low to satisfy the demand of services for the elderly, then some predictions in section 3 of the paper are no longer valid. The result in this case is somewhat paradoxical: a society that is in great need of immigrants to satisfy the demand of services for the elderly tend to be very averse to any positive level of immigration of specialized workers. Additional details and results about this extension of the model are available in the supplementary online material.

[18]Notice that multiplicity of equilibria is possible in this case.

4.5 Welfare Analysis

In the previous section I have proved that a rise in the longevity or a fall in the birth rate of the native population generates a political pressure towards more restrictive immigration policy. This does not necessarily imply that this change is desirable on the point of view of the society as a whole. In this section I present a welfare analysis which shows that, if a society has certain demographic characteristics, a marginal increase in the restrictions to immigration is unambiguously harmful for the society. I define a measure of the wellbeing of the society in the form of a *Social Welfare Function* (*SWF*). The idea that is exploited in this section is the following. If at an equilibrium policy the marginal effect of an increase in a policy dimension $x_{j,t}$ on the *SWF* is greater than than the one of the median voter (and at the equilibrium $x^*_{j,t} < \bar{x}_{j,t}$), then there exists a policy with $x'_{j,t} > x^*_{j,t}$ which is welfare improving. This implies in turn that if, as a consequence of a shock, a certain policy dimension j is such that $x^*_{j,t-1} > x^*_{j,t}$, then $x_{j,t}$ has moved in the "wrong direction" on a social welfare point of view and that the society would benefit, *ceteris paribus*, from a marginal change in the direction of $x^*_{j,t-1}$. In other words, the society is harmed by the change in policy at the margin. Consider a SWF that is a weighted average of the utility of each individuals of the working age generation (y), of the retired generation (o) at time t and the expected future utility of the children (ch), where $\mu^a_t(\theta^i_s)$ represents the Pareto weight assigned to an individual i of generation a at time t. Notice that I am not ruling out either the possibility that the *SWF* attributes zero weight to the immigrants or the possibility that some or all the immigrants have positive weight[19]. The *SWF* has form:

$$\begin{aligned} SWF(x_t,;\varphi_t,g_t) = & \int_0^{\bar{\theta}_t} \mu^y_t(\theta^i_t)V^y(x_t;\theta^i_t,\varphi_t,g_t)q(\theta^i_t)d\theta^i_t + \\ & + \int_0^{\bar{\theta}_{t-1}} \mu^o_t(\theta^i_{t-1})V^o(x_t;\theta^i_{t-1},\varphi_t,g_t)q(\theta^i_{t-1})d\theta^i_{t-1} + \\ & + \int_0^{\bar{\theta}_{t+1}} \mu^y_{t+1}(\theta^i_{t+1})E_t\left[V^y(x^{**}_{t+1};\theta^i_{t+1},\varphi_{t+1},g_{t+1})\right]q(\theta^i_{t+1})d\theta^i_{t+1} \end{aligned} \tag{4.24}$$

Most welfare implications of this analysis depend on the Pareto weights assigned to each individual in the *SWF*. For instance, some results that can be obtained using a specific *SWF* (e.g. Utilitarian or Rawlsian) are presented in the supplementary material. Nevertheless an interesting general result can be stated under relative weak restrictions on the *SWF*. Specifically, I analyse the welfare effects of changes in the immigration policy keeping the other policy dimension constant at the equilibrium level. This analysis is also consistent with the extended model presented in section 4.3.

[19] One has to specify the objective function of an immigrant in this case.

4.5.1 Welfare Effects of a Marginal Opening in the Immigration Policy

Assume that $c'(M_t) < \infty$ for all $x_t \in X_t$ and that at the equilibrium $0 < M_t < \overline{M}_t$, i.e. the solution is internal for the immigration policy. Then I can state the following result.

Theorem 21. *For any Social Welfare Function $SWF(x_t; \varphi_t, g_t)$ that assigns a strictly positive weight to each native individual of working age, there exist a threshold $\breve{g}_t \in [0,1]$ such that if $g_t \geq \breve{g}_t$ then a marginal tightening in the immigration policy caused by a change in the equilibrium outcome reduces the Social Welfare.*

Proof. See Appendix 4.8.2.3.

The intuition that underpins this result is that - as g_t tends to 1 - the parameter θ_t^v that identifies the pivotal voter get close to 0. On one hand, the benefits for the individuals of working age from a marginal opening of the immigration policy increase rapidly as M_t approaches 0. On the other hand, the cost of immigration becomes increasingly small at low levels of M_t. If $\theta_t^v = 0$, then $M_t = 0$, which implies that the marginal social gains from immigration are very large relative to the marginal social costs. Also notice that the converse of the statement in Theorem 21 is not always true. Specifically, a threshold $\breve{g}_t \in [0,1]$ such that if $g_t \leq \breve{g}_t$ then the society would benefit from a marginally more restrictive immigration policy may not exists for all the *SWF*s with the features stated above. Nevertheless, such threshold $\breve{g}_t$ exists for Utilitarian and Rawlsian *SWF*. The result in Theorem 21 suggests that societies characterized by high income inequality and/or by a high share of elderly in the total population (which have a g_t close to 1 or larger) are likely to adopt excessively restrictive immigration policies. Moreover, it implies that a tightening in the immigration law - for instance the one caused by population ageing - reduces the Social Welfare. In other words, the policy adjustment of the immigration quota is harmful for the society. This result is suggestive in the light of the increasingly and rather controversial restrictions to immigrations that have been progressively introduced in countries characterized by a rapidly ageing population and by a high degree of income inequality, such as the UK and the USA, or in countries that feature a very large elderly population, such as Japan or Italy. In the supplementary material I propose a welfare analysis about the effects of a change in the public spending in the imperfect Public Good and in education. These results are less general because they rely on more restrictive assumptions about the *SWF* (e.g. Utilitarianism). Nevertheless, they suggest that the allocation of public spending may be too generous for the imperfect Public Good and perhaps insufficient for education in society characterized by high income inequality and by a large share of elderly.

4.6 Empirical Evidence

In this section I investigate the determinants of the attitudes towards immigration and public spending of adult residents in Great Britain using data from the British Social Attitude Survey, and in particular from the rounds of data 2009 - 2011 - 2013 that includes a specific section about immigration. The dataset accounts for a total of 6639 observations. The explanatory variables are the age of the respondent, the income decile of the household and the highest educational qualification attained by the respondent, on a scale from 1 (postgraduate degree) to 8 (no qualification). Observations of individuals with foreign qualifications have been omitted. Dummy variables capture whether the household includes children, and if the respondent is a woman, if she lives in rural areas, if she is born abroad and if she is not part of any religion. Characteristics related to the employment status and type are captured by dummies. In particular, I include the effects of being employed in a manual job, unemployed or retired.

4.6.1 Determinants of Attitude towards Immigration

The outcome variable LETIN captures the attitude towards further immigration in the country. The question is “Do you think the number of immigrants to Britain nowadays should be increased a lot, increased a little, remain the same as it is, reduced a little or reduced a lot?” and the respondents must choose a value on a discrete scale from 1 (“increased a lot”) to 5 (“reduced a lot”). The variable LETIN measures therefore the degree of aversion towards further immigration. I use an ordered Logit model because of the discrete and ordered nature of the outcome variable. Table 1 presents the results of this analysis. In line with what is observed in the literature (Dustmann and Preston 2007, Facchini and Mayda 2007 and Card *et al.* 2012) and with what is implied by the model proposed in this paper, the age of the respondent exhibit a significant positive relationship with the hostility towards immigration. Moreover, the parameter on household income is negative and significant in all the specifications. This means that high income individuals tend to be less averse to immigration relative to the low income, and this is consistent with the implications of the model. Similarly, low level of education tend to be associated with a stronger aversion to immigrants. Lastly, the presence of children in the household, the location in a urban area and the birth of the respondent outside of the UK are all significantly related to a more positive attitude towards immigrants.

4.6.2 Determinants of Attitude towards Public Spending

The outcome variable TaxSpend is a measure of the attitude towards public spending financed through taxation. This variable capture a fundamental trade-off that drives the results in section 3. Namely, it measures the degree of aversion to higher taxes in exchange of more social spending. The question is "Suppose the government had to choose between the three options on this card: reduce taxes and spend less on health, education and social benefits, Keep taxes and spending on these services at the same level as now, Increase taxes and spend more on health, education and social benefits. Which do you think it should choose?" and the respondents must choose a value on a discrete scale from 1 ("spend less") to 3 ("spend more"). I use an ordered Logit model for the same reasons explained in the previous section. Table 2 shows the results of this analysis. The relationship between the outcome variable and the age and the income of the respondent are both significant and the signs are consistent with the implication of the model and in line with the previous literature (Brook, Hall and Preston, 1998). Unemployment is also related with a more favorable attitude towards public spending. It is somewhat surprising that low levels of education are associated with a stronger aversion to taxes and public spending. This may be due to factors that are not considered in the theoretical analysis and that are likely to vary across different education level, such as knowledge of the structure of the fiscal system, awareness of the demographic and economic structure of the country and degree of altruism.

4.6.3 Discussion

The analysis in this section provides a strong support for two crucial implications of the model regarding voters' preferences in Britain. Namely, the analysis of the attitudinal data in the BSA suggests that older age tend to be associated with stronger aversion towards immigration and with a higher propensity to increase the size of public intervention in public spending policies, even if this implies higher taxes. Moreover, the analysis implies that (conditional and unconditional on the level of education), higher levels of income tend to correspond to a more positive attitude towards immigrants and to a stronger propensity to cut taxes and public spending. It may be worth underlining that this analysis does not make any claim about a causal relationship between the variables of interest. The results in section 6.1 are consistent with other similar studies in the literature that use alternative dataset and analyse other countries or group of countries. For instance Dustmann and Preston (2007), Facchini and Mayda (2007) and Card *et al*. (2012), using respectively data from the British Social Attitude Survey, the International Social Survey Programme and the European Social Survey, all support these findings. Thus, one can conclude that there is substantial empirical evidence in support of the patterns of attitudes induced by age and income that are implied by the model proposed in this paper, even if no causal relationship can be claimed. A more general question concern the empirical support to the main predictions of the paper, which concern the comparative statics of the policy outcome. Specifically, it would be critical for this stream of literature to assess in future research if population ageing tend to be associated to more restrictive immigration policies and, if so, to what extent this is due to a causal link between these two variables. The answer to this question is not straightforward. First of all, population ageing is a demographic phenomenon that produces effects on a very long time span and it is likely to be associated with a number of other economic and political transformations. Thus, it is not an easy task to disentangle its effect on specific policies, such as immigration, from other endogenous processes that may induce correlation between the the variables of interest. Moreover, immigration policies are not easy to measure. The "tightness" of an immigration policy is a multidimensional concept, in the sense that such policies can be restricted in various ways, targeting different kinds of immigrants, etc. Moreover, its relationship with the number of immigrants that legally enter a country in a given period of time may be highly endogenous. For instance, on one hand it is reasonable to expect that a country with a more restrictive immigration policy allows, *ceteris paribus*. a smaller number of immigrants to enter the country relative to one with a more liberal set of rules. On the other hand, a country that is subject to a more intense *immigration pressure*, for instance because it is more attractive for potential immigrants, may tend to experience a larger inflow of immigrants even if its immigration policy is more restrictive in comparison with a less attractive country. Similarly, an increase in the immigration pressure due to exogenous factors may translate into a more restrictive immigration law and to a larger inflow of foreigners in the country. In other words, immigration choices and policy choices are two interdependent endogenous processes, and this must be accounted for if one aims to study the latter in isolation from the former. Lastly, immigration policies are often formulated in terms of *qualitative* requirements, which may not be easy to translate into an objective measure of "tightness". For instance, the immigration law often assigns different status to potential immigrants that possess different education levels, or that come from specific countries. Attempts to measure the "tightness" of immigration policies have been made by Boeri and Brucker (2005) for 15 European countries countries and by Ortega and Peri (2009) for 14 OECD countries. Their measures consist of a number of indexes constructed under different definitions of "tightness" of an immigration policy. The limitations in the use of these data are not negligible. Specifically, the low number of observations, the limited extent of time variation that can be exploited and the robustness of the findings to different concepts of "tightness" are important issues. Thus, this literature did not provide so far enough evidence in support or against the predictions of the model proposed in this paper. This remains an open and challenging question for future research.

4.7 Concluding Remarks

This paper investigates the interaction between two crucial demographic, economic and social processes in our society: ageing and immigration. The aim is to analyse how these two processes shape policy choices in democratic countries, and how such policy choices may affect the demographic profile of the society. In particular, I study the effects on immigration policies of two major demographic changes that have caused population ageing in western societies, namely increasing life expectancy and decreasing birth rates. The main finding concerns the fiscal consequences of population ageing. That is, if the share of elderly population is large enough, population ageing increases the political pressure to restrict the inflow of immigrant workers into the country and to arise public spending. This result implies that the negative effects of population ageing on public finances - due to increasing costs for public pensions - may be exacerbated by the endogenous political effects on immigration and public spending policies. Direct and indirect effects of the ageing phenomenon may affect the overall fiscal soundness of the public sector in the long run. The second result looks at the demographic consequences of ageing. In particular, I show that the effects of a demographic shock on the age profile of the population tend to worsen with time because of the endogenous political effects on the immigration policies. Specifically, I find that an ageing society tends to support increasingly restrictive immigration policies. This translates into a reduced number of immigrants and - in some cases - into further population ageing in the future. The third finding is about social welfare. I show that the changes in the immigration policy induced by population ageing tend to harm the society, in particular the young individuals and future generations. One element that emerges from this analysis is that the way in which costs and benefits generated by immigration are divided up in the society is crucial to determine the attitudes towards immigration of different demographic groups. This implies that an analysis of the political processes that lead to the division of these net gains is essential in order to assess the political effects of ageing on immigration policies. Thus, the study of the latter cannot abstract from how fiscal policies are determined.

There are anyway some limitations in this analysis that one has to consider. First, in this study the endogenous adjustment of wages has no effect on the equilibrium policy choices. This is due to the assumption that the individual labour supply is perfectly inelastic both at the extensive and at the intensive margin. This modelling choice is justified by theoretical (Ben-Gad, 2004) and empirical considerations (Dustmann and Preston, 2006, 2007; Boeri, 2010) and can be relaxed to some extent (see additional material). Nevertheless this aspect is likely to play a role in shaping immigration policies. Thus, this is a topic that calls for further research. Secondly, I do not fully investigate the effects of the heterogeneity in the productivity of immigrants. This aspect is likely to be relevant given that such heterogeneity may be - at least to some extent - endogenous in the political process. For instance, simple theoretical models suggest that countries with a generous welfare system may attract relatively low skilled immigrants

(Borjas, 1999), and that the attitude towards different types of immigration may vary with the composition of skills of the native population (Benhabib, 1996). Even if the empirical literature provide limited support for either of these channels (see Preston, 2014), they represent important elements to enrich the study of the determinants of immigration policy. Lastly, a deeper analysis of the determinants of the aversion to immigration due to concerns related to the effects on the "compositional amenities" of the society is needed in order to better understand what other factors shape immigration policies. This aspect has been shown to play a major role in attitudinal studies (Card *et al.*, 2012) and it is an active field of research in other disciplines (see Brettell and Hollifield, 2007), but it has not been sufficiently analyzed with the tools of economic theory. A more general remark should be made about the model of political interaction and the equilibrium concept adopted in this paper. This framework represents a tool that does not only serves for the purposes of this analysis, but it is sufficiently general to be used in many other applications in Political Economy. There are many other questions in Political Economy for which the multidimensionality of the policy space represents a major obstacle in the analysis, and that therefore represent a promising field of application for the voting model presented in this paper. Examples of these potential new applications are described in chapter 2 of this work.

Lastly, I emphasize that this analysis delivers an essentially pessimistic message about the evolution of our society in the immediate future and its consequences for the young generations. If population ageing means an increasing power for the elderly to shape public policies according to their needs, the main victims of this process are going to be the young, both the ones born in rich countries and the ones native of poorer regions. On one hand the former will have to support the fiscal burden of an increasingly large and long-living elderly population through high tax rates on their income. On the other hand the latter are going to be prevented from searching for better employment opportunities by the excessively restrictive immigration policies that are going to be implemented in the high income countries.

4.8 Appendix

4.8.1 Proofs: Main Results

Appendix A includes the proofs to the main results of the paper. Specifically, in appendix A.1 I prove the Lemmas related to the existence of a Markov-Perfect coalitional equilibrium. In appendix A.2 I provide proofs of the main comparative statics results.

4.8.1.1 Existence of Equilibrium

4.8.1.1.1 Markov Property of Ideal Policies

Lemma 1. *In a Markov-Perfect coalitional equilibrium - if it exists - (i) each individual's ideal policy x_t^i and (ii) the equilibrium policy x_t^* at time t are invariant - conditional on g_t - to the history up to time $t-1$, i.e. $x_t^i(g_t,\varphi_t,h_{t-1}) = x_t^i(g_t,\varphi_t,h'_{t-1})$ and $x_t^*(g_t,\varphi_t,h_{t-1}) = x_t^*(g_t,\varphi_t,h'_{t-1})$ $\forall t$ and $\forall h_{t-1},h'_{t-1} \in H_{t-1}$.*

Proof. Part (ii) is simply a consequence of rational beliefs and of the Markov property. Part (i) follows the F.O.C.s for each individual i at time t:

$$\begin{aligned}\widetilde{V}^{i,y}_{M_t} = \quad & -c'(M_t) + \theta_t^i\,(\alpha+\gamma)\,\tfrac{l_{t-1}}{(1-\widehat{M}_t)} + \\ & -\theta_t^i\lambda_t + \left(\tfrac{\beta l_t}{(1-\widehat{M}_{t+1})}(\alpha/\theta_t^i+\gamma) - e_t\right)(\sigma_t^m-\sigma_t^n)\theta_t^i + \\ & +\beta l_t \textstyle\sum_{j=1}^{3} \tfrac{\partial V^{i,o}_{t+1}}{\partial x_j}\tfrac{dx^{**}_{j,t+1}}{dM_t}\end{aligned} \tag{4.25}$$

$$\widetilde{V}^{i,y}_{Y_t} = -\theta^i + b'(Y_t) + \beta l_t \sum_{j=1}^{3} \frac{\partial V^{i,o}_{t+1}}{\partial x_j}\frac{dx^{**}_{j,t+1}}{dY_t} \tag{4.26}$$

$$\widetilde{V}^{i,y}_{e_t} = -E(\sigma_t)\theta_t^i + \delta\sigma_t^n f'(e_t)E(\bar{\omega}_{t+1}) + \beta l_t \sum_{j=1}^{3} \frac{\partial V^{i,o}_{t+1}}{\partial x_j}\frac{dx^{**}_{j,t+1}}{de_t} \tag{4.27}$$

as x^{**}_{t+1} only depends upon $g_{t+1} = \frac{l_t}{\bar{\sigma}_t(M_t)}$, (and in the case of endogenous education, the pivotal voter is unaffected by e_t) then the above reduces to:

$$\widetilde{V}^{i,y}_{M_t} = \quad -c'(M_t)+\theta^i_t(\alpha+\gamma)\tfrac{l_{t-1}}{(1-\widehat{M}_t)}-\theta^i_t\lambda_t+ \\ +\left(\tfrac{\beta l_t}{(1-\widehat{M}_{t+1})}(\alpha/\theta^i_t+\gamma)-e_t\right)(\sigma^m_t-\sigma^n_t)\theta^i_t+ \\ +\beta l_t\textstyle\sum_{j=1}^3\tfrac{\partial V^{i,o}_{t+1}}{\partial x_j}\tfrac{dx^{**}_{j,t+1}}{dg_{t+1}}\tfrac{l_t(\sigma^m_t-\sigma^n_t)}{\bar{\sigma}^2_t} \tag{4.28}$$

$$\widetilde{V}^{i,y}_{Y_t} = -\theta^i + b'(Y_t) \tag{4.29}$$

$$\widetilde{V}^{i,y}_{e_t} = -E(\sigma_t)\theta^i_t + \delta\sigma^n_t f'(e_t)E(\bar{\omega}_{t+1}) \tag{4.30}$$

Given that the F.O.C.s are invariant to h_s for all $s<t$ and that the optimum is unique, then the individual ideal policy must be invariant to h_s for all $s<t$ as well. Q.E.D.

4.8.1.1.2 *SM* and *SID*

Lemma 2. *The function $V(x_t;\theta^i_t,\varphi_t,g_t)$ satisfies SM and SID in $(x_t;\theta^i_t)$ for all $\theta^i_t\in\Theta$ and for all $\varphi_t\in\Phi_t$ for any given state g_t.*

Proof. Given the definition of V^i_t ($\widetilde{V}^i_t$ for the full model), using formula (4.14) one gets:

$$V^i_t = V(x_t;\theta^i_t,\varphi_t,g_t) = \begin{cases} V^{i,y} & if \quad young \quad then \quad \theta=\theta^i_t \\ \kappa V^o & if \quad old \quad then \quad \theta=-1 \end{cases} \tag{4.31}$$

With $x_{1t}=M_t$, $x_{2t}=-Y_t$, $x_{3t}=-e_t$ and for an arbitrarily large $\kappa>0$. Notice that κ represents a strictly increasing transformation of the original objective function of the elderly therefore κV^o implies the same preferences as V^o. First I need to show that each component $V^{i,y}_t$, V^o_t ($\widetilde{V}^{i,y}_t$, $\widetilde{V}^o_t$) satisfies the required properties and then I will show that it also holds for the overall function V^i_t ($\widetilde{V}^i_t$). Recall that formula (4.16) implies that the objective function of a young individual in the baseline model is:

$$V^{i,y}_t = (1-\tau_t)\omega^i_t + b(Y_t) - c(M_t) + \beta l_t\left(\frac{(\alpha+\gamma\theta^{i,y}_t)\bar{\sigma}_t}{(1-\widehat{M}_{t+1})} + d(Y^{**}_{t+1}) - c(M^{**}_{t+1})\right) \tag{4.32}$$

and in the full model is:

$$\begin{aligned}\widetilde{V}_t^{i,y} = \quad & (1-\tau_t) f(e_{t-1})\omega_t^i + b(Y_t) - c(M_t) + \\ & +\beta l_t \left((\alpha + \gamma\theta_t^{i,y}) \tfrac{\bar{\sigma}_t}{(1-\widehat{M}_{t+1})} + d(Y_{t+1}^{**}) - c(M_{t+1}^{**}) \right) + \\ & +\delta^y \sigma_t^n f(e_t) \bar{\omega}_{t+1}\end{aligned} \tag{4.33}$$

Below I derive the conditions for the full model with endogenous education. Given these conditions, the ones for the baseline model are straightforward. Given that the function $\widetilde{V}_t^{i,y}$ is twice differentiable under the assumption stated in section 2.3, sufficient conditions for *SM* and *SID* are simply related to the sign of the cross derivatives and in particular: $\widetilde{V}_{e_tM_t}^{i,y}, \widetilde{V}_{e_tY_t}^{i,y} \leq 0$, $\widetilde{V}_{M_tY_t}^{i,y} \geq 0$ for all $x \in X$ and all $\theta_i^t \in \Theta$ and $\widetilde{V}_{e_t\theta_t^i}^{i,y}, \widetilde{V}_{Y_t\theta_t^i}^{i,y} < 0$, $\widetilde{V}_{M_t\theta_t^i}^{i,y} > 0$ for all $x \in X$ and all $\theta_i^t \in \Theta$. The first derivatives are:

$$\begin{aligned}\widetilde{V}_{M_t}^{i,y} = \quad & -c'(M_t) + \theta_t^i (\alpha+\gamma) \tfrac{l_{t-1}}{(1-\widehat{M}_t)} - \theta_t^i \lambda_t + \\ & + \left(\tfrac{\beta l_t (\alpha/\theta_t^i + \gamma)}{(1-\widehat{M}_{t+1})} (\alpha/\theta_t^i + \gamma) - e_t \right) (\sigma_t^m - \sigma_t^n)\theta_t^i + \\ & +\beta l_t \left(d'(Y_{t+1}^{**}) \tfrac{\partial Y_{t+1}^{**}}{\partial M_t} - c'(M_{t+1}^{**}) \tfrac{\partial M_{t+1}^{**}}{\partial M_t} \right)\end{aligned} \tag{4.34}$$

$$\widetilde{V}_{Y_t}^{i,y} = -\theta^i + b'(Y_t) \tag{4.35}$$

$$\widetilde{V}_{e_t}^{i,y} = -\bar{\sigma}_t \theta_t^i + \delta \sigma_t^n f'(e_t) \bar{\omega}_{t+1} \tag{4.36}$$

Notice that the expectations M_{t+1}^{**} and Y_{t+1}^{**} are solely affected by M_t (through g_{t+1}) because of the Markov assumption. Calculate the cross derivatives of $\widetilde{V}_t^{i,y}$ with respect to each two policy dimensions:

$$\widetilde{V}_{e_tM_t}^{i,y} = -\theta_t^i (\sigma_t^m - \sigma_t^n) \leq 0 \tag{4.37}$$

$$\widetilde{V}_{e_tY_t}^{i,y} = V_{Y_tM_t}^{i,y} = 0 \tag{4.38}$$

And with respect to each policy dimension and the parameter θ_t^i (recall that $x_t^{**}$is a function of solely g_{t+1} and it is therefore invariant to θ_t^i):

$$\widetilde{V}^{i,y}_{e_t\theta_t^i} = -E(\sigma_t) < 0 \tag{4.39}$$

$$\widetilde{V}^{i,y}_{M_t\theta_t^i} = \frac{(\alpha+\gamma)\,l_{t-1}}{(1-\widehat{M}_t)} - \lambda_t - \left(e_t - \frac{\beta l_t \gamma}{(1-\widehat{M}_{t+1})}\right)(\sigma_t^m - \sigma_t^n) > 0 \tag{4.40}$$

$$\widetilde{V}^{i,y}_{Y_t\theta_t^i} = -1 < 0 \tag{4.41}$$

Notice that the FOCs with respect to M_t imply that an interior solution with a partially open migration policy $M_t > 0$ can exist even if immigrants "contribute less than what they take out" in the current period, or more precisely if at a given policy (e_t, Y_t, M_t) a marginal increase in the number of migrants at constant e_t, Y_t implies a rise in the income tax rate. This is true because a native individual of working age will have a future benefit from immigration $\frac{\beta l_t(\sigma_t^m - \sigma_t^n)}{(1-\widehat{M}_{t+1})}(\alpha+\gamma\theta_t^i)$ which incorporates the fact that he will partially internalize the positive effect of immigration today on the governmental budget constraint in the following period through the adjustment in the pension system. This implies that this model is not affected by the dichotomy between "skilled migration" and "unskilled migration" in the patterns of attitude towards immigration and income that is typical of traditional models such as Facchini and Mayda (2008). In my model the attitude towards immigration may improve with income even if the immigrants are a net burden for the society in the short run, because if the Bismarkian component of the pension system is positive ($\gamma > 0$), then the future benefits of current immigration are increasing with income. The next step is to state the elderly's objective function and calculate its first derivatives. Using the formulas for C^o_{t+1} (4.9) one gets:

$$\widetilde{V}^o = l_{t-1}\left[d(Y_t) - c(M_t)\right] + \delta^o E(k_t^i) f(e_t)\bar{\omega}_{t+1} \tag{4.42}$$

First derivatives are:

$$\widetilde{V}^o_{e_t} = E(k_t^i) f'(e_t) E(\omega_{t+1}) > 0 \tag{4.43}$$

$$\widetilde{V}^o_{M_t} = -l_{t-1}c'(M_t) < 0 \tag{4.44}$$

$$\widetilde{V}^o_{Y_t} = l_{t-1}d'(Y_t) > 0 \tag{4.45}$$

and the cross derivatives are given by:

$$\widetilde{V}^o_{e_tM_t} = \widetilde{V}^o_{e_tY_t} = \widetilde{V}^o_{Y_tM_t} = 0 \tag{4.46}$$

Notice that the preferences for (M_t, Y_t, e_t) are the same for all elderly individuals. Now I can show that the function $\widetilde{V}(x_t; \theta^i_t, \varphi_t, g_t)$ satisfies (i) *SM* and (ii) *SID* in $(x_t; \theta^i_t)$.

(i) *SM*. It follows from *SM* of $\widetilde{V}^{i,y}_t$ and $\widetilde{V}^o_t$. (ii) *SID*. I need to show that if $x'_t \geq x''_t$, $x'_t \neq x''_t$ and $\theta'_t > \theta''_t$ then

$$\widetilde{V}(x'_t; \theta''_t, \varphi_t, g_t) - \widetilde{V}(x''_t; \theta''_t, \varphi_t, g_t) > \widetilde{V}(x'_{t},; \theta'_t, \varphi_t, g_t) - \widetilde{V}(x''_t; \theta'_t, \varphi_t, g_t)$$

(ii) (a) $\theta'_t, \theta''_t \neq -1$. *SID* follows from *SID* of $V^{i,y}_t$ and V^o_t. (ii) (b)$\theta' \neq -1$, $\theta''_t = -1$. Notice that $\widetilde{V}(x'_t; \theta''_t, \varphi_t, g_t) - \widetilde{V}(x''_t; \theta''_t, \varphi_t, g_t) > 0$ is always true under the assumption previously stated so it is sufficient to choose κ large enough such that *SID* holds trivially. (ii) (c) $\theta'_t, \theta''_t = -1$. Straightforward. Also notice that under the restriction the parameter set Θ_t defined in (4.13) is a totally ordered set. Q.E.D.

4.8.1.1.3 Median Voter Theorem and Comparative Statics

Theorem 3. *(Median Voter Theorem). If conditions 1-2-3 are satisfied, then (i) A Markov-Perfect coalitional equilibrium of the voting game exists; (ii) in any equilibrium the set of winning policies is a subset of the set of ideal points of the median voter v_t; (iii) if the median voter has a unique ideal policy, then the set of equilibrium policies is a singleton.*

Proof. Consider an objective function $v(x_t, x^{**}_{t+1}(g_t, \varphi_t); \theta^i_t, \varphi_t)$ for some (common) expectations $x^{**}_{t+1}(g_t, \varphi_t)$. Notice that notion of coalitional equilibrium implies that rational expectations must exists at time t, because the political process implied by such equilibrium concept always delivers a policy outcome (it can be an equilibrium outcome in the form of a *Condorcet winner*, or, in case such outcome does not exist, a default policy x^0). Thus, rational expectations exist even if there is no Markov-Perfect coalitional equilibrium at time $t+1$. Moreover, given that in each period $t+s$ voters' indirect utility v - conditional on g_{t+s} and x_{t+1+s} is unaffected by history up to time $t-1+s$, then there must be rational expectations $x^{**}_{t+1}(g_t, \varphi_t)$ that satisfy *MP*. Thus, choose a function x^{**}_{t+1} such that the expectations are rational and satisfy the *MP*. These two ensure that conditions (a) and (b) of the definition of Markov-Perfect coalitional equilibrium (Definition 1) are satisfied. The Markov assumption (*MP*) implies $x^{**}_{t+1}(g_t, \varphi_t) = x^{**}_{t+1}(g_{t+1}(x_t, \varphi_t), \varphi_{t+1})$. Using such rational expectations, define the function $V(x_t; \theta^i_t, \varphi_t, g_t) = v(x_t, x^{**}_{t+1}(g_{t+1}(x_t, \varphi_t), \varphi); \theta^i_t, \varphi_t)$. This is the objective function that corresponds to the static case of coalitional equilibrium. Theorem 1 in

Dotti (2015) states that, if condition 1-2-3 are satisfied, then a coalitional equilibrium exists in the form $(\mathbb{P}_t, A_t, x_t^*)$ (see Ch.2, Theorem 1). This implies that condition (a) of Definition 1 is also satisfied. Then a Markov-Perfect coalitional equilibrium exists. Results (ii) and (iii) in Theorem 3 follow directly from Theorem 1 in Dotti (2015) (see Chapter 2). Q.E.D.

Theorem 4. *(Monotone Comparative Statics). If conditions 1-2-3 are satisfied, then the set of equilibrium policies of the voting game is (i) a sublattice of X_t which is (ii) monotonic nondecreasing in θ_t^v.*

Proof. Theorem 3 implies that if conditions 1-2-3 are satified, then $(\mathbb{P}_t, A_t, x_t^* \{x_{t+s}^{**}\}_{s=0}^{\infty}; g_t)$ is a Markov-Perfect coalitional equilibrium and $(\mathbb{P}_t, A_t, x_t^*)$ is a coalitional equilibrium given the objective function $V(x_t; \theta_t^i, \varphi) = v(x_t, x_{t+1}^{**}(g_{t+1}(x_t, \varphi_t), \varphi); \theta_t^i, \varphi_t)$, in which x_{t+1}^{**} satisfy Rational Expectations and *MP*. Thus, the results in Theorem 2, chapter 2 apply. Q.E.D.

Theorem 5. *(Monotone Comparative Statics 2). If conditions 1-2-3 are satisfied, then the set of equilibrium policies of the voting game is monotonic nondecreasing in φ.*

Proof. Theorem 3 implies that if conditions 1-2-3 are satified, then $(\mathbb{P}_t, A_t, x_t^* \{x_{t+s}^{**}\}_{s=0}^{\infty}; g_t)$ is a Markov-Perfect coalitional equilibrium and $(\mathbb{P}_t, A_t, x_t^*)$ is a coalitional equilibrium given the objective function $V(x_t; \theta_t^i, \varphi_t, g_t) = v(x_t, x_{t+1}^{**}(g_{t+1}(x_t, \varphi), \varphi); \theta_t^i, \varphi_t)$, in which x_{t+1}^{**} satisfy Rational Expectations and *MP*. Thus, the results in Theorem 3, chapter 2 apply. Q.E.D..

4.8.1.1.4 Equilibrium Existence and Characterization

Lemma 6. *If $|\sigma_{t+s}^m - \sigma_{t+s}^n| \leq \hat{\sigma}$ for some $\hat{\sigma} > 0$ and all $s \geq 0$, then (i) a Markov-Perfect coalitional equilibrium for the voting game exists. Moreover, (ii) in any Markov-Perfect coalitional equilibrium at time t the equilibrium policy is the unique ideal point of the median voter $x_t^v = x_t^* \in I_t(v)$. (iii) The parameter θ_t^v that identifies the median voter is weakly decreasing in g_t. If $\sigma_t^m - \sigma_t^n$ is arbitrarily small, then (iv) there is a unique equilibrium policy that is chosen in any Markov-Perfect Coalitional Equilibrium in period t.*

Proof. (i) Consider expectations $\{x_{t+s}^{**}(g_t, \varphi_t)\}_{s=0}^{\infty}$ that are consistent with a MCE such that $x_{t+s}^{**}(g_{t+s}, \varphi_{t+s})$ is unique and differentiable, and $\left|\frac{dx_{k,t+s}^{**}}{dx_{j,t+s}}\right| \leq c_{t+s}(k, j)$ for all k, j and all $s \geq 0$, in which $c_{t+s}(k, j)$ are numbers that are arbitrarily close to 0. I need to show that such expectations are rational for $\hat{\sigma}$ close enough to zero. Start with $x_t^{**}(g_t, \varphi_t)$. As stated, $x^{**}_{t+1}(g_t, \varphi_t)$ satisfies the Markov property, hence $x^{**}_{t+1}(g_{t+1}(x_t, \varphi_t), \varphi_{t+1}) = x^{**}_{t+1}(g_t, \varphi_t)$. Moreover, it is differentiable and consistent with a unique MCE. This means $x^{**}_{t+1}(g_{t+1}, \varphi_{t+1}) = x^{*}_{t+1}(g_{t+1}, \varphi_{t+1}) = x^{v}_{t+1}(g_{t+1}, \varphi_{t+1})$, i.e. must be the unique ideal point of the median voter v_{t+1}. Thus $\frac{dx^{**}_{k,t+1}}{dx_{j,t}} = \frac{\partial x^{**}_{k,t+1}}{\partial g_{t+1}} \frac{\partial g_{t+1}}{\partial x_{jt}} = \frac{\partial x^{**}_{k,t+1}}{\partial \theta^{v}_{t+1}} \frac{\partial \theta^{v}_{t+1}}{\partial x_{j,t}}$. Notice that $\frac{\partial \theta^{v}_{t+1}}{\partial x_{j,t}} = 0$ for all j except for the one such that $x_{j,t} = M_t$. In such case, $\frac{\partial \theta^{v}_{t+1}}{\partial M_t} = -\frac{l_t(\sigma^m_t - \sigma^n_t)}{[(\sigma^m_t - \sigma^n_t)M_t + \sigma^n_t]^2} \frac{1}{q(\theta^{v}_{t+1})}$, which is finite for all M_t and tends to 0 as $\sigma^m_t - \sigma^n_t \to 0$ (notice that if v_{t+1} is the median voter, then $q(\theta^{v}_{t+1}) > 0$, and continuity of q implies that his must be true in a neighborhood of θ^{v}_{t+1}). Moreover, $\frac{\partial x^{**}_{k,t+1}}{\partial \theta^{v}_{t+1}} = 0$ if k is in a corner solution of the maximization problem of the median voter, else $\frac{\partial x^{**}_{k,t+1}}{\partial \theta^{v}_{t+1}} = -\frac{V_{x_{k,t+1}\theta_{t+1}} + \sum_{j \neq k} V_{x_{k,t+1}x_{j,t+1}} V_{x_{j,t+1}\theta^{v}_{t+1}}}{V_{x_{k,t+1}x_{k,t+1}}}$. The numerator is finite (see A.1.2). About the denominator, it is finite if $\frac{dx^{**}_{k,t+2}}{dx_{j,t}}$ is finite for all k, j. But this is true because expectations are such that $\left|\frac{dx^{**}_{k,t+2}}{dx_{j,t+1}}\right| \leq c_{t+2}(k, j)$ for all k, j. One gets $\frac{dx^{**}_{k,t+1}}{dx_{j,t}}$ is the product of a finite factor times a factor that is continuous in $\sigma^m_t - \sigma^n_t$ and tends to zero as $\sigma^m_t - \sigma^n_t \to 0$. Hence, there exists $\hat{\sigma} > 0$ such that if $|\sigma^m_t - \sigma^n_t| \leq \hat{\sigma}$, then $\left|\frac{dx^{**}_{k,t+1}}{dx_{j,t}}\right| \leq c_{t+1}(k, j)$ for all k, j. Because $c_{t+1}(k, j)$ are arbitrarily close to zero, this implies that $V(x_t; \theta^i_t, \varphi_t, g_t)$ is strictly concave, thus $x^*_t(g_t, \varphi_t)$ is consistent with a MCE, unique and differentiable and $\frac{dx^{*}_{k,t}}{dx_{j,t}} \leq c_t(k, j)$ for all k, j. As rational expectations are assumed, then $x^{**}_t(g_t, \varphi)$ must also satisfy those properties. Similarly, one can show that x^{**}_{t+1} is consistent with MCE, unique, differentiable and satisfies $\frac{dx^{**}_{k,t+1}}{dx_{j,t+1}} \leq c_{t+1}(k, j)$ for all k, j given such expectations. Thus, recursively, one can show that this is true for all $x^{**}_{t+s}(g_{t+s}(x_{t+s-1}, \varphi_{t+s-1}), \varphi_{t+s})$ with $s \geq 0$ and, because of the Markov assumption, for $x^{**}_{t+s}(g_t, \varphi_t)$ for all $s \geq 0$. This means that $V(x_{t+s}, \theta^i_{t+s}, \varphi_{t+s}, g_{t+s})$ is continuous and strictly concave in x_{t+s} (it satisfies *SM* and *SID* because of Lemma 2) and that the expectations $\{x^{**}_{t+s}(g_t, \varphi_t)\}^{\infty}_{s=0}$ are rational and satisfy the Markov property. Summarizing, (i) Lemma 2 and the definitions of the policy space X_t and of the parameter space Θ_t, plus the result above imply that all the conditions for the existence of a coalitional equilibrium in Theorem 1 are satisfied. (ii) The strict concavity of the objective function of each working age individual and the convexity of X imply that the pivotal voter has a unique ideal policy, and therefore that is the only policy vector that can be implemented in any coalitional equilibrium of the voting game. (iii) If $g_t \leq 1$, then the median individual in the totally ordered set Θ_t solves $Q(\theta^v_t)n_t + l_{t-1}(m_{t-1} + n_{t-1}) = [1 - Q(\theta^v_t)]\, n_t$[20]. Rearranging and solving for θ^v_t one gets $\theta^v_t = Q^{-1}\left(\frac{1-g_t}{2}\right)$ which is weakly positive and weakly decreasing in g_t. If $g_t > 1$, then the parameter of the pivotal voter is fixed at $\theta^i_t = -1$. Lastly, for (iv), if $\sigma^m_t \to \sigma^n_t$ then $\frac{dx^{**}_{k,t+1}}{dx_{j,t}} \to 0$ for any rational expectations. Thus $V(x_t; \theta^i_t, \varphi_t, g_t)$ is strictly concave in x_t and in period t there is a unique policy vector x^*_t that is chosen in any MCE given g_t.

[20] The tie-breaking rule assumed in section 2.1.2 ensures that this formula is correct even if the number of voters is even.

Given that g_t is known at time t, then x_t^* must be unique. Q.E.D.

4.8.1.2 Comparative statics

4.8.1.2.1 Unanticipated Rise in the Longevity of the Retired Population

Theorem 7. *The effects of an increase in the life expectancy l_{t-1} is weakly positive on the spending policy and ambiguous on the immigration policy. Moreover, there exists a threshold $\hat{g} \in [0,1]$ such that if $g_t \geq \hat{g}$ then the effect on immigration policy is unambiguously (weakly) negative and the effect on the tax rate is strictly positive.*

Proof. Calculate the cross derivatives of $V_t^{i,y}$ ($\widetilde{V}_t^{i,y}$) with respect to each policy dimension M_t, Y_t, e_t and the parameter l_{t-1} using (4.34), (4.35) and (4.36).

$$\widetilde{V}^{i,y}_{M_t l_{t-1}} = \frac{\theta_t^v(\alpha+\gamma)}{(1-M_t^*)} > 0 \tag{4.47}$$

$$\widetilde{V}^{i,y}_{Y_t l_{t-1}} = 0 \tag{4.48}$$

$$\widetilde{V}^{i,y}_{e_t l_{t-1}} = 0 \tag{4.49}$$

(i) Effects at fixed g_t. Consider a totally ordered subset $\Phi_t^j := \{\varphi_t \in \Phi_t | \varphi_{i,t} = \hat{\varphi}_{i,t} \forall i \neq j\}$ where j is the position of the longevity parameter in the vector φ_t, i.e. $\varphi_{j,t} = l_{t-1}$. Notice that $\widetilde{V}(x_t; \theta_t^i, \varphi_t, g_t)$ in Φ_t^j satisfies *SM* in (x_t) and *SID* in $(x_t; \varphi_t)$, it also satisfies *SM* in (z_t) and *SID* in $(z_t; \varphi_t)$ for $z_t = (x_{1t}, -x_{2t}, -x_{3t})$. Using Theorem 3, one gets $\triangle M_t \geq 0$, $\triangle Y_t = 0$, $\triangle e_t = 0$, $\triangle \tau_t \leq 0$. (ii) Recall that

$$g_t = \frac{l_{t-1}}{\bar{\sigma}_{t-1}} \tag{4.50}$$

which is increasing in l_{t-1}. Hence a rise in l_{t-1} corresponds to a change in the voter distribution such that the new median voter is lower than before. Hence $\triangle M_t \leq 0$, $\triangle Y_t \geq 0$, $\triangle e_t \geq 0$, $\triangle \tau_t \geq 0$. Total effect: ambiguous for M_t. But $\triangle e_t \geq 0$, $\triangle Y_t \geq 0$. Finally notice that if $g_t = 1$ then $\theta_t^v = 0$ and $\widetilde{V}^{i,y}_{M_t l_{t-1}} = 0$, which means that the "budget effect" is equal to zero and therefore the the political effect (weakly) dominates. Hence there exists a threshold $\hat{g} \in [0,1]$ (possibly $\hat{g} = 1$) such that if $g_t \geq \hat{g}$ then the effect on immigration policy is unambiguously (weakly) negative. Q.E.D.

4.8.1.2.2 Unanticipated Fall in the Natural Growth Rate of the Native Population

Theorem 8. *The effects of a decrease in the growth rate of the working age population is a weak decrease in the openness of the immigration policy and and a weak increase in spending in the imperfect Public Good and in the tax rate.*

Proof. Calculate the cross derivatives of $V_t^{i,y}$ ($\widetilde{V}_t^{i,y}$) with respect to each policy dimension M_t, Y_t, e_t and the parameter σ_{t-1}^n.

$$\widetilde{V}^{v,y}_{e_t\sigma^n_{t-1}} = \widetilde{V}^{v,y}_{Y_t\sigma^n_{t-1}} = \widetilde{V}^{v,y}_{M_t\sigma^n_{t-1}} = 0 \tag{4.51}$$

The share of "old" voters decreases at each point in time because of the formula for g_t in (4.50), which implis that g_t is decreasing in σ_{t-1}^n. Using Theorem 3, a fall in σ_{t-1}^n implies $\triangle M_t \leq 0$, $\triangle Y_t \geq 0$, $\triangle e_t \geq 0$, $\triangle \tau_t \geq 0$. Q.E.D.

4.8.1.2.3 Rise in the Life Expectancy of the Working Age Population

Theorem 9. *The effects of an increase in the life expectancy l_t is ambiguous on the immigration policy. If voters are "naive" then the effect is weakly positive. If the birth rate of the native is the same as the one of the immigrants, then there is no effect.*

Proof. One needs to analyze the cross derivative of $\widetilde{V}_t^{i,y}$ with respect to M_t, Y_t, e_t and the parameter l_t. Define $\tilde{\pi}_{t+1}^{i,o} = (\alpha + \gamma\theta_t^i)\frac{\bar{\sigma}_t}{1-\widehat{M}_{t+1}} + d(Y_{t+1}) - c(M_{t+1})$ (this is only relevant for the case of endogenous public education).

$$\begin{aligned}\widetilde{V}^{i,y}_{M_t l_t} = & \underbrace{\frac{\beta(\alpha+\gamma\theta_t^i)}{(1-\widehat{M}_{t+1})}(\sigma_t^m - \sigma_t^n)}_{preferences \quad effect} - \underbrace{\frac{\beta 2 l_t}{\bar{\sigma}_t{}^2}\left[\sum_{j=1}^{3}\frac{d\tilde{\pi}^{i,o}_{t+1}}{dx^{**}_{j,t+1}}\frac{\partial x^{**}_{j,t+1}}{\partial\theta^v_{t+1}}\right]\frac{d\theta^v_{t+1}}{dg_{t+1}}(\sigma_t^m - \sigma_t^n)}_{sophisticated \quad effect} \\ & \underbrace{-\frac{\beta l_t^2}{\bar{\sigma}_t{}^2}\left[\sum_{j=1}^{3}\frac{d^2\tilde{V}^{i,o}_{t+1}}{dx^{**}_{j,t+1}}\frac{\partial x^{**}_{j,t+1}}{\partial\theta^v_{t+1}}\frac{dx^{**}_{j,t+1}}{dl_t} + \frac{d\tilde{\pi}^{i,o}_{t+1}}{dx^{**}_{j,t+1}}\frac{d}{dl_t}\left(\frac{\partial x^{2**}_{t+1}}{\partial\theta^v_{t+1}}\right)\right]\frac{d\theta^v_{t+1}}{dg_{t+1}}(\sigma_t^m - \sigma_t^n)}_{sophisticated \quad effect}\end{aligned} \tag{4.52}$$

$$\widetilde{V}^{i,y}_{Y_t l_t} = 0 \tag{4.53}$$

$$\widetilde{V}^{i,y}_{e_t l_t} = 0 \tag{4.54}$$

First of all notice that if $\sigma_t^m = \sigma_t^n$, then the cross derivatives are equal to zero and g_{t+1} is unaffected by changes in l_t, therefore a shock on l_t has no effects on the equilibrium outcome. If $\sigma_t^m \geq \sigma_t^n$ the sign of $\widetilde{V}^{i,y}_{M_t l_t}$ is ambiguous. The reason is that two different effects enter the formula. On one hand an increase in life expectancy increase the relative weight of consumption after retirement in the utility function of a working age individual, increasing the desirability of better future pensions and therefore of an increase in the number of immigrants at time t ("preferences effect"). On the other hand there is a "sophisticated effect" that concerns the effect of current political choices on future outcomes. If the "preferences" effect dominates, then using the same procedure as in C.5.1 I can show that $\widetilde{V}(x_t;\theta_t^i,\varphi_t,g_t)$ satisfies *SM* in (x_t) and *SID* in $(x_t;\varphi_t)$ in Φ_t^j where $\varphi_{j,t} = l_t$, it also satisfies *SM* in (z_t) and *SID* in $(z_t;\varphi_t)$, for $z_t = (x_{1t}, -x_{2t}, -x_{3t})$, which by Theorem 3 implies $\triangle M_t \geq 0, \triangle \tau_t \leq 0$ and no effect on the other variables. If the "sophisticated" effect dominates in a similar way one can show that $\triangle M_t \leq 0$, $\triangle \tau_t \geq 0$. If agents are "naive" then there is no "sophisticated effect" because $\frac{d\theta^v_{t+1}}{dg_{t+1}} = 0$ and therefore an increase in l_t has a weakly positive effect on the openness of the immigration policy. Q.E.D.

4.8.1.2.4 Decrease in the Birth Rate of the Natives

Theorem 10. *The effects of a decrease in the birth rate of the native population σ_t^n is a weak increase in the openness of the immigration policy and a fall in the tax rate. The effects of a decrease in the birth rate of the native population σ_t^n is ambiguous on the immigration policy. If voters are "naive", then the effect is weakly positive.*

Proof. Calculate the cross derivatives of $V_t^{i,y}$ ($\widetilde{V}_t^{i,y}$) with respect to each policy dimension M_t, Y_t, e_t and the parameter σ_t^n. $\tilde{v}^{i,o}_{t+1}$ is defined as in 4.8.1.2.3.

$$\begin{aligned}\widetilde{V}^v_{M_t\sigma_t^n} = \underbrace{e_t\theta_t^v}_{b.e.} - \underbrace{\frac{\beta l_t(\alpha+\gamma\theta_t^v)}{(1-\widehat{M}_{t+1})}}_{preferences\quad effect} + \\ + \underbrace{\frac{\beta l_t}{\bar{\sigma}_t{}^2}\left[d'(Y^{**}_{t+1})\frac{\partial Y^{**}_{t+1}}{\partial\theta^v_{t+1}} - c'(M^{**}_{t+1})\frac{\partial M^{**}_{t+1}}{\partial\theta^v_{t+1}}\right]\frac{d\theta^v_{t+1}}{dg_{t+1}}\left[1+\frac{2(1-M_t)(\sigma_t^m-\sigma_t^n)}{\bar{\sigma}_t}\right]}_{sophisticated\quad effect} +\end{aligned} \tag{4.55}$$

$$\underbrace{+\frac{\beta l_t(\sigma_t^m-\sigma_t^n)}{\bar{\sigma}_t{}^2}\left\{\left[\sum_{j=1}^{3}\frac{d^2\tilde{\pi}_{t+1}^{i,o}}{dx_{j,t+1}^{**}}\left(\frac{\partial x_{j,t+1}^{**}}{\partial\theta_{t+1}^v}\right)^2\frac{dx_{j,t+1}^{**}}{d\sigma_t}+\right.\right.}_{sophisticated\quad effect}$$

$$\underbrace{\left.+\frac{d\tilde{\pi}_{t+1}^{i,o}}{dx_{j,t+1}^{**}}\frac{\partial x_{t+1}^{2**}}{\partial(\theta_{t+1}^v)^2}\right]\left(\frac{d\theta_{t+1}^v}{dg_{t+1}}\right)^2+}_{sophisticated\quad effect}$$

$$\underbrace{\left.+\left[\sum_{j=1}^{3}\frac{d\tilde{\pi}_{t+1}^{i,o}}{dx_{j,t+1}^{**}}\frac{\partial x_{j,t+1}^{**}}{\partial\theta_{t+1}^v}\right]\frac{d^2\theta_{t+1}^v}{d(g_{t+1})^2}\right\}}_{sophisticated\quad effect}$$

Notice that in this case the effect of σ_t^n on future outcomes affect the pivotal voter. Also notice that $d'(Y_{t+1}^{**})\frac{\partial Y_{t+1}^{**}}{\partial\theta_{t+1}^v}-\hat{c}'(M_{t+1}^{**})\frac{\partial M_{t+1}^{**}}{\partial\theta_{t+1}^v}\leq 0$ because Theorem 11. Hence for $\sigma_t^m=\sigma_t^n$ the sophisticated effect is weakly positive hence the overall effect is ambiguous. If agents are naive then $\frac{d\theta_{t+1}^v}{dg_{t+1}}=0$ and the overall sign is negative if and only if:

$$\frac{l_t p_{t+1}^v}{e_t}\geq\frac{\theta_t^v}{\beta} \tag{4.56}$$

i.e. the total transfer in pensions to the median voter at time $t+1$ is sufficiently large in comparison with his tax expenditure in education per pupil (notice that this is always true in the basic model with no public education).

$$\widetilde{V}_{Y_t\sigma_t}^v=0 \tag{4.57}$$

$$\widetilde{V}_{e_t\sigma_t^n}^v=-\theta_t^v(1-M_t)+\delta f'(e_t)\bar{\omega}_{t+1}>0 \tag{4.58}$$

as long as $e_t>0$ at the equilibrium (this condition is only relevant for the extended model). $\widetilde{V}(x_t;\theta_t^i,\varphi_t,g_t)$ satisfies *SM* in (z_t') and *SID* in $(z_t';\varphi_t)$ in Φ^j ($\varphi_{j,t}=\sigma_t^n$), where $z_t'=(-x_{1t},-x_{2t},-x_{3t})$, it also satisfies *SM* in (z_t'') and *SID* in $(z_t'';\varphi_t)$ where $z_t''=(-x_{1t},-x_{2t},x_{3t})$. By Theorem 3 a fall in σ_t^n implies: $\triangle M_t\geq 0$, $\triangle Y_t=0$, $\triangle e_t\leq 0$, $\triangle\tau_t\leq 0$. Q.E.D.

4.8.1.2.5 Steady-State Equilibrium

Lemma 13. *If there exists a Markov-Perfect Coalitional Equilibrium in each period $t+s$,* *for all $s \geq 0$, then (i) an equilibrium for the OLG model at time t exists . Moreover, if $\varphi_{t+s} = \varphi$ for all $s > 0$, then (ii) there is an equilibrium that always converges to a steady-state. Lastly, if $\sigma_t^m = \sigma_t^n = \sigma_t$, then (iii) the political equilibrium at time t is independent of the previous political choices and the economy converges immediately to the steady state after a shock.*

Proof. Fix the value of the parameters. (i) Notice that if a Markov-Perfect coalitional equilibrium exists in each period $t+s$ (Lemma 6), then an equilibrium of the OLG model also exists because it is simply a sequence of such temporary equilibria. Q.E.D. (ii) The equilibrium political choice at time t depends uniquely on the value of the state g_t. Notice that g_t depends on the parameters l_{t-1}, σ_t^m, σ_t^n and on the choice variable M_{t-1} but is is independent of anything else. This implies that the evolution of g depends uniquely on the evolution of M if l, σ^m, σ^n are constant over time. Notice that if $\sigma_t^m = \sigma_t^n = \sigma_t$ then the political equilibrium at time t is independent of the previous political choices because the state g_t is independent of history: $g_t = \frac{l_{t-1}}{\sigma_t} = g^*$, which implies in turn that the economy converges immediately to the steady state after a shock. Also notice that in this case the equilibrium is independent of the lagged value M_{t-1}, hence the steady-state is unique. If $\sigma_t^m > \sigma_t^n$ then this is no longer true and the convergence may take several periods. Finally notice that at constant parameters if $g_{t+s} = g_{t+s+1}$ for some $t+s$, then $g_{t+s+u} = g_{t+s}$ for all $s > 0$, i.e. $g_{t+s} = g_{t+s+1}$ is sufficient for a steady state. Suppose a steady state does not exists, i.e. $g_{t+s} \neq g_{t+s+1}$ for all $s \geq 0$. If $g_{t+1} > g_t$ $(<)$ then the pivotal voter $\theta_{t+1}^v \leq \theta_t^v$ $(\geq)$ which using Theorem 9 implies $M_{t+1}^* \leq M_t^* (\geq)$. This implies in turn that $g_{t+2} \geq g_{t+1} (\leq)$. If $g_{t+2} = g_{t+1}$ then we have reached a steady state. If instead $g_{t+2} > g_{t+1} (<)$ the process continues recursively. There are three possibilities. Either (1) the process stops because $g_{t+s} = g_{t+s+1}$ and a steady state is achieved, or (2) the process converges to some g^{ss}. Else, (3) suppose that $g_{t+s+1} - g_{t+s} > 0$ $(<)$ for all $s \geq 0$. if this is true, then the process implies $M_{t+s+1}^* < M_{t+s}^* (>)$ for all $s \geq 0$. Because the direction of this iterative process is monotonic (increasing or decreasing), if it does not converge to some M^{ss}, then it this implies that if M is unbounded it will diverge to $-\infty$ $(+\infty)$. But $M_t \in [\underline{M}, \overline{M}]$ by assumption, hence the process must stop at $M_{t+s}^* = \underline{M}$ $(\overline{M})$ for some $s \geq 0$. Notice that monotonicity under case (3) implies $g_{t+s+1} - g_{t+s} > 0$ $(<)$ and therefore $M_{t+s+1}^* < M_{t+s}^* (>)$, but this is impossible because $M_{t+s}^* = \underline{M}$ $(\overline{M})$. Hence, $M_{t+s+1}^* = \underline{M}$ $(\overline{M})$, which means $M_{t+s+1}^* = M_{t+s}^*$ and implies $g_{t+s+1} = g_{t+s}$. Hence the system has achieved a steady state, and this leads to a contradiction. Q.E.D. (iii) Straightforward from (ii) and Lemma 6. Q.E.D.

Theorem 7b. *The long-run effect of an increase in l_{t-1} on the immigration policy has same sign as the short-run effect and a weakly larger magnitude. If $g_t \geq \hat{g}$ then the effect on immigration policy is (weakly) negative and the effect on the public spending and the tax rate is strictly positive.*

Proof. If at time t the "Budget Effect" prevails, i.e. $M_t \geq M_{t-1}$, then $g_{t+1} \leq g_t$ and $\theta^v_{t+1} \geq \theta^v_t$ by Lemma 6. Using Theorem 2 one gets $M_{t+1} \geq M_t$ and $Y_{t+1} \leq Y_t$. Notice that this is implies recursively $\theta^v_{t+s+1} \geq \theta^v_{t+s}$ and therefore $M_{t+s+1} \geq M_{t+s}$ and $Y_{t+s+1} \leq Y_{t+s}$ for all $s > 0$. Hence I can conclude that at the new steady state $M^{ss} \geq M_t \geq M_{t-1}$ and $Y^{ss} \leq Y_t$ but $Y^{ss} \gtreqless Y_{t-1}$, which means that the long run effect of an increase in l_{t-1} is positive on the openness of the immigration policy and ambiguous on the public spending variable, which increases at the time in which the shock occurs and falls in the following periods. Similarly one can show that if at time t the "Political Effect" dominates, then at the new steady state $M^{ss} \leq M_t \leq M_{t-1}$ and $Y^{ss} \geq Y_t \geq Y_{t-1}$.

Theorem 8b. *The long-run effect of a decrease in the growth rate of the native population is a weak decrease in the openness of the immigration policy and a weak increase in spending in the imperfect Public Good and in the tax rate.All the effects have weakly larger magnitude relative to the short-run effects.*

Proof. Similar to the previous case.

4.8.2 Proofs: Extensions and Welfare Analysis

Appendix 4.8.2 includes the proof to the results regarding extensions in section 4 of the paper and of the Welfare results in section 5. Moreover, it provides a formal description of the setup in the case of "Elderly goods" informally described in section 4.4.

4.8.2.1 Partially Funded Pension System

Theorem 16. *The effect of a marginal decrease in the size of the public pension system in the short run is an increase in the restrictiveness of the immigration policy. In the long run, the effect is an increase in restrictions to immigration and an increase in public spending in the imperfect Public Good. The total effect on the tax rate is ambiguous.*

Proof. It is sufficient to show that the objective function $V^{i,y}_t$ ($\widetilde{V}^{i,y}_t$) satisfies ID in α (γ). Calculate the cross derivatives of $V^{i,y}_t$ ($\widetilde{V}^{i,y}_t$) with respect to M_t,Y_t (e_t) and the parameter α (γ).

$$\widetilde{V}^{i,y}_{M_t\alpha} = \frac{\theta^i_t l_{t-1}}{(1-M^*_t)} + \frac{\beta l_t(\sigma^m_t - \sigma^n_t)}{(1-M^*_{t+1})} \geq 0 \tag{4.59}$$

$$\widetilde{V}^{i,y}_{Y_t\alpha} = 0 \tag{4.60}$$

$$\widetilde{V}^{i,y}_{e_t\alpha} = 0 \tag{4.61}$$

Hence given a subset Φ^j_t defined as in 4.8.1.2.1 with $\varphi_{j,t} = \alpha$ (γ), one can show that $\widetilde{V}^i_t$ satisfies *ID* with respect to $(x_t;\varphi_t)$ and to $(z_t;\varphi_t)$ with $z_t = (x_{1t}, -x_{2t} - x_{3t})$. Using Theorem 3 this implies that the short-run effect of a fall in α (γ) is $\triangle M_t \leq 0$. In the long run the effect of a weak fall in M_t is a rise in g_t, which implies in turn a "political effect" at time $t+1$ with $\triangle M_t \leq 0 \triangle Y_t \geq 0$, which implies recursively the same effect for all the periods after $t+1$ until the economy converges to a new steady state. Notice that the effect on the tax rate is ambiguous at time t because of a simultaneous reduction of the total cost of pension (as α falls) and of the workforce (because of the fall in M_t), while from time $t+1$ the tax rate increases until a new steady state is achieved, because of the fall in the workforce and the rise in public spending. Therefore the overall long-run effect is ambiguous. Q.E.D.

Theorem 18. *The effects of an increase in the longevity of the retired population l_{t-1} and /or of a decrease in the growth rate of native population σ^n_{t-1} is a weak increase in the public spending in education per child e_t.*

Proof. Straightforward from 4.8.1.2.1 and 4.8.1.2.2.

4.8.2.2 Services for the Elderly ("Elderly Goods")

Suppose that the elderly consume a different private good, for instance home care, denoted by O_t while the young consume the private good C_t. The good O_t is produced with the same technology as the consumption good C_t and the imperfect public good Y_t, but only the immigrant workers are capable of producing it. For simplicity I assume that there is no cost of immigration, i.e. $\lambda_t = 0$, that the default immigration is $\widehat{M}_t = 0$ and I analyse the case in which $\sigma^m_t - \sigma^n_t$ is arbitrarily small. Also assume that the functions $a(Y)$ and $d(Y)$ are such that $-\frac{a''}{a'}Y \geq 1$ and $-\frac{d''}{d'}Y \geq 1$ for all Y in the policy space. There are two possibilities. If at the equilibrium there are enough immigrant workers, then the segmentation of the labour market is irrelevant and the results are identical to the baseline model. The perfect substitutability in production and the perfect competition ensure that all prices are are unaffected by immigration choices. The implications change dramatically if in the proximity of an equilibrium there are not enough immigrant workers to satisfy the demand at the constant price. In detail, the total demand of services for the elderly is given by:

$$O^{TD}_t = \frac{\bar{p}_{t-1}l_{t-1}n_t}{P^o_t} = \frac{(\alpha+\gamma)l_{t-1}n_t}{P^o_t} \tag{4.62}$$

Suppose that all the immigrants endogenously select themselves into the sector that produces O_t (this is the case if wages are higher in this sector), then the total supply is given by: $O_t^{TS} = \xi m_t \bar{\varepsilon}_t$, and the equilibrium price of the elderly good P_t^O is $P_t^O = \frac{(\alpha+\gamma)l_{t-1}n_t}{\xi m_t \bar{\varepsilon}_t}$. The zero profit condition implies that the total revenue in the elderly good sector must be equal to the total cost, thus one gets a different wage w_t^O in this sector, namely $w_t^O = \frac{(\alpha+\gamma)l_{t-1}n_t}{m_t \bar{\varepsilon}_t}$, such that the total nominal income of the workers in the elderly good sector is $w_t^O \bar{\varepsilon}_t m_t = (\alpha+\gamma)l_{t-1}n_t$. Notice that the perfect substitutability in production between the consumption good and the imperfect public good, together with the zero profit condition still imply $P_t^C = P_t^Y = P_t$ (else only one of the two would be produced and the result would still hold). Hence, in order to solves for the wage of the native workers, we can use the total demand of consumption and imperfect public good. Using the budget constraint one can show that $(C_t + Y_t)^{TD} = \frac{w_t^C \bar{\varepsilon}_t n_t}{P_t}$. Because the total supply is $\xi n_t \bar{\varepsilon}_t$ one can solve for the price $P_t = w_t^C/\xi$. The zero profit condition for the production of the consumption good holds for all prices P_t, namely $P_t \xi n_t \bar{\varepsilon}_t - w_t^C \bar{\varepsilon}_t n_t = 0$. Hence I can normalize $P_t = 1$ (this means that good C is the *numéraire*) and I get the wage $w_t^C = \xi$. A competitive equilibrium of this kind exists only if $w_t^O \geq w_t^C$. In this problem this condition is equivalent to: $\bar{p}_t \left(\frac{1-M_t}{M_t}\right) \geq \xi_t \bar{\varepsilon}_t$. Notice that as long as positive pensions are paid, one can always find M_t small enough that such inequality is satisfied. I can now state the formulas for the consumption of young and old individuals.

$$C_t^{i,y} = (1-\tau_t)\xi \varepsilon_t^i \tag{4.63}$$

and

$$C_t^{i,o} = \frac{[\alpha + \gamma(\varepsilon_{t-1}^i/\bar{\varepsilon}_{t-1})]\bar{\sigma}_{t-1}\xi}{(\alpha+\gamma)l_{t-1}} \frac{M_t}{1-M_t} \tag{4.64}$$

Finally notice that the government budget constraint is now different because the immigrants have different wages relative to the natives. In order to keep the problem tractable it is useful to define a new variable $\widetilde{Y}_t = \frac{Y_t}{(1-M_t)}$. In detail:

$$\tau_t = \frac{\widetilde{Y}_t}{[\xi \bar{\varepsilon}_t + (\alpha+\gamma)l_{t-1}]} + \frac{(\alpha+\gamma)l_{t-1}}{\xi \bar{\varepsilon}_t + (\alpha+\gamma)l_{t-1}} \tag{4.65}$$

The objective function of a young individual becomes:

$$\begin{aligned} V_t^{i,y} = \quad & (1-\tau_t)\xi \varepsilon_t^i + a[\widetilde{Y}_t(1-M_t)] - c(M_t) + \\ & + \beta l_t \left\{ \frac{[\alpha+\gamma(\theta_t^{i,y})]\bar{\sigma}_t \xi}{(\alpha+\gamma)} \frac{M_{t+1}^{**}}{1-M_{t+1}^{**}} + d[Y_{t+1}^{**}] - c(M_{t+1}^{**}) \right\} \end{aligned} \tag{4.66}$$

where $\theta_t^{i,y} = y_t^i/\bar{y}_t$. Notice that the assumption of $\sigma_t^m - \sigma_t^n$ arbitrarily small implies that M_{t+1}^{**} is unaffected by current policy choices. Thus, the first derivatives are:

$$V_{M_t}^{i,y} = -a'[\widetilde{Y}_t(1-M_t)]\widetilde{Y}_t - c'(M_t) < 0 \tag{4.67}$$

$$V_{\widetilde{Y}_t}^{i,y} = a'[\widetilde{Y}_t(1-M_t)](1-M_t) - \frac{\xi\theta_t^{i,y}}{\xi + (\alpha+\gamma)l_{t-1}/\bar{\varepsilon}_t} \tag{4.68}$$

Regarding the elderly, they have an objective function in the form:

$$V_t^{i,o} = \frac{[\alpha + \gamma(\varepsilon_{t-1}^i/\bar{\varepsilon}_{t-1})]\bar{\sigma}_{t-1}\xi}{(\alpha+\gamma)l_{t-1}}\frac{M_t}{1-M_t} - c(M_t) + d[\widetilde{Y}_t(1-M_t)] \tag{4.69}$$

Notice that

$$V_{M_t}^{i,o} = \frac{[\alpha + \gamma(\varepsilon_{t-1}^i/\bar{\varepsilon}_{t-1})]\bar{\sigma}_{t-1}\xi}{(\alpha+\gamma)l_{t-1}(1-M_t)^2} - \hat{c}'(M_t) - d'[\widetilde{Y}_t(1-M_t)]\widetilde{Y}_t \tag{4.70}$$

and

$$V_{\widetilde{Y}_t}^{i,o} = d'[\widetilde{Y}_t(1-M_t)](1-M_t) > 0 \tag{4.71}$$

One can notice that in this case the young individuals are more hostile to immigration and to public spending than the elderly. Using the same method presented in the paper, one can define a common objective function $V_t^i = V(x_t;\theta_t^i,\varphi_t,g_t)$ by setting $\theta_t^i = \theta_t^{i,y} = \varepsilon_t^i/\bar{\varepsilon}_t$ for the young individuals and $\theta_t^i = -\varepsilon_{t-1}^i/\bar{\varepsilon}_{t-1}$ for the elderly. Moreover I apply the increasing transformation $V_t^i = (1+\theta_t^i)V_t^{i,y}$ for all young individuals and $(\kappa - \theta_t^i)V_t^{i,o}$ with κ arbitrarily large (these transformation do not affect the preferences). Define $z_t = (-M_t, -\widetilde{Y}_t)$. I can show the following results.

Lemma 23. *(i) If l_{t-1} is small enough, the function V_t^i satisfies SM and SID in $(z_t;\theta_t^i)$. Therefore (ii) a coalitional equilibrium exists.*

Proof. (i) It is easy to show that $V_{M_t\theta_t^i}^{i,o} < 0$ and $V_{\widetilde{Y}_t\theta_t^i}^{i,y} > 0$ for all M_t, Y_t, θ_t^i. Because $V_{M_t}^{i,y} < 0$ and $V_{\widetilde{Y}_t}^{i,o} > 0$ for all M_t, Y_t, θ_t^i, then the *SID* is satisfied within the young and within the elderly respectively. Lastly, one need to show that $V(z_t';\theta_t^i,\varphi_t,g_t) - V(z_t'';\theta_t^i,\varphi_t,g_t) > V(z_t';\theta_t^j,\varphi_t,g_t) - V(z_t'';\theta_t^j,\varphi_t,g_t)$ for all $z_t' \geq z_t''$ and $z_t' \neq z_t''$and whenever i is a young individual and j is an elderly. Notice that for l_{t-1} arbitrarily small $V_{M_t}^{i,o} > 0$ for all M_t, Y_t, θ_t^i. Hence $V(z_t';\theta_t^j,\varphi_t,g_t) - V(z_t'';\theta_t^j,\varphi_t,g_t)$ is strictly negative and because κ is large enough, the condition is satisified for all M_t, Y_t, θ_t^i. (ii) Straightforward from Theorem 3.

Theorem 20. *If $g_t \leq 1$ then at the equilibrium, if it exists, the immigration policy is $M_t = 0$, else a positive level of immigration is possible.*

Proof. If $g_t \leq 1$ and an equilibrium exists, then the pivotal voter is a young individual with $V^{i,y}_{M_t} < 0$. Hence her ideal policy is $M_t = 0$.

Further details and additional results for this extension are provided in the supplementary online material.

4.8.2.3 Welfare Analysis: Immigration Policy

Theorem 21. *For any Social Welfare Function $SWF(x_t;\varphi_t,g_t)$ that assigns a strictly positive weight to each native individual of working age, there exist a threshold $\breve{g}_t \in [0,1]$ such that if $g_t \geq \breve{g}_t$ then a marginal tightening in the immigration policy caused by a change in the equilibrium outcome reduces the Social Welfare.*

Proof. Notice that the theorem above is stated for the baseline model without endogenous education. Here I show the proof for the full model with *SWF* denoted by $SWF(x_t;\varphi_t,g_t) = \widetilde{SWF}(M_t,Y_t,e_t;\varphi_t,g_t)$ for $x_t = (M_t,-Y_t,-e_t)$. The proof of the baseline model is straightforward. Define the overall weight of each generation as follows:

$$\int_0^{\bar{\theta}_t} \mu_t^y(\theta_t^i) q_t(\theta_t^i) d\theta_t{}^i = \mu^y \tag{4.72}$$

$$\int_0^{\bar{\theta}_{t-1}} \mu_t^o(\theta_{t-1}^i) q_{t-1}(\theta_{t-1}^i) d\theta_{t-1}{}^i = \mu^o \tag{4.73}$$

$$\int_0^{\bar{\theta}_{t+1}} \mu_{t+1}^y(\theta_{t+1}^i) q_{t+1}(\theta_{t+1}^i) d\theta_{t+1}{}^i = \mu^c \tag{4.74}$$

Normalize $\mu^y = 1$ and assume $\mu^y + \mu^o + \mu^c = \mu$ with $0 < \mu < \infty$. This can be done without loss of generality under the assumption that $\mu_t^y(\theta_t^i) > 0$ for each native individual of working age. Suppose the equilibrium policy x_t^* is such that $\underline{M}_t < M_t < \overline{M}_t$, which implies that a marginal opening in the immigration policy is feasible. If the difference between the marginal social benefit for the society from an increase in M_t and the marginal utility of M_t for the pivotal voter evaluated at the equilibrium policy vector is strictly positive, i.e.

$$\widetilde{WD}_{M_t}(M_t^*,Y_t^*,e_t^*;\varphi_t,g_t) = \widetilde{SWF}_{M_t}(M_t^*,Y_t^*,e_t^*;\varphi_t,g_t) - V_{M_t}^{v,y}(M_t^*,Y_t^*,e_t^*;\varphi_t,g_t) > 0 \tag{4.75}$$

then a marginal increase in the openness of the immigration policy M_t is, ceteris paribus, beneficial for the society. Notice that if $\underline{M}_t < M_t < \overline{M}_t$, then $V_{M_t}^{v,y}(M_t^*,Y_t^*,e_t^*;\theta_t^v,\varphi_t,g_t)=0$ from the F.O.C. The social benefit for the society from an increase in M_t is given by:

$$\begin{aligned}\widetilde{SWF}_{M_t} = & \int_0^{\bar{\theta}_t}\mu_t^y(\theta_t^i)\tilde{V}_{M_t}^y(M_t^*,Y_t^*,e_t^*;\theta_t^i,\varphi_t,g_t)q_t(\theta_t^i)d\theta_t^i + \\ & \int_0^{\bar{\theta}_{t-1}}\mu_t^o(\theta_{t-1}^i)\tilde{V}_{M_t}^o(e_t,M_t,Y_t;\theta_t^i,\varphi_t,g_t)q_{t-1}(\theta_{t-1}^i)d\theta_{t-1}^i + \\ & \int_0^{\bar{\theta}_{t+1}}\mu_{t+1}^y(\theta_t^i)E[\tilde{V}_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1};\theta_{t+1}^i,\varphi_{t+1},g_t)]q_{t+1}(\theta_{t+1}^i)d\theta_{t+1}^i\end{aligned} \tag{4.76}$$

First of all notice that the linearity in consumption of the utility function implies $E[V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1},;\theta_{t+1}^i,\varphi_{t+1},g_{t+1})] = V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1};\bar{\theta}_{t+1},\varphi_{t+1},g_{t+1})]$ hence $\int_0^{\bar{\theta}_{t+1}}\mu_{t+1}^y(\theta_{t+1}^i)E[V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1};\theta_{t+1}^i,\varphi_{t+1},g_{t+1})]q_{t+1}(\theta_{t+1}^i)d\theta_{t+1}^i =$ $E[\mu_{t+1}^y(\theta_{t+1}^i)]V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1};\bar{\theta}_{t+1},\varphi_{t+1},g_{t+1})$. Moreover, notice that a change in x_t only affects the future generation through a fall in g_{t+1}, which has no effects neither on the budget constraint at time $t+1$ nor on the preferences of an individual (it only affects the political equilibrium at time $t+1$). Therefore $V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1},;\theta_{t+1},\varphi_{t+1},g_{t+1})$ is independent of M_t and therefore *SID* implies: $V_{M_t}^y(M_{t+1},Y_t,e_{t+1};\bar{\theta}_{t+1},\varphi_{t+1},g_{t+1}) \geq$ $V_{M_t}^y(M_{t+1},Y_{t+1},e_{t+1};\theta_{t+1}^v,\varphi_{t+1},g_{t+1})$ as long as $\theta_{t+1}^v \leq \bar{\theta}_{t+1}$. I use the latter result and I substitute the formulas for $V_{M_t}^{i,y}$, $V_{M_t}^{i,o}$ into $\widetilde{WD}_{M_t}$, and I can write the following inequality:

$$\begin{aligned}\widetilde{WD}_{M_t} \geq & \left[\tfrac{(\alpha+\gamma)l_{t-1}}{(1-M_t^*)} - \lambda_t + \left(\tfrac{\beta l_t\gamma}{(1-M_{t+1}^*)} - e_t\right)(\sigma_t^m - \sigma_t^n)\right]\left(\int_0^{\bar{\theta}_t}\theta_t^i\mu_t^y(\theta_t^i)q_t(\theta^i)d\theta_t{}^i - \theta_t^v\right) + \\ & -c'(M_t)\int_0^{\bar{\theta}_{t-1}}\mu_t^o(\theta_{t-1}^i)g(\theta_{t-1}^i)d\theta_{t-1}^i\end{aligned} \tag{4.77}$$

Notice that:

$$\begin{aligned}V_{M_t}^{v,y} = & -c'(M_t) + \theta_t^v(\alpha+\gamma)\tfrac{l_{t-1}}{(1-\widehat{M}_t)} - \theta_t^v\lambda_t + \\ & + \left(\tfrac{\beta l_t}{(1-\widehat{M}_{t+1})}(\alpha/\theta_t^v + \gamma) - e_t\right)(\sigma_t^m - \sigma_t^n)\theta_t^v\end{aligned} \tag{4.78}$$

also represent the FOC of the optimization problem of the pivotal individual. This implies that if at the equilibrium $\underline{M}_t < M_t$ then:

$$\frac{(\alpha+\gamma)l_{t-1}}{(1-\widehat{M}_t)} - \lambda_t + \left(\frac{\beta l_t \gamma}{(1-\widehat{M}_{t+1})} - e_t\right)(\sigma_t^m - \sigma_t^n) \geq$$
$$\geq \frac{1}{\theta_t^v}\left(c'(M_t) - \frac{\alpha\beta l_t(\sigma_t^m - \sigma_t^n)}{(1-\widehat{M}_{t+1})}\right) \tag{4.79}$$

Define the weighted average

$$E_{g_t}(\mu_t^y \theta^i) = \int_0^{\bar{\theta}_t} \lambda_t^y(\theta_t^i)\theta_t^i g_t(\theta_t^i) d\theta_t^i = h_{g_t}\int_0^{\bar{\theta}_t} \theta_t^i \dot{g}_t(\theta_t^i) d\theta_t^i = h_{g_t} E_{\dot{g}_t}(\theta_t) \tag{4.80}$$

for some p.d.f $\dot{g}_t$. Notice that $h_{g_t}E_{\dot{g}_t}(\theta_t) > 0$ under the assumption that $\mu_t^y(\theta_t^i) > 0$ for each native individual of working age. Therefore we can state the following inequality:

$$\widetilde{WD}_{M_t} \geq \left(c'(M_t) - \frac{\alpha\beta l_t(\sigma_t^m - \sigma_t^n)}{(1-\widehat{M}_{t+1})}\right)\frac{h_{g_t}E_{\dot{g}_t}(\theta_t) - \theta_t^v}{\theta_t^v} - c'(M_t)\mu^o \tag{4.81}$$

The F.O.C.s of the pivotal individual plus the assumption that immigrants are not net beneficiaries (in expectation) of the fiscal system imply $c'(M_t) - \frac{\alpha\beta l_t(\sigma_t^m - \sigma_t^n)}{(1-\widehat{M}_{t+1})} > 0$ for $\underline{M}_t < M_t < \overline{M}_t$. Finally notice that because of a previous assumption $c'(M_t) < \infty$ and that $\mu^o < 0$ imply:

$$\lim_{\theta_t^v \to 0^+}\left(c'(M_t) - \frac{\alpha\beta l_t(\sigma_t^m - \sigma_t^n)}{(1-\widehat{M}_{t+1})}\right)\frac{h_{g_t}E_{\breve{g}_t}(\theta) - \theta_t^v}{\theta_t^v} - c'(M_t)\mu^o = +\infty \tag{4.82}$$

Therefore, given a certain distribution of weights, either $\widetilde{WD}_{M_t}(M_t^*, Y_t^*, e_t^*; \varphi_t, g_t) > 0$ for all $\theta_t^v > 0$, else the Intermediate Value Theorem implies the existence of a threshold $0 < \check{\theta}_t < \bar{\theta}$ such that $\widetilde{WD}_{M_t}(M_t^*, Y_t^*, e_t^*; \varphi_t, g_t) = 0$. This threshold is always meaningful because I have previously assumed that the distribution of θ_t is such that $q(0) > 0$ and therefore $\theta_t^j = 0$. Moreover, $\widetilde{WD}_{M_t}(M_t^*, Y_t^*, e_t^*; \varphi_t, g_t)$ is strictly decreasing in θ_t^v because $\widetilde{WD}_{M_t}$ is independent of θ_t^v and $V_{M_t}^{v,y}$ is strictly decreasing in θ_t^v because of *SID*. Therefore if the wage distribution is such that $\theta_t^v < \check{\theta}_t$ then $\widetilde{WD}_{M_t}(M_t^*, Y_t^*, e_t^*; \varphi_t, g_t) > 0$ which implies that it would be welfare improving to increase M_t. Lastly, because of Lemma 6 (iii), a threshold $\breve{g}_t \in [0,1]$ exists, such that if $g_t \geq \breve{g}_t$ iff $\theta_t^v < \check{\theta}_t$, which implies the result stated. Q.E.D.

4.8.2.4 Alternative Assumption about the Default Policy: Status Quo

One may want to assume that the default platform is the policy implemented in the previous period (if feasible) In such case, $x_t^0 = x_{t-1}^*$. Following the same steps described in the proofs to Theorem 1 in chapter 2 one can show that there is no equilibrium in which a platform $x_t \in X_t$ such that $x_t \in M(v_t)$ is implemented. Neverthless, given that the default policy under this alternative assumption may not be the least preferred option for some players, then there may additional possible outcomes. Specifically, there may be (i) equilibria in which no coalition is active and the default policy is implemented and (ii) situations of instability, in which some coalitions are active only in order to prevent the victory of some other candidate. This may be possible because of the assumption that, if no Condorcet Winner exists in the final stage of the voting game, then x_t^0 is implemented. Suppose that is the case. The characterizations of all the equilibria given in Theorem 3 is no longer valid. Neverthless, the compartative statics results in Theorem 4-5 still apply for those equilibria in which a platform other than x^0 is chosen. Thus, the main results are still valid, with the possible exception of those cases in which the comparative statics exercise induces a change from an equilibrium in which x_t^0 is implemented to an equilibrium in which a platform different from the default one is chosen (or the opposite).

4.8.2.5 Maximum Tax Rate

In section 3 we have restricted the policy space in such a way that for all $x_t \in X$ the tax rate is internal $0 < \tau_t < k < 1$. Suppose that this assumption fails and at an equilibrium $\tau_t = k$. In this case it is not straightforward to derive results in the full model. Nevetheless, some results can be obtained in the baseline model with $x_t = (M_t, -Y_t)$ under the assumption that $d'(Y_t) \leq b'(Y_t)$ for all $Y_t \in [0, \overline{Y}]$ and $\hat{c}'(M_t) \geq c'(M_t)$ for all $M_t \in [0, \overline{M}]$. If $\tau_t = k$ the policy space is unidimensional, thus the traditional Median Voter Theorem applies if voter preferences satisfy the *Spence-Mirrlees* condition. Consider the slope of the indifference curve of an working age individual i:

$$MRS_{M_t,Y_t}^{i,y} = -\frac{1}{b'(Y_t)} \left[\frac{\beta l_t(\sigma_t^m - \sigma_t^n)}{(1 - \widehat{M}_{t+1})} (\alpha/\theta_t^i + \gamma))\theta_t^i - c'(M_t) \right] \tag{4.83}$$

and its derivative with respect to θ_t^i:

$$\frac{\partial MRS_{M_t,Y_t}^{i,y}}{\partial \theta_t^i} = -\frac{1}{b'(Y_t)} \frac{\beta l_t(\sigma_t^m - \sigma_t^n)\gamma}{(1 - \widehat{M}_{t+1})} \leq 0 \tag{4.84}$$

Moreover, notice that the MRS of any retired individual is given by $MRS_{M_t,Y_t}^{o} = \frac{\hat{c}'(M_t)}{d'(Y_t)}$, which implies that $MRS_{M_t,Y_t}^{o} \geq MRS_{M_t,Y_t}^{i,y}$ for all i. Thus, preferences satisfy the *Spence-Mirrlees* condition, and standard results can be applied to make predictions about the effects of changes in the pivotal voter on the equilibrium outcome. The results differ from the ones of most Benefit Adjustment Models. Specifically, an increase in the relative share of the elderly implies, *ceteris paribus*, an fall in public spending and a reduction of the immigration quota. In this framework I cannot derive analytical results about the effects of a rise in life expectancy, because this kind of shock typically involves not only a change in the pivotal voter but also in the position and slope of the budget constraint, such that the sign of the overall effect cannot be determined using the *Spence-Mirrlees* condition only.

Tables

Table 4.1. Determinants of Attitudes Towards Immigration

Ordered Logit - Dependent Variable: LETIN

VARIABLES	(1) Robust SE	(2) Cluster SE	(3) Robust SE	(4) Cluster SE	(5) Cluster SE
Age	0.00949***	0.00949***	0.0117***	0.0117***	
	(0.00196)	(0.00202)	(0.00269)	(0.00310)	
Income	-0.0258**	-0.0258***	-0.0235*	-0.0235***	-0.0241***
	(0.0114)	(0.00917)	(0.0120)	(0.00655)	(0.00884)
LowEdu	0.248***	0.248***	0.238***	0.238***	0.257***
	(0.0177)	(0.0206)	(0.0185)	(0.0220)	(0.0195)
BornAbroad	-0.790***	-0.790***	-0.768***	-0.768***	-0.806***
	(0.154)	(0.162)	(0.157)	(0.169)	(0.164)
Children	-0.267***	-0.267***	-0.247***	-0.247***	-0.225***
	(0.0650)	(0.0822)	(0.0668)	(0.0888)	(0.0785)
Rural	0.172***	0.172***	0.167***	0.167***	0.182***
	(0.0323)	(0.0624)	(0.0324)	(0.0620)	(0.0615)
Sex			0.0778	0.0778	
			(0.0625)	(0.0744)	
Unemployed			-0.0793	-0.0793	
			(0.143)	(0.110)	
ManualWork			0.139*	0.139	
			(0.0720)	(0.0893)	
Retired			-0.0813	-0.0813	
			(0.112)	(0.0943)	
Religion			0.0779	0.0779	
			(0.0635)	(0.0516)	
year	-0.0635**	-0.0635**	-0.0615**	-0.0615**	-0.0668**
	(0.0298)	(0.0314)	(0.0298)	(0.0308)	(0.0323)
Age65					0.209**
					(0.0857)
Observations	4,421	4,421	4,421	4,421	4,421

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

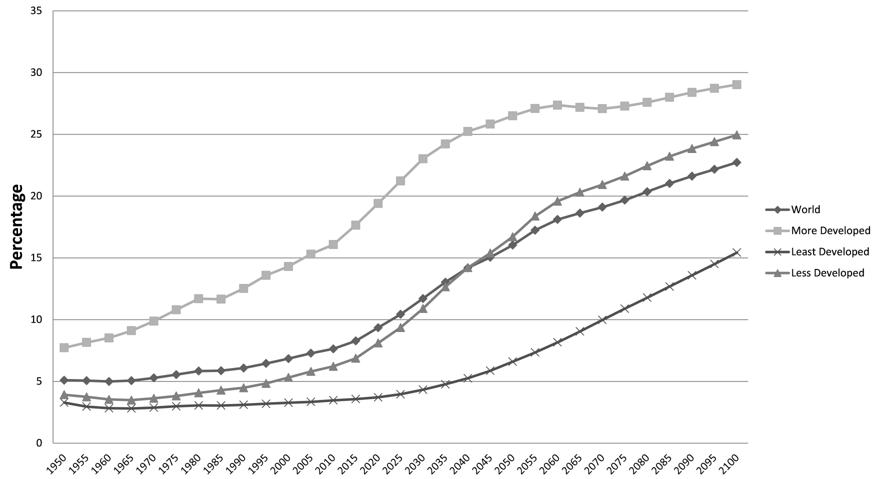
Table 4.2. Determinants of Attitude towards Public Spending

Ordered Logit - Dependent Variable: TaxSpend					
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Robust SE	Cluster SE	Robust SE	Cluster SE	Cluster SE
Age	0.0102***	0.0102***	0.0124***	0.0124***	
	(0.00161)	(0.00178)	(0.00270)	(0.00282)	
Income	-0.0320***	-0.0320***	-0.0422***	-0.0422***	-0.0385***
	(0.00983)	(0.00996)	(0.0126)	(0.0123)	(0.0129)
LowEdu	-0.0451***	-0.0451***	-0.0669***	-0.0669***	-0.0473***
	(0.0134)	(0.00997)	(0.0175)	(0.0145)	(0.0150)
BornAbroad			-0.0424	-0.0424	-0.0760
			(0.111)	(0.0702)	(0.0733)
Children			-0.0765	-0.0765	-0.0262
			(0.0677)	(0.0828)	(0.0816)
Rural			-0.0103	-0.0103	0.00376
			(0.0323)	(0.0310)	(0.0281)
Sex	0.180***	0.180***	0.188***	0.188***	0.161***
	(0.0504)	(0.0476)	(0.0637)	(0.0492)	(0.0490)
Unemployed	0.251**	0.251**	0.202	0.202	0.144
	(0.116)	(0.125)	(0.142)	(0.138)	(0.127)
ManualWork			0.202***	0.202***	0.185***
			(0.0691)	(0.0506)	(0.0483)
Retired			-0.173	-0.173*	0.230**
			(0.106)	(0.0946)	(0.117)
Religion	0.218***	0.218***	0.235***	0.235***	0.167***
	(0.0517)	(0.0486)	(0.0645)	(0.0598)	(0.0594)
year	0.00639	0.00639	0.00614	0.00614	0.000110
	(0.0298)	(0.0314)	(0.0300)	(0.0310)	(0.0310)
Age65					-0.136
					(0.112)
Observations	6,639	6,639	4,421	4,421	4,421

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

Figures

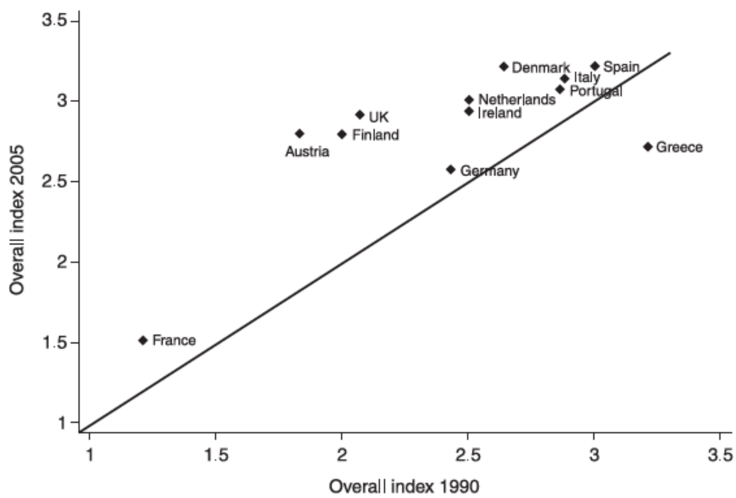
Figure 4.1. Share of Population of Age 65 or Older



Source: United Nations, 2015

Evolution of the share of population of age above 65 from 1950 to 2015 and the forecast for the next decades (source: United Nations, 2015).

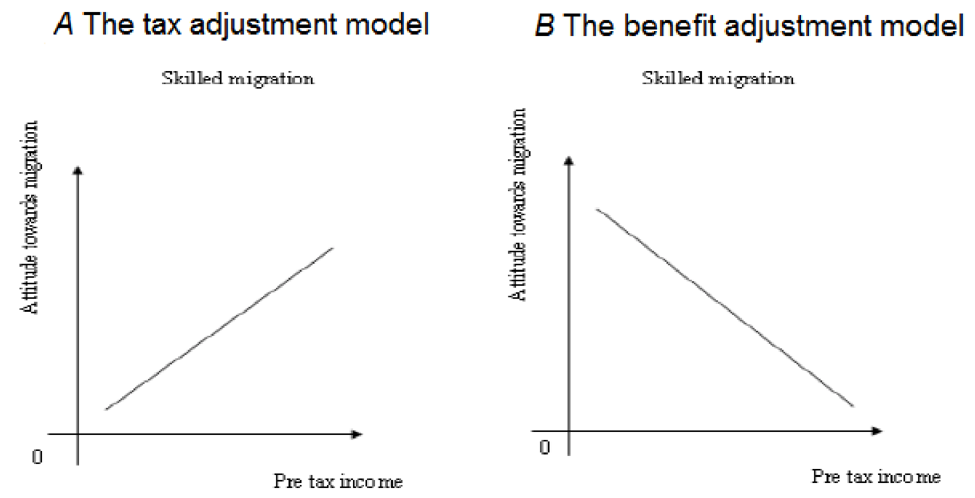
Figure 4.2. Trends in Migration Policies



Source: Boeri and Brucker, 2005

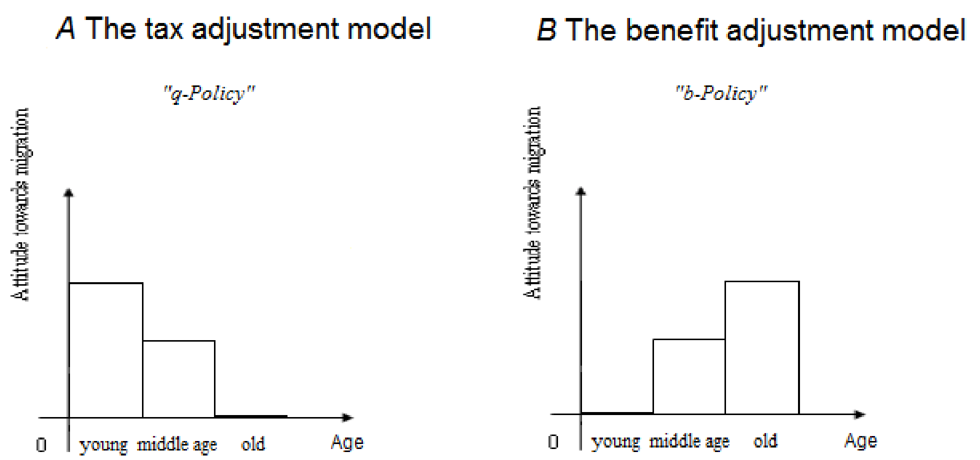
Comparison of the value of the index of tightness of immigration policies proposed by Boeri and Brucker (2005) in 1990 and 2005 for 12 European countries.

Figure 4.3 Effects of income on the attitudes towards immigration



Relationship between income and attitude towards immigration (preferred number of immigrants) in a Tax Adjustment Model (A) and in a Benefit Adjustment Model (B). Based on Facchini and Mayda (2008).

Figure 4.4. Effects of age on the attitudes towards immigration



Attitude towards immigration (preferred number of immigrants) of different generations of voters in a Tax Adjustment Model (A) and in a Benefit Adjustment Model (B). Based on Haupt and Peters (1998).

Figure 4.5. Structure of Overlapping Generations

time	$t-1$			t			$t+1$	
born $t-3$	OLD	(o)	$\rightarrow$	$\times$				
born $t-2$	NATIVE (n) Immigrant (m)	(y)	$\rightarrow$	OLD	(o)	$\rightarrow$	$\times$	
born $t-1$	Children	(ch)	$\rightarrow$	NATIVE (n) Immigrant (m)	(y)	$\rightarrow$	OLD	(o)
born t				Children	(ch)	$\rightarrow$	NATIVE (n) Immigrant (m)	(y)
born $t+1$							Children	(ch)

The categories of individuals that possess voting rights are highlighted in capital letters.

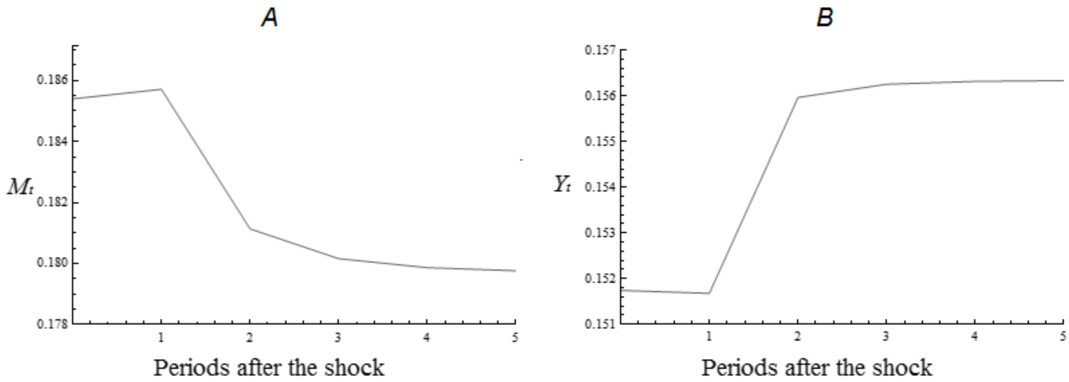
Figure 4.6. Size of each generation

$t-1$			t			$t+1$	
$l_{t-2}(n_{t-2}+m_{t-2})$	(o)	$\rightarrow$	$\times$				
$n_{t-1}+m_{t-1}$	(y)	$\rightarrow$	$l_{t-1}(n_{t-1}+m_{t-1})$	(o)	$\rightarrow$	$\times$	
$\sigma^n_{t-1}n_{t-1}+\sigma^m_{t+1}m_{t-1}$	(ch)	$\rightarrow$	n_t+m_t	(y)	$\rightarrow$	$l_t(n_t+m_t)$	(o)
		born	$\sigma^n_t n_t+\sigma^m_t m_t$	(ch)	$\rightarrow$	$n_{t+1}+m_{t+1}$	(y)
					born	$\sigma^n_{t+1}n_{t+1}+\sigma^m_{t+1}m_{t+1}$	(ch)

The arrow denotes the transition of a group of individuals into the next period

Figure 4.7. Long-Run Effects of an Increase in Life Expectancy

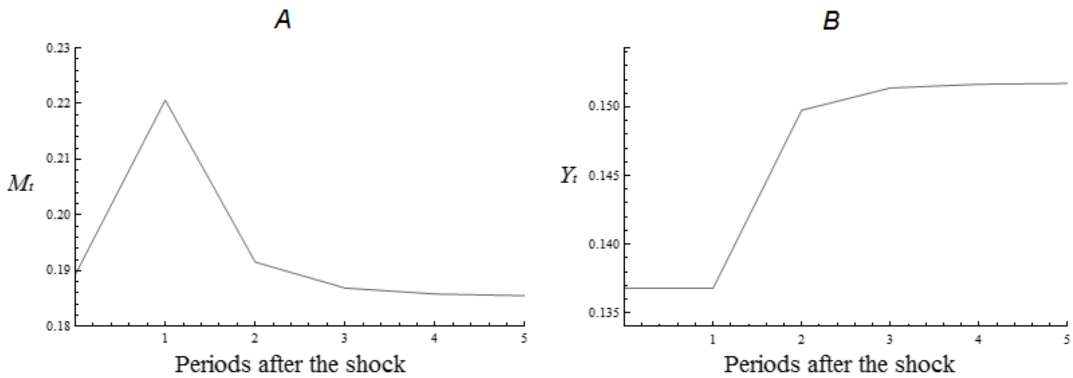
Parameters: $\sigma^n = 1$, $\sigma^m = 1.5$, before shock $l = 0.6$, after shock: $l = 0.62$.



Effects of a positive shock on life expectancy on the immigration quota M_t (A) and on public spending per worker Y_t (B).

Figure 4.8 Long-Run Effects of a Decrease in the Birth Rate of the Natives

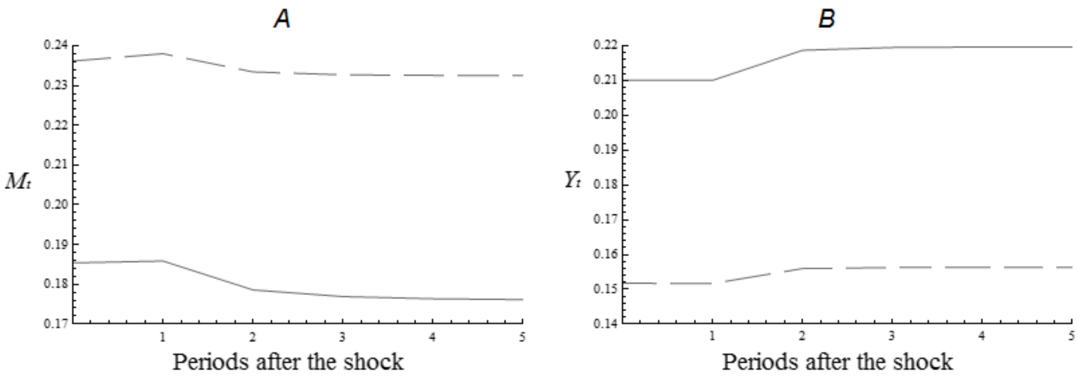
Parameters: $\sigma^n = 1.2$, $\sigma^m = 1.5, l = 0.6$, after shock: $\sigma^n = 1$.



Effects of a negative shock on the birth rate of the native population on the immigration quota M_t (A) and on public spending per worker Y_t (B).

Figure 4.9. "Naive" vs. "Sophisticated" agents

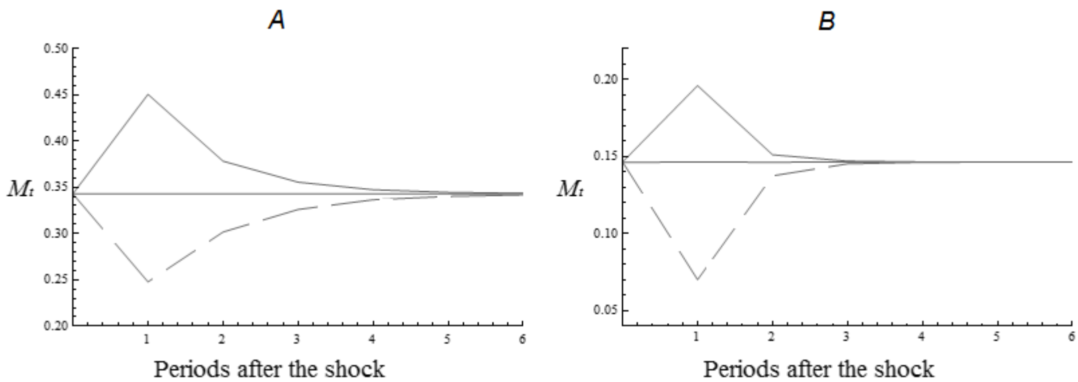
Parameters: $\sigma^n = 1$, $\sigma^m = 1.5$, before shock $l = 0.6$, after shock: $l = 0.62$.



Effects of a positive shock on life expectancy on the immigration quota M_t (A) and on public spending per worker Y_t (B) for "naive" (dashed line) and "sophisticated" voters (solid line).

Figure 4.10. Convergence to the Steady-State

Parameters: $\sigma^n = 1$, $l = 0.6$.



Effects of a temporary negative shock on g_t (solid lines) and of a temporary negative shock on g_t (dashed line) for $\sigma^m = 1.5$ (A) and $\sigma^m = 2$ (B).

5 The Political Economy of Public Education

I study the relationship between income inequality and public intervention in education in a probabilistic voting model. Traditional Political Economy models typically imply a positive relationship between income inequality and public intervention in redistributive policies. Empirical evidence suggests that this may hold true only for certain kinds of policies, such as public education, but it may not hold true for other forms of public intervention. I propose a method to study the sign of this this relationship in the case in which forms of redistribution other than the public provision of education are available to voters. Moreover, I allow consumers to opt-out of the public education system and get private education. This feature of the public provision of education plays a crucial role in shaping the results. I show that an increase in income inequality causes a rise in governmental intervention in education if the expected marginal returns to education are larger for children of low income parents. This finding is is consistent with the results in the empirical literature about public investment in education. Moreover, I show that the policy adjustment tends to reduce future inequality. Lastly, I show that for other kind of publicly provided goods, such as Health care, the relationship has ambiguous or opposite sign.

JEL classification: D72, H42, I21, I22.

Keywords: Probabilistic Voting, Education, Inequality.

5.1 Introduction

What is the effect of an exogenous increase in income inequality on the level of public intervention in public education in a democratic country? Does such effect mitigate income inequality of future generations? This paper attempts to provide a theoretical framework to answer these questions. The relationship between the degree of governmental intervention in the provision of good and services and the features of the population in democratic political systems has been a major topic of research in Political Economy. Traditional models typically imply a positive relationship between the size of the intervention and income inequality (Meltzer and Richards, 1981). The reason is that such policies tend to have redistributive effects[1], thus an increase in the public provision favors the relatively low income part of the voting population. This has important consequence in a voting model because of two factors. First, an increase in income inequality is associated with an increase in the share and the political power of the relatively low income voters. Second, traditional models do not allow voters to access other redistributive policies such as lump-sum grants because of technical constraints. Thus, unsurprisingly, such models usually imply a positive relationship between income inequality and the size of any kind of policy with redistributive effects. Empirical evidence suggests that this relationship may hold true only for certain kinds of policies, for instance public education, but it may not hold true for other policies with redistributive effects such as social security and public health. In this paper I attempt to disentangle voters' preferences for redistribution from their demand for public education by allowing them to choose both the size of in-cash redistribution - through a flexible tax system - and the quality of public education. This implies that the policy space is multidimensional. In this setting, the specific features of the public provision of education play an important role in determining the relationship between the size of public intervention and the degree of income inequality. Specifically, the presence of private alternatives to public education and the possibility of *opting-out* of the public sector are crucial in shaping the results. Because of these reasons, both a multidimensional policy space and the possibility of *opting-out* are essential features of this analysis. Unfortunately, both such modeling choices are source of well-know problems of existence of a voting equilibrium in the traditional deterministic Downsian framework. Thus, I adopt a Probabilistic Voting framework that allows one to tackle both issues, and I use it to study voters' behavior in a model of parental investment in education. I find that public intervention in education may be affected by income inequality not because of its redistributive effects, but because of the peculiar way in which the provision is delivered. I derive analytical conditions for a positive relationship between income inequality and quality of the publicly provided education. I find that the sign of this relationship is positive if the expected marginal returns to public education are decreasing in parental income. This is consistent with recent empirical evidence, and can be due to credit constraints that induce relative low income parents to underinvest in their children. Moreover, I show that if this condition is met, then an increase in the quality of public education reduces income inequality in the next generation.

The paper is structured as follows. In section 2 I describe the findings of the empirical literature about the relationship between public provision of education and income inequality and how the theoretical literature has tackled this question. In section 3 I present the voting model and the methodology I propose to study the sign of the relationship between the equilibrium level of public provision of a good of interest and

[1]Such effects are typical consequences of in-kind policies and are achieved even if no income redistribution occurs. For a definition see Appendix 5.7.1.

the degree of income inequality in the population of voters. In section 4 I apply these results to a model of parental investment in education in order to provide an answer to the main question of the paper. In section 5 I compare the predictions in section 4 with the one that the same framework would deliver for other kinds of publicly provided goods such as pure public goods and health insurance. Section 6 concludes highlighting the achievements and the limitations of this analysis.

5.2 Facts and Literature

There is a large empirical literature about the relationship between income inequality and governmental spending in redistributive policies (see de Mello, Tiongson 2006 for a review of this literature). On one hand the traditional theoretical literature typically predicts a positive relationship between income inequality and size of redistribution (Meltzer and Richard, 1981). On the other hand, empirical evidence provides mixed results. Perotti (1996) finds no relationship between inequality and redistribution in democracies. Using data from the U.S. General Social Survey, Lind (2007) finds that inequality between different groups reduces redistribution, while within group inequality increases it. A number of papers have found that support for redistribution and public goods provision is weaker in more unequal or more heterogeneous societies (Goldin and Katz 1997, Alesina *et al.* 1999, 2001, Luttmer 2001). A more recent paper by Boustan *et al.* (2010) finds that rising inequality in cities and districts is associated with higher local revenue collection and expenditures. The question becomes even more challenging if one in interested in modeling the degree of public intervention in a specific policy with redistributive effects, such as public education. The literature about the relationship between income inequality and public spending in education is limited and provides mixed evidence. A majority of empirical studies find evidence of a positive correlation between income inequality and public intervention in schooling in cross-sectional studies about US states. Easterly and Rebelo (1993) using cross sectional country data show that high level income inequality tend to be associated with future high level of public spending in education in the period 1970-1988. Sylwester (2000) also finds a weak but significant positive correlation between income inequality and future public spending in education, even if the issue of reverse causality in the relationship is not completely addressed in his paper. Conversely, Corcoran and Evans (2010) using a panel of U.S. school districts spanning 1970-2000 find a negative relationship between inequality and local spending in public education. Figure 5.1 shows a small positive correlation between the pre-tax Gini index of income inequality in 2013 and the public expenditure per capita in education in the 50 U.S. States (American Community Survey, 2013). The observed correlation is weak and may be due to several sources of endogeneity. In particular, one that has been suggested in the literature is reverse causality. On one hand Political Economy models typically imply an important role for income inequality in shaping the degree of public intervention in certain policies. On the other hand, part of the literature in public economics suggests that a uniform public investment in education may induce a fall in the degree of future income inequality (see Coons *et al.*, 1970 and Sylwester, 2002). Nevertheless, the theoretical foundations of a positive effect of uniform public education on future income inequality have been challenged by the theoretical literature (Glomm and Ravikumar, 2003), and the empirical evidence about this second channel is mixed (see Abdullah *et al.*, 2015 for a review of this literature).

The theoretical literature in Political Economy has relied mostly on the traditional Downsian framework (Downs, 1957) to answer this question. Such theoretical framework has many appealing features described in chapter 2 of this thesis. In particular, the *pivotal voter* result described in section 2.2 implies that comparative static results are easy to derive (see Appendix 5.7.1 for an example). The effect of a rise in income inequality on public investment in education in Downsian models typically depends on the features of voters' preferences and on how public intervention interacts with market choices. For instance, Fernández and Rogerson (1995) show that in a model in which education is partially subsidized, poorer individuals may be excluded from obtaining an education and that increased inequality in the income distribution makes this outcome more likely. Glomm (2004) adopts the Downsian framework and finds that the relationship between inequality and the amount of redistribution through public education services depends on the elasticity of substitution between consumption and the quality of education in the parents' utility. He argues that for empirically relevant value of this parameter, higher inequality generates less redistribution. Stiglitz (1974) has pointed out that the use of Downsian models to study this question may be prone to some relevant theoretical issues. Namely, he has shown that if consumers are allowed to *opt-out* of the public service in presence of private alternatives, then a a *Condorcet Winner* may fail to exist. In detail, the existence of a Condorcet Winner relies on the assumption that individual preferences satisfy some ordinal condition, such as *single peakedness*. Such condition often fails to apply if opting-out occurs. In the cases in which the Downsian framework is successful in characterizing a Political equilibrium (for instance in Ireland 1990, Epple and Romano 1996a, Gouveia 1997, Glomm and Ravikumar 1998, Naito and Nishida 2012), the opting-out assumptions is shown to play a crucial role in shaping the relationship between the shape of the income distribution and the equilibrium level of public spending in education (Epple and Romano, 1996b). Neverthless, this kind of analysis may deliver some paradoxical results. For instance, a change in income inequality typically has non-zero effect on the equilibrium level of public education even if, in absence of public intervention, all voters choose exactly the same level of education on the private market. The reason is that in order to achieve single peakedness such models assume a unidimensional policy space. This means that the degree of redistribution provided by the tax system is assumed to be exogenous. The uniform provision of a good financed by tax revenues has redistributive effects[2]. Thus, low-income individuals support a larger amount of public provision relative to the high-income simply because the model does not allow for other endogenous forms of redistribution. In other words, a relatively poor voter can only achieve redistribution through the provision of the good, thus she votes for larger level of provision relatively to a high income individual. Lastly, higher income inequality increases the political power of the less well-off, and this translates into larger governmental intervention in education. Another way to think about the same mechanism is to notice that that in pivotal voter models - abstracting from possible externalities - the collective demand for public provision of the good is equal to the private demand for the good of an individual characterized by an income level and by a specific marginal tax-price for the good[3]. If the identity of the pivotal voter changes, the marginal tax-price faced by the pivotal individual also changes. Thus, the sign of the overall effect depends on the relative size of income and price elasticities. This example suggests that in those models the relationship between income inequality and degree of public intervention in education is driven - at least to some extent - by the redistributive effects of the provision rather than by the specific features of the good. An intuitive way to tackle this problem is to include in the analysis at least another endogenous redistributive policy variable (for instance a uniform in-cash grant). This would allow one to disentangle the social demand for redistribution from the one for public intervention in education. Unfortunately, such modelling choice implies a second theoretical issue, that worsens the problems of existence of a *Condorcet Winner* induced by the *opting-out* assumption mentioned above. Namely, it determines an increase in the dimensionality of the policy space. In chapter 2 of this thesis I have described extensively the important issues that arise in the analysis of collective choices over a multidimensional policy space if the traditional Downsian framework is employed. Because of these issues, a vast majority of Political Economy papers that study voters' choices over public investment in education employ a unidimensional choice space. In chapter 2 I have also proposed a new theoretical framework that can tackle the problem of multidimensionality of the policy space if voter preferences possess specific ordinal properties. A model of electoral competition with similar features can be employed to answer questions regarding the Political Economy of public education. An example of this approach is in Levy (2005). She studies how democratic societies choose the level of public intervention in education if redistribution in-cash is also available to voters. She allows individuals to differ in their income and age and show that positive levels of provisions are possible in equilibrium.

[2]In the sense that low-income individuals pay a lower tax-price for the good relatively to its market price. The tax-price is defined as total taxes paid by the individual divided by the size of the public provision.

[3]Defined as the increase in taxes induced by a marginal increase in the provision.

Nevertheless, Levy's analysis cannot be extended to answer the specific question of this chapter, because she abstracts from the other aspect that - as previously mentioned - is deemed to be crucial in the traditional literature. Namely, she does not allow for *opting-out*. More generally, a theoretical approach like the one proposed in chapter 2 of this thesis would not prove useful to tackle such a question. The reason is that the ordinal conditions on voter preferences stated in section 2.3.4 are usually not satisfied in presence of non-convexities induced by the *opting-out* assumption. In order to allow both for *opting-out* and for a multidimensional policy space I employ a more traditional model of electoral competition, namely a Probabilistic Voting Model. This choice relies on three appealing features. First - as extensively described in section 2.5.5 of this thesis - the Probabilistic Voting framework delivers existence and uniqueness of an equilibrium under relatively mild restrictions even if the choice domain is multidimensional. Secondly, the probabilistic nature of voters' choice helps to smooth out the potential non-convexities in individual preferences induced by opting-out. Third, departing from a pivotal voter equilibrium in favor of a concept in which the equilibrium policy depends - in principle - on the entire distribution of voters' preferences, allows one to link the predictions of the model directly to some measure of income inequality, such as the variance of the income distribution. The latter aspect differs from traditional deterministic voting models, in which the feature of income distribution that is relevant for comparative statics is a measure of *skewness*, such as the mean-to-median ratio. The shortcoming of this approach is that - for the reasons described in section 2.5.5 of this thesis - analytical comparative statics exercises are not as straightforward to perform as in Downsian models, thus the results one can derive are limited. For instance, in two recent papers de la Croix and Doepke (2009) and Arcalean and Schiopu (2012) employ a probabilistic voting model to study the relationship between income inequality and public intervention in education. They assume a parametric specification of the income distribution and of consumer preferences and a unidimensional policy space. They find that higher inequality decreases public spending per student and increases enrollment in public schools in poor economies, while the opposite holds in the rich ones. In this paper I propose a more general analytical result about the relationship between the variance of the income distribution and the equilibrium level of public intervention in a publicly provided good. I do not impose strong parametric restriction on voters' direct utility function other than quasilinearity in consumption of a composite private good and additive separability in other goods. Moreover, I allow for a more general income distribution, namely income is the sum of a continuously distributed variable (with no parametric restrictions) and a uniform i.i.d. component. Details about the voting models are described in the next section.

5.3 Probabilistic Voting with Non-Convex Preferences

In this section I will present a relatively simple model of Probabilistic Voting that is substantially similar to the ones that are prevalent in the literature, such as the one proposed by Lindbeck and Weibull (1987), Enelow and Hinich (1989), Banks and Duggan (2004). The key feature of these models is that the vote of every individual (or type of individual) is not deterministic. This assumption eases dramatically the conditions for the existence of a Political Equilibrium when the policy space is multidimensional in comparison with Downsian models. The shortcoming is that the characterization of the equilibrium outcomes is not as simple as in Downsian models. In the next subsections I describe the setup of the voting model and I provide sufficient conditions for existence and uniqueness of a political equilibrium in the case in which the interaction between public and private provision of a good leads to possible non-convexities in voters' preferences. Then I derive the sign of comparative statics of interest in such environment.

5.3.1 Setup

The voting population consists of a continuum of size 1 of consumer-voters. They differ from each other only in a unidimensional parameter w that is continuously distributed with c.d.f. $\hat{R}(\theta, w)$ and p.d.f. $\hat{r}(\theta, w)$ for some parameter $\theta \in [0,1]$. A feasible policy is a N-dimensional vector $x \in X$ where $X \in R^N$ is a convex set such that $X := \{x \in R^n : B(x) \leq 0\}$ and $B(x) \leq 0$ is a constraint that ensures the feasibility of the policy. There are 2 parties: A and B. Before the election the two parties simultaneously choose a feasible policy x^A and x^B, respectively. Denote with $v(x,w)$ the indirect utility induced by policy x to an individual with parameter w. Following Banks and Duggan (2005), I define the expected vote share of type w voters for party A given policies x^A, x^B as follows:

$$P^A(x^A, x^B, w) = \mathbb{P}[v(x^A, w) - v(x^B, w)] \tag{5.1}$$

where $\mathbb{P}(\cdot)$ is an increasing C^2 function. Hence the expected vote share for party A is:

$$V^A(x^A, x^B, \theta) = \int_{\underline{w}}^{\overline{w}} [\mathbb{P}(v^n(x^A, w) - v(x^B, w))]\hat{r}(\theta, w)dw \tag{5.2}$$

The expected vote share for party B given policies x^A, x^B is simply $V^B(x^A, x^B, \theta) = 1 - V^A(x^A, x^B, \theta)$. Each party maximizes the expected share of votes[4]. Notice that a large numbers of voters implies that the actual vote share is equal to the expected share. So far, this setting is relatively standard and resembles the one in Banks and Duggan (2005). In the next paragraph I am going to impose additional restrictions to this model to allow for the interaction of public and private provision of a good.

5.3.1.1 Interaction between Public and Private Provision

Denote with x_i the i-th element of the policy vector x and suppose that x_i represents the degree of uniform public provision of a good that is also available on the private market. Examples of such goods are Education, Health Care, social security, etc. Publicly provided goods may differ in the way in which the provision is delivered. In particular one can distinguish the following cases. (i) Exclusive provision (socialization of commodities). The publicly provided good is not available on the private market. This is typical in the case of some pure public goods such as national defense (an example is Usher, 1977). (ii) *Top-up* goods. For this kind of goods the nature of the consumer choice is *quantitative*. Individuals receiving a certain level of public provision can decide to supplement this quantity with private purchases. A typical example is Health insurance (see Epple and Romano 1996a and Gouveia, 1997). (iii) *Opting-out* goods. The nature of the consumer choice is *qualitative* in this case, meaning that individuals can either enjoy the publicly provided good or purchase a different level of quality on the private market (no supplementation occurs). This case is often claimed to represent a good description of the way in which public education is provided in several countries (Stiglitz, 1974; Epple and Romano, 1996b), although some supplementation may occur. In this section I propose a general setting that allows for the interaction of Public and Private provision of a good in the Probabilistic Voting Model described in the previous section. This setting applies for all cases (i); (ii) (iii) mentioned above. In sections 4 and 5 I describe the different implications of these three cases. First, consider the indirect utility of an individual with income w:

$$v(x, w) = \max\left[v^n(x, w), v^m(x, w)\right] \tag{5.3}$$

Where $v^m(x, t)$ is the indirect utility if the individual decides to purchase some positive amount of the good on the private market and $v^n(x, t)$ is the indirect utility of an individual that does make any private purchase for the good of interest. Notice that, even

[4]Aranson, Hinich and Ordeshook (1974) have shown that in Probabilistic Voting models this is equivalent to maximizing the expected plurality and, as the number of voters approaches infinity, it is also equivalent to maximize the probability of winning the elections.

if $v^n(x,w), v^n(x,w)$ are concave functions, the function $v(x,w)$ may be neither differentiable in all the points of his domain nor concave. In order to keep the problem tractable I will assume that $v^m(x,w) - v^n(x,w)$ is monotone weakly increasing in w $\forall x, w$. This assumption implies that for each vector of policies x either $v^m(x,w) \leq v^n(x,w)$ for all z -i.e. no opting-out occurs-, or $v^m(x,w) > v^n(x,w)$ for all w -i.e. all individuals opt-out-, or there exists $\hat{w}(x)$ such that:

$$v(x,w) = \begin{cases} v^n(x,w) & if \quad w \leq (\geq)\hat{w}(x) \\ v^m(x,w) & if \quad w > (\leq)\hat{w}(x) \end{cases} \tag{5.4}$$

Party A's objective function in formula (5.1) becomes

$$\begin{aligned} V^A(x^A, x^B, \theta) = \quad & \int_{\underline{w}}^{\hat{w}(x^A)} [\mathbb{P}(v^n(x^A,w) - v(x^B,w))]\hat{r}(\theta,w)dw + \\ & + \int_{\hat{w}(x^A)}^{\overline{w}} [\mathbb{P}(v^m(x^A,w,m) - v(x^B,w))]\hat{r}(\theta,w)dw \end{aligned} \tag{5.5}$$

Following Banks and Duggan (2005), in order to show that the voting game has a unique Nash equilibrium in pure strategies, one has to show that (i) X is compact and convex, (ii) for each w, $\mathbb{P}[v(x^A,w) - v(x^B,w)]$ is jointly continuous in (x^A,x^B), (iii) for each x^A and x^B , $V^A(x^A,x^B,\theta)$ is strictly concave in x^A and $V^B(x^A,x^B)$ is strictly concave in x^B. In the setting proposed in this paper, (ii) is ensured because $\mathbb{P}$ is continuous and $v(x^A,w), v(x^B,w)$ are jointly continuous in (x^A,x^B). So one need to show the condition under which (i) and (iii) are satisfied. Regarding (i) the condition is not trivially satisfied if one of the good is publicly provided. Specifically, X may not be a convex set because $B(x)$ may fail to be a convex function. This can be the case, for example, for opting-out goods (see Epple and Romano, 1996b). In the next sections, I am going to show that convexity holds in the applications of this paper. Regarding (iii), $V^A(x,x^B)$ ($V^B(x^A,x,\theta)$) is strictly concave in x for all $x \in X$ if the Hessian matrix $H_V^A(x)$ ($H_V^B(x)$) is negative definite. Linbeck and Weibull (1987) have shown in a slightly different setting that in the case of concave indirect utility function this condition is satisfied if the distribution of $\mathbb{P}$ is such that $p'[v(x^A,w) - v(x^B,w)] \leq \bar{p}'$ for some positive $\bar{p}'$, where p' denotes the second derivative of the function $\mathbb{P}$. This result simply means that the function $\mathbb{P}$ is sufficiently "flat". Here we have an additional condition to be satisfied. One can show the following.

Theorem 1. *(Existence and Uniqueness). If there exist positive $\bar{r}$ and $\bar{p}'$ such that the distributions $\hat{R}(\theta,w)$ and $\mathbb{P}(d)$ satisfy $\hat{r}(\theta,w) \leq \bar{r}$ for all w and $p'[v(x^A,w) - v(x^B,w)] \leq \bar{p}'$, then there is a unique equilibrium in pure strategies. The unique electoral equilibrium is such that the two parties choose the same policy.*

Proof. See Appendix 5.7.2.1.

The additional condition simply states that the distribution of the individual parameter w

does not have peaks with excessively high density. Intuitively, this additional condition is required because a marginal change in the choice vector x in this case has an additional consequence relatively to the standard case described in Linbeck, Weibull (1987). On one hand, there is a direct effect of the change in x due to the change in the expected voting behavior of each type w, similar to the one of the standard analysis. On the other hand, there is also an indirect effect. Namely, the threshold $\hat{w}(x)$ may change as a consequence of the change in policy, and the size of the effect of such change on the objective function depends on the density of the distribution in a neighborhood of $\hat{w}(x)$. If such density is sufficiently low, then the effect of the change in $\hat{w}(x)$ is dominated by the direct effect. If the additional condition $\hat{r}(\theta, w) \leq \bar{r}$ for all w is satisfied, not only existence and uniqueness of a Nash equilibrium in pure strategies are ensured, but also other properties of the standard framework hold. Specifically, it is possible to show that an important welfare result holds in this setting as much as in the convex case. Denote with $V(x,\theta) = \int_{\underline{w}}^{\overline{w}} v(x,w)\hat{r}(w,\theta)dw$ the utilitarian social welfare function and with with $x^*(\theta)$ the policy vector that maximizes $V(x,\theta)$ subject to the governmental budget constraint, i.e. $x^*(\theta) = \arg\max_{B(x)\leq 0} V(x,\theta)$. One can state the following Theorem.

Theorem 2. *(Utilitarian outcome). The policy chosen by both parties in equilibrium is the same as the policy that would be chosen by an omnipotent Benthamite government, i.e.* $x^A = x^B = x^*(\theta)$.

Proof. See Appendix 5.7.2.2.

This result is very useful for the purposes of this paper because it reduces the study of the comparative statics of the equilibrium policy outcome to the one of the utilitarian social optimum.

5.3.2 Comparative Statics

In this section I describe the way in which a change in income inequality is defined in this paper. I assume that individual productivity is given by the sum of two independent random variables $Z \perp \Sigma$ such that $w = z + \varepsilon$. One can interpret this as the sum of a component due to parental and public investment in early life plus an idiosyncratic i.i.d. component. Variables z and ε have joint p.d.f. $r(z,\varepsilon,\theta)$ in the form[5]:

$$r(z,\varepsilon,\theta) = \{f(z) + \theta\,[g(z) - f(z)]\}\,\sigma(\varepsilon) \tag{5.6}$$

In order to impose an exogenous variation to the degree of inequality, I adopt the concept of Mean Preserving Spread (MPS). Consider the marginal distributions of z at $\theta = 0$ and

[5]Notice that the distribution of w would have the following p.d.f. $\hat{r}(w,\theta) = \int_{-\infty}^{+\infty} \sigma(w-z) r(z,\theta)dz$

$\theta = 1$, given by $g(z)$ and $f(z)$ respectively. Denote with $G(z)$, $F(z)$ the correspondent marginal c.d.f.s. Distribution G is a MPS of F if and only if $E_g(z) = E_f(z)$ and $VAR_g(z) > VAR_f(z)$. Notice how this concept is much more general and easy to interpret in comparison with the mean-to-median ratio that typically drives the comparative statics in Downsian models. A mean preserving spread is imposed as follows: F, G are two c.d.f.s such that $\int_{\underline{z}}^{z} [G(z) - F(z)]\,dz \geq 0$ for all $z \leq \bar{z}$ (i.e. the distribution F Second Order Stochastically Dominates G), and $\int_{\underline{z}}^{\bar{z}} [G(z) - F(z)]\,dz = 0$ (i.e. z has same mean under G and F). The expected value of productivity is given by

$$E(w) = \int_{\underline{z}}^{\bar{z}} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} (z+\varepsilon) r(z,\varepsilon,\theta) d\varepsilon dz =$$

$$= (1-\theta) \int_{\underline{z}}^{\bar{z}} z f(z) dz + \theta \int_{\underline{z}}^{\bar{z}} z g(z) dz + \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \varepsilon \sigma(\varepsilon) d\varepsilon dz = \tag{5.7}$$

$$= E_f(z) + E_\sigma(\varepsilon)$$

Notice that a change in θ preserves the average of w. Moreover, independence implies that $Var(z+\varepsilon) = Var(z) + Var(\varepsilon)$. Thus, the effect of moving θ in a neighborhood of $\theta = 0$ corresponds to the effect of increasing the variance of w keeping the mean constant. Moreover, the derivative of the equilibrium value x_i^* of a policy dimension i with respect to θ at $\theta = 0$ corresponds to the comparative statics of interest. Lastly, in order to understand the effect of a marginal mean preserving spread in the distribution of z on the equilibrium level of one policy variable, say x_i, one can use the simple monotone comparative statics result that follows. Consider a subset of policy dimensions with index $i \leq L < N$. Suppose that the utilitarian social welfare function can be written in the form in the form $V(x,\theta) = a(x,\theta) + \sum_{i<L} e_i(x_i,\theta)$ and the government budget constraint $B(x,\theta) = b(x,\theta) + \sum_{i \in L} \delta_i(x_i)$ for some twice differentiable functions $a, b, \{e_i, \delta_i\}_{i=1}^{L}$. Suppose that a, b are constant functions of x_i for all $i \leq L$.

Lemma 3. *(Monotonicity): If there exists at least one x_j with $N \geq j > L$ such that (i) the solution of the maximization problem is interior for x_j, (ii) $b(x,\theta)$ is such that $\frac{\partial b(x,\theta)}{\partial x_j} = \alpha \frac{\partial a(x,\theta)}{\partial x_j}$ for some constant α, and if (iii) $\frac{\partial^2 e_i(x_i,\theta)}{\partial x_i \partial \theta} \geq 0$ (≤ 0), then x_i is weakly increasing (decreasing) in θ in a neighborhood of $x^*(\theta)$.*

Proof. See Appendix 5.7.2.3.

This result is relatively restrictive, but it will prove useful for the purposes of this paper. In the next section I show that a simple model of public provision of a good financed by tax revenues satisfies the conditions of Lemma 3.

5.3.2.1 General Publicly Provided Good

Suppose that w is a scalar individual parameter capturing some characteristics that are positively related to income, such as productivity. Income is a weakly increasing function y of w. P_i^m represents the unitary cost for the public sector of a level of public provision of good x_i for each individual that purchase a positive amount of the good on the private market. P_i^n is the cost for an individual who consumes exclusively the public provision. Notice that if an individual cannot supplement the public provision, then $P_i^m = 0$. Conversely, if the public provision can be perfectly supplemented and the government also purchases the good on the private market, then $P_i^m = P_i^n$. Intermediate cases are possible (see section 4.2). Denote with $\pi(x) \in [0,1]$ the share of individuals that enjoys exclusively the public provision. Notice that $\pi(x)$ is not affected by θ under the assumptions stated in section 3.2. Lastly x_l is the amount of another good that is provided by the government at price P_l per unit. This captures other public spending, such as provision of public goods. The choice of opting-out is endogenous, thus the governmental budget constraint can be modeled in the form proposed here:

$$B(x) = -E_w[\tau(x,w) - \lambda(x,w)] + P_i^n x_i \pi(x) + P_i^m x_i (1-\pi(x)) + P_l x_l \leq 0 \tag{5.8}$$

where x is a $N \times 1$ vector of policy variables and $\tau(x,w) - \lambda(x,w)$ is strictly convex in x. In this setting $\tau(x,w)$ represents the amount of taxes paid by an individual with productivity w under policy x and $\lambda(x,w)$ is a function capturing losses from taxation with $E_w[\lambda_{x_j}(x,w)] = \gamma E_w[\tau_{x_j}(x,w)]$ for some $j \neq i,l$ and some constant γ. Lastly x_i is the level of provision of the private good of interest, and x_l represents the public spending on other goods and services. For instance, one may consider a tax system with a linear component and a lump-sum tax in the form $\tau(x,w) = x_1 w - x_2$ for $x_1 \in [0,1]$ and $x_2 \in [0, E(w)]$, and loss function in the form $\lambda(x,w) = \hat{\lambda}(x_1 w) + \alpha x_2$ for some convex function $\hat{\lambda}$. Because of the presence of $\pi(x)$ the budget constraint may not be linear. I study a simple model with quasilinear utility in the form $U^n(c, x_i, x_l, w) = c + u(x_i, w) + d(x_l, w)$ for an individual that chooses to consume only the public provision, and in the form $U^m(c, x_i, x_l, w) = c + \upsilon(x_i, w) + d(x_l, w)$ for an individual that purchase some positive amount on the private market, in which c is the consumption of a composite private good. Notice that u, υ are (possibly non-constant) functions of the parameter w. The reason of this assumption will become clear in the next sections. The corresponding indirect utility functions conditional on choice of provision are $v^n(x,w) = y(w) - \tau(x,w) + u(x_i,w) + d(x_l,w)$ and $v^m(x,w) = y(w) - \tau(x,w) + \upsilon(x_i,w) + d(x_l,w)$ respectively. Thus, the indirect utility of a voter is given by:

$$v(x,w) = \max[v^n(x,w), v^m(x,w)] \tag{5.9}$$

Lastly, assume that in the neighborhood of the equilibrium, either (i) $\bar{z} - \underline{\varepsilon} \leq \hat{w}(x) \leq \underline{z} + \bar{\varepsilon}$

or (ii) $\hat{w}(x) \in \{\underline{w}, \overline{w}\}$. Now in case (i) one can define a threshold level $\tilde{\varepsilon}(x,z)$ that satisfies $v^n(x, z+\tilde{\varepsilon}(x,z)) = v^m(x, z+\tilde{\varepsilon}(x,z))$. In case (ii) one gets $\tilde{\varepsilon}(x,z) = \bar{\varepsilon}$ if $v^n(x, z+\varepsilon) < v^m(x, z+\varepsilon)$ for all ε; $\tilde{\varepsilon}(x,z) = \underline{\varepsilon}$ if $v^n(x, z+\varepsilon) >< v^m(x, z+\varepsilon)$ for all ε. The utilitarian social welfare is given by:

$$\begin{aligned} V(x,\theta) = \quad & E[y(z+\varepsilon) - \tau(x, z+\varepsilon)] + \int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x,z)} u(x_i, z+\varepsilon) kr(z,\theta) d\varepsilon dz + \\ & + \int_{\underline{z}}^{\overline{z}} \int_{\tilde{\varepsilon}(x_i,z)}^{\overline{\varepsilon}} \upsilon(x_i, z+\varepsilon) kr(z,\theta) d\varepsilon dz + \\ & - P x_i \pi(x,\theta) + \int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} d(x_l, w) kr(z,\theta) d\varepsilon dz \end{aligned} \tag{5.10}$$

Assume that $r(z,\theta)$ and k are such that the conditions of Theorem 1 are satisfied. Theorem 2 implies that the unique equilibrium of the voting game corresponds to the vector of policies that maximizes $V(x,\theta)$. The quasilinearity of the utility function simplifies the analysis because it implies that the threshold $\tilde{\varepsilon}(x,z)$ is a constant function of x_j for all $j \neq i$. Moreover it is easy to show that under the assumption stated $\pi(x) = k \int_{\underline{z}}^{\overline{z}} [\tilde{\varepsilon}(x,z) - \underline{\varepsilon}] r(z,\theta) dz$ is also a constant function of x_j for all $j \neq i$. An immediate consequence of Theorem 3 is that if the solution to the optimization problem is internal for x_i, then the sign of the effect of a marginal increase in θ on the equilibrium level of x_i is given by the following Theorem.

Theorem 4. *If the solution to the utilitarian social welfare maximization problem is interior x_k, and (ii) $\int_{\underline{z}}^{\overline{z}} [\upsilon_{122}(x_i, z+\bar{\varepsilon}) - u_{122}(x_i, z+\underline{\varepsilon})] \left(\int_{\underline{z}}^{z} [G(s) - F(s)]\, ds \right) dz \geq 0$, then x_i is weakly increasing in θ in a neighborhood of $x^*(\theta)$.*

Proof. Notice that $V(x,\theta)$ and $G(x,\theta)$ satisfy all the assumptions of Lemma 3 for $a(x,\theta) = E[y(z+\varepsilon) - \tau(x, z+\varepsilon)]$, $b(x,\theta) = -E[\lambda(x,w)]$, $e_i(x_i,\theta) = \int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x,z)} u(x_i, z+\varepsilon) kr(z,\theta) d\varepsilon dz + \int_{\underline{z}}^{\overline{z}} \int_{\tilde{\varepsilon}(x_i,z)}^{\overline{\varepsilon}} \upsilon(x_i, z+\varepsilon) kr(z,\theta) d\varepsilon dz$, $\delta_i(x_i) = P_i x_i \pi(x,\theta)$, $e_l(x_l,\theta) = \int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} d(x_l, w) kr(z,\theta) d\varepsilon dz$ and $\delta_l(x_l) = P_l x_l$. Thus then x_i is weakly increasing in θ in a neighborhood of the equilibrium policy x^* if $\frac{\partial^2 e_i(x_i,\theta)}{\partial x_i \partial \theta} \geq 0$, and weakly decreasing if $\frac{\partial^2 e_i(x_i,\theta)}{\partial x_i \partial \theta} \leq 0$.

Notice that $\int_{\underline{z}}^{z} [G(s) - F(s)]\, ds$ is weakly positive by assumption, therefore the sign of the above depends on $\upsilon_{122}(x_i, z+\bar{\varepsilon}) - u_{122}(x_i, z+\underline{\varepsilon})$.

5.3.2.2 Interpretation as weighted average

Define function h over the support $z \in [\underline{z}, \bar{z}]$. Specifically, $h(z) = \frac{2\int_{\underline{z}}^{z} [G(s)-F(s)]ds}{VAR_g(z) - VAR_f(z)} \geq 0 \; \forall z \in [\underline{z}, \bar{z}]$. It is easy to show that $h(z)$ is weakly positive for all $z \in [\underline{z}, \bar{z}]$, it integrates to 1 and it is inverse U-shaped. Thus, it can be interpreted as the p.d.f. of a distribution with support $[\underline{z}, \bar{z}]$ and with higher density for central values of z.

Denote with E_h the expectation under such distribution and with $\tilde{c} = 0.5[VAR_g(z) - VAR_f(z)] > 0$. One can rewrite $\int_{\underline{z}}^{\bar{z}} [\upsilon_{122}(x_i, z+\bar{\varepsilon}) - u_{122}(x_i, z+\underline{\varepsilon})] \left(\int_{\underline{z}}^{z} [G(s) - F(s)]\, ds \right) dz = \tilde{c} E_h[\upsilon_{122}(x_i, z+\bar{\varepsilon}) - u_{122}(x_i, z+\underline{\varepsilon})]$, thus the condition above can thus be restated as follows.

Corollary 5. *If (i) the solution to the utilitarian social welfare maximization problem is interior for x_k, and (ii) $E_h[\upsilon_{122}(x_i, z+\bar{\varepsilon}) - u_{122}(x_i, z+\underline{\varepsilon})] \geq 0$, then x_i is weakly increasing in θ in a neighborhood of x^*.*

Proof. Straightforward from Theorem 4 and the definition of h.

Corollary 5 delivers the conditions for local monotonicity of the outcome of interest. In the next section I show how these conditions have a useful economic interpretation if the model is applied to the study of public intervention in education.

5.4 Publicly Provided Opting-out Good: Public Education

In this section I assume that the good of interest is an opting-out good, such as public education. It is publicly provided at a uniform quality level x_i and is also available on the private market at a continuum of different quality levels $q \in \mathcal{Q}$ at price Pq, with $\mathcal{Q} = [0, \bar{q}]$ for sufficiently large $\bar{q}$ (one may assume a discrete number of quality levels with no changes in the results, see Appendix C.1). Notice that with opting-out $P_i^m = 0$, i.e. an opting-out individual has no cost for the government. For simplicity I assume $P_i^n = P$, i.e. the price of quality on the private market is equal to the one faced by the government. The first consequence of the opting-out assumption is that the governmental budget set may not be linear. In order to understand why this is the case, consider the following way of modeling uniform provision of an opting out good. The total cost of providing the quality x_i is equal to the price per unit of quality $P = P^n$ times the quality x_i times the number of individuals that use the public service and the price per unit of quality P. Because the consumers-voters are a continuum of size 1, this means that the total spending on public education is given by $Px_i\pi(x_i) = x_i P \int_{\underline{z}}^{\bar{z}} \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x,z)} kr(z, \varepsilon, \theta) d\varepsilon dz$. Lastly assume that w is the marginal productivity of a worker and income $y(w) = w$. This would be the case in a model with labor supply if the labor market is perfectly competitive and labor supply is perfectly inelastic. The governmental budget constraint is also more complex in comparison with the convex utility case because of the endogeneity of the threshold $\tilde{\varepsilon}(x,z)$ as I described in section 3.1. The budget constraint has form:

$$\int_{\underline{z}}^{\bar{z}}\left(\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} -[\tau(x,z+\varepsilon)-\lambda(x,z+\varepsilon)]kd\varepsilon + Px_i \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x,z)} kd\varepsilon\right) r(z,\theta)dz + P_l x_l \leq 0 \quad (5.11)$$

Consider the following simplified version of the Becker-Tomes (1979) model of parental investment in children's education, in which the utility of a parent with income w is a function of parents' consumption c, of public spending in public goods x_l, and of the expected income of their children w^s, i.e.:

$$U(c,x_i,w) = c + d(x_l,w) + \beta E\left[w^s(x_i,w)|w\right] \quad (5.12)$$

Suppose a child's future productivity w^s is a function of the quality of education and of the endowment $e(w)$ she receives from her parents, plus an idiosyncratic i.i.d. ability ν^s. For simplicity, I assume $e = w$. This formulation describes a simple transmission mechanism of human capital from parents to children. Education is provided by the government at uniform quality level x_i but other levels of quality q are available on the private market. Lastly, one may want to allow for positive spillovers of education. This can be the case, for instance, if there are *peer effects*. The formula for the productivity of a child with parents of type w and public education of quality x_i is given by the following formula:

$$w^s(x_i,w) = \begin{cases} \rho_{t+1}\left[H(x_i,w) + s(x_i,\phi) + \nu^s\right] & if \quad public \\ \rho_{t+1}\left[\check{H}(\check{q}(w),w) + s(x_i,\phi) + \nu^s\right] & if \quad private \end{cases} \quad (5.13)$$

where $\check{q}(w)$ represents the quality of private education chosen among the levels available in the set $\mathcal{Q}$, and H, $\check{H}$ and s are twice differentiable. The human capital production function of the child is allowed to differ between a child that attends public school (H) relative to one that receive private education ($\check{H}$). Assume that the first derivative $H_1(x_i,w)$ is finite for all x,w and that ν^s is independent of w,x and of the choice of the kind of education. Lastly, $s(x_i,\phi)$ represents the spillovers from other children's education. For instance, $s(x_i,\phi)$ could be the average human capital of other children $s(x_i,\phi) = \phi E_{w,\nu}[w^s(x_i,w)]$ and therefore $s(x_i,\phi) = \frac{\phi}{1-\phi}\left[\int_{\underline{z}}^{\bar{z}}\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x_i,z)} H(x_i,z+\varepsilon)kd\varepsilon r(z,\theta)dz + \int_{\underline{z}}^{\bar{z}}\int_{\tilde{\varepsilon}(x_i,z)}^{\bar{\varepsilon}} \check{H}(\check{q},z+\varepsilon)kd\varepsilon r(z,\theta)dz\right]$ for some $\phi \in [0,1)$. The objective functions of a parent choosing public education or private education respectively have form:

$$v^n(x,w) = \xi w - \tau(x,w) + d(x_l,w) + \beta\rho_{t+1}\left[H(x_i,w) + s(x_i,\phi)\right] \quad (5.14)$$

and

$$v^m(x,w) = \max_{q\in\mathcal{Q}} \xi w - \tau(x,w) - Pq + d(x_l,w) + \beta\rho_{t+1}\left[\check{H}(q,w) + s(x_i,\phi)\right] \quad (5.15)$$

Formulas (5.14) and (5.15) correspond to $u(x_i,w) = H(x_i,w) + s(x_i,w)$ and $\upsilon(x_i,w) = \check{H}[\check{q}(w),w] + s(x_i) - P\check{q}(w)$ respectively. Notice that because H is differentiable, then $\upsilon(x_i,w)$ is continuous and differentiable with respect to x_i,w under the assumption that $\mathscr{Q}$ is a continuum. The conclusions are unchanged if one allows for a discrete number of alternatives, see Appendix 5.7.3. First of all, one can notice that equilibria with a positive level of public intervention are possible even if there are no externalities in consumption , i.e. $\phi = 0$ and the tax system allows for uniform lump-sum transfers. The reason is that, if some opting out occurs, the amount that can be rebated to the voters if the government stops providing the good is lower than the cost of purchasing the same amount on the private market (if available). Nevertheless, if the tax system fully flexible, a positive level of provision may emerge in equilibrium only if there are positive externalities in consumption. If a positive level of provision is chosen in equilibrium, then the sign of the comparative statics of interest can be derived using Corollary 5. It is straightforward to show that in this case $\upsilon_{12}(x_i,z+\bar{\varepsilon}) = 0$, thus the condition (ii) in Corollary 5 reduces to $E_h[H_{12}(x_i,z+\underline{\varepsilon})] \leq 0$. Again, the sign of the comparative statics depends on the sign of H_{12}. An interesting interpretation of this condition follows. Define the Expected Individual Marginal Returns to Public Education: $MRE(x_i,w) = \frac{\partial E[w^s(x_i,w)|w]}{\partial x_i} = \rho_{t+1}[H_1(x_i,w)+\omega s(x_i)]$. Moreover, define the public spending per capita in education as the total spending divided by the size of the population, i.e. $PCE(x_i,\theta) = Px_i\pi(x_i,\theta)$. One can show the following.

Proposition 6. *If $MRE(x_i,w)$ is decreasing in income for all x,w, then (i) the quality of public education x_i and (ii) the public spending per capita in education $PCE(x_i,\theta)$ are weakly increasing in θ* in a neighborhood of $x^*(\theta)$.

Proof. (i) One can derive the Expected Individual Marginal Returns to Public Education: $MRE(x_i,w) = \rho_{t+1}[H_1(x_i,w)+\omega s(x_i)]$, hence $\rho_{t+1}H_{12}(x_i,z+\underline{\varepsilon}) = \frac{\partial MRE(x_i,w)}{\partial w}$. Substituting in condition (ii) of Corollary 5 one gets $E_h[u_{12}(x_i,z+\underline{\varepsilon})] = E_h\left[\frac{\partial MRE(x_i,z+\underline{\varepsilon})}{\partial w}\right]$, which is negative if $MRE(x_i,w)$ is decreasing in w for all x_i,w. *(ii)* The derivative including the political equilibrium change is: $\frac{dPPE(x_i,\theta)}{d\theta} = \frac{\partial PPE(x_i,\theta)}{\partial\theta} + \frac{\partial PPE(x_i,\theta)}{\partial x_i}\frac{dx_i}{d\theta} = P\frac{\partial x_i}{\partial\theta}\left[\int_{\underline{z}}^{\overline{z}}\left(\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x,z)} kd\varepsilon r(z,\theta)dt + \frac{\partial\tilde{\varepsilon}(x,z)}{\partial x_i}k\right)r(z,\theta)dz\right] + Px_i\frac{\partial\pi(x_i,\theta)}{\partial\theta} = P\frac{\partial x_i^*}{\partial\theta}\left(\pi(x_i,\theta) + \frac{\partial\tilde{\varepsilon}(x,\bar{z})}{\partial x_i}\right)$. Result (i) implies $\frac{\partial x_i^*}{\partial\theta} \geq 0$. Moreover, $\frac{\partial\tilde{\varepsilon}(x,z)}{\partial x_i} = \frac{H_1(x_i,z+\tilde{\varepsilon}(x,z))}{\check{H}_2(q^*,z+\tilde{\varepsilon}(x,z))-H_2(x_i,z+\tilde{\varepsilon}(x,z))} \geq 0$ because the denominator is positive by assumption. Thus, $PPE(x_i,\theta)$ is weakly increasing in θ in a neighborhood of $x^*(\theta)$.

The result in Proposition 6 has an intuitive interpretation. That is, an increase in the political weight of relatively low income individuals imply a rise in the quality of public education if a better education reduces the intensity of the transmission mechanism of income from parents to children. This implication is potentially testable and can be the object of future empirical research. There are both theoretical and empirical argument in favor and against a negative value for such derivative. Some are related to aspects not explicitly modeled in this paper, such as credit constraints, partial supplementation, complementarities between parental education and children's returns to education, etc. As explained in Becker, Kominers, Murphy, Spenkuch (2015), the question is strictly related to the one if government spending and parental investments are substitutes or complements in the production of human capital. Another aspect that is highlighted in the literature is the possibility that parents that wish to supplement public education with private spending face credit constraints. To understand why this may be important, it is useful to show a simple example. Suppose parents can supplement public education with private spending $s \in [0, \bar{s}]$, which is a substitute (perfect or imperfect) of the public spending in education in the form: $H(x,w) = \max_{s \in [0,\bar{s}]} \tilde{H}[x_i + a(s)] - \tilde{p}s$ subject to $w - \tau(x,w) - s \geq 0$. Say X is such that $\tau(x,w) = -\underline{b}$ for all $w \leq w^{min}$ and some $\underline{b} \geq 0$ which means that low income households do not pay taxes, but they may receive a grant that ensure a minimum level of private consumption. Denote with $s^{int}(x_i)$ the optimal level of private spending of a parent that is not credit constrained. Lastly, suppose that $s^{int}(x_i) \leq \underline{b} + w^{min}$, which means that all liquidity constrained individuals are not positive taxpayers under any policy $x \in X$. These two assumptions imply that the optimal level of private investment $s^*(x_i, w) = s^{int}(x_i)$ for $w + \underline{b} - s^{int}(x_i) \geq 0$ and $s^*(x_i, w) = w + \underline{b}$ otherwise. Then $H_{12}(x_i, w) = \tilde{H}''[x_i + a(w)]a'(w) < 0$ for all $w < s^*$ and $H_{12}(x_i, w) = 0$ for all $w \geq s^{int}(x_i)$. Because there are theoretical arguments in both directions, the question if the marginal returns to education are decreasing in income can only be addressed by empirical analysis. The estimation of returns to education is a classical exercise in applied economics that involves several issues, the analysis of which is beyond the scope of this paper. Nevertheless, it is interesting to mention the results of a few recent papers that have tried to disclose the relationship between returns to education and parental income in the data. Brenner and Rubinstein (2012) estimate a model of returns to a year of additional education and find that controlling for individual and family characteristics the returns to educations are decreasing in the quintile of parental income. Their findings suggest that individuals from low-income families have lower levels of educational attainment because they face higher costs of schooling, not because they cannot gain from further education. In particular a strong and statistically significant difference between the 1st and the 5th quintile is observed in all specifications, such that the returns to educations at the top quintile are less than 50% the ones in the lowest quintile. Although the concept of returns to education in their paper is not directly comparable with the one implied by this analysis, this result suggests that the sign of the above may be indeed negative. Conversely, other papers (Altonji and Dunn, 1996) find a positive relationship between parental education and marginal returns to education. Lastly, Card (2001) reviewing the literature about returns to education, find some support for the partial supplementation hypothesis by comparing OLS and IV estimates of the returns to education in several studies.

5.4.1 Effects on future generations

It is interesting to analyse how a change in the level of public intervention in education affects the income levels and income inequality of the future generations. This analysis provides a further possible interpretation of the condition derived in the previous section. In this section, I derive results about how the expected income, the variance and the coefficient of variation of the distribution of w^s change if there is a marginal increase in x_i. I can state the following.

Proposition 7. *If $Var(\Sigma)$ is large enough, then (i) the expected income of the next generation $E(w^s)$, (ii) the variance $Var(w^s, \theta)$ and (iii) the coefficient of variation $CV(w^s, \theta)$ of the income distribution of the next generation are all weakly increasing in x_i. Moreover, if the expected marginal returns to public education are weakly decreasing in parental income, then (iv) the total effect of a marginal increase in θ on $Var(w^s)$ and $CV(w^s)$ is ambiguous.*

Proof. See Appendix 5.7.3.2.

The interpretation of result (i) is simple. A marginal increase in the quality of public education has two effects in this setting. On one hand, it causes an increase in the future productivity of the individuals that choose public education for their children. On the other hand, it implies that some individuals with ε close to the cutoff $\tilde{\varepsilon}$ may switch from private to public education, reducing the future income of their children. If the marginal density k of ε is sufficiently low, then the share of switching parents is low and the first effect dominates. Results (ii), (iii), (iv) imply that if a society has sufficient income inequality, then an exogenous shock that further increases such inequality is going to be mitigated in its effects on future generations. The intuition is that a rise in current income inequality has two opposite effects. On one hand, the mechanism of transmission of human capital imply that the direct effect may increase the inequality of the next generation. On the other hand, if marginal returns to education are higher - on average - for children from poorer families, then the increase in the degree of public intervention in education is going to cause a decrease in future income inequality.

5.4.2 Vouchers

Suppose that the government provides vouchers that cover the cost of education in public institutions, but can be also spent to partially cover the fees of private schools if parents decide opt-out of the public service. This setting is similar to the one proposed by Ireland (1990). In this case the indirect utility of an opting-out individual becomes $v^m(x,w) = w - \tau(x,w) - P(\check{q}(w) - x_i)1[\check{q}(w) - x_i \geq 0] + d(x_l,w) + \check{H}(\check{q}(w),w) + s(x_i,\phi)$, while the one of an individual that does not opt-out is unchanged. The government budget constraint is now simplified in the form $\int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\overline{\varepsilon}} [-\tau(x,z) + \lambda(x,z)]\, kd\varepsilon r(z,\theta)dz + P_i x_i + P_l x_l \leq 0$ because the voucher is provided to everybody. Notice that in this case the presence of positive spillovers $s(x_i,\phi)$ is crucial to ensure a positive level of public provision. Specifically, for any positive level of public provision x_i, if (i) $s(x_i,\phi) = 0$, (ii) the tax system allows for uniform lump-sum transfers with no losses and if (iii) the same level of provision is also available on the private market, i.e. $\bar{q} \geq x_i$, then each individual can be weakly better off with a policy x' such that $x'_i = 0$ and $\tau(x',w) = \tau(x,w) - Px_i$. This implies in turn that no positive level of public intervention is chosen by voters in equilibrium[6]. If uniform lump-sum transfers are not allowed, or they are costly, or the private market does not provide all the quality levels that can be provided by the public sector, then the equilibrium may exhibit positive x_i even in absence of positive spillovers. This result suggests that the introduction of a voucher system may have unexpected effects and lead to a fall in public spending in education per capita. Regarding the comparative statics induced by an increase in θ, it is easy to show that, if $x_i > 0$ at the equilibrium, then the sign of the comparative statics is the sign of $-E_h[H_{12}(x_i, z + \underline{\varepsilon})]$, as in the baseline case.

5.5 Comparison with other kinds of Publicly Provided Goods

In this section I compare the results in section 4 with the ones that this voting framework delivers if used to analyse other kinds of publicly provided goods. The aim is to show that the way in which the good is provided and consumed plays a crucial role in determining how income inequality affects the degree of public intervention in the provision of such good.

5.5.1 Exclusive Public Provision (Pure Public Good)

Consider the case in which the provision of a certain good is exclusively public, either because of legal restrictions or because of a market failure. A typical example is National Defense. Following the structure of the previous section, one can model this case as follows. The individual indirect utility $v(x,w)$ is given by (a) $v(x,w) = v^n(x,w) = w -$

[6] A small positive level equal to the lowest optimal private purchase at $x_i = 0$ may still prevail only if rising revenues does not imply net losses.

$\tau(x,w)+u(x_i,w)+d(x_l,w)$ and $\tilde{\varepsilon}(x,z)=\bar{\varepsilon}$ for all z,x. The government budget constraint is simply linear in x_i in the form $P_i x_i + P_l x_l - E_w[\tau(x,w)] \leq 0$. Substitute (a) and (b) into condition (ii) of Corollary 5. The condition for a positive sign for the comparative statics of interest becomes (b) $E_h[u_{12}(x_i,z+\bar{\varepsilon}) - u_{12}(x_i,z+\underline{\varepsilon})] \geq 0$. Thus, one can state the following result.

Proposition 8. *The effect of a marginal increase in income inequality on the equilibrium level of a publicly provided good with exclusive public provision is ambiguous. If $u(x_i,w)=a(x_i)c(w)$, then the effect is weakly positive if c is concave, and weakly negative if c is convex.*

Proof. Straightforward from Corollary 5.

Proposition 8 suggests that the strong relationship between income inequality and size of public spending in public goods usually implied by traditional Downsian models may not survive if one departs from the traditional deterministic framework. Moreover, it shows that - differently from traditional models - imposing restrictions on the sign of the cross derivative u_{12} is not sufficient to deliver a monotone comparative statics. As an example, compare this result with the one of the correspondent unidimensional Downsian prediction, and for simplicity set $\tau(x,z)=x_j\check{\tau}(w)$. It is easy to show that if $\check{\tau}$ is such that an individual with median income is a positive taxpayer and $u_{12}(x_i,w)\leq 0$ for all x_i,w, then $v(x,w)$ satisfies the *Spence-Mirrlees* condition and the level of public provision of the good would be weakly increasing in the median income, and strictly increasing if $x_j\check{\tau}(w)$ is not a lump-sum tax. This means that the skewness of the income distribution increases (i.e. the median income decreases at constant mean), this would translate into an increase in public spending in the good in equilibrium. Lastly, notice that the function u may exhibit different features depending on the kind of public good considered. For instance policing and other services that improve the protection of property rights may be more desirable by individuals with high productivity, who are likely to accumulate larger wealth. Conversely, for other public goods the direction of the relationship may be reversed.

5.5.2 Top-up goods

Now consider the case of a good that is uniformly provided by the government and such that consumers can supplement the public provision with private purchases from a set of available market options $\mathcal{Q}$. The set $\mathcal{Q}$ can have a discrete number of elements or it can be a continuum. In the case of *top-up* goods, the interaction between the public provision and private purchases of the good has a *quantitative* nature. Specifically, consumers care only about the total quantity of the goods they can consume, independently on the source of provision. Because of this, I am going to assume in this section that the (direct) utility from consuming the top-up good is solely a function of the quantity consumed. Thus, the indirect utility of individuals that does not supplement the public provision is given by $v^n(x,w) = w - \tau(x,w) + \hat{u}(x_i,w) + d(x_l,w)$, while the one of an individual that does supplement the public provision with a positive amount of public purchases is $v^m(x,w) = \max_{q\in\mathscr{Q}} w - \tau(x,w) - Pq + \hat{u}(x_i+q,w) + d(x_l,w)$, for some twice differentiable function $\hat{u}$ that is increasing and concave in its first argument. Because no opting-out occurs, the government budget constraint is linear in x_i, in the form $-E_w[\tau(x,w) - \lambda(x,w)] + P_i x_i + P_l x_l \leq 0$. Thus, the condition (ii) in Corollary 5 for a positive sign of the comparative statics becomes $E_h[\hat{u}_{12}(x_i+q, z+\bar{\varepsilon}) - \hat{u}_{12}(x_i, z+\underline{\varepsilon})] \geq 0$. In the next subsections I will analyse the consequences of top-up for the level of provision and the comparative statics at the political equilibrium.

5.5.2.1 Pure Private Goods: undifferentiated consumption good

If the good of interest is a pure private good, then each consumer's utility is affected only by its own consumption. Denote with $\bar{x}_i$ the maximum level of quality that can be provided by the government and with $q^*(x_i,w)$ the optimal quantity purchased on the private market by an individual with productivity w facing public provision x_i. Define $\breve{q}(w) = \min_{w\in[\underline{z}+\underline{\varepsilon}, \bar{z}+\bar{\varepsilon}]} q^*(0,w)$.

Proposition 9. *If (i) $[0,\bar{x}_i] \subseteq \mathscr{Q}$ and (ii) the tax system allows for costless lump-sum transfers, then in equilibrium $x_i \leq \breve{q}(w)$. Moreover, if (iii) rising tax revenues is costly, i.e. $E_w[\lambda(x,w)] > 0$ whenever $E_w[\tau(x,w)] > 0$ and $E_w[\lambda(x,w)] = 0$ whenever $E_w[\tau(x,w)] = 0$, then no positive public provision of a pure private good occurs.*

Proof. Consider any level of spending Px_i that would emerge if a feasible policy vector $x \in X$ with $x_i > 0$ is chosen. Under assumption *(ii)*, a policy vector such that $\tau(x',w) = \tau(x,w) - Px_i$ for all w and $x'_i = 0$ is feasible. Such policy makes all voters weakly better off relative to x because all of them can choose a bundle that implies the same amount of consumption of both the *numeraire* and the pure private good. Thus all consumers are as well off as under policy x' only if $x_i \leq \breve{q}(w)$. This implies that if x is such that $x_i > \breve{q}(w)$, then $V(x',\theta) > V(x,\theta)$, thus x cannot be the policy chosen in a voting equilibrium. Lastly, if *(iii)* also applies, then $x_i > 0$ implies in equilibrium $x_i < \breve{q}(w)$. Thus, $V(x',\theta) > V(x,\theta)$, and therefore no positive provision occurs.

This result implies that in this setting a private good is provided by the government only if the private market is not capable of providing all the the level of consumption per capita that would be enjoyed if the good is provided by the public sector in some positive amount and/or the tax system does not allow for costless lump-sum transfers. In other words, if the private market is effective in providing the good and the tax system is sufficiently flexible, no pure private good should be publicly provided. If inefficiencies in the private provision or an excessively restrictive tax system imply a positive level of provision, notice that the sign of the comparative statics would be positive if and only if (a) $E_h[\hat{u}_{12}(x_i, z+\underline{\varepsilon})] \leq 0$ in the case in which a continuum of alternatives is available on the private market (i.e. $\mathscr{Q} = [0, \bar{q}]$ for sufficiently high $\bar{q}$) and if and only if (b) $E_h[\hat{u}_{12}(x_i+\check{q}, z+\bar{\varepsilon}) - \hat{u}_{12}(x_i, z+\underline{\varepsilon})] \geq 0$ in the case of a discrete number of alternatives on the private market. About case (a), notice that the sign of the comparative statics is weakly negative if the publicly provided private good is a normal good. Regarding case (b), because the sign of $\hat{u}_{12}(x_i+\check{q}, z+\bar{\varepsilon}) - \hat{u}_{12}(x_i, z+\underline{\varepsilon})$ does not have a straightforward economic interpretation. Thus, one can conclude that the sign of the comparative statics is ambiguous. These results suggest that, if there is governmental intervention in the provision of a pure private good, one should not expect the size of this provision to be increasing in the degree of income inequality.

5.5.2.2 Imperfect Public Good with Supplementation: Public Health Insurance

It is well known that a positive level of public intervention in the provision of a good may be socially desirable if the good of interest exhibits positive externalities in consumption or it is an imperfect public good. The reason is that in such cases the level of consumption that is chosen by self-interested private agents tends to be suboptimally low on a social welfare point of view. Thus, if the loss due to underprovision is large relatively to the one induced by a uniform public provision, then a positive level of public intervention in the provision of the good characterizes the political equilibrium. A typical example of a good that is described in the literature as a top-up (see, Epple and Romano 1996a and Gouveia, 1997) is health insurance. Typically, the government can provide a certain level of service or insurance coverage, and consumer can purchase additional insurance on the private market. In this case the availability on the private market may influence the comparative statics, thus it is worth to analyse two cases. The first possibility is that a discrete number of options is available on the private market, i.e. $\mathscr{Q} = \{q_1, q_2, ..., q_n\}$. In such case, similarly to what shown for a opt-out good, the sign of the comparative statics is the same as in the case in which only one private option is available, provided that such option is selected from $\mathscr{Q}$ as the one that maximize the objective function of an individual with income $w = \hat{w}$ when the policy chosen is the equilibrium policy (see Appendix C.1). In this case, I assume that individual utility is given by $U(c, x_i, w) = c + d(x_l, w) + \beta E[Health(x_i, w, \phi)|w]$, where $Health$ is a function of the level of health insurance, of the individual parameter w, of the externality produced by the health of other individuals and by an i.i.d shock. It has form: $Health(x_i, w, \phi) = I(x_i + q, w) + s(x_i, \phi) + \nu$ where $q = 0$ if an individual does not top-up. I is a function that captures the effect of insurance on voters' health. The interaction between the level of health insurance and the parameter capturing productivity in the function I may be due to various reasons. For instance, individual with higher productivity may have different costs of illness relative to low productivity ones, or they may have a different probability of getting sick. Lastly, $s(x_i, \phi)$ represents the average level of health in the whole population, i.e. $s(x_i, \varphi) = \frac{\varphi}{1-\varphi}\left[\int_{\underline{z}}^{\overline{z}}\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}(x_i,z)} H(x_i, z+\varepsilon)kd\varepsilon r(z,\theta)dz + \int_{\underline{z}}^{\overline{z}}\int_{\tilde{\varepsilon}(x_i,z)}^{\overline{\varepsilon}} H(x_i+\check{q}, z+\varepsilon)kd\varepsilon r(z,\theta)dz\right]$. The indirect utilities of individuals that top-up and do not top-up the public provision are given by:

$$v^n(x,w) = w^f - \tau(x,w^f) + I(x_i,w) + s(x_i,\varphi) \tag{5.16}$$

and

$$v^m(x,w) = \max_{q\in\mathcal{Q}} w^f - \tau(x,w^f) - pq + I(x_i+q,w) + s(x_i,\varphi) \tag{5.17}$$

respectively. The condition (ii) in Corollary 5 for a weakly positive sign of the comparative statics of interest reduces to $E_h\left[I_{12}(x_i+q, z+\bar{\varepsilon}) - I_{12}(x_i, z+\underline{\varepsilon})\right] \geq 0$ if a discrete number of options is available on the private market (i.e. $\mathcal{Q} = \{q_1, q_2, ..., q_n\}$) and to $E_h\left[I_{12}(x_i, z+\underline{\varepsilon})\right] \leq 0$ if there is a continuum of of private insurance available (i.e. $\mathcal{Q} = [0, \bar{q}]$ for sufficiently large $\bar{q}$).

Proposition 10. *(a) (Continuum of choices on the private market). If (i) $\mathcal{Q} = [0,\bar{q}]$ for sufficiently large $\bar{q}$ and (ii) health insurance is a normal good for all income levels, then a marginal rise in income inequality has a weakly negative effect on the equilibrium level of the public provision of health insurance. (b) (Discrete set of choices on the private market). If (i) a marginal rise in income inequality has an ambiguous effect on the equilibrium level of the public provision of health insurance. If (ii) $I(x_i+q,w) = b(x_i+q)c(w)$ for some monotone functions b, c then the effect is weakly positive if c is concave, and weakly negative if c is convex.*

Proof. (a): denote with $q^*(x_i, w)$ the demand of private health insurance of an individual with productivity w given a level of public insurance x_i. Using the F.O.C. of the consumer's optimization problem, an individual that purchases a positive amount on the private market has income elasticity of demand $\frac{\partial q^*(x_i,w)}{\partial w}\frac{w}{q^*(x_i,w)} = -\frac{I_{12}(x_i+q^*(x_i,w),w)}{I_{11}(x_i+q^*(x_i,w),w)}\frac{w}{q^*(x_i,w)}$. This is positive if $I_{12}(x_i+q^*, w) \geq 0$. Thus, if q is a normal good for all w, then $E_h\left[I_{12}(x_i, z+\underline{\varepsilon})\right] \geq 0$ which implies a weakly negative sign of the comparative statics. (b): (i), straightforward from the condition (ii) Theorem 5; (ii) notice that the condition becomes $E_h[b'(x_i+q)c'(z+\bar{\varepsilon}) - b'(x_i)c'(z+\underline{\varepsilon})]$ where b must be concave by assumption. Thus the expectation is weakly negative if c is concave.

Proposition 10 implies that the sign of the relationship between income inequality and degree of public intervention in health insurance is ambiguous, and that under additional restrictions it is negative. This results shows that, if voters' preferences over a publicly provided good are separated from redistributive motives, then the effect of a shock on income inequality on the degree of public intervention depends solely on the characteristics of the good. Moreover, this result has important consequences for the traditional theoretical literature that analyses link between income inequality and size of public intervention in redistributive policies (e.g. Meltzer and Richards, 1981). If the marginal effect of an increase in income inequality may have different sign for different kinds of public intervention, then the total effect on the size of public spending in policies with redistributive effects may depends on various factors, including the relative size of public intervention in different kinds of policies at the political equilibrium. For instance, one may expect to observe a positive link between inequality and redistribution - as implied by the traditional literature - if in-cash policies and public education absorb a large share of the governmental budget, but this may not hold true if other form of public spending - such as public health insurance - are substantial.

5.6 Concluding Remarks

This paper provides a theoretical framework to analyse the effects of marginal shocks in income inequality on public spending on education in democratic countries. In order to separate voters' demand for public provision of education from their preferences for redistribution I assume a multidimensional policy space that includes as choice variables the quality of public education, the parameters of the tax system and the public spending in other policies. Moreover, following the literature, I assume that education is a good that is characterized by the possibility *opting-out*, which means that individuals can either enjoy the quality of public education, or purchase private education at one of the quality levels available on the market. The *qualitative* aspect of the consumer's choice is crucial, because the choice of *opting-out* implies the loss of the private benefits from the public provision. This implies in turn that preferences may exhibit non-convexities and that standard results in deterministic voting models may not hold. I adopt a probabilistic voting model similar to the one in Lindbeck and Weibull (1987) and Banks and Duggan (2005) to study how a marginal increase in income inequality - at constant mean - affects the equilibrium quality of education. This choice allows one to tackle the problems of existence of a political equilibrium induced both by the multidimensionality of the policy space (Grandmont, 1975) and by the presence of non-convexities in the objective function (Stiglitz, 1975), that are well-known in the theoretical literature. I show that the sign of the effect of a marginal mean preserving spread of the income distribution on the equilibrium quality of public education is positive if the expected marginal returns to public education are larger for children that have relatively low-income parents, which is in line with the findings in the empirical literature. Such literature suggests that this may be the case if public education can be partially supplemented by parental private investment and low income parents face credit constraint. Moreover, this condition is equivalent in the model to the case in which the degree at which income is transmitted across generation is lower if better public education is provided. Unsurprisingly, I can also show that, under the same conditions, better quality of public education implies lower income inequality in the next generation. This result suggests that exogenous shocks on income inequality may be mitigated in future generations by the endogenous adjustment in the degree of public intervention in education. Such predictions rely on the particular way in which this good is publicly provided and do not hold for other kinds of publicly provided goods, such as pure public goods and *top-up* goods. Specifically, I show that for a typical *top-up* good such as health insurance, a marginal increase in income inequality has ambiguous effects on the level of insurance provided by the government, and that under some mild additional assumption such effect is weakly negative. This suggests that the direction of the relationship between income inequality and total size of governmental intervention in redistributive policies may differ from the one implied by traditional models, and may depend on the relative size of the public spending on different kinds of public intervention. In other words, if one can write a sufficiently flexible model, then voters choose to redistribute in the most effective way - through the tax system - and redistribution motives do not strongly affect the size of the public intervention in other policies. Public education is an exception in this framework, because it is a policy that allows relatively low income voters to achieve income redistribution in the generation of their children. Because this kind of redistribution cannot be achieved through the tax system, relatively low income voters support high levels of public education in order to ensure higher consumption levels to their children. If that is the case, then the effects of a positive shock on income inequality are mitigated by the endogenous political choices both in the short run - through the tax system - and in the long run - through better public education, thus the consequences of shocks on income inequality are less dramatic for the society.

5.7 Appendix

5.7.1 Simple Downsian Model

Consider the following simple Downsian model with n voters. Agents differ only in a unidimensional parameter $w \in W$ (income). There are two choice variables x_1, x_2 related by a convex governmental budget set X such that $X \equiv \{(x_1,x_2)\,|\,(x_1,x_2) \in R^2_+ \cap \tilde{B}(x_1,x_2,E(w)) \leq 0\}$. Individual preferences are represented by the indirect utility function $u(x_1,x_2,w)$, which continuous, increasing in x_1,x_2,w and twice differentiable in each argument. The Spence-Mirrlees condition states $M(x_1,x_2,t) = \frac{\partial}{\partial t}\left(\frac{u_1(x_1,x_2,t)}{u_2(x_1,x_2,t)}\right) > (<) 0\, \forall x_1,x_2,t$. If such condition is satisfied, then the voting game has a Condorcet Winner, which is the individual with median w. Thus the social choice is given by $(x_1,x_2) = \arg\max_{(x_1,x_2)\in X} u(x_1,x_2,w^m)$ where w^m is the median of w. Because of the Spence-Mirrlees condition the preferred choice of an individual with parameter w is such that x_1 is increasing (decreasing) in w and x_2 is decreasing (increasing) in w. Hence the equilibrium social choice would also change in this way if the distribution of w is changed in such a way that the median voter has higher t and $E(t)$ is unchanged. This implies a monotone link between a measure of the skewness of the income distribution (in this case the difference the mean to median ratio). For instance, suppose $x_2$is a private good that in uniformly publicly provided (no private purchases are allowed in this simple example) , after tax income is given by $w - x_1\hat{\tau}(w)$ and the indirect utility function is $u(x_1,x_2,w) = u(w - x_1\hat{\tau}(w),x_2))$. The government budget constraint is in the form $-x_1E(\hat{\tau}(w)) + Px_2 \leq 0$ where $P_2$is the price of one unit of the good. Notice that in an interior solution the budget constraint is binding hence the problem is equivalent to $\max_{x_2\in X} u(w - p(w,P)x_2,x_2)$ where $p(w,P)$ is the tax-price of the good defined as total amount of tax paid divided by the size of the provision, i.e.: $p(w,P) = \frac{x_1\hat{\tau}(w)}{x_1E[\hat{\tau}(w)]/P} = \frac{P\hat{\tau}(w)}{E[\hat{\tau}(w)]}$. The formula shows that the tax-price of the good is lower than the market price for all individuals that pay less taxes than average. If this is the case for the median income voter, this results in a positive level of public intervention in equilibrium. Moreover, this suggest the provision has redistributive effects, in the sense that after a positive provision is implemented, relatively lower income individuals can afford new consumption bundles, while for relatively high income individuals some bundles are not affordable anymore. The total effect of a marginal increase in income on the demand for public provision is given by:

$$\frac{dx_2^*(w,P)}{dw} = \frac{\partial x_2^M(w,\tilde{P})}{\partial w} + \frac{\partial x_2^M(w,\tilde{P})}{\partial P}p'(w,P) \tag{5.18}$$

where $x_2^*(w,P)$ is the demand for public provision of and individual with income w and $x_2^M(w,\tilde{P})$ is the private Marshallian demand of the same individual at price $\tilde{P} = p(w,P)$. Denote with $\eta_w^M(w,P)$ the income elasticity of the Marshallian demand for an individual with income w at price P, and with $\eta_P^H(w,P)$ the corresponding price elasticity of the Hicksian demand. Also, denote with $\eta_w^{pub}(w,P)$ the income elasticity of the demand for public provision of and individual with income w. Using Slutsky equation one gets:

$$\frac{dx_2^*(w,P)}{dw}\frac{w}{x_2^*(w,P)} = \underbrace{\frac{\partial x_2^m(w,\tilde{P})}{\partial w}\frac{w}{x_2^m(w,\tilde{P})}[1-p'(w,P)x_2^m(w,\tilde{P})]}_{A}+ \\ +\underbrace{\frac{\partial x_2^h(w,\tilde{P})}{\partial P}\frac{p(w,P)}{x_2^m(w,\tilde{P})}\frac{wp'(w,\tilde{P})}{p(w,P)}}_{B} = \tag{5.19}$$

which rewrites $\eta_w^{pub}(w,P) = \eta_w^M(w,\tilde{P})(1-p'(w,P)x_2^*(w,\tilde{P}))+\eta_P^H(w,\tilde{P})\frac{wp'(w,P)}{p(w,P)}$. Because $\eta_P^H(w,P) < 0$ by the law of compensated demand, this implies that for any tax system in which the tax paid is increasing in income (i.e. $\hat{\tau}'(w) > 0$ which implies $p'(w) >\geq 0$), the income elasticity of the demand for public provision of the good is strictly lower than the income elasticity of the private demand at price $\tilde{P}$. Thus implies that the demand for public provision is going to be decreasing in w unless the good has sufficiently high income elasticity of private demand. Moreover, the higher is the progressivity of the tax system, the larger $p'(w,P)$. Hence if the tax system is progressive, $\eta_w^{pub}(w,P)$ tends to be negative implying that a decrease in the income of the median voter at constant mean will lead to a higher level of public provision in equilibrium. This may lead to paradoxical results. For instance, suppose that all individuals in the economy demand the same amount on the private market, i.e. $x_2^M(w,P) = \bar{x}_2^M(P)$ for all w. Nevertheless, the demand for public provision of the good will be decreasing in the income of the pivotal voter. The intuition is that a larger public provision is desirable for low income voters because implies more redistribution. Because such voters cannot achieve redistribution in cash because of the restrictions in the tax system, they support a larger public provision.

5.7.2 Existence and Uniqueness with opting-out or top-up

Recall voters choose a n-dimensional policy vector and have indirect utility $v(x,w) = \max\{v^n(x,w), v(x,w)\}$ where $v^n(x,w), v^m(x,w)$ are two differentiable functions of x,w and concave in x. Suppose $v_w^n(x,w) - v_w^m(x,w) \leq 0 \forall x,w$. Then for given x there is at most one $\hat{w}$ such that $v^n(x,\hat{w}) = v^m(x,\hat{w})$. This implies the existence of an endogenous threshold in w such that Party A's objective function becomes:

$$V(x^A,x^B,\theta) = \int_{\underline{w}}^{\hat{w}(x^A)}[\mathbb{P}(v^n(x^A,w)-v(x^B,w))]\hat{r}(\theta,w)dw+ \\ +\int_{\hat{w}(x^B)}^{\overline{w}}[\mathbb{P}(v^m(x^A,w)-v(x^B,w))]\hat{r}(\theta,w)dw \tag{5.20}$$

5.7.2.1 Existence

Theorem 1. *(Existence, uniqueness and policy convergence). If there exist $\bar{r}$ and $\bar{p}'$ such that if the distributions $\hat{R}(\theta,w)$ and $\mathbb{P}(d)$ is such that $\hat{r}(\theta,w) \leq \bar{r}$ for all w and $p'[v(x^A,w,n) - v(x^B,w)] \leq \bar{p}'$for some positive $\bar{r}, \bar{p}'$, then (i) there is a unique equilibrium in pure strategies. The unique electoral equilibrium is such that (ii) the two parties choose the same policy.*

Proof. (i) Existence and uniqueness. Sufficient conditions for existence and uniqueness imply $V^A(x,x^B)$ being a (strictly) concave function of x and the inequality constraint $B(x,\theta) \leq 0$ is a continuously differentiable convex function. Define $V^A_{jk}(x^A,x^B)$ an element of the Hessian H_V, i.e. $H_V(j,k) \equiv V^A_{jk}(x^A,x^B) = \frac{\partial^2 V^A(x,x^B)}{\partial x_j \partial x_k}$. Denote with $d^l(x^A,x^B,w) \equiv v^l(x^A,w) - v(x^B,w)$ with $l \in \{n,m\}$. Then one can show that:

$$\begin{aligned} V^A_{jk}(x^A,x^B,\theta) = \; & \frac{\partial \hat{w}}{\partial x_k} p(d(x^A,x^B,\hat{w}))[d^n_j(x^A,x^B) - d^m_j(x^A,x^B,\hat{w})]\hat{r}(\theta,\hat{w}) + \\ & + \int_{\underline{w}}^{\hat{w}(x^A)} \frac{\partial^2}{\partial x_j \partial x_k}[\mathbb{P}(d^n(x^A,x^B,w)]\hat{r}(\theta,w)dw + \\ & + \int_{\hat{w}(x^A)}^{\overline{w}} \frac{\partial^2}{\partial x_j \partial x_k}[\mathbb{P}(d^m(x^A,x^B,w)]\hat{r}(\theta,w)dw \end{aligned} \tag{5.21}$$

Define a matrix $M(x^A,x^B)$ such that each element is $M_{jk}(x^A,x^B) = \frac{\partial \hat{w}}{\partial x_k} p(d(x^A,x^B,\hat{w}))[d^n_j(x^A,x^B) - d^m_j(x^A,x^B,\hat{w})]\tilde{\varepsilon}_j(x,z)\hat{r}(\theta,\hat{w})$ and two matrices $H^l(x^A,x^B,w)$ for $l \in \{n,m\}$ such that each element is $H^l_{jk}(x^A,x^B,w) = \frac{\partial^2 \mathbb{P}(d(x^A,x^B,w))}{\partial x_j \partial x_k}$. Recall that the sum of negative semidefinite matrices is negative semidefinite. Hence one needs $x^T H_V x \leq 0$ for negative semidefiniteness. Using the matrices defined above can be written as:

$$\begin{aligned} x^T H_V x = \; & x^T M(x^A,x^B)x + \int_{w}^{\hat{w}(x)} \left[x^T H^n(x^A,x^B,w)x\right] \hat{r}(\theta,w)dw + \\ & + \int_{\hat{w}(x)}^{\overline{w}} \left[x^T H^m(x^A,x^B,w)x\right] \hat{r}(\theta,w)dw \leq 0 \end{aligned} \tag{5.22}$$

That can be written as $x^T[M(x^A,x^B) + E_w(H^n|w \leq \hat{w}(x))\pi(x) + E_w(H^m|w \geq \hat{w}(x))(1-\pi(x))]x \leq 0$. Define $H^l_v \equiv D^2[d^l(x^A,x^B,w)]$ as the Hessian of individual indirect utility and $\nabla v^n(x^A,w)$ the gradient vector. Following Enelow, Hinrich (1989) for the second and third component of $x^T H_V x$ we need for any $n \times 1$ vector y:

$$\begin{aligned} y^T H^n y = \; & \int_{\underline{w}}^{\hat{w}(x^A)} p'(d(x^A,x^B,w))y^T \left[\nabla v^n(x^A,w)\right] \left[\nabla v^n(x^A,w)\right]^T y\,\hat{r}(\theta,w)dw + \\ & + \int_{\underline{w}}^{\hat{w}(x^A)} p(d(x^A,x^B,w))y^T H^n_v y\,\hat{r}(\theta,w)dw \leq 0 \end{aligned} \tag{5.23}$$

And similarly one can derive $y^T H^n y$. Thus sufficient conditions for uniqueness are

$$\frac{p'(d^l(x^A,x^B,w))}{p(d^l(x^A,x^B,w))} \leq -y^T H_v^l y \left[\left[\nabla d^l(x^A,x^B,w) \right] \left[\nabla d^l(x^A,x^B,w) \right]^T \right]^{-2} \tag{5.24}$$

for $l = n,m$ and for all w and $x^T M(x^A,x^B)x \leq 0$. Notice that as $\mathbb{P}$ becomes close to uniform this condition is equivalent to the matrix $H_v(i)$ to be negative semidefinite, which is equivalent to a concave utility function. But in comparison with Enelow, Hinrich (1989) we have an additional element: $M(x^A,x^B)$. The definition of $\hat{w}(x^A)$ in formula (5.4) implies:

$$v^m(x^A,\hat{w}) = v^n(x^A,\hat{w}) \tag{5.25}$$

Differentiate this w.r.t. x_k and rearrange to get:

$$\frac{\partial \hat{w}(x^A)}{\partial x_k} = -\frac{v_k^m(x^A,\hat{w}) - v_k^n(x^A,\hat{w})}{v_w^m(x^A,\hat{w}) - v_w^n(x^A,\hat{w})} \tag{5.26}$$

Substituting into $M_{jk}(x^A,x^B)$ one gets:

$$\begin{gathered} M_{jk}(x^A,x^B) = \\ p(d(x_A,\hat{t},n))\frac{[v_j(x_A,\hat{t};n)-v_j(x_A,\hat{t};m)]}{v_t(x_A,\hat{t};n)-v_t(x_A,\hat{t};m)} \quad [v_k(x_A,\hat{t};n)-v_k(x_A,\hat{t};m)]\hat{r}(\theta,\hat{w}) \end{gathered} \tag{5.27}$$

Hence $M(x^A,x^B) =$

$$\begin{gathered} M(x^A,x^B) = \\ \frac{p(d(x^A,x^B,\hat{w}))\hat{r}(\theta,\hat{w})}{v_w^n(x^A,\hat{w})-v_w^m(x^A,\hat{w})} \quad \left[\nabla v^n(x^A,\hat{w}) - \nabla v^m(x^A,\hat{w})\right] \left[\nabla v^n(x^A,\hat{w}) - \nabla v^m(x^A,\hat{w})\right]^T \end{gathered} \tag{5.28}$$

Hence the sufficient conditions for existence of a Political Equilibrium are the same as in Enelow, Hinrich (1989), plus the additional condition stated above. Given that $\left[\nabla v^n(x^A,\hat{w}) - \nabla v^m(x^A,\hat{w})\right] \left[\nabla v^n(x^A,\hat{w}) - \nabla v^m(x^A,\hat{w})\right]^T$ is the product of the same vector it is positive semidefinite, hence $x^T M(x^A,x^B)x \leq 0$ for all x is not satisfied under the assumption $v_w^n(x^A,\hat{w}) - v_w^m(x^A, \tilde{}_j(\hat{x},z)\mathfrak{e}) < 0 \; \forall x$. i.e. if individuals with relatively high w choose to enjoy the private provision. On the other hand, it is possible that even if $v_w^n(x^A,\hat{w}) - v_w^m(x^A,\hat{w}) < 0$ the second and the third elements of $V_{jk}^A(x^A,x^B,\theta)$ are sufficiently concave to guarantee concavity of the whole function. Specifically, notice that as the variance of w increases (with $\overline{w} - \underline{w}$ increasing), $\hat{r}(\theta,\hat{w}) \to 0$, thus the conditions for existence become similar to the ones in Enelow, Hinich (1989). Specifically, under the assumption of Section 4 that $W = Z + \Sigma$, with Σ being a uniformly distributed random variable independent of Z, one can find a threshold k such that if

$\bar{\varepsilon} - \underline{\varepsilon} \geq k$ then the conditions are satisfied.

(ii) Policy convergence. Notice that the game described above can be modeled as a zero sum game because the expected plurality for Party B: is equal to $1 - V(x_A, x_B)$. Suppose (x^A, x^B) is an equilibrium strategy with $x^A \neq x^B$ and delivering expected plurality $V(x^A, x^B)$ to party A. Party A can always achieve a certain value $\overline{V}^A$ by playing $x^A = x^B$. Hence if $V(x^A, x^B) < \overline{V}^A$ (a) then x^A cannot be a best response for Party A because it can profitably deviate to $\hat{x}^A = x^B$. If $V^A(x^A, x^B) > \overline{V}^A$ (b), then $V^B(x^B, x^A) = 1 - V^A(x^A, x^B)$. Then x^B cannot be a best response for Party B because it can deviate to $\hat{x}^B = x^A$ and get $V^B(\hat{x}^B, x^A) = 1 - \overline{V}^A$. Inequality (b) implies that this deviation is profitable. Hence in equilibrium it must be true that $V^A(x^A, x^B) = \overline{V}^A$ and $V^B(x^B, x^A) = 1 - \overline{V}^A$ and $x^A = x^B$.

5.7.2.2 Utilitarian outcome

Theorem 2. *(Utilitarian outcome). The policy chosen by both parties in equilibrium is the same as the policy that would be chosen by an omnipotent Benthamite government, i.e.* $x^* = \arg\max_{x \in X} V(x, \theta)$ *where* $V(x, w) = \int_{\underline{w}}^{\overline{w}} v(x, w)\hat{r}(w, \theta)dw.$

Proof. The Lagrangian for this problem is:

$$\begin{aligned} L = \quad & \int_{\underline{w}}^{\hat{w}} [\mathbb{P}(v^n(x, w) - v(x^B, w))] r(\theta, w)dw + \\ & + \int_{\hat{w}}^{\overline{w}} [\mathbb{P}(v^m(x, w) - v(x^B, t))] \hat{r}(\theta, w)dw - \mu B(x) \end{aligned} \tag{5.29}$$

First order conditions are:

$$\begin{aligned} [x_i]: \quad & \int_{\underline{w}}^{\hat{w}} p\left[d^n(x^A, w)\right] v^n_{x_i}(x^A, w) r(\theta, w)dw + \\ & + \int_{\hat{w}}^{\overline{w}} p\left[d^m(x^A, w)\right] v^m_{x_i}(x^A, w)\hat{r}(\theta, w)dw \leq \mu B_{x_i}(x) \end{aligned} \tag{5.30}$$

with respect to each policy dimension x_i and $B(x) \leq 0$ with respect to μ. Hence for any $i \neq j$ such that $L_{x_i} = L_{x_j} = 0$ one gets:

$$\frac{\int_{\underline{w}}^{\hat{w}} p\left[d^n(x^A, w)\right] v^n_{x_i}(x^A, w) r(\theta, w)dw + \int_{\hat{w}}^{\overline{w}} p\left[d^m(x^A, w)\right] v^m_{x_i}(x^A, w)\hat{r}(\theta, w)dw}{\int_{\underline{w}}^{\hat{w}} p\left[d^n(x^A, w)\right] v^n_{x_j}(x^A, w) r(\theta, w)dw + \int_{\hat{w}}^{\overline{w}} p\left[d^m(x^A, w)\right] v^m_{x_j}(x^A, w)\hat{r}(\theta, w)dw} = \frac{B_{x_i}(x)}{B_{x_j}(x)}$$

Notice that at an equilibrium point $x^A = x^B$ (see proof to Theorem 1) hence $d^n(x^A, w) = d^m(x^A, w) = 0 \forall w$ and $\hat{w}(x^A) = \hat{w}(x^B)$. Given that $p(d(x^A, w))$ is independent of w in this case, i.e. $p(d(x^A, w)) = p(0) \neq 0$, then the previous equation becomes:

$$\frac{\int_{\underline{w}}^{\hat{w}} v^n_{x_i}(x^A, w) r(\theta, w)dw + \int_{\hat{w}}^{\overline{w}} v^m_{x_i}(x^A, w)\hat{r}(\theta, w)dw}{\int_{\underline{w}}^{\hat{w}} v^n_{x_j}(x^A, w) r(\theta, w)dw + \int_{\hat{w}}^{\overline{w}} v^m_{x_j}(x^A, w)\hat{r}(\theta, w)dw} = \frac{B_{x_i}(x)}{B_{x_j}(x)} \tag{5.31}$$

which is the same condition that one can derive for the problem:

$$\max_{x\in X, G(x)\leq 0} \int_{\underline{w}}^{\hat{w}} v^n(x,w)\hat{r}(\theta,w)dw + \int_{\hat{w}}^{\overline{w}} v^m(x,w)\hat{r}(\theta,w)dw \tag{5.32}$$

which is the Utilitarian Social Optimum. Notice that for this result to hold it is crucial that the function $\mathbb{P}$ is the same for all income levels w.

5.7.2.3 Comparative Statics

Lemma 3. *(Monotonicity): If there exists at least one x_j with $N \geq j > L$ such that (i) the solution of the maximization problem is interior for x_j, (ii) $b(x,\theta)$ is such that $\frac{\partial b(x,\theta)}{\partial x_j} = \alpha \frac{\partial a(x,\theta)}{\partial x_j}$ for some constant α, and if (iii) $\frac{\partial^2 e_i(x_i,\theta)}{\partial x_i \partial \theta} \geq 0$ (≤ 0), then x_i is weakly increasing (decreasing) in θ in a neighborhood of $x^*(\theta)$.*

Proof. The Lagrangian for the maximization problem is $L = a(x,\theta) + \sum_{i\in L} e_i(x_i,\theta) - \mu[\kappa a(x,\theta) + b(x,\theta) + \sum_{i\in L}\delta(x_i)]$. The F.O.C.s with respect of each $i \leq M$ are given by the following: $e_{1i}(x_i,\theta) - \mu\delta'(x_i) \leq 0$ and for each j such that $M < j \leq N$, they are given by $a_{x_j}(x,\theta) - \mu\kappa a_{x_j}(x,\theta) - \mu b(x,\theta) \leq 0$. Lastly, $\kappa a(x,\theta) + b(x,\theta) + \sum_{i\in L}\delta_i(x_i) \leq 0$. If the solution is interior for at least one j, and $b_{x_j}(x,\theta) = 0$, this implies $\mu = 1/\kappa$, i.e. the Lagrangian multiplier is a constant in a neighborhood of x^*. Now either one gets a corner solution for the i policy dimension, in which case x_i is unaffected by marginal changes in θ , or the solution is interior for such dimension. In the latter case, one can differentiate the F.O.C. with respect to θ to get:

$$\frac{dx_i(0)}{d\theta} = -\frac{1}{L_{x_jx_j}(x,\theta)}\left[L_{x_j\theta} + \sum_j L_{x_ix_j}\frac{dx_i(0)}{d\theta} + \sum_j L_{x_i\mu}\frac{d\mu}{d\theta}\right] \tag{5.33}$$

Notice that additive separability of x_i in the indirect utility and in the budget constraint implies $L_{x_ix_j} = 0$ for all $j \neq i$. Moreover, because μ is constant in a neighborhood of x^*, one gets $\frac{d\mu}{d\theta} = 0$. Lastly, strict concavity of V and convexity of the budget set B imply $L_{x_jx_j}(x,\theta) < 0$, and $L_{x_i\theta} = e_{x_i\theta}(x_i,\theta)$ because B is a constant function of θ. Thus, $sign\left(\frac{dx_i^*(0)}{d\theta}\right) = sign\left(\frac{\partial^2 e(x_i,\theta)}{\partial x_i \partial\theta}\right)$.

5.7.3 Income Inequality and Public Education

This section presents the proofs about the comparative statics in presence of a discrete set of choices on the private market with $n > 1$ elements and of the effects on changes in the level of public provision of education on the features of the income distribution of the next generation.

5.7.3.1 Multiple Discrete Options on the Private Market

Suppose there is more than one option available on the private market $\mathscr{Q}=\{q_1,q_2,...,q_n\}$ and that $k \geq 2$ options are chosen with positive probability at an equilibrium. Denote with $\upsilon^{(j)}(x_i,z+\varepsilon)$ the indirect utility of an individual that purchase option q_j. Define new thresholds $\tilde{\varepsilon}_j(x,z)$ such that $\upsilon^{(j)}(x_i,z+\tilde{\varepsilon}_j(x,z)) = \upsilon^{(j+1)}(x_i,z+\tilde{\varepsilon}_j(x,z))$ for $j=1,2,...,k$. Assume $\upsilon^{(j)}(x_i,w)-\upsilon^{(j+1)}(x_i,w)$ to be decreasing in w, that is, individuals with higher income choose higher levels of quality on the private market. Also, assume that the distribution of ε has enough variance to ensure that for any $z\in[\underline{z},\bar{z}]$ there exists $\tilde{\varepsilon}_j(x,z)$ defined as above for all $\underline{k}\leq j\leq\bar{k}$, where $\underline{k}$ and $\bar{k}$ are the lowest and the highest quality levels that are chosen by a positive share of individuals. Then the objective function becomes:

$$\begin{aligned}\tilde{V}(x,\theta) = & \int_{\underline{z}}^{\bar{z}}\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}_0(x,z)} u(x_i,z+\varepsilon)kd\varepsilon r(z,\theta)dz+ \\ & +\textstyle\sum_{i=\underline{k}}^{\bar{k}}\int_{\tilde{\varepsilon}_{j-1}(x,z)}^{\tilde{\varepsilon}_j(x,z)} \upsilon^{(j)}(x_i,z+\varepsilon)kd\varepsilon r(z,\theta)dz+ \\ & +\int_{\tilde{\varepsilon}_{k-1}(x,z)}^{\bar{\varepsilon}} \upsilon_k(x_i,z+\varepsilon)kd\varepsilon r(z,\theta)dz+E_{z,\varepsilon}(z+\varepsilon)+ \\ & +\int_{\underline{z}}^{\bar{z}}\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} d(x_l,w)kr(z,\theta)d\varepsilon dz+E_{z,\varepsilon}[y(z+\varepsilon)-\tau(x,z+\varepsilon)]\end{aligned} \tag{5.34}$$

And the government budget constraint is unchanged in comparison with the baseline case. In the same way shown in the proof of Theorem 4, one can show that x_i is weakly increasing in θ in a neighborhood of $x^*(\theta)$ if $\int_{\underline{z}}^{\bar{z}}\left[\upsilon_{122}^{(\bar{k})}(x_i,z+\bar{\varepsilon})-u_{122}(x_i,z+\underline{\varepsilon})\right]$ $\left(\int_{\underline{z}}^{z}[G(s)-F(s)]\,ds\right)dz\geq 0$. Hence the comparative statics is not affected in the proximity of the political equilibrium if one analyses a simpler problem in which $\widetilde{\mathscr{Q}}=\{q_j\}$ where $q_j\in\arg\max_{q\in\mathscr{Q}}\upsilon^{(\bar{k})}(x_i,z+\varepsilon)$. Q.E.D.

5.7.3.2 Effects of policy changes on the next generation

Proposition 7. *If $Var(\Sigma)$ is large enough, then (i) the expected income of the next generation $E(w^s)$, (ii) the variance $Var(w^s,\theta)$ and (iii) the coefficient of variation $CV(w^s,\theta)$ of the income distribution of the next generation are all weakly increasing in x_i. Moreover, if the expected marginal returns to public education are weakly decreasing* *in parental income, then (iv) he total effect of a marginal increase in θ on $Var(w^s)$ and $CV(w^s)$ is ambiguous.*

Proof. *(i)* Differentiate $E(w^s)$ with respect to θ. $\frac{\partial E[w^s(x,w)|w]}{\partial x_i} = \rho_{t+1}\left\{\int_{\underline{z}}^{\overline{z}} \frac{\partial \tilde{\varepsilon}(x,z)}{\partial x_i} k\left[H(x_i, z+\tilde{\varepsilon}) - \check{H}(\check{q}, z+\tilde{\varepsilon})\right] + \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \left[H_1(x_i, z+\varepsilon)\right] kd\varepsilon r(\theta, z)dz + \omega s'(x_i)\right\}$ which using the formula for $\frac{\partial \tilde{\varepsilon}(x,z)}{\partial x_i}$ becomes $\frac{\partial E[w^s(x,w)|w]}{\partial x_i} = \rho_{t+1}\left[\int_{\underline{z}}^{\overline{z}} -kH_1(x_i, z+\tilde{\varepsilon}) + \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} H_1(x_i, z+\varepsilon)kd\varepsilon r(\theta, z)dz + \omega s'(x_i)\right]$. Notice that $\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \left[H_1(x_i, z+\varepsilon)\right] kd\varepsilon r(\theta, z)dz + \omega s'(x_i) > 0$ for all x if $s'(x_i) \geq 0$. Recall $Var(\Sigma) = (\bar{\varepsilon} - \underline{\varepsilon})^2/12$. As $Var(\Sigma) \to \infty$ it must be true that $k = 1/(\bar{\varepsilon} - \underline{\varepsilon}) \to 0$. Thus, because $H_1(x_i, z+\tilde{\varepsilon})$ is finite by assumption, then there exists $\hat{k} \in [0, \infty)$ such that if $k \leq \hat{k}$ then $\frac{\partial E[w^s(x,w)|w]}{\partial x_i} \geq 0$. *(ii)* The derivative of the variance the income of the next generation $Var(w^s, \theta)$ is given by the following:

$$\frac{\partial Var(w^s)}{\partial x_i} = \underbrace{\int_{\underline{z}}^{\overline{z}} \frac{\partial \tilde{\varepsilon}}{\partial x_i} k\left[H(x_i, z+\tilde{\varepsilon})^2 - \check{H}(\check{q}, z+\tilde{\varepsilon})^2\right]}_{A}$$
$$+2\underbrace{\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \left[H(x_i, z+\varepsilon)H_1(x_i, z+\varepsilon) - E(u)\frac{\partial E(w^s)}{\partial x_i}\right] kd\varepsilon g(z)dz}_{B} = \tag{5.35}$$

which rewrites as follows:

$$\frac{\partial Var(w^s)}{\partial x_i} = \underbrace{\frac{\partial \tilde{\varepsilon}}{\partial x_i} k\left[H(x_i, z+\tilde{\varepsilon})^2 - \check{H}(\check{q}, z+\tilde{\varepsilon})^2\right]}_{A}$$
$$\underbrace{\left[H(x_i, z+\tilde{\varepsilon}) + \check{H}(\check{q}, z+\tilde{\varepsilon}) - 2E(w^s)\right]}_{A} + \tag{5.36}$$
$$+2\underbrace{\int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} H_1(x_i, z+\varepsilon)\left[H(x_i, z+\varepsilon) - E(w^s)\right] kd\varepsilon g(z)dz}_{B}$$

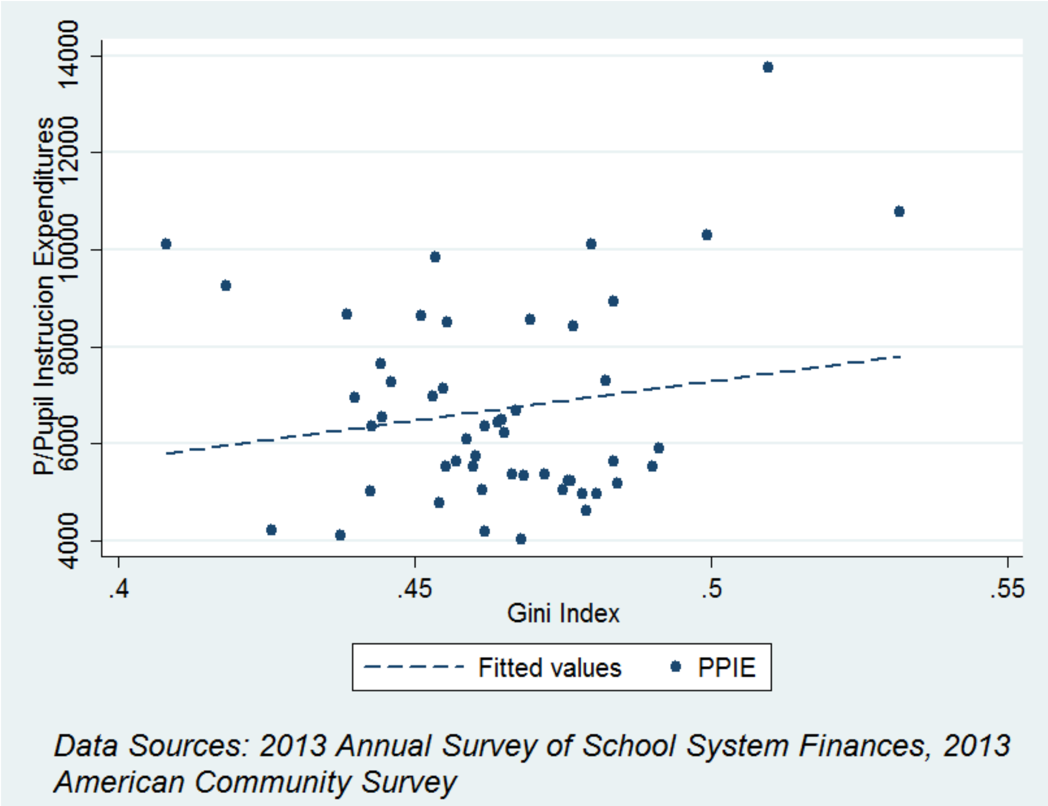
Thus, $B \leq \int_{\underline{z}}^{\overline{z}} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} H_1(x_i, z+\varepsilon)] \left[H(x_i, z+\varepsilon) - E(w^s)\right] kd\varepsilon g(z)dz$ if $z + \tilde{\varepsilon}(x_i, z) \geq E(z+\varepsilon)$ and H is concave in w, because of Jensen's inequality. Lastly, recall $H_1(x_i, w)$ is decreasing in w, which implies:

$$\int_{\underline{z}}^{\bar{z}} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} H_1(x_i, z+\varepsilon)] \left[H(x_i, z+\varepsilon) - E(w^s)\right] k d\varepsilon g(z) dz \leq$$
$$\leq E_{z,\varepsilon}[H_1(x_i, z+\varepsilon)] \int_{\underline{z}}^{\bar{z}} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[H(x_i, z+\varepsilon) - E(w^s)\right] k d\varepsilon g(z) dz = 0 \quad (5.37)$$

Thus B is negative. The sign of A is ambiguous, but the magnitude tend to zero as k becomes large (i.e. the variance increases). Thus the variance of the income distribution of the next generation is decreasing in x_i. (iii) Regarding the coefficient of variation notice that it is defined as $CV(w^s) = \sqrt{Var(w^s)}/E(w^s)$. Recall from Proposition 5 that $E(w^s)$ is increasing in x_i for sufficiently small k. Thus it is decreasing in x_i if $Var(w^s)$ is. (iv) Notice that $\frac{dVar(w^s)}{d\theta} = \frac{\partial Var(w^s)}{\partial\theta} + \frac{\partial Var(w^s)}{\partial x_i}\frac{dx_i}{d\theta}$. The first part $\frac{\partial Var(w^s)}{\partial\theta}$ may be positive, but as the second part has negative sign if $MRE(x,w)$ is decreasing in w for all x, w, then the sign of the total effect is ambiguous.

Figures

Figure 5.1. Income Inequality vs Public Spending in Education in US States



6 Conclusions

This thesis provides a number of important contributions to the theoretical analysis of how electoral competition shapes public policy in democratic countries. The first contribution is methodological. It consists of a simple theoretical tool that allows one to characterize the policy outcome of an electoral process and to assess the relationship between such outcome and some features of the populations of voters. Specifically, in chapter 2 I propose a model of electoral competition in the spirit of the one proposed by Levy (2004) and an equilibrium concept to tackle an important problem that affects the traditional literature. That is, traditional models lack predictive power with respect to the policy outcome if voters face a multidimensional choice domain, because a *Condorcet winner* usually fails to exist in such case. The tool I propose can deal with such cases and - under suitable restrictions - delivers a sharp characterization of the policy that is implemented in equilibrium. The second contribution consists in employing the new theoretical tool to address two popular questions in the literature. In detail, in chapter 3 I study the relationship between income inequality and size of the government. I augment the model proposed by Meltzer and Richard (1987) - in which the government redistributes income uniquely through a lump-sum grant - introducing a second endogenous spending policy. The second policy dimension is the amount of a public good provided by the government. I show that the positive relationship between income inequality and total size of the government that prevails in the traditional analysis does not survive in this augmented model. Moreover, the prediction has opposite sign if - in the proximity of an equilibrium - the degree of progression of the tax system is not too strong. I also show that the degree of progression of the tax system - and not the total size of the government - is increasing in income inequality in this economic environment. Lastly, I show that the predictions of this augmented model are more consistent with the result in the empirical literature, in comparison with the one in Meltzer and Richard's paper. In chapter 4 I study the effect of population ageing on immigration policies and fiscal policies in a simple overlapping generation model. Similar attempts in the literature abstract from the endogeneity of the fiscal policy in order to avoid the problems induced by a multidimensional choice domain. Thus, such studies overlook an important economic channel. That is, in modern democracies citizens choose - with their voting behavior - not only the immigration policy, but also how costs and benefits of immigration are divided up among the natives through taxes and public spending. Existing papers about the political economy of immigration policies typically assume that fiscal policies are exogenous to the political process. As a result, such papers derive predictions that are uniquely driven by the assumptions regarding the exogenous structure of the fiscal system, which determines how the net gains from immigration are allocated to different groups of voters. Unsurprisingly, these analyses deliver predictions about the determinants of immigration policies that are often empirically controversial. In my analysis I allow voters to choose both the immigration policy - in the form of a quota - and the fiscal policy. I find that - in the model - the elderly and the low income voters are relatively more hostile to immigration and favorable to public spending relative to the young and the high income. The intuition is simple. On one hand, the elderly and the poor have stronger preferences for large public spending, because they are less affected by the high income taxes needed to finance such policy. On the other hand, they are always better off by financing such public spending through high tax rates - that affects mostly the young and high income individuals - rather than by increasing the tax base by allowing more immigrants. Such preference patterns are consistent with most empirical evidence based on survey data. Because of this channel, population ageing - which implies an increase in the political power of the elderly - translates into more restrictive immigration policies, more public spending and higher tax rates. Lastly, in chapter 5 I analyse a problem that requires a different theoretical approach. Specifically, I study the relationship between income inequality and public spending in schooling, in a model of parental investment in education. I allow voters to redistribute income directly - through a lump-sum grant - or indirectly - investing in the future productivity of their children. Thus, the policy space is two-dimensional. Moreover, I allow parents to *opt-out* of the public school system and choose private education. The latter assumption is known to generate non-convexities in voter preferences, which often imply problems of existence of a *Condorcet winner* in traditional models (see Stiglitz, 1974). Because of such issue, the framework proposed in chapter 2 of this thesis cannot be successful in characterizing the equilibrium policy outcome. Thus, I employ a probabilistic model of electoral competition to derive the direction of the relationship of interest. I find that, if a rise in the quality of public education implies a reduction in the degree at which income is transmitted from parents to children, then a mean-preserving spread in the income distribution translates into higher public spending in education per pupil. This result suggests that, if public education helps in reducing future income inequality, then public intervention in education should be particularly large in economies that experience very high level of inequality.

A common goal of chapters 3, 4 and 5 is to show that the interaction among multiple endogenous policy dimensions is often crucial in shaping the trade-offs faced by voters. Because of that, the theoretical approach proposed in chapter 2 and - to a certain extent -

the one employed in chapter 5, may prove useful to revisit many other questions that have been studied in the literature abstracting from such interaction. A typical example - extensively mentioned in the previous chapters - is the question of whether wealth inequality affects the relative tax rates on labour and capital income. The attempts in the literature to answer such intrinsically two-dimensional question rely on very strong restrictions on the policy space (Benhabib and Bassetto 2006), and because of that, such attempts have delivered a very limited set of predictions. In other cases, the analysis of two-dimensional policies is performed imposing very strong restrictions on voter preferences (Borge and Ratsø 2004). The latter approach gives rise to questions about the robustness of the results if a more flexible theoretical environment is assumed, and is extremely difficult to apply to other similar questions. Thus, the theoretical tools proposed in this thesis represent an opportunity to derive new testable predictions from these existing models. Another interesting feature of the approach presented in chapter 2 is that it is based on a set of preference restrictions borrowed from the literature about *monotone comparative statics* and *supermodular games* (Milgrom and Shannon 1994). This implies that the model of electoral competition I propose may admit extensions to analyse other interesting questions, such as the ones regarding the role played by the interaction among different constituencies, or countries, or elective bodies in shaping public policies. The attempt to extend the baseline framework to tackle such questions is likely to be successful whenever the policy choices made by voters in different constituencies are *strategic complements*. Because of that, it may prove useful to extend the analysis of the determinants of immigration policies. Strategic complementarities may emerge in such analysis if the choices of a community regarding the immigration policy and the size of welfare spending affect voters' trade-offs in neighbour countries. In such case, the proposed extension of the theoretical framework may prove useful to analyse the cross-country interdependencies in the choice of immigration policies, and to assess the cost and benefits of policy coordination across different states.

On one hand this thesis represents an important contribution to the analysis of how electoral processes shape public policies. On the other hand, several questions remain open. On the theoretical side, my analysis focuses on two particular cases in which traditional models of electoral competitions usually fail to deliver a sharp characterization of the policy outcome. Specifically, I study Political Economy questions that require the analysis of voting over multidimensional choice domains and/or in presence of non-convexities in the voter objective function. Throughout this thesis, I focus on cases in which voters can be ordered along a single preference dimension, such as income, productivity, age, etc., or on cases in which multiple preference dimensions can be shown to reduce to a single one. Conversely, several interesting cases studied in the literature (e.g. Levy 2005, Lee and Roemer 2006) are characterized by voters that differ in multiple orthogonal dimensions. In such cases it is not easy to derive a general analytical characterization of the policy outcome that prevails in equilibrium using the tools available in the literature. Moreover, the framework proposed in this thesis does not seem promising to achieve such goal, because of the strong ordinal preference restrictions that underpin the monotone comparative statics result. Thus, further research is needed in order to obtain a sufficiently general theoretical environment to study this kind of questions. On the empirical side, the analysis in this thesis provides new testable predictions about the determinants of some of the most important policies choices faced by modern democratic countries, namely (i) fiscal policies, (ii) immigration policies and (iii) public spending in education. These predictions represent a challenge for future empirical research. On one hand there is a vast body of empirical evidence regarding the determinants of fiscal policies and public spending, including specific studies about public intervention in education (see chapter 3 and 5). Thus, the test of the new predictions regarding these two types of policies can start from a solid base in terms of data availability and methodology. On the other hand, an empirical analysis of the relationship between population ageing and immigration policies may prove a difficult task. The issue of measurement of the outcome variable and the many potential sources of endogeneity described in chapter 4 represent a big challenge for future empirical research. If successful, such analysis would help to shed light on one of the most important - and somewhat controversial - objects of political debate in many democratic countries.

In conclusion, this thesis provides both an important contribution to the literature in Political Economy, and a set of promising inputs for future theoretical and empirical research.

Bibliography

[1] Abdullah, A., H. Doucouliagos, H., E. Manning, E. (2015). “Does education reduce income inequality? A meta-regression analysis”, *Journal of Economic Surveys*, 29 (2), 301-316.

[2] Alesina, A., Baqir, R. and W. Easterly, W. (1999). “Public Goods And Ethnic Divisions”, *The Quarterly Journal of Economics*, 114(4), 1243-1284.

[3] Alesina, A., E. Glaeser, and B. Sacerdote (2001). “Why Doesn’t the US Have a EuropeanStyle Welfare System?”, *Brookings Papers on Economic Activity*, (2), 187-277.

[4] Altonji, J.G. and Dunn, T.A. (1996). “The Effects of Family Characteristics on the Return to Education”, *The Review of Economics and Statistics*, MIT Press, vol. 78(4), 692-704.

[5] Aranson, P., Hinich, M., and Ordeshook, P. (1974). “Election goals and strategies: Equivalent and nonequivalent candidate objectives”, *American Political Science Review*, 68: 135-152.

[6] Arrow, K.J., (1950). “A Difficulty in the Concept of Social Welfare”, *Journal of Political Economy,* 58(4), 328–346.

[7] Banks, J.S., and J. Duggan (2005). “Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-Motivated Candidates,” in: Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks, ed. David Austen-Smith and John Duggan, 15-56. New York: Springer.

[8] Banks, J., and Emmerson, C. (2003). “Public and private pension spending: Principles, practice and the need for reform,” in: The economics of public spending, ed. D. Miles, G. Myles and I. Preston, Oxford University Press, 2003.

[9] Bassett, W.F., Burkett, J.P. and Putterman, L. (1999). “Income distribution, government transfers, and the problem of unequal influence”, *European Journal of Political Economy*, 15, 207-28.

[10] Bassetto, M. and Benhabib, J. (2006). “Redistribution, Taxes, and the Median Voter,” *Review of Economic Dynamics*, vol.9, n.2, 211- 223.

[11] Becker, G.S., Tomes, N. (1986). "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, University of Chicago Press, vol. 4(3), pages S1-39, July.

[12] Becker, G.S., Kominers, S.D., Murphy, K.M., Spenkuch, J.L. (2015). "A Theory of Intergenerational Mobility". Working Paper.

[13] Bénabou, R. (1996). "Inequality and Growth", CEPR Discussion Papers 1450, C.E.P.R. Discussion Papers.

[14] Bénabou, R. (2000). "Unequal Societies: Income Distribution and The Social Contract", *The American Economic Review*, 90(1), 96–129.

[15] Ben-Gad, M. (2012). "On deficit bias and immigration", City University London Department of Economics Discussion Paper Series No 12-09.

[16] Bergson, A. (1976). "Social Choice and Welfare Economics under Representative Government", *Journal of Public Economics*, 6, 171-190.

[17] Besley, T., Coate, S. (1991). "Public provision of private goods and the redistribution of income," *American Economic Review*, 81, 979–984.

[18] Besley, T., Coate, S. (1997). "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 108(1), 85-114.

[19] Birkhoff, H. (1967). Lattice Theory. Vol. 25 of Colloquium publications. American Mathematical Soc.

[20] Black, D. (1948). "On the Rationale of Group Decisionmaking", *Journal of Political Economy*, 56: 23-34.

[21] Boeri, T. and Brücker, H. (2005). "Why are Europeans so tough on migrants?," *Economic Policy*, vol. 20(44), pages 629-703, October.

[22] Borge, L.-E. and Rattsø, J.J. (2004). "Income distribution and tax structure: Empirical test of the Meltzer-Richard hypothesis," *European Economic Review*, Elsevier, vol. 48(4), pages 805-826, August.

[23] Borjas, G. (1999). "The economic analysis of immigration", in (O. Ashenfelter and D. Card, eds.), Handbook of Labor Economics, pp. 1697-1760, vol. 3A, chap. 28, Amsterdam: Elsevier.

[24] Boustan, L., Ferreira, F., Winkler, H., and Zolt, E.M. (2010). "Inequality and Local Government: Evidence from U.S. Cities and School Districts, 1970-2000", Law and Economics Workshop, Berkeley Program in Law and Economics, UC Berkeley.

[25] Brenner, D., and Rubinstein, Y. (2012). The Returns to Education and Family Income. Working Paper.

[26] Brook, L., Hall, J., and Preston, I. (1998). "What Drives Support for Higher Public Spending?" in: P. F. Taylor-Gooby ed., "Choice and Public Policy", London: Macmillan.

[27] Card, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, Vol. 69, No. 5., pp. 1127-1160.

[28] Card, D., Dustmann, C. and Preston, I. (2012). "Immigration, wages and compositional amenities", *Journal of the European Economic Association*, vol. 10, pp. 78-119.

[29] Condorcet, M. de (1785). Essai sur l'application de l'analyse á la probabilité des décisions rendues à la pluralité des voix, Paris.

[30] Coons, J. E., Clune, W. H., and Sugarman, S. D. (1970). Private Wealth and Public Education. Cambridge, MA: Harvard Univ. Press.

[31] Corcoran, S. and W. N. Evans (2009). "Income Inequality, the Median Voter, and the Support for Public Education", Department of Economics, University of Notre Dame.

[32] Davis, O.A., DeGroot, M.H., Hinich, M.J. (1972). "Social Preferences Orderings and Majority Rule", *Econometrica*, Vol. 40, No. 1: 147-157.

[33] de la Croix, D. and M. Doepke (2009). "To Segregate or to Integrate: Education Politics and Democracy". *Review of Economic Studies*, 76 (2), 597–628.

[34] De Mello, L. and Tiongson, E.R. (2006), "Income Inequality and Redistributive Government Spending", *Public Finance Review*, 34, 282-305.

[35] Downs, A. (1957). An Economic Theory of Democracy. New York: Harper Collins.

[36] Dotti, V. (2014). "The Political Economy of Publicly Provided Private Goods", Working Paper.

[37] Dotti, V. (2015). "Generalized Comparative Statics for Voting Models", Working Paper.

[38] Dustmann, C. and Preston, I. (2006). "Is immigration good or bad for the economy? analysis of attitudinal responses", *Research in Labor Economics*, vol. 24, pp. 3-34.

[39] Dustmann, C. and Preston, I. (2007). "Racial and economic factors in attitudes to immigration", *Berkeley Electronic Journal of Economic Analysis and Policy*, vol. 7, Art. 62.

[40] Dustmann, C. and Frattini, T. (2014), "The Fiscal Effects of Immigration to the UK," *The Economic Journal*, 124: F593–F643.

[41] Enelow, J.M., Hinich, M.J. (1989). "A general probabilistic spatial theory of elections", *Public Choice,* 61: 101-113.

[42] Epple, D., Romano, R.E. (1996a). "Public provision of private goods", *Journal of Political Economy*, 104, 57–84.

[43] Epple, D. and Romano, R.E. (1996b), "Ends against the middle: Determining public service provision when there are private alternatives", *Journal of Public Economics*, Elsevier, vol. 62(3), pages 297-325, November.

[44] Easterly, W. and Rebelo, S. (1993). "Fiscal policy and economic growth: An empirical investigation", *Journal of Monetary Economics*, Elsevier, vol. 32(3), pages 417-458, December.

[45] Facchini, G. and Mayda, A.M., (2008). "From individual attitudes towards migrants to migration policy outcomes: Theory and evidence," *Economic Policy*, vol. 23, pages 651-713, October.

[46] Facchini, G. and Mayda, A.M., (2009). "Does the Welfare State Affect Individual Attitudes toward Immigrants? Evidence across Countries," *The Review of Economics and Statistics*, MIT Press, vol. 91(2), pages 295-314, May.

[47] Fernández, R. and Rogerson, R. (1995). "On the Political Economy of Education Subsidies", *Review of Economic Studies*, 62 (2), 249–262.

[48] Galasso, V. and Profeta, P. (2004). "Lessons for an ageing society: the political sustainability of social security systems," *Economic Policy*, CEPR;CES;MSH, vol. 19(38), pages 63-115.

[49] Gibbard, A. (1973). "Manipulation of voting schemes: a general result", *Econometrica,* 41 (4), 587–601.

[50] Glomm, G. and Ravikumar, B. (1998). "Opting Out of Publicly Provided Services: A Majority Voting Result", *Social Choice and Welfare*, 15 (2), 197–199.

[51] Glomm, G. and Ravikumar, B. (2003). "Public education and income inequality", *European Journal of Political Economy*, Vol. 19 289 –300.

[52] Glomm, G. (2004). "Inequality, Majority Voting And The Redistributive Effects Of Public Education Funding", *Pacific Economic Review*, 9(2), 93-101.

[53] Goldin, C. and L. F. Katz (1997). "Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940", NBER Working Papers.

[54] Gouveia, M. (1997). “Majority rule and the public provision of health care”, *Public Choice*, 93, 221–244.

[55] Gouveia, M., and Masia, N.A. (1998). “Does the median voter explain the size of government? Evidence from the states”, *Public Choice*, 97, 159-77.

[56] Grandmont, J.-M. (1978). “Intermediate preferences and the majority rule”, *Econometrica*, 46, 317–330.

[57] Haupt, A. and Peters, W. (1998). “Public Pensions and Voting on Immigration,” *Public Choice*, Springer, vol. 95(3-4), pages 403-13, June.

[58] Hinich, M. J., J. Ledyard, and P. C. Ordeshook (1972). “A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games”, *Journal of Politics* 35, 154-93.

[59] Haupt, A. and Peters, W. (1998). “Public Pensions and Voting on Immigration,” *Public Choice*, Springer, vol. 95(3-4), pages 403-13, June.

[60] Hotelling, H. (1929). “Stability in Competition”, *The Economic Journal*, Vol. 39, 41–57.

[61] Kemnitz, A. (2003). “Immigration, unemployment and pensions”, *Scandinavian Journal of Economics*, vol. 105, pp. 31-47.

[62] Krieger, T. (2003). “Voting on low-skill immigration under different pension regimes”, *Public Choice*, vol. 117, pp. 51-78.

[63] Krouwel, A. (2012). Party tranformations in European democracies, SUNY Press: Albany, NY.

[64] Lambert, P.J. (1989). The Distribution and Redistribution of Income: A Mathematical Analysis. First edition. Oxford: Basil Blackwell.

[65] Lee, W. & Roemer, J. E. (2006). “Racism and redistribution in the United States: A solution to the problem of American exceptionalism”, *Journal of Public Economics*, Elsevier, vol. 90(6-7), 1027-1052.

[66] Leers, T., Meijdam, L. and Verbon, H.A.A. (2003). “Ageing, migration and endogenous public pensions”, *Journal of Public Economics*, vol. 88, pp. 131-159.

[67] Levy, G. (2004). “A model of political parties”, *Journal of Economic Theory*, 115, 250-277.

[68] Levy, G. (2005). “The Politics of Public Provision of Education”, *The Quarterly Journal of Economics*, Vol. 120, No. 4, pp. 1507-1534.

[69] Lind, J.T. (2007). “Fractionalization and the size of government”, *Journal of Public Economics*, 91(1-2), 51-76.

[70] Lindbeck A., Weibull, J.W. (1987). “Balanced-budget redistribution as the outcome of political competition”, *Public Choice*, 52, Number 3, 273-297.

[71] Luttmer, E.F.P. (2001). “Group Loyalty and the Taste for Redistribution”, *Journal of Political Economy*, 109(3), 500-528.

[72] Maskin E. and Dasgupta P. (2008). “On the Robustness of Majority Rule”, *Journal of the European Economic Association*, 6 (5), 949-973.

[73] Meltzer, A.H., Richard, S.F. (1981). “A Rational Theory of the Size of Government”, *The Journal of Political Economy*, 89, Issue 5, 914-927.

[74] Meltzer, A.H., Richard, S.F. (1983). “Tests of a rational theory of the size of government”, *Public Choice*, 41, 403-418.

[75] Milanovic, B. (2000). “Do more unequal countries redistribute more? Does the median voter hypothesis hold?”, *European Journal of Political Economy*, 16, 367-410.

[76] Milgrom, P., Shannon, C. (1994). “Monotone Comparative Statics”, *Econometrica*, Vol. 62(1): 157-180.

[77] Musgrave, R.A. and Thin, T. (1948). “Progressive taxation in an inflationary economy”, *Journal of Political Economy*, 56, 498-514.

[78] Naito, K., Nishida, K. (2012). “The Effects of Income Inequa lity on Education Policy and Economic Growth”, *Theoretical Economics Letters*, 2, 109-113

[79] Nash, J.F. (1950). “The bargaining problem”, *Econometrica*, 18(2), 155-162.

[80] OECD (2015), Pension spending (indicator). doi: 10.1787/a041f4ef-en (Accessed on 25 August 2015).

[81] OECD (2015b), Public spending on education (indicator). doi: 10.1787/f99b45d0-en (Accessed on 25 August 2015).

[82] Ordershook, P.C. (1986). Game Theory and Political Theory: an Introduction. Cambridge University Press.

[83] Ortega, F. and Peri, G. (2009). “The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005,” NBER Working Papers 14833, National Bureau of Economic Research, Inc.

[84] Osborne, M.J., Slivinski, A. (1996). "A Model of Political Competition with Citizen-Candidates", *The Quarterly Journal of Economics*, Vol. 111(1), 65-96.

[85] Panizza, U. (1999). "Income inequality and economic growth: Evidence from American data", Inter-American Development Bank Working Paper WP-404, Washington, DC.

[86] Perotti, R. (1992). "Income distribution, politics, and growth", *American Economic Review*, 82, 311-316.

[87] Perotti, R. (1994). "Income distribution and investment", *European Economic Review*, 38, 827-35.

[88] Perotti, R. (1996). "Growth, income distribution, and democracy: What the data say", *Journal of Economic Growth*, 1, 149-87.

[89] Persson, T. and G. Tabellini (1994). "Is Inequality Harmful for Growth?", *American Economic Review*, 84 (3), 600–621.

[90] Persson, T. and Tabellini, G. (2002). "Political economics and public finance," in: A. J. Auerbach & M. Feldstein (ed.), Handbook of Public Economics, edition 1, volume 3, chapter 24, pages 1549-1659 Elsevier.

[91] Piketty, T. (1995). "Social Mobility and Redistributive Politics", *Quarterly Journal of Economics*, 110(3), 551-584.

[92] Plott, C.R. (1967). "A notion of equilibrium and its possibility under majority rule", *American Economic Review*, 57 (Sept.): 787-806.

[93] Poole, K.T. and Rosenthal, H. (1997). Congress: A Political-Economic History of Roll Call Voting. New York: Oxford University Press.

[94] Preston, I. (2014), "The Effect of Immigration on Public Finances," *The Economic Journal*, 124: F569–F592.

[95] Quah, J.K.-H. (2004). "The Comparative Statics of Constrained Optimization Problems", Working Paper.

[96] Quah, J.K.-H. (2007). "The Comparative Statics of Constrained Optimization Problems", *Econometrica*, Vol. 74, No. 2: 401-431.

[97] Ray, D., Vohra, R. (1997). "Equilibrium Binding Agreements", *Journal of Economic Theory*, 73, 30-78.

[98] Razin, A. and Sadka, E. (1999). "Migration and Pension with International Capital Mobility," *Journal of Public Economics* 74(1): 141-150.

[99] Razin, A. and Sadka, E. (2000). "Unskilled Migration: A Burden or a Boon for the Welfare State," *Scandinavian Journal of Economics* 102(3): 463-479.

[100] Razin, A., Sadka, E. and Swagel, P. (2002). "The Aging Population and the Size of the Welfare State", *Journal of Political Economy*, 110(4), 900-918.

[101] Razin, A., Sadka, E., and Suwankiri, B. (2011). Migration and the Welfare State: Political-Economy Policy Formation. MIT Press.

[102] Razin, A., Sadka, E. and Suwankiri, B. (2014). "The Welfare State and Migration: a Dynamic Analysis of Political Coalitions", NBER Working Paper Series, WP 20806.

[103] Riker, W.H. (1982). "The Two-Party System and Duverger's Law: An Essay on the History of Political Science", *The American Political Science Review*, Vol. 76, No. 4, 753-766

[104] Rodriguez, F.C. (1999). "Does distributional skewness lead to redistribution? Evidence from the United States", *Economics and Politics*, 11, 171-99.

[105] Roemer, J.E. (1999). "The Democratic Political Economy of Progressive Taxation", *Econometrica*, 67, 1-19.

[106] Roemer, J.E. (2001). Political Competition: Theory and Applications Harvard University Press, Boston, MA.

[107] Saint-Paul, G. (1994). "The dynamics of Exclusion and Fiscal Conservatism", Centre for Economic Policy Research (London), Discussion Paper No. 998.

[108] Satterthwaite, M.A. (1975). "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", *Journal of Economic Theory*, 10: 187–217.

[109] Scholten, U. and Thum, M. (1996). "Public pensions and immigration policy in a democracy", *Public Choice*, vol. 87, pp. 347-361.

[110] Soares, J. (1998). "Altruism and Self-interest in a Political Economy of Public Education", IGIER working paper, Bocconi University.

[111] Stiglitz, J.E. (1974). "The demand for education in public and private school systems", *Journal of Public Economics*, 3, 349–385. 61

[112] Sylwester, K. (2000). "Income inequality, education expenditures, and growth", *Journal of Development Economics*, Vol. 63, 379–398.

[113] Sylwester, K. (2002). "Can education expenditures reduce income inequality?", *Economics of Education Review*, 21, 43–52.

[114] Tavits, M. (2006) "Party system change: testing a model of new party entry", *Party Politics* 12(1), 99–119.

[115] Tullock, G. (1967). "The General Irrelevance of the General Impossibility Theorem," *The Quarterly Journal of Economics,* Vol. 81, No. 2, pp. 256-270.

[116] United Nations (2000). Replacement migration: Is it a solution to declining and ageing populations? New York: UN Population.

[117] U.S. Census Bureau (2013). Annual Survey of School System Finances.

[118] Usher, D. (1977). "The welfare economics of the socialization of commodities", *Journal of Public Economics*, Elsevier, vol. 8(2), pages 151-168, October.