Public Education for the Minority, Private Education for the Majority*

Gilat Levy
London School of Economics and Political Science
and Tel Aviv University

Contents:

Discussion Paper
No.TE/04/470
March 2004

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7955 6698

---

* I am thankful to Tim Besley who suggested to me to work on this problem. I thank Tim Besley and Raquel Fernandez for their comments, and the Sapir Center for Development for financial support.

Abstract

Public provision of private goods such as education is usually viewed as a form of redistribution in kind. However, does it arise when income redistribution is feasible as well? In this paper I analyse a two-dimensional model of political decision making. Society has to choose both the tax rate and the allocation of the revenues between income redistribution and public provision of education. The political process that I analyse involves endogenous parties. Parties have a unique role in the model; I assume that parties increase the commitment ability of politicians and, as a result, increase the ability of different groups in society to compromise with one another. I find that public provision of education arises as an anti-majoritarian outcome; public provision of education arises only when those who benefit from education, e.g., voters with children, are a minority. The reason is that when education is consumed only by a minority, such redistribution in kind is 'cheap' relative to universal income redistribution, i.e., it can be effectively provided even with low taxes. Public provision of education arises then as a political compromise offered by the party of the poor who benefit from education and the rich voters who prefer low taxes. Thus, when those who benefit from education are a minority, it is publicly provided. When those who benefit from education are a majority, they have to buy private education, since there is no public provision of this good.

Keywords: Education, redistribution, political parties.
JEL Nos.: H42, H52, D72.

Contact address: Dr Gilat Levy, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Email: g.levy1@lse.ac.uk

1 Introduction

Economists have long been puzzled by the question of public provision of private goods, such as education.[1] In the normative literature, the reasons that are put forward for government intervention in the provision of education are externalities or other market failures such as imperfect information. In the positive analysis though, the focus is on the view that public provision of education arises as a form of *redistribution*. For example, Epple and Romano (1996a) or Glomm and Ravikumar (1998) view it as redistribution from the rich to the poor since the poor do not have enough means to finance private education.[2] In the context of high education, Fernandez and Rogerson (1995) show that public provision of education is actually a form of redistribution from the poor to the rich, where the former are financially constrained from attending universities. Gradstein and Kaganovich (2003) perceive government provision of education as a redistribution from the old (who do not benefit from education) to the young (whose future income is positively correlated with education).

All these papers analyze models in which the unique possible form of redistribution available to society is redistribution in kind, i.e., public provision of education. However, income redistribution may be a more efficient tool for shifting resources from one group of voters to the other, so that it can substitute for redistribution in kind while creating a Pareto improvement.[3] By disregarding income redistribution as a possible policy tool, these descriptive models may predict an excessive level of public provision of education.

This paper engages in the a positive analysis of public provision of education. But in contrast to previous literature cited above, I allow society to use income redistribution as a possible policy tool. Questions that arise in this context are as follows. When income redistribution is feasible, is education publicly provided as well? What are the factors determining the level of public provision of education, if it is provided, and how is it related to the distribution of preferences in society, or to parameters such as income inequality? What is the size of government, i.e., its aggregate resources or the total amount of both redistribution in kind and income redistribution, when society can choose both the size of the government and how to target its resources? These are the questions I address in the paper.

[1]Other examples are health care, police protection or refuse collection. For an argument why education should be considered as a private and not a public good, see Barzel (1973).

[2]This is also the view in the normative work of Besley and Coate (1991).

[3]This is clearly the case in the absence of labour supply incentives.

I analyze a two-dimensional political-economy model. I assume that agents are differentiated according to their income (with the poor being the majority) as well as according to their attitudes towards education, i.e., some agents benefit from education, where some do not. For example, it is usually perceived that education can increase future income. Thus, old or retired agents are considered as those who do not benefit from education whereas agents with children do benefit from education. Education is consumed only by those who benefit from it. Following Epple and Romano (1996a), I assume that such agents can also increase their consumption of education by buying private education. Society chooses, through a political process, both the size of the government and how to allocate its resources between redistribution in kind (public provision of education) and income redistribution. The political process that I analyze has realistic institutional features; it allows both for endogenous entry of politicians and for endogenous political parties.

The key finding is that public provision of education exists only when those who care for education (and would therefore consume it) are a *minority*. The intuition is as follows. When public education is consumed only by the few, i.e., a minority, the effective price of publicly financing education is low compared to the resources needed to finance a universal income redistribution. Public provision of education is relatively cheap in the sense that even a low tax rate can provide a generous per capita level of education. It then arises as a political compromise between the rich who vouch for low taxes, and the poor who indeed benefit from education, but cannot afford to buy it privately. The policy they offer reduces the size of government but targets most of its resources to public provision of education. This registers wide enough support in the electorate as it pleases the rich who are 'rewarded' with a relatively high income as well as the poor who benefit from education.

On the other hand, when those who benefit from education are a majority, there would be no public provision of education; the rich would view such redistribution in kind as too expensive and inefficient, and would cooperate with the poor who do not benefit from education in order to reduce the size of the government and target all its revenues to income redistribution.[4]

[4]The results that education is provided publicly if and only if the share of the voters who consume it constitutes a minority is consistent with the observation that the level of public provision of education has not decreased in Western democracies in the last decades, whereas the share of the old voters in the population, who do not benefit from education, has increased. For example, in the years 1970-1990, the share of individuals aged 65 and higher increased in the US by 25% whereas the per-pupil spending in public elementary schools increased by 75% (see the Digest of Educational Statistics (2000) and also Fernandez and Rogerson (2001)). It

The main insight in my analysis uncovers the strategic interplay between income redistribution and redistribution in kind. It links the size of the group that benefits from education to the relative costs of the two different types of redistribution and to the ability of different groups in society to compromise with one another. Such an insight can arise only in a model in which the tax level and the type of redistribution are jointly determined, and agents differ in their attitudes towards education. The paper focuses on education but the analysis may apply to other private goods which are publicly supplied or also to local public goods. The model predicts then that minority groups who advocate specific or local public goods, can be rather successful in the political process.[5]

Additional results relate public provision of education to the degree of income inequality in society. Higher income inequality means more cohesiveness among the rich groups. This arises because when income inequality is high, even the rich who do benefit from education view the public provision of it as a too costly venture. They therefore vouch for no taxation, in line with the rich who do not benefit from education. This implies a stronger political power for the rich groups and hence in some cases, a smaller government and less public provision of education.

The literature on the positive analysis of public provision of education was pioneered by Stiglitz (1974), who highlighted the fact that a median voter equilibrium may not exist. Epple and Romano (1996a,1996b), Gloom and Ravikumar (1998), and Fernandez and Rogerson (1995) provided conditions for the existence of a median voter result in models in which tax revenues are fully targeted towards public provision of education, whereas Bearse, Glomm and Janeba (2001) show that the median voter result fails when the tax rate is fixed but voters decide between income redistribution and redistribution in kind.

My paper is the first to derive analytical results about public provision of education in a model that combines both the choice of the size of government and the choice between

is also consistent with the empirical analysis of Poterba (1997).

[5]There is a more general literature in political economy about the ability of special interest groups to attain public resources. In Besley and Coate (2000) anti-majoritarian outcomes arise as well, but for another reason; it is due to the issue being non-salient for the majority. Grossman and Helpman (1994) show that interest groups which are organized are likely to induce policies which are more favorable to them, whereas in Dixit and Londregran (1995), politicians redistribute income to voters who have little ideological bias, since these are the type of voters who are easily swayed and therefore willing to 'sell' their vote. See Persson and Tabellini (2002) for a summary of this literature.

redistribution in kind and income redistribution. Although my analysis is two-dimensional, since it builds on the 'citizen candidate' model, which restricts the set of policies offered by politicians, it results in a stable political outcome (which is not the median voter's preferred choice, even when such a voter exists). Other papers that attempt a two-dimensional analysis use a two-stage voting process and numerical simulations to solve the model, as in Glomm and Patterson (2003) or Bearse, Glomm and Janeba (2000). But in these models equilibria do not necessarily exist in some cases and hence the numerical analysis is further complicated by the need to rule out these cases.

In Fernandez and Rogerson (1995) and Epple and Romano (1996a,1996b), as in my model, a voting 'coalition' arises between the rich and the poor. In these papers though, the rich and the poor vote together in an 'ends against the middle' type of coalition, i.e., the poor and the rich collude against the middle class. In my analysis the coalition between the rich and the poor is against another group of poor voters, which, depending on parameters, is either the group that benefits from education or the group that does not benefit from it. This coalition divides the poor voters and allows the rich to reduce the size of government.[6]

The rest of the paper is organized as follows. In the next section I present the model. Section 3 presents a benchmark: the political outcome in the absence of parties. Section 4 presents the main result, i.e., when is education publicly provided. In section 5, I present the political predictions of the model, about the size and composition of parties. Section 6 discusses the main assumptions of the model. I conclude in section 7 and all proofs are in an appendix.

2 The Model

2.1 The economic environment

The economic environment builds on the model of Epple and Romano (1996a).[7] There are two types of goods in the economy, education, denoted by e, and a numeraire good, denoted by x. There are two types of agents in the economy; those who benefit from education, and those

[6] Roemer (1998) also uses a two-dimensional analysis to explain why the poor do not expropriate the rich, i.e., why the tax level is not set at its maximum level (he does not address the issue of education or redistribution in kind). In his analysis, as opposed to my model, parties are exogenous.

[7] There are two notable differences however. First, in my analysis income redistribution is a possible policy tool. Second, there are some agents in the economy who do not benefit from education.

who do not. The utility function of an 'education type', denoted by type 1, is:

$$u_1(e,x) = u(e,x)$$

where u is strictly quasi-concave and twice differentiable. For the type-1 agents, both education and the numeraire are assumed to be normal goods.[8] The agents who do not benefit from education, denoted as type 0, care essentially only for their disposable income, which is assumed to be a normal good, and hence for simplicity their utility function is

$$u_0(e,x) = x.$$

The share of type-1, the 'education types', in the economy is $\theta \in (0,1)$. The case of a completely homogenous society ($\theta = 1$ or $\theta = 0$) is analyzed in section 6. An important assumption is that only those who benefit from education consume it, i.e., the 0-types do not consume education even if it is provided publicly (I discuss this assumption in section 6).

Society can choose a tax level t. With the revenues it may either redistribute income, denoted by T, or finance the provision of education, denoted by g. The price of education in terms of the numeraire x is q. The budget constraint per capita is therefore

$$ty = T + \theta q g$$

where y is average income.[9]

Consumers who benefit from education, may supplement the public provision by buying education in the market, through private tutors for example, for the same price q.[10] The appendix shows however that the price q has no effect on the results; for the purpose of exposition, I normalize $q = 1$. Denote the additional education consumption by s.[11]

There are two levels of income in the economy. The rich have the high income y_h whereas the poor have the low income y_l. The share of the poor in the population is π (there is no correlation between income and the preferences for education). The average income y is therefore

$$y = \pi y_l + (1-\pi) y_h$$

[8] See Epple and Romano (1996a).

[9] For simplicity, there are no labour decisions and the tax is therefore not distortive.

[10] Epple and Romano (1996a) also analyze the case in which there is a difference in the productivity of the private and public market.

[11] An alternative model for the analysis of public versus private consumption of education is a model in which they are mutually exclusive, i.e., a household can either consume private education or public one but not both. In a multidimensional environment, this is a much harder model to analyze.

Without loss of generality, let us set $y_l = 1$. For tractability (and since this is the case in most countries) I focus the analysis on the case of the poor being the majority, that is, $\pi > \frac{1}{2}$. Also, for the sake of interest, assume that no group in the population composes a strict majority.

The parameters of the model are therefore

$$< \pi, \theta, y_h >,$$

and the four groups are denoted by:

$$r_0, r_1, p_0, p_1,$$

that is, rich of type 0 (r_0), rich of type 1 (r_1), poor of type 0 (p_0), and poor of type 1 (p_1).

The policy space that society is facing is $\{t, g\}$, since the 3 dimensional problem of $\{g, T, t\}$ reduces to a two dimensional problem by the budget constraint, $T = ty - \theta g$. Thus, the policy space is bounded by a triangle, i.e., $t \in [0, 1]$ and $g \leq \frac{ty}{\theta}$.

I now characterize the ideal policies and indifference curves of the different groups in society, in the policy space $\{t, g\}$. The 0-types care only for income. They are therefore indifferent between all policies which give them the same income. This implies, for $y_i \in \{y_l, y_h\}$:

$$\begin{aligned} y_i(1-t) + ty - \theta g &= const \rightarrow \\ \frac{\Delta g}{\Delta t} &= \frac{(y - y_i)}{\theta}. \end{aligned}$$

Hence, the slope of their indifference curve is linear, it is positive for the poor and negative for the rich. Figure 1 describes the policy space $\{t \in [0, 1], g \leq \frac{ty}{\theta}\}$, and depicts the indifference curves of r_0 and p_0 :

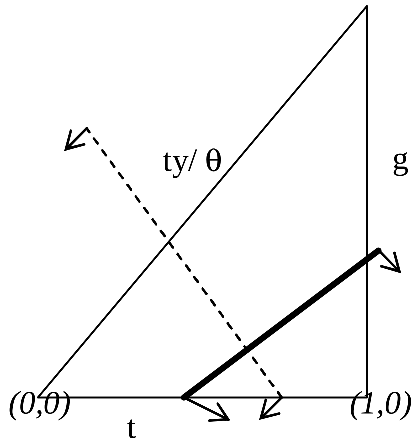

Figure 1: Indifference curves for r_0 (the dashed line) and for p_0 (the bold line). Arrows shows

direction of increase in utility.

It is also easy to see from Figure 1 that the ideal policy of r_0 is $\{g = 0, t = 0\}$, and that of p_0 is at $\{t = 1, g = 0\}$, i.e., $T = y$ and equal income for all.

We can now describe the indifference curves of the 1-types in the $\{t, g\}$ space, as depicted in the following figure:

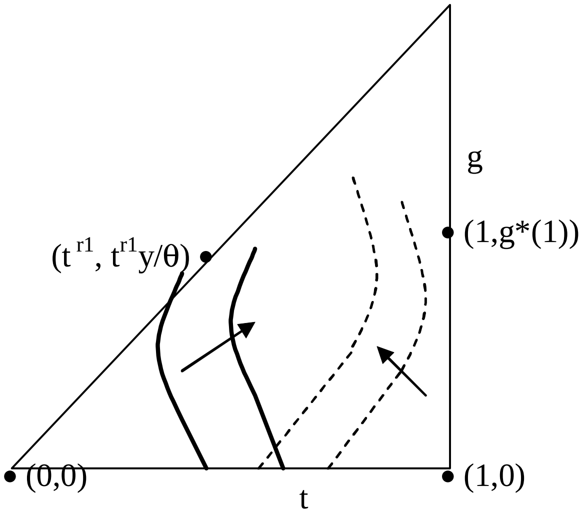

Figure 2: Indifference curves for r_1 (dashed) and for p_1 (bold) with arrows denoting direction of increase in utility. The figure also depicts the possible ideal policies for r_1 and p_1, which are explained below.

To understand the shape of the indifference curves, note that given the publicly decided (t, g), each household chooses how much private education (s) to buy, being constrained by $s \geq 0$ and its budget constraint. When g is relatively low, then both the rich and the poor who benefit from education need to supplement it by buying private education (i.e., $s > 0$ in the optimal solution). This implies that when g is sufficiently low, any additional g is seen as a pure money subsidy and substitutes private consumption. As a result, the indifference curves are linear for low values of g. When g is high enough though, there is no need in private education (that is, $s = 0$). The indifference curves become concave (given the strict quasi-concavity of u).

In terms of ideal policies, the poor obviously prefer the highest tax level, $t = 1$, and only have to consider how to divide it between public provision of education and income redistribution. Both are viewed as a form of redistribution from other groups in society to themselves. Denote by $g^*(1)$ their optimal public provision of education given $t = 1$.

The rich, r_1, clearly prefer not to redistribute any income. But, as opposed to those who do not care for education, they view public provision of education as a redistribution from the 0-types to the 1-types and may therefore favour public provision of education. If θ, the share of those who benefit from education, is relatively low, then such redistribution is beneficial for r_1 (technically, this arises when the slope of the linear part of their indifference curve is less steep than society's budget constraint). If on the other hand θ is relatively high, it is too costly for r_1 to finance the public provision of education, and they prefer no taxation.

Lemma 1 summarizes the above discussion (the proof is in the appendix):

LEMMA 1 *In the* $\{t, g\}$ *policy space: (i) The ideal policy of* p_1 *is* $\{t = 1, g = g^*(1)\}$ *and the ideal policy of* r_1 *has T=0 and is* $\{t = 0, g = 0\}$ *if* $\theta > \frac{y}{y_h}$ *and otherwise it is* $\{t = t^{r_1}, g = \frac{yt^{r_1}}{\theta}\}$ *for some* $t^{r_1} \in (0,1)$*. (ii) The indifference curves of* r_1 *and* p_1 *are weakly concave and differentiable, with a slope that is everywhere less than or equal to* $\frac{y_i - y}{1-\theta}$ *for* $y_i \in \{y_l, y_h\}$*. (iii) The ideal policy of* r_0 *is* $\{g=0,t=0\}$*, and that of* p_0 *is at* $\{t=1,g=0\}$*. (iv) The indifference curves of* r_0 *and* p_0 *are linear, with a slope* $\frac{y - y_i}{\theta}$ *for* $y_i \in \{y_l, y_h\}$*.*

To conclude the description of the economic environment, I make some assumptions about its parameters. Let $u_i(t, g)$ denotes the utility of type $i \in \{p_0, p_1, r_0, r_1\}$ from a policy (t, g). I assume that the poor 'stick' together, i.e., that:

$$\textbf{A1} \; u_{p_1}(1, 0) = u_{p_1}(t^{r_1}, \frac{yt^{r_1}}{\theta}) + \delta, \text{ and } u_{p_0}(1, g^*(1)) = u_{p_0}(0, 0) + \mu,$$

where $\delta \geq \delta_0 > 0$ and $\mu \geq \mu_0 > 0$ for δ_0 and μ_0 that are defined in the appendix. This assumption is designed only for insuring the existence of pure strategy equilibria in the political game defined below. Focusing on pure strategy equilibria makes the results more stark but does not change their qualitative nature (see section 6 for a discussion of this assumption).

2.2 The political game

The political process translates the economic preferences into a policy (t, g). Those who take part in this process are the voters (which are all the citizens), and the politicians. For simplicity, I assume that the set of politicians is fixed and is composed of a representative from each group of voters.[12] That is, there are 4 politicians, with politician i having the ideological preferences of group $i \in \{r_0, p_0, r_1, p_1\}$. The actual candidates or parties running for election will be

[12]This assumption is not important. Alternatively, one can assume that all voters can run as politicians. For a more general model using this assumption see Levy (2003).

endogenously determined though. In the political process, these politicians would offer policy platforms to voters, in a way that is described below, and voters would simply vote for their favorable platform to determine the winning platform and hence the political outcome (t, g).

Politicians can remain *individuals*, or join together in *parties*. A *party* is a coalition of heterogeneous politicians, that is, a party is formed when representatives of different groups join together. A *party structure* is a description of how the politicians are organized, either into parties or as individuals, i.e., it is a partition on the set of politicians, denoted by ρ. For example, the partition $p_0|p_1|r_0|r_1$ is the party structure in which each politician can only run as an individual, and the partition $p_0p_1|r_0|r_1$ is the party structure in which the poor representatives join together in one party and each of the rich politicians is an individual. A party or an individual politician in a party structure are denoted by $R \in \rho$ and $i \in R$ means that a politician i is a member of R (R can be a singleton).

Assume for now that the party structure ρ is fixed, that is, it is already 'determined' which politician is in which party or which politician is an individual. I now describe the election, given a fixed party structure ρ.

The main assumption about the election, is that politicians cannot commit (see Besley and Coate (1997) and Osborne and Slivinski (1996)). Consequently, a politician who runs for election as an individual, can offer only his or her ideal policy in the election. For example, r_0 can only offer the policy $(t = 0, g = 0)$ to the voters. Following Levy (2003), I assume that a party can commit to implement policies in the Pareto set of its members. The Pareto set is the set of all feasible policies, (t, g), such that there are no other policies which make all the party members better off (and some of them strictly better off). Trivially, when a politician runs as an individual candidate his Pareto set is his ideal policy. But this changes when heterogeneous politicians join together in a party. For example, the type-0 party, r_0p_0, can offer to the voters all policies with $g = 0$ and $t \in [0, 1]$. Parties increase therefore the commitment ability of politicians; they allow for different factions to reach an internal compromise and to offer these compromises to the voters on election day (I discuss my assumptions on the role of parties in section 6).

Thus, in the election, all parties or individual politicians simultaneously choose either not to run or to offer a platform from their Pareto set, where the Pareto set of party/individual R is denoted by Q_R. Given the platforms offered in the election, the voters vote - sincerely - for the platform they like most.[13] The election's outcome is the policy which receives the maximum number of votes.[14] If no policy is offered by any party, a default status quo is implemented. Following Osborne and Slivinski (1996), I assume that the utility from the default policy is worse for all players than any other outcome (such as a government shut-down). This insures that there is always one party which contests the election. As a tie-breaking-rule, I also assume that if all party members are indifferent between running and not running, the party or candidate choose not to run.

It is then easy to show that typically in this environment, only one party/candidate will run for office and hence win. For simplicity of exposition I therefore present here the definition of a pure strategy equilibrium, in which one party or individual run for election (the analysis in the appendix is general and allows for all types of pure-strategy equilibria. See also section 6). In such an equilibrium, the party that offers a platform will not withdraw given the assumption that the status quo is worse for all. We therefore need to be concerned only with the incentives of the parties that choose not to run (recall that politicians care about policies):[15]

Definition 1 An equilibrium in ρ is a platform $(t, g) \in Q_R$ offered by a party $R \in \rho$, such that there is no $R' \in \rho$ and $(t', g') \in Q_{R'}$ which can win against q_R in the election and for all $i \in R'$, $u_i(t, g) \geq u_i(t', g')$, with a strict inequality at least for one $i \in R'$.

Let $(\mathbf{t}^{\rho}, \mathbf{g}^{\rho})$ be an equilibrium winning policy given ρ. So far I assumed that the party structure ρ is fixed. However, politicians may not be satisfied with their party membership. The final details of the political model estbalish which parties can endogenously arise.

A politician or a group of politicians will induce a *party fragmentation* when they break away from their party so that the original party is divided into two parties/candidates, and all the other politicians remain in their original parties. Politicians will fragment their party if in the resulting party structure, the winning policy provides them a higher utility. Stable political outcomes, and as a result, also stable parties, are defined as equilibrium outcomes which are immune to party fragmentations.

Definition 2 A *stable political outcome* is the platform $(\mathbf{t}^{\rho}, \mathbf{g}^{\rho})$ which is an equilibrium

[13] If some platforms give them the same utility, they mix fairly between them. The assumption of sincere voting is for simplicity.

[14] Where a fair lottery is held if there is more than one such policy.

[15] For a general definition of equilibria see Levy (2003). I also show there that an equilibrium in mixed strategies exists in such games for all partitions. In the simplified model analyzed here, pure strategy equilibria exist as well.

winning policy in ρ, such that there is no $R' \subset R \in \rho$ that can induce a party fragmentation so that all $i \in R'$ (weakly) prefer the equilibrium winning policy $(\mathbf{t}^{\rho'}, \mathbf{g}^{\rho'})$ in the new party structure ρ'.

Note that multidimensional policy making models tend to result in cycles in political decisions; in this model, cycles are avoided due to two reasons. First, the set of policies that can be offered by parties or candidates are restricted. Second, politicians are restricted in how they can deviate from parties. For a discussion of this restriction, see section 6.

2.3 Summary of the model

To summarize the model, society has four groups of citizens, and a conflict on two dimensions, the rate of tax and how to redistribute tax gains. The four groups are represented in the political process by politicians. Politicians care about the implemented policy and can induce different political outcomes when in different parties. These parties offer policies - in their Pareto set - on which the voters vote. The political outcome is the platform which wins the election. Parties are endogenous in the model in the sense that we identify the array of political parties and outcomes such that no politician (or a group of politicians) wishes to quit her party and thereby induce a different political outcome. The prediction of the model is the set of the stable political outcomes.

3 A benchmark

In the model, parties are essentially defined as coalitions of heterogenous politicians. In other words, parties can increase the commitment ability of politicians. As a benchmark though, let us consider first the political outcomes when such parties, for some reason, cannot form. That is, when each politician can only run by himself. In this case, the stable political outcomes are simply the equilibria in the partition $p_0|p_1|r_0|r_1$.[16] This is then the 'citizen-candidate' model; each politician decides whether to enter or not, and if she enters, she can only offer her ideal policy.

When these are the only feasible platforms that can be offered to voters, results are *majoritarian*. That is, the representative of the majority wins (recall that the poor are the majority):

[16] No politician can fragment his party and hence equilibrium outcomes are also stable.

Proposition 1 *When those who benefit from education are a minority, the representative of the poor who do not benefit from education wins the election. The political outcome is the maximum tax level and no public provision of education. When those who benefit from education are a majority, the representative of the poor who benefit from education wins the election. The political outcome is the maximum tax level and positive public provision of education.*[17]

The appendix provides the full proof whereas here I just illustrate why the equilibria described in the proposition indeed hold. The policies that can be offered to the voters are the ideal policies of the different groups; the ideal policy of r_0 is $(t, g) = (0, 0)$, that of p_0 is $(1, 0)$, that of p_1 is $(1, g^*(1))$ and that of r_1 is either $(0, 0)$ or $(t^{r_1}, \frac{t^{r_1}y}{\theta})$. When $\theta < \frac{1}{2}$, i.e., the 1-types are the minority, then if p_0 runs alone he clearly wins the election, and prefers not to withdraw. Can another politician successfully challenge p_0? If p_1 challenges p_0 then r_0 would still vote for p_0. The reason is that he prefers the policy of $t = 1$ and $g = 0$ to that of $t = 1$ and $g > 0$, which leaves him with lower income. Thus, p_0 would win against p_1 because the 0-types are a majority. Also, r_0 or r_1 do not win against p_0 because of the assumption that the poor 'stick together'; p_1 prefers the ideal policy of p_0 to that of r_0 or r_1.[18] This implies that p_0 would register a majority of the votes because the poor are the majority. Thus, no one can challenge p_0.

When the 1-types are a majority, i.e., when $\theta > \frac{1}{2}$, the winner of the election becomes p_1; if p_0 challenges him he can only attract the votes of the 0-types but the 0-types are minority. If r_1 or r_0 challenge p_1, as above, they can only attract the votes of their fellow rich, which are a minority. In this case, the implemented policy is $(1, g^*(1))$.

4 Public provision of education

The above section has characterized the election winners when there are no parties involved. In particular, with respect to public provision of education, it is the group size which matters; education is provided if and only if those who benefit from it are a majority. When I allow for endogenous parties, however, outcomes are reversed. Parties, and their ability to facilitate compromise between different factions, will actually endow minorities with strong political

[17] The case of $\theta = \frac{1}{2}$, i.e., when society is equally divided between those that benefit from education and those who do not, does not add much to our understanding on top of the other two cases. See the discussion in section 6.

[18] This is obvious if the ideal policy of r_1 is $(0, 0)$ and assumption A1 is redundant in this case.

power. I now present the main result, i.e., when is education publicly provided in the presence of parties:

PROPOSITION 2 *When those who benefit from education are a minority, then all stable political outcomes with parties are characterized by a positive but not a maximum tax rate, a positive level of income redistribution and a positive level of public provision of education. The poor never buy private education, whereas the rich may do so. When those who benefit from education are a majority, then all stable political outcomes with parties are characterized by a positive but not a maximum tax rate, a positive level of income redistribution and no public provision of education. Both the rich and the poor buy private education.*

In all stable political outcomes, income redistribution is positive. This is intuitive, since the poor are the majority and both poor groups vouch for some positive level of income redistribution. The counter-intuitive result in Proposition 2 is that public provision of education is an *anti-majoritarian* result. That is, public provision of education arises when those who care for education are actually a minority and vice versa. Despite being there in numbers, the political power of the majority is eroded. Moreover, the tax level is not set at its maximum level, despite the poor being a majority.

This anti-majoritarian result, as I now illustrate, is due to two factors, both economic and political. The first factor is the group size. In particular, I isolate an unusual positive effect for being a minority. When the group that consumes some good such as education is a minority, it becomes relatively 'cheap' to redistribute via redistribution in kind versus income redistribution. Even a relatively low tax rate could support a generous provision of per capita public education. Group size would therefore determine the relative 'costs' of the different types of redistribution. The second factor is the institution of parties; the ability of parties to facilitate compromise between different factions allows the rich to collude with a segment of the poor, and to take advantage of the divergent views among the poor with regard to the provision of education. I now explain this intuition in a more detailed manner.

Let us start with the case in which those who benefit from education are a minority, i.e., $\theta < \frac{1}{2}$, and let us focus on a particular party, that of p_1 and r_0, i.e., when the rich who do not benefit from education collude with the poor who do benefit from education. To see what this party can offer to the voters, we have to find its Pareto set. For this party, denoted by r_0p_1, the shape of the Pareto set depends on the parameters of the model, as illustrated in figures

3a and 3b:

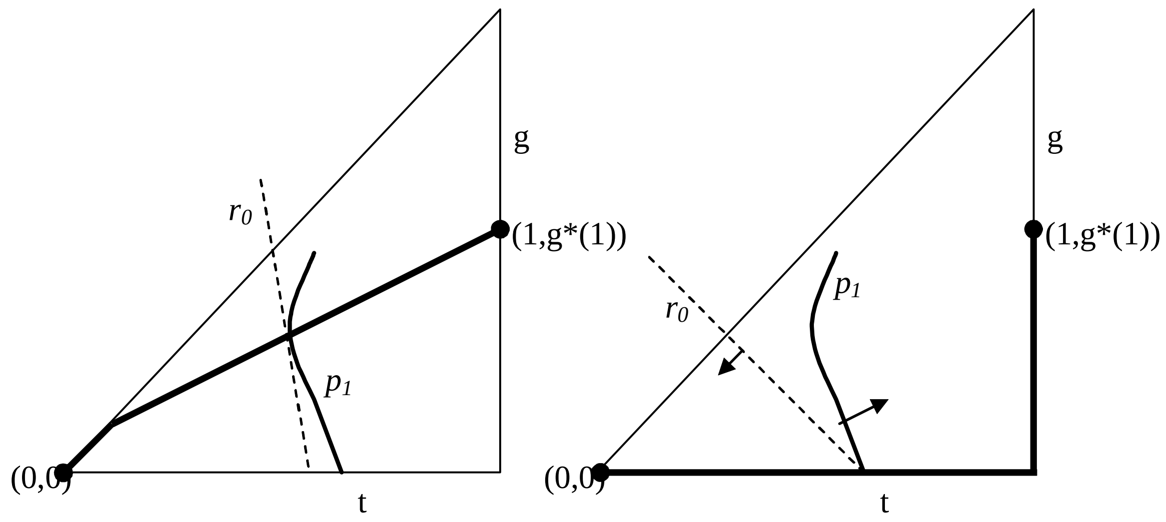

Figure 3a (left) and Figure 3b (right). The figures depict the ideal policies of r_0 and p_1, and the different types of Pareto sets (the bold lines).

In figure 3a, the Pareto set of r_0 and p_1 is termed 'interior'. It arises when the slope of the indifference curve of r_0 is steeper than that of p_1 (its linear part), i.e., when:

$$\theta < \frac{y_h - y}{y_h - 1}.$$

Thus, a policy with $t > 0$ but $g = 0$ cannot be in the Pareto set of $r_0 p_1$; when θ is low enough, education is relatively 'cheap' since only a few would consume it. This implies that both factions can be better off by lowering the tax rate and increasing public provision of education. On the other hand, when θ is relatively high, this does not hold. The Pareto set is then on the boundaries, as in figure 3b, and includes all the policies with $t \in [0, 1]$ and $g = 0$, as well as policies with $t = 1$.

Now, when $\theta < \frac{1}{2}$, two forces combine to produce the result. First, when there are no parties, p_0 wins the election with the policy of $t = 1$ and $g = 0$. Second, when $\theta < \frac{1}{2}$, it is also the case that the Pareto set of $r_0 p_1$ is 'interior', as described in figure 3a, since whenever $\theta < \frac{1}{2}$, also $\theta < \frac{y_h - y}{y_h - 1}$ holds.[19] This implies that p_0 is not part of the Pareto set of r_0 and p_1 and as a result, there are some policies that $r_0 p_1$ can advocate, with $t < 1$ and $g > 0$, *that are better for both party members than* p_0. These policies, are described in figure 4 below:

[19]This result does not rely on any parameter restrictions.

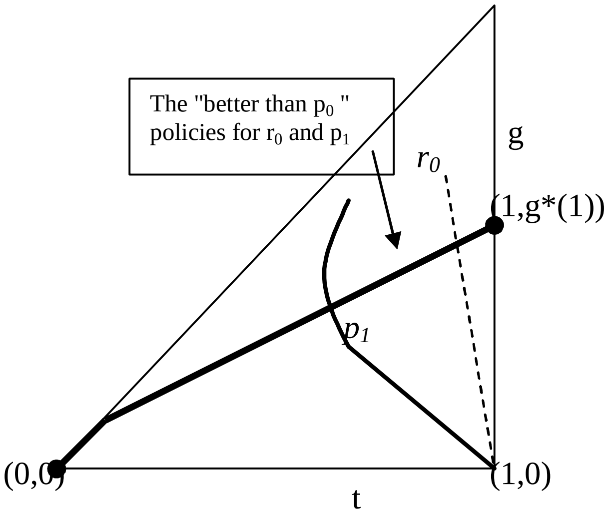

Figure 4: The indifference curve of r_0 (the dashed line) that goes through p_0 and that of p_1 that goes through p_0 define the set of policies on the Pareto set of r_0 and p_1 (the bold line) which are better for both relative to p_0.

When $\theta < \frac{1}{2}$, r_0 and p_1 can reach an internal compromise which is better for both of them relative to p_0, whose ideal policy advocates $t = 1$ and $g = 0$. By reducing and shifting tax revenues from a costly universal income redistribution to the public provision of education, this compromise increases the utility of p_1 (who is in need of positive provision of education) as well as the income of r_0 because redistribution in kind is relatively 'cheap' when θ is low.

When $\theta < \frac{1}{2}$, I therefore find that in the partition $r_0p_1|r_1|p_0$, at least some of the policies described above are stable political outcomes, when offered by the party r_0p_1. In the appendix I first show that some of these policies are indeed an equilibrium (that is, no other candidate can challenge this party and win).[20] It is then easy to check that these equilibrium policies are also stable, because if either r_0 or p_1 break their party, the unique equilibrium in the resulting partition, $r_0|p_1|r_1|p_0$, is that p_0 wins the elections, which is worse for both.

[20]When $\theta < \frac{y}{y_h}$ all of these policies are equilibria policies. When $\theta > \frac{y}{y_h}$, r_1 can upset the equilibrium by running for election. The assumption on the economic parameters of the model (that $\delta \geq \delta_0$) assures that some of these policies can win against r_1. Otherwise, the party still runs but wins only with some probability in a mixed strategy equilibrium.

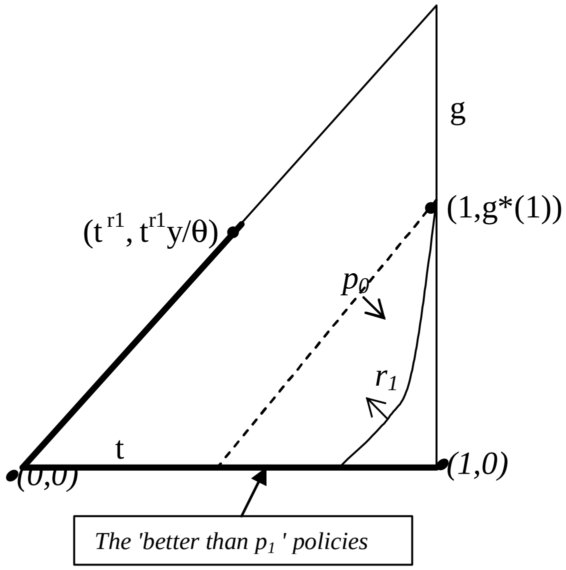

Figure 5: The figure depicts the ideal policies of the groups in society, for the case of $\theta > \frac{y}{y_h}$, i.e., when r_1 prefers positive provision of education. The indifference curve of r_1 that goes through p_1 and that of p_0 that goes through p_1 (the dashed line) define the set of policies on the Pareto set of r_1 and p_0 (the bold line) which are better for both relative to p_1.

On the other hand, when $\theta > \frac{1}{2}$, the scenario is reversed. First, the benchmark shifts, that is, p_1 wins the election when there are no parties. His ideal policy has $t = 1$ and $g > 0$. The party that can win is then r_1p_0. When $\theta > \frac{1}{2}$, the Pareto set of r_1 and p_0 is on the boundaries. That is, it includes all the policies with $g = 0$, disregarding the ideal policy of r_1, as illustrated in figure 5. This implies that p_1 is not in the Pareto set of r_1 and p_0 and they can reach a compromise that is better for both compared to the ideal policy of p_1. In particular, when θ is relatively high education becomes expensive compared to income redistribution since too many consume it. It is then possible to reduce the tax rate from a maximum of $t = 1$ while shifting the resources and supporting a relatively large lump sum transfer. This increases the utility of r_1 (who benefits from education but still sees it as a too costly form of redistribution) as well as the income of p_0. When $\theta > \frac{1}{2}$, I therefore find that in the partition $r_1p_0|r_0|p_1$, the political outcomes described in figure 5 are stable political outcomes.

To summarize, the main insight of the analysis is as follows. Redistribution in kind and income redistribution are viewed differently from the point of view of the rich voters. In particular, if those who benefit from education are minority, then the rich voters view redistribution in kind as 'cheaper' than a universal income redistribution. If in addition parties are allowed to form, these relative prices of redistribution in kind versus income redistribution induce the rich to compromise with the segment of the poor that represents this cheaper form of redistribution. This results in the minority 'winning' in the sense that public provision of education is positive if and only if those who benefit from it constitute a minority. Moreover, the rich who are a minority 'win' as well, since the tax rate is not set at its maximum level.

So far I have explained the main insight of the analysis, which illustrates why policies can be anti-majoritarian. The appendix as well as the next section shed some more light as to why these type of policies are the unique stable policies and fully characterizes the set of stable parties.

5 Political representation

The main result established in Proposition 2 is that public provision of education arises if and only if those who benefit from education are a minority. The next result describes the political implications of the model, i.e., the size and composition of parties which win the election:

PROPOSITION 3 *(i) When those who benefit from education are a minority, then the party of the rich who do not benefit from education and the poor who do benefit from education is stable for all parameters. If the degree of income inequality is relatively high, then the large party which encompasses both rich representatives and the poor who benefit from education may also be stable. Finally, for some parameters and more often when income inequality is relatively high, then also the minority party of the rich and poor who benefit from education is stable. (ii) When those who benefit from education are a majority, the party of the rich who benefit from education and the poor who do not benefit from education is stable for all parameters, as well as the minority party of the rich and poor who do not benefit from education.*

The proposition insures that the parties identified in the previous section are indeed always stable. In other words, when $\theta < \frac{1}{2}$ the party structure $r_0p_1|r_1|p_0$ is stable and the party r_0p_1 wins the election, with the policies described in the previous section (such that $g > 0$), and similarly, when $\theta > \frac{1}{2}$, the party structure $r_1p_0|p_1|r_0$ is stable, and the party r_1p_0 wins the election with the policies described in the previous section (such that $g = 0$).

But aside from these parties, other parties may be stable. To explain the intuition for the existence of these other types of parties, I focus on the case of $\theta < \frac{1}{2}$, i.e., the 0-types who do not benefit from education are a majority.

Minority parties

Parties allow for compromising policies; but these type of compromises, such as the policies illustrated in the previous section, are more likely to be stable political outcomes when the 'most bitter enemies' join forces and collude in one party (as do r_0 and p_1 when $\theta < \frac{1}{2}$). To see why this is the case, note that in order to win against p_0 (which is the default winner when $\theta < \frac{1}{2}$ in the absence of parties), any party would need to attract the votes of r_0 and p_1 as well. Otherwise, if p_0 would attract one of these groups, he would win, since both the 0 types are a majority and the poor are a majority. But the party which is most proficient at providing policies which are better for r_0 and p_1 relative to p_0 is indeed the r_0p_1 party.

Let us consider whether the minority 1-types party, r_1p_1 can win the election, in the partition $r_1p_1|r_0|p_0$. In order to win, the party's Pareto set must contain policies which are better for both p_1 and r_0 (who is not a party member) relative to p_0. Nothing ensures that such policies exist in their Pareto set. Thus, for some parameters, this would be impossible to obtain. The degree of income inequality plays a role in this case; when $\theta < \frac{y}{y_h}$, and $r_1's$ ideal policy favours public provision of education and positive tax level, then the conditions stated above are harder to meet. That is, the preferences of r_1 become further apart from those of r_0 and thus the Pareto set of r_1p_1 would be less likely to provide r_0 with his 'better than p_0' policies. On the other hand, if this party is stable when $\theta < \frac{y}{y_h}$, then its policies are characterized by an even higher public expenditure on education, since r_1 is more keen on education in this case. In other words, policies may be characterized by lower taxes and lower level of public education when income inequality is high.

Large parties

Let us consider the party structure $p_1r_0r_1|p_0$ in which both segments of the rich cooperate with the 1-type poor against p_0 (we continue to focus on the case of $\theta < \frac{1}{2}$). The Pareto set of $r_1r_0p_1$ contains the Pareto set of r_0p_1 and r_1p_1. In particular, it contains the policies that r_0 and p_1 prefer to p_0. However, when it is only r_0 and p_1 that form a party, some of these 'better than p_0' policies cannot be implemented. In particular, when $\theta > \frac{y}{y_h}$, r_1 has the ideal policy of $(0, 0)$. He can then 'threat' to run against r_0p_1 and attract p_0. Therefore, in the partition $r_0p_1|r_1|p_0$, the party has to offer policies which are better for p_0 relative to $(0, 0)$ and thus can win the election only with a subset of the policies that both party members prefer to p_0. These policies have relatively high tax rate and high level of public education.

However, when r_1 joins this party, his membership allows him to commit not to run against

r_0p_1. Thus, policies that cannot be implemented when the party is r_0p_1 can be implemented as stable political outcomes when the large party is formed. These are policies which favour the rich, i.e., have lower tax rates.[21] The model provides therefore some justification for large parties.

Party competition

So far I have illustrated (for $\theta < \frac{1}{2}$) why the stable political outcomes and parties characterized in Propositions 2 and 3 can arise. It is left to show why other parties and political outcomes cannot be stable. First, it is clear that if p_0 cooperates with another politician, and no other coalitional parties exist, this cannot be stable; p_0 would always be better off breaking this party, inducing the benchmark situation in which he wins the election alone. But what if several coalitional parties arise? for example, if indeed p_1 and r_0 cooperate together, can p_0 'fight back' and collude with r_1 against the coalition of r_0p_1? I find however that for all parameters of the model such party competition between two coalitional parties is not a stable equilibrium phenomenon.

Let us consider for example the partition $r_0p_1|r_1p_0$, when $\theta > \frac{y}{y_h}$, i.e., $r_1's$ ideal policy is the same as of r_0. In this case, party competition is too fierce; for each policy offered by one of the parties, the other party can offer a policy which is better for both its members and would attract a strict majority of the voters. The only equilibrium policy in the party structure $r_0p_1|r_1p_0$ is therefore that one party offers the ideal policy of the rich, $(0,0)$. In this case, the other party would not deviate since the rich member of this party gets his ideal policy and would block this deviation. But this cannot be stable. For example, p_1 would rather break his party since in the remaining partition, $r_0|p_1|r_1p_0$, the winning political outcome $(1,0)$ is better for him than $(0,0)$. For other party structures with two coalitional parties, the intuition is similar. The main reason in these cases as well is some lack of polarization in society. When two coalitional parties exist, their Pareto sets are too similar; each can cater to most voters in society. This implies that party competition is too fierce, inducing at least one of the politicians to break his party.[22]

[21] Note however that this party does not form or is not stable when $\theta < \frac{y}{y_h}$. In this case, r_1 vouches for positive taxes and public provision of education. He then poses no threat to the party r_0p_1 and thus his commitment not to run is useless.

[22] Note that this effect holds also when $\theta = \frac{1}{2}$, that is, it does not relate to whether both parties can or cannot get 50% of the votes in equilibrium.

6 Discussion

The main result shows that public provision of education is an anti-majoritarian outcome. This is a consequence of both the assumptions about the economic environment and of the assumptions about the political process, i.e., the role of parties. More specifically, group size has two effects in the model. First, voters who are a majority may have more political power by voting together as one block. Second, the size of those who consume education is inversely related to the price of public provision of education and hence size decreases the political power of this group. In the political process that I analyze, the second effect dominates for all parameters.

The main goal of this paper was to identify this effect; i.e., to show that political power may be decreasing in group size. In other models of political power, this effect would still exist (although will not always be the dominant effect). In the remaining of this section, I discuss the main assumptions of the model. This sheds light on the robustness of the results, and also suggests some possible extensions.

Public education is consumed only by those who benefit from education Suppose to the contrary that even the 0 types, who do not benefit from education, consume public education when it is provided (they may still vote against it). This would wipe out the relative benefits of redistribution in kind. In such a case, it is easy to see that public provision of education is never positive. The rich - both types - would see it as a too expensive type of redistribution and would not cooperate with the poor who benefit from it in order to facilitate a public provision of education.

However, it seems unlikely that those who do not benefit from education will still consume it. The consumption of education entails the burden of sitting in the classroom or downloading lecture notes from the internet. Such cost, in terms of time and effort, must deter the types who do not benefit from education, from consuming it.

No differentiation in attitudes towards education ($\theta = 1$ and $\theta = 0$) In the extreme cases of an homogenous society, it is easy to see that the stable political outcome is majoritarian and coincides with the ideal policy of the poor. The intuition is simple. For example, when $\theta = 1$, there are only two groups in the population, the rich and the poor. Thus, if the poor are the majority, their ideal policy can be the only stable political outcome.

This result can also be generalized for societies with more than two levels of income. In this case, all agents with income less than the mean income would share the same ideal policy, that of the poor in my analysis, and all the agents with income more than the mean income would share the same ideal policy of no taxation (even the rich who benefit from education prefer no taxation when $\theta = 1$). The result is then that if the median income is less than the mean, the 'poor' win and there is redistribution according to the preferences of the poor. If the median income is higher than the mean, there is no taxation and no redistribution.[23]

The fact that outcomes are majoritarian at the extreme cases of $\theta = 0$ and $\theta = 1$ implies that political outcomes exhibit discontinuity when $\theta \to 1$ or $\theta \to 0$. This however is a feature of the somewhat simplified political process that I use, which allows each group to be equally represented in the political process, no matter how small it is. A more detailed political process may endow groups with the ability to be represented if they are larger than some threshold. Although some discontinuities may arise still, these are reasonable since indeed it is possible that the utility of a group of voters jumps when it becomes represented in the political process.

The poor 'stick' together In the model I assume that the poor 'stick' together and that this preference is relatively strong. If the poor do not 'stick' together, then already in the benchmark case, i.e., without parties, it is not guaranteed that pure strategy equilibria exist. This makes the analysis, as well as the predictions, much harder to derive. If in addition the preferences of the poor for one another are not strong enough, then also pure strategy equilibria are not certain in the analysis with parties. The main insights uncovered in this paper though would still hold, i.e., the political outcomes would be anti-majoritarian with some positive probability, but not with certainty.

The role of parties In the model I assume that parties increase the commitment ability of politicians. It builds on the assumption in the citizen candidate literature. In this literature, the assumption is that individual politicians can only commit to implement their ideal policy after the election. It is indeed reasonable that individual politicians are not able to commit to voters because the mass of voters will find it hard to coordinate its actions and to monitor politicians by punishing them in the ballot box. However, it is relatively easy for a *small*

[23] The results accord therefore with the previous literature on redistribution, notably Roberts (1977), which shows that the relation between the mean income and the median income determines whether there is redistribution.

group of politicians to monitor *one another*. In particular, when heterogenous politicians join together in the same party, they should be able to commit to offer platforms which are compromises between their ideal policies. At least some of the politicians would then have an incentive to ensure that other politicians in the party do not deviate from the agreed policy, since such deviations result in a lower utility for these party members. The public can then trust promises which are an internal compromise between different factions within one party.

The definition of stability The political model assumes a particular definition of stable parties. In this definition, players can deviate alone or together with other players. I restrict however the way politicians can deviate. They can only break an existing party into two. This is the simplest way in which one can insure the existence of stable outcomes in the multidimensional policy space, since without any limits on the ability of politicians to deviate, no outcome would be stable. In reality it is probably also the case that party fragmentation is perceived as the easiest form of deviation in terms of time and other resources. In any case, the nature of the results is maintained even if one allows for more complex deviations.[24]

Equilibria with more than one platform In citizen candidate models with sincere voting and discrete distribution of preferences, it is typically the case that only one platform is offered in equilibrium. For more than one platform to be offered it has to be the case that each platform can win with a positive probability. For example, when two platforms are offered, it has to be possible to divide society into two equally measured groups. This happens for example when society can then be equally divided between those who benefit from education and those who do not. The qualitative nature of the results is maintained but is less stark.

7 Conclusion

The main insight in this paper is that public provision of education is an anti-majoritarian result. Although composing a larger share of the population may provide voters with more political power, it can also be a disadvantage. In the model, those who favour public education are also those that consume it. If their share is relatively large, education is expensive to

[24] If for example politicians play some 'membership game' in which they announce first which party they wish to join, taking into consideration the equilibria that follow, the same stable outcomes arise. See also a similar discussion in Levy (2003).

provide relative to income redistribution and hence the rich prefer income redistribution and can collude with the segment of the poor who do not care for education in order to implement this outcome. On the other hand, when the share of those who favour public education is relatively small, redistribution in kind becomes relatively cheap. The rich would then collude with the segment of the poor which prefers redistribution in kind, so as to support a relatively low tax rate.

Appendix

In the proofs I maintain the price of education in terms of the numeraire as q; as will be evident from the proofs, however, the level of q does not change the results as reported in Lemma 1, Propositions 1, 2 and 3.

Proof of Lemma 1:

The derivation of the preferences of r_0 and p_0 is done in the text. Here I focus on r_1 and p_1. First, it is easy to check that the strict quasi-concavity of u will induce quasi concave preferences on the (t, g) space. Hence, the indifference curves are (weakly) concave.

Consider first the rich types, r_1. Given any (t, g) they choose s to maximize:

$$\max_{s \geq 0} u((g+s), (y_h(1-t) + ty - \theta qg - qs))$$

The first order condition is:

$$u_e - qu_x + \lambda = 0$$

where the solution is:

$$\begin{aligned} s &> 0 \text{ if } g \leq g(t) \\ s &= 0 \text{ otherwise.} \end{aligned}$$

Note that u_e decreases with g, whereas u_x increases with g. Thus, when g is high enough, $u_e < qu_x$ and therefore the optimal solution has the constraint binding at $s = 0$. Also, $g(t)$ has a negative slope; when t increases and g is fixed, income decreases for the rich voters. Thus, for the same g, it must be that one prefers $s = 0$. Thus, $g(t)$ has a negative slope.

The slope of the indifference curve, using the indirect utility and the envelope theorem is characterized by:

$$\begin{aligned} (u_e - \theta qu_x)dg + u_x(y - y_h)dt &= 0 \rightarrow \\ \frac{dg}{dt} &= \frac{u_x(y_h - y)}{u_e - \theta qu_x}. \end{aligned}$$

When $s > 0$, $qu_x = u_e$, then the slope is linear and positive:

$$\frac{dg}{dt} = \frac{u_x(y_h - y)}{u_e - \theta qu_x} = \frac{u_x(y_h - y)}{qu_x - \theta qu_x} = \frac{(y_h - y)}{q(1-\theta)}$$

When $s = 0$, the slope is

$$\frac{u_x(y_h - y)}{u_e - \theta qu_x},$$

which is positive first by continuity, and at some point becomes negative. The magnitude of the slope is always larger than $\frac{(y_h - y)}{q(1-\theta)}$.

It is then easy to see graphically from Figure 2 that if

$$\frac{(y_h - y)}{q(1-\theta)} > \frac{y}{q\theta} \Leftrightarrow \theta > \frac{y}{y_h}$$

then the ideal point is $(t = 0, g = 0)$. On the other hand, if this condition is not satisfied, then the ideal point is on the boundaries where

$$g = \frac{ty}{q\theta}$$

at a point where $s = 0$. Thus, equating the mrs with the economy budget constraint:

$$\frac{u_x(y_h - y)}{u_e - \theta q u_x}_{|g=\frac{ty}{q\theta}} = \frac{y}{q\theta}$$

defines t^{r_1}. We now turn to the analysis of the poor. This is similar to the analysis of the rich, where the solution for the optimal s is:

$$\begin{aligned} s &> 0 \text{ if } g \leq g'(t) \\ s &= 0 \text{ otherwise.} \end{aligned}$$

This implies a threshold function $g'(t)$ with a positive slope such that if g is below this threshold for t, then $s > 0$ whereas above it $s = 0$. The slope of the indifference curve is

$$\frac{dg}{dt} = \frac{-(y-1)}{q(1-\theta)}$$

when $s > 0$ (a negative slope), and

$$\frac{u_x(1-y)}{u_e - \theta q u_x}$$

when $s = 0$, which is negative and then positive. The magnitude of the slope is always larger than $\frac{-(y-1)}{q(1-\theta)}$. It is easy to see graphically that the ideal policy must be on the budget constraint where $t = 1$ and $s = 0$. Hence, it is at the point g which is optimal given $t = 1$, denoted by $g^*(1)$.■

Proof of Proposition 1:

The text establishes the existence of the equilibria described in the Proposition. It is left to establish uniqueness. Clearly, since $p_0(p_1)$ bits every candidate when $\theta < (>)\frac{1}{2}$, there are no other one candidate equilibria. Consider two candidate equilibria. If $p_0(p_1)$ is one of them in the case of $\theta < (>)\frac{1}{2}$, then he wins so the other candidate should drop. If $p_0(p_1)$ is not one of the two candidates in the case of $\theta < (>)\frac{1}{2}$, then he must vote for either $p_1(p_0)$ or $r_0(r_1)$ where one of them must be a candidate. This implies that the candidate he votes for wins, and therefore the other candidate must drop. Consider equilibria with three candidates. If p_1 and p_0 do not both run, then it must be that one of them runs. Suppose it is p_1. Then p_1 must win and others lose and have to drop. If they do run together, then it must be that neither drops from the race if all win with the same probability, i.e., that the share of the population of p_0 and p_1 is equal but this cannot be when $\theta \neq \frac{1}{2}$. Finally, since it cannot be that all the groups are equal in their size, there is no equilibrium with four candidates.■

Proof of Propositions 2 and 3:

The proof is in several steps. I first find the Pareto set for each possible party, and prove some preliminary results regarding these Pareto sets. I then characterize the equilibria for each partition, for each set of parameters. I then find which party structures are stable.

Step 1: Characterization of the Pareto sets.

Given the preferences characterized in Lemma 1, we can now characterize the Pareto set of the different groups in society (this is most easily done using the indifference curves). Once we characterize the Pareto set of any two groups, the rest (i.e., the Pareto set of three groups) follows from the union of all bilateral Pareto sets.

The Pareto set of r_0p_0 is $\{t \in [0,1], g = 0\}$. The Pareto set of p_0p_1 is $\{t = 1, g \leq g^*(1)\}$.

The Pareto set of r_0p_1 is as follows. First, note that the indifference curve of r_0 is linear and that of p_1 is linear for all $s > 0$. Then, if the slope of r_1 is less steep (in absolute value) relative to that of p_1, i.e., if:

$$\frac{y_h - y}{q\theta} < \frac{y-1}{q(1-\theta)} \Leftrightarrow \frac{y_h - y}{y_h - 1} < \theta,$$

then the Pareto set is

$$\{t \in [0,1], g = 0\} \cup \{t = 1, g \leq g^*(1)\}.$$

Otherwise there is a part of the Pareto set in which the indifference curves are tangent, when $s = 0$. That is, when

$$\frac{\frac{\partial u_{p_1}(t,g)}{\partial x}(1-y)}{\frac{\partial u_{p_1}(t,g)}{\partial e} - \theta q \frac{\partial u_{p_1}(t,g)}{\partial x}} = \frac{y - y_h}{\theta q}$$

This defines an increasing function $\tilde{g}(t)$. The 'interior' Pareto set is therefore

$$\{t \leq t', g = \frac{ty}{q\theta}\} \cup \{t \geq t', g = \tilde{g}(t)\} \text{ where } t' \text{ is defined by } \tilde{g}(t') = \frac{t'y}{q\theta}.$$

The Pareto set of $r_0 r_1$ is simple to derive and is:

$$\begin{aligned} \{t &= 0, g = 0\} \text{ if } \theta > \frac{y}{y_h}, \\ \{t &\leq t^{r_1}, g = \frac{ty}{q\theta}\} \text{ otherwise.} \end{aligned}$$

Let us now analyze the Pareto set of $r_1 p_1$. Let us denote by $g_h^*(t)$ and $g_l^*(t)$ the optimal provision of g given a fixed t, for the rich and the poor respectively. Obviously, $g_h^*(t) > g_l^*(t)$ since the rich have higher income. Given that, it is easy to see graphically that this implies that the indifference curves of r_1 and p_1 can be tangent to one another only when the slope of each is positive. When $\theta < \frac{y}{y_h}$ the Pareto set is fully characterized by the policies (t, g) which satisfy:

$$\frac{\frac{\partial u_{p_1}(t,g)}{\partial x}(1-y)}{\frac{\partial u_{p_1}(t,g)}{\partial e} - \theta q \frac{\partial u_{p_1}(t,g)}{\partial x}} = \frac{\frac{\partial u_{r_1}(t,g)}{\partial x}(1-y)}{\frac{\partial u_{r_1}(t,g)}{\partial e} - \theta q \frac{\partial u_{r_1}(t,g)}{\partial x}}.$$

This defines a function $g'(t)$. When $\theta < \frac{y}{y_h}$, let $t' = \max(t | s_{r_1}(t, g'(t)) = 0)$ and let $g''(t)$ be defined by

$$\frac{\frac{\partial u_{p_1}(t,g)}{\partial x}(1-y)}{\frac{\partial u_{p_1}(t,g)}{\partial e} - \theta q \frac{\partial u_{p_1}(t,g)}{\partial x}} = \frac{(y_h - y)}{q(1-\theta)}.$$

Let t'' be defined by $g''(t) = \frac{ty}{q\theta}$. The Pareto set is therefore

$$\begin{aligned} \{t &\geq t^{r_1}, g = g'(t)\} \text{ If } \theta < \frac{y}{y_h} \\ \{t &\in [0, t''], g = \frac{ty}{q\theta}\} \cup \{t \in [t'', t'], g = g''(t)\} \cup \{t \in [t', 1], g = g'(t)\} \text{ otherwise} \end{aligned}$$

Finally, the Pareto set of $r_1 p_0$ is as follows. If $\theta > \frac{y}{y_h}$ then it is trivially on the $g = 0$ line. If $\theta < \frac{y}{y_h}$, there are two possibilities. Either it is on the boundaries of the policy space, which is the case when the slope of the indifference curve of p_0 is less steep than that of r_1, or that it is interior:

$$\begin{aligned} \{g &= 0, t \in [0,1]\} \text{ if } \theta > \frac{y}{y_h}, \\ \{g &= 0, t \in [0,1]\} \cup \{t \leq t^{r_1}, g = \frac{ty}{q\theta}\} \text{ if } \frac{y-1}{y_h - 1} < \theta < \frac{y}{y_h} \\ \{g &= \bar{g}(t), t \in [t^{r_1}, 1]\} \cup \{t = 1, g \leq \bar{g}(1)\} \text{ otherwise,} \end{aligned}$$

where $\bar{g}(t)$ is defined by

$$\frac{\frac{\partial u_{r_1}(t,g)}{\partial x}(1-y)}{\frac{\partial u_{r_1}(t,g)}{\partial e} - \theta q \frac{\partial u_{r_1}(t,g)}{\partial x}} = \frac{y-1}{q\theta}.$$

I now prove some results regarding the Pareto sets.

LEMMA A1: *When $\theta < \frac{1}{2}$, the Pareto set of $r_0 p_1$ is interior.*

To see this note that:

$$\frac{1}{2} < \frac{y_h - y}{y_h - 1}$$

because

$$\begin{aligned} y_h - 1 &< 2y_h - 2y \Leftrightarrow \\ 2\pi + 2(1-\pi)y_h - 1 &< y_h \Leftrightarrow \\ 1 &< y_h \end{aligned}$$

and therefore when $\theta < \frac{1}{2}$ it is also the case that $\theta < \frac{y_h - y}{y_h - 1}$ and the Pareto set of $r_0 p_1$ is interior (more generally, this holds for $y_h > y_l$).□

LEMMA A2: *When $\theta > \frac{1}{2}$, the Pareto set of $r_1 p_0$ is on the boundaries.*

To see this note that

$$\frac{1}{2} > \frac{y-1}{y_h - 1}$$

and hence when $\theta > \frac{1}{2}$, the Pareto set of $r_1 p_0$ is on the boundaries.□

LEMMA A3: *The Pareto set of $r_1 p_1$ is above that of $r_0 p_1$ when the latter is interior.*

As established, the Pareto set is the tangency of the indifference curves of r_1 and p_1 when the slope of the indifference curve of p_1 is positive, whereas the Pareto set of $r_0 p_1$ has the tangency when the slope is negative (since the slope of r_0 is always negative). This implies the above.□

LEMMA A4: *When $\frac{y}{y_h} < \theta < \frac{1}{2}$, the indifference curves of r_1 cross those of p_0 only once.*

The slope of the indifference curve of p_0 is $\frac{y-1}{q\theta} < \frac{y}{q\theta}$ whereas the slope of the linear part of the indifference curve of r_1 in this case is $\frac{u_x(y_h - y)}{u_e - \theta q u_x} \geq \frac{(y_h - y)}{q(1-\theta)} > \frac{y}{q\theta}$. This establishes the result.□

REMARK *Note that q, the price of acquiring education in terms of the numeraire, will play no role in the results because of Lemma 1, and Lemmata A1-A4 which hold for all q. From now on I therefore assume that $q = 1$.*

Step 2: Characterization of equilibria for each partition.

Note first that party structures that I do not analyze will turn out not to be relevant for the stability analysis. Second, the proof of Proposition 1 demonstrated that typically there are only one candidate equilibrium in the set up, hence I focus on these. Finally, note that in the case of the grand coalition, any feasible policy can be an equilibrium, for all parameter values.

Definitions and notations:

Denote the Pareto set of r_0p_1 when it is interior as $\tilde{Q}_{r_0p_1}$. Define t^*, g^* which is the policy on $\tilde{Q}_{r_0p_1}$ that makes r_0 indifferent to p_0 :

$$(t^*, g^*) \in \tilde{Q}_{r_0p_1}|u_{r_0}(t^*, g^*) = u_{r_0}(1, 0)$$

Similarly, let $(t^{**}, 0)$ be defined by:

$$(t^{**}, 0) \in \tilde{Q}_{r_0p_0}|u_{r_1}(t^{**}, 0) = u_{r_1}(1, g_l^*(1))\}$$

Define also:

$$\delta_0 \equiv u_{p_0}(1, g_l^*(1)) - u_{p_0}(t^*, g^*)$$

where I assume that $\delta_0 > 0$, and let:

$$\mu_0 \equiv u_{p_1}(1, 0) - u_{p_1}(t^{**}, 0)$$

where μ_0 is clearly positive. Hence, by assumption A1:

$$u_{p_0}(t^*, g^*) > u_{p_0}(0, 0) \text{ and } u_{p_1}(t^{**}, 0) > u_{p_1}(t^{r_1}, \frac{t^{r_1}y}{\theta})$$

Case 1:

$$\theta < \frac{1}{2}, \theta > \frac{y}{y_h}$$

This is therefore the case in which those that benefit from education are a minority, and r_1 has the ideal policy $(0, 0)$.

One party, two members:

In the partition $r_0r_1|p_0|p_1$ it is the case that p_0 wins alone since neither player's Pareto set has changed compared to the partition without parties. In the partition $p_0r_0|r_1|p_1$, the 0-type wins with all policies better for p_0 than p_1 (and similarly when the party is r_1p_0). To see why others cannot run, note that if p_1 runs for election then the party can deviate and run as well and improve the utility of both the 0 types even by offering the ideal policy of p_0. If r_1 runs for election, then p_1 can run against him and win the votes of p_0.

In the partition $p_0p_1|r_0|r_1$, the poor party wins with all their policies since each of their members prefers this to the ideal policy of the rich.

I now focus on the partition $r_1p_1|p_0|r_0$ and $r_0p_1|p_0|r_1$. By lemma A1, there are some policies in the Pareto set of r_0p_1 such that both r_0 and p_1 prefer them to p_0. By assumption A1 it is also the case that some of these policies are such that p_0 prefers them to r_0. In equilibrium of $r_0p_1|p_0|r_1$, the party wins therefore the election with the subset of policies described in the text (those with relatively high tax rates). Clearly nothing else can be an equilibrium since the party can win against p_0 and p_0 can win against r_0.

In the partition $r_1p_1|p_0|r_0$ there are two cases of pure strategy equilibrium. If the Pareto set of r_1p_1 has no policies that both p_1 and r_0 prefer to p_0, then p_0 wins the election, since no one can contest him successfully. Otherwise, the party can win, in particular with a subset of these policies that p_0 prefer to r_0 (otherwise it is a mixed strategy equilibrium).

Two parties, two members each:

In the partition $r_0r_1|p_0p_1$ the poor always win with all their policies. Consider the partition $r_0p_1|r_1p_0$: for any policy in the Pareto set of p_1r_0, then there is a policy with $g = 0$ that wins against it (attracts all the rich and p_0). On the other hand, for any policy with $g = 0$, there is a policy in the Pareto set of p_1r_0 which can win against it by lemma A1. Thus, the unique equilibrium is that one of the parties offers $g = 0, t = 0$. Finally, in the partition $r_0p_0|r_1p_1$ the 0-type party must win. In particular, $g = 0, t = 0$ is a pure strategy equilibrium.

One party, three members:

In the partition $r_0r_1p_1|p_0$, the party wins with policies that are better for p_1 and r_0 relative to p_0. In the partition $r_0|r_1p_0p_1$, either r_0 wins, or the party wins with the policies that are better for p_0 than r_0. Finally, in the partition $r_1|r_0p_0p_1$, again r_1 can win or the party can win with all policies that are better for p_0 than the ideal policy of the rich.

Case 2:

$$\theta < \frac{1}{2}, \theta < \frac{y}{y_h}$$

In this case those who benefit from education are still a minority but the ideal policy of r_1 is $(t^{r_1}, \frac{t^{r_1}y}{\theta})$.

One party, two members:

Everything is as in case 1, besides the partitions $r_1p_1|r_0|p_0$ and $r_0p_1|r_1|p_0$. In the latter partitions, all policies in the pareto set which are better for both party members than p_0 can now be an equilibrium because r_1 cannot attract the votes of p_0 anymore. When the party is r_1p_1, then again equilibria may arise in which the party wins. If equilibria exist in this case, they also exist in case 1, because the Pareto set of r_1p_1 in case 2 is above the Pareto set of this party in case 1. When equilibria exist in case 2, they are with higher g than in case 1.

Two parties, two members each:

In the partition $r_0r_1|p_0p_1$ the poor win with all the policies in their Pareto set. In $r_0p_1|r_1p_0$, equilibria must be on the common parts of the Pareto set (otherwise, one party can always deviate and attract enough votes). If the Pareto set of r_1p_0 is on the boundaries, then the common part is when $g = \frac{ty}{\theta}$ for some $t \in [0, t']$. Otherwise, there is a unique common point which is interior. If the partition is $r_0p_0|r_1p_1$ then the party of the 0-types must win.

One party, three members:

In the partition $r_1r_0p_1|p_0$, the party wins with all the policies that are better for both r_0 and p_1 relative to p_0.

Case 3:

$$\theta > \frac{1}{2}, \theta > \frac{y}{y_h}$$

In this case, those who benefit from education are a majority and r_1 has an ideal policy of $(0, 0)$.

One party with two members:

Analogously to the cases above, the interesting partitions are the ones that include p_0. In the partition $r_0p_0|r_1|p_1$ the party of the 0-types can win with policies that are better for both p_0 and r_1 relative to p_1. The reason is that if p_1 attracts any of the groups p_0 or r_1 he then wins the election. Such policies exist since the Pareto set of r_1p_0 includes all policies with $g = 0$ and only these, which implies that there must be some policy which both prefer to p_1. The partition $r_0|r_1p_0|p_1$ yields exactly the same set of equilibria.

Two parties, two members each:

In the partition $r_0r_1|p_0p_1$, the poor win with all the policies in their Pareto set, as in the case in which the rich are separated. In the partition $r_0p_0|r_1p_1$, the unique equilibrium is that one of the parties offers the ideal policy of the rich - $g = 0, t = 0$. Otherwise, for any policy of the 0 party, the 1 type party can always attract its own voters, a majority. However, since the Pareto set of r_1 and p_0 is on the boundaries when $\theta > \frac{1}{2}$, this implies that the 0 party can always find policies to attract r_1, r_0 and p_0 given the platform of the 1 type party. In the partition $r_0p_1|r_1p_0$, the analysis is the same if

$$\theta < \frac{y_h - y}{y_h - 1}$$

and the Pareto set of r_0p_1 is interior. The reason is that because r_1p_1 have a policy on their Pareto set which is better than any policy of the other party that has $g = 0$, it must imply that they also have some policies on the Pareto set of r_0p_1 which are better for them than any policy with $g = 0$, by Lemma A3. Thus, each party can attract enough votes given any other policy beside the ideal policy of the rich. On the other hand, if the Pareto set of r_0p_1 is on the boundaries, then any party offering any policy with $g = 0$ can win, since this policy set is common for the Pareto set of both parties.

One party, three members:

In the partition $r_0r_1p_0|p_1$, the party wins with policies that are better for p_0 and r_1 than p_1. These are exactly the same policies that win the election when either r_0 or r_1 is in a coalition with p_1.

Case 4:

$$\theta > \frac{1}{2}, \theta < \frac{y}{y_h}$$

In this case those that benefit from education are a majority and r_1 has the ideal policy of $(t^{r_1}, \frac{t^{r_1}y}{\theta})$.

One party with two members:

In the partition $r_0p_0|r_1|p_1$ the party can win with policies that are better for both p_0 and r_1 relative to p_1 and are better for p_1 than r_1, which exist by assumption A1. In the partition r_1p_0 the Pareto set is on the boundaries. Then the party wins with policies that are better for both p_0 and r_1 than p_1.

Two parties, two members each:

In the partition $r_0r_1|p_0p_1$, the poor win with all the policies in their Pareto set, as in the case in which the rich are separated. In the partition $r_0p_0|r_1p_1$ the 1-types must win in a pure strategy equilibrium. In the partition $r_1p_0|r_0p_1$, any party can win on the common part of the pareto set which is $g = 0$ when the Pareto set of r_0p_1 is on the boundaries, and on $g = \frac{yt}{\theta}$ otherwise, for all $t < \tilde{t} < t^{r_1}$, for some $\tilde{t}$, because the Pareto set of r_0p_1 is below that of r_1p_1.

One party, three members:

In the partition $r_0r_1p_0|p_1$, the party wins with policies that are better for p_0 and r_1 than p_1.

Step 3: Stable political outcomes.

Case 1

Any party structure with one party and two members such that p_0 is one of them is not stable because p_0 will break. The partition in which the only party is r_0r_1 is not stable as well because in this partition p_0 wins. Consider now the party r_0p_1. But if these party members break they get p_0 which is worse for both. It is therefore stable. Similarly, if r_1p_1 win the election they do not break and otherwise they do.

Consider now two parties, each with two members. The partition $r_0p_0|r_1p_1$ is not stable because p_1 can break and get at least the utility from $(1, 0)$ in the equilibrium of $r_0|p_0|r_1p_1$. If it is $r_0p_1|r_1p_0$ it is not stable because p_0 will break. In this partition the equilibrium is the ideal policy of r_0 whereas if he breaks he gets something which is better than $(0, 0)$. Also $r_1p_1|r_0p_0$ is not stable because in this partition the 0-type wins so p_1 can break and get p_0, the best equilibrium of the 0 types.

Consider one party with three members. If p_0 is together with the rich, it cannot be stable since he breaks to win alone. If $r_0p_0p_1$ or $r_1p_0p_1$ are together or in the grand coalition, then the poor can always weakly improve by deviating together.

Finally, the partition $r_1r_0p_1|p_0$ can be stable. Any equilibrium outcome that can be achieved by r_1p_1 or by r_0p_1 is not stable since politicians prefer to win in smaller parties. However, consider outcomes that are not achievable by r_0p_1, that is, they are better than p_0 for both p_1 and r_0 but worse for p_0 than r_0. If p_1 deviates alone he is worse of since then the outcome is p_0. But also neither r_0 or r_1 deviate alone or with p_1 so than these outcomes are stable. If r_0 deviates alone then in the partition $r_1p_1|p_0|r_0$ it is a worse outcome for him; either p_0 wins or the party wins. But when the party r_1p_1 wins, it has to win with policies which are better for p_0 than r_0. Such a Pareto improvement for p_0 must be damaging the utility of r_0 then. Also, r_0 cannot deviate together with p_1 because the current outcome is on their Pareto set. Consider now r_1. However, a deviation with p_1 implies policies which are better for p_0 than the current policies (since they are better for p_0 than r_0). By the single crossing property identified in Lemma A4, and the fact that the Pareto set of r_1 and p_0 is for policies with $g = 0$, this must be damaging the utility of r_1. The same argument holds for a deviation alone, which implies that policies of r_0p_1 are implemented. Thus, the party r_1r_0p is stable with policies on the Pareto set of r_0p_1, that are better than p_0 for both p_1 and r_0 but worse for p_0 than r_0.

In a similar way, it is possible to show that the large party may for some parameters, be stable when implementing policies in the Pareto set of r_1p_1 which are better for all others than p_0.

Case 2

It is clear that $r_0p_1|p_0|r_1$ is stable and also that r_1p_1 is stable if the party wins. Let us consider now $r_0p_1|r_1p_0$. If the Pareto set of r_1p_0 is on the boundaries, then equilibria are worse for p_0 than r_0. Thus, he has an incentive to break the party since in the partition $r_0p_1|r_1|p_0$ there are some equilibria which provide him higher utility than the policy $(0, 0)$. If the Pareto set of r_1p_0 is interior, then the equilibrium is a unique interior point on the Pareto set of both parties. If this point is worse for r_0 than p_0 then he breaks it since in r_1p_0 he can get p_0 as an equilibrium. If this point is worse for p_1 than p_0 he breaks it due to the same reason. If it is better for both, then the equilibrium is interior in the set of policies with which r_1p_0 win the election in the partition $r_1|p_0|p_1r_0$. Clearly then there must be some policies in this set that p_0 (weakly) prefers, so he breaks his party.

In the partition $r_1r_0p_1|p_0$ there are no stable political outcomes. The party r_0p_1 can win alone with all the policies that their members prefer to p_0. This implies that if the large party wins with these policies they break away from it. If the large party wins with policies on the Pareto set of r_1p_1, then still r_0p_1 can deviate and win alone with policies that both prefer.

Case 3

The partition $r_0r_1|p_0|p_1$ is not stable since the rich would break and achieve the same outcome. No partition with one party and two members such that p_1 is a party member is stable since p_1 would break it. Similarly, p_1 breaks $r_0r_1|p_0p_1$ and $r_0r_1p_1|p_0$ to get his ideal policy. The partition $r_0p_0|r_1p_1$ is not stable because this is the worst outcome from the point of view of p_1 and if he breaks he receives at least positive income redistribution. Similarly, in the partition $r_1p_0|r_0p_1$ if the outcome is the ideal policy of the rich it is not stable. For any other policy, either r_0 or p_1 prefer to break it since the policies are on their Pareto sets. The partition $r_0r_1p_0|p_1$ is not stable since the same equilibria can be achieved in a smaller party. The partitions $r_0|r_1p_0p_1$, $r_1|r_0p_0p_1$ and the grand coalition are not stable because the poor can always deviate and get something weakly better.

We are therefore left with $r_0p_0|r_1|p_1$ and $r_0|r_1p_0|p_1$. But if any of the party members break their party, they achieve lower utility from p_1 winning. These are therefore stable.

Case 4

The partition $r_0p_1|r_1p_0$ is not stable; when the Pareto set of r_0p_1 is interior, then p_1 will break: some policies which r_1p_0 implement are better for p_1 than r_0 and as a result from all the policies in the common part of the Pareto set which are the equilibria in $r_0p_1|r_1p_0$. When the Pareto set of r_0p_1 is on the boundaries, then for one of them, r_0 or p_1, there is a better equilibrium if they break, since the equilibria are also on the Pareto set of r_0 and p_1. The partition $r_0p_0|r_1p_1$ is not stable. For it to be stable, p_0 must get something better than p_1 which is an equilibrium if he breaks. On the other hand, if r_1 breaks he can get as an equilibrium in r_0p_0, the policy which makes p_0 indifferent to p_1. Since this policy is in the Pareto set of r_1p_0, and provides both players with the minimum utility they get if they break their parties, it must be that the current equilibrium policy in $r_0p_0|r_1p_1$ cannot be a Pareto improvement for both r_1 and p_0, a contradiction. All other partitions are as in case 3.

This completes the proof of Propositions 2 and 3.■

References

[1] Y. Barzel, Private Schools and Public School Finance, Journal of Political Economy 81 (1973), 174-186.

[2] P. Bearse, G. Glomm and E. Janeba, Why Poor countries rely mostly on redistribution in kind, Journal of Public Economics 75 (2000), 432-481.

[3] P. Bearse, G. Glomm and E. Janeba, Composition of Government Budget, Non-Single Peakedness, and Majority Voting, Journal of Public Economic Theory 3 (2001), 471-481.

[4] T. Besley and S. Coate, Issue Unbundling via Citizens' Initiatives, mimeo (2000).

[5] T. Besley and S. Coate, An Economic Model of Representative Democracy, Quarterly Journal of Economics (1997), 85-114.

[6] T. Besley and S. Coate, Public Provision of Goods and the Redistribution of Income, American Economic Review 81 (1991), 979-984.

[7] A. Dixit and J. Londregan, The Determinants of Success of Special Interests in Redistribute Politics, Journal of Politics 58 (1996), 1132-1155.

[8] D. Epple and R. Romano, Public Provision of Private Goods, Jounral of Political Economy 4 (1996a), 57-84.

[9] D. Epple and R. Romano, Ends Against the Middle: Determining public service provision when there are private alternatives, Journal of Public Economics 62 (1996b), 297-325.

[10] R. Fernandez and R. Rogerson, On the Political Economy of Education Subsidies, Review of Economic Studies 62 (1995), 249-262.

[11] R. Fernandez and R. Rogerson, The Determinants of Public Education Expenditures: Longer-run Evidence from the States: 1950-1990, Journal of Education Finance 27 (2001), 567-583.

[12] G. Glomm and D. Patterson, Endogenous Public Expenditures on Education, working paper, Indiana University (2003).

[13] G. Glomm and Ravikumar, Opting out of publicly provided services a majority voting results Social Choice and Welfare 15 (1998) , 187-193.

[14] M. Gradstein and M. Kaganovich, Aging Population and Education Finance, CEPR discussion paper 3950 (2003).

[15] G. Grossman and E. Helpman, Protection for Sale, American Economic Review (1994).

[16] G. Levy, A Model of Political Parties, forthcoming, Journal of Economic Theory.

[17] M. Osborne and A. Slivinski , A Model of Political Competition with Citizen-Candidates, Quarterly Journal of Economics (1996), 65-96.

[18] T. Persson and G. Tabellini, Political Economics - Explaining Economic Policy, MIT Press: Cambridge, 2000.

[19] J. Poterba, Demographic Structure and the Political Economy of Public Education, Journal of Policy Analysis and Management 16 (1997). 48-66.

[20] K. Roberts, Voting over Income Tax Schedule, Journal of Public Economics 8 (1977), 329-340.

[21] J. Roemer, Why Don't The Poor Expropriate the Rich, Journal of Public Economics 70 (1998), 399-424.

[22] J. Stiglitz, The Demand for Education in Public and Private School Systems, Journal of Public Economics 3 (1974), 349-385.