

NBER WORKING PAPER SERIES

POLITICAL ACTIVISM AND THE PROVISION OF DYNAMIC INCENTIVES

Antoine Camous
Russell Cooper

Working Paper 26654
<http://www.nber.org/papers/w26654>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

Thanks to Andrea Mattozzi, Guido Tabellini, Fritz Sager and seminar participants at University of Mannheim, European University Institute and Harvard CES. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Antoine Camous and Russell Cooper. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Political Activism and the Provision of Dynamic Incentives
Antoine Camous and Russell Cooper
NBER Working Paper No. 26654
January 2020, Revised April 2020
JEL No. D72,D74,E62,H31

ABSTRACT

This paper studies the determination of income taxes in a dynamic setting with human capital accumulation. The goal is to understand the factors that support an outcome without complete redistribution, given a majority of relatively poor agents. In the analysis, the internal dynamics of income are not sufficient to prevent complete redistribution under majority rule without commitment. However, a political influence game limits the support for expropriatory taxation and preserves incentives. In some cases, the outcome of the game corresponds with the optimal allocation under commitment.

Antoine Camous
University of Mannheim
Department of Economics
L7 3-5 68131
Mannheim,
Germany
camous@uni-mannheim.de

Russell Cooper
Department of Economics
European University Institute
Villa La Fonte
Via delle Fontanelle 18
I-50014 San Domenico di Fiesole (FI)
ITALY
and NBER
russellcoop@gmail.com

Political Activism and the Provision of Dynamic Incentives: Growing the Pie in the Battle for Redistribution*

Antoine Camous[†], Russell Cooper[‡]

April 8, 2020

Abstract

This paper studies the determination of income taxes in a dynamic setting with human capital accumulation. The goal is to understand the factors that support an outcome without complete redistribution, given a majority of relatively poor agents. In the analysis, the internal dynamics of income are not sufficient to prevent complete redistribution under majority rule without commitment. However, a political influence game limits the support for expropriatory taxation and preserves incentives. In some cases, the outcome of the game corresponds with the optimal allocation under commitment.

Keywords: *activism, electoral competition, commitment, redistribution, human capital.*

JEL classification: D72, D74, E62, H31.

1 Introduction

Income distributions are asymmetric, with a majority of agents earning below average income. However, in contrast to predictions of simple models of electoral competition, labor income taxes are progressive but far from confiscatory. Accordingly, this paper asks: what factors constrain the relatively poor from expropriating the income and wealth of the relatively rich?

In a seminal contribution, Bénabou and Ok (2001) provides two key conditions, based upon the *promise of upward mobility* (POUM), that alter the incentives of the relatively poor to tax the rich. First, a change in the tax structure must be permanent, or at least difficult to undo. Second, there must be sufficient mobility in the income distribution over time: the poor of today recognize that they may be the wealthy of tomorrow. Hence, they are not in favor of high tax rates on the rich today since these same rates are likely to apply to them in the future.¹

*Thanks to Andrea Mattozzi, Guido Tabellini, Fritz Sager and seminar participants at University of Mannheim, European University Institute and Harvard CES.

[†]Department of Economics, University of Mannheim, camous@uni-mannheim.de

[‡]Department of Economics, European University Institute, russellcoop@gmail.com

¹A third necessary condition for POUM to hold is that individuals are not too risk averse so that poor agents find it attractive to be rich in the future.

If, to the contrary, there is either insufficient mobility or sufficient flexibility in the tax system, then the relatively poor will favor the immediate and complete taxation of the rich.

The analysis of Bénabou and Ok (2001) imposes an exogenous dynamic on the income distribution and thus misses adverse incentive effects stemming from redistribution. In contrast, our analysis considers an explicit human capital decision that is forward looking and thus dependent on future taxation. This choice may be interpreted as formal education or more broadly as the accumulation of experience and the generation of ideas. The interaction of individual education decisions with the underlying distribution of ability implies that the dynamics of the income distribution do not satisfy the POUM conditions.²

Along with the consideration of incentives, timing matters: are taxes set prior to the human capital decision? If we maintain commitment to *ex ante* tax choices, as in Bénabou and Ok (2001), then simple majority voting implements a social optimal level of taxes, not full redistribution. However, in the absence of commitment, taxes are set after the education choice: majority voting implements full redistribution, which in turns eliminates private incentives to invest in human capital.

Thus in our environment with endogenous mobility, no commitment and majority voting over taxes, the POUM hypothesis does not hold: the relatively poor will vote for complete redistribution. This first result motivates our analysis of alternative political institutions that might restrain full redistribution.

This study deviates from standard electoral competition in two ways: probabilistic voting and activism. As in Lindbeck and Weibull (1987), probabilistic voting introduces stochastic elements into individual political preferences, interpreted as evaluations of politicians beyond their economic platforms. This framework allows other factors to influence voters' perceptions of a candidate, including the persuasive efforts of activism.

Our notion of activism relates to the analysis of “political competition among pressure groups” conceptualized by Becker (1983): *Political influence is not simply fixed by the political process but can be expanded by expenditures of time and money on campaign contributions, political advertising and in other ways that exert political pressure.*³ Hence, activism embodies the idea that individuals, acting through groups, take joint actions to influence political opinions. This may take a variety of forms, ranging from direct persuasion within social circles to influencing public opinion through media and online campaigns. Overall, activism by large income groups “distorts” the political ideal of one person, one vote.

We embed these political institutions in a dynamic model to study the interaction between human capital accumulation, electoral competition and activism. In our analysis, activism, viewed as a continuous process, shapes individual policy preferences prior to the education decision, and these effects persist over the lifetime of an individual.⁴ Through the persuasion of voters, activism impacts electoral outcomes, and in turn influences

²In Bénabou and Ok (2001), an essential condition for POUM to hold is that the income transition function displays negative skewness in the future: for the relatively poor median voter today to vote against redistribution tomorrow, it must be that she becomes richer than average tomorrow. Even with uncertainty, this is an unattractive description of income distributions. In our environment, the evolution of the income distribution is endogenous, but positive skewness is preserved.

³Similarly, building on the pioneering work by Olson (1965) on collective action and the paradox of coordination, Maisel, Berry, Edwards, and Schlozman (2010) argue that formal and institutionalized interest groups are only a part of the set of organizations that represent collective interests in politics: influence activities are undertaken by a multitude of organizations with diverse memberships, structures and purposes, generating a *heavenly chorus* of interests.

⁴The central idea is that voter's policy preferences are shaped by social context and continuous information rather than specific electoral communication, as argued by Beck, Dalton, Greene, and Huckfeldt (2002). Using a panel dataset around elections in OECD countries, Pennec and Pons (2019) find that changes in vote choice during electoral seasons are mostly driven by changes in beliefs about candidates and policy salience, while policy preferences remain stable over time. The social origin of policy preferences is illustrated by Alesina and Fuchs-Schündeln (2007): they contrast East and West German policy preferences and find that *45 years of heavy state*

education choices and candidates' redistributive platforms: our results point to the social gains from a political system with activism under a lack of commitment.

Indeed, when income groups organize as activists to influence the political preferences of voters, the equilibrium outcome no longer coincides with full redistribution. As the analysis makes clear, there are conflicts across influence groups: rich income activists support the low tax candidate while poor income agents promote the high tax one. But all influence groups internalize the benefits of preserving dynamic incentives. This provides a basis for relatively rich households to contribute more in favor of low taxes and, at the same time, motivates the relatively poor to advocate less in favor of high taxes.⁵ Overall, activism induces the relatively poor to vote for lower tax rates.

Candidates, in deciding on their policy platforms, internalize how activism will tilt voters preferences toward (or away from) their proposed redistribution rate. In equilibrium, activism eliminates the incentive for candidates to form a coalition around full redistribution. When the impact of activism is sufficiently strong on voter's preferences, the disciplining effect on candidates is powerful enough that the *ex ante* social optimum allocation with commitment is supported in equilibrium.

There are three important elements of our findings. First, our environment does not impose any asymmetry across income groups in terms of participation or activism technology.⁶ The outcome of less than full redistribution reflects the fact that all groups benefit from the dynamic incentives created by lower taxes. Activism provides a mechanism for the expression of these preferences.⁷ Second, the level of activism that arises is the outcome of strategic interaction between interest groups and politicians. Each group activism choice is a function of competing policy platforms and also a best response to all other groups influence decision. Third, as explained in detail, the beneficial effects of activism arise even though along the equilibrium path, the level of activism is zero by all groups. Thus, it is the credible threat of activism that creates the incentive for candidates to lower taxes below full redistribution.⁸ A deviation from the equilibrium level of taxation towards more redistribution would incentivize activism and lead to a lower probability of winning the election.

Section 2 discusses related literature. Section 3 presents the economic environment. Section 4 derives a policy benchmark under commitment and discusses limits to political decentralization. Section 5 studies how political activism shapes economic outcomes under electoral competition. All derivations and proofs are detailed in an Appendix.

2 Literature

This section contains a review of the related literature with a focus on the political economy of redistribution, dynamic incentives under a lack of commitment and activism. Specific papers that relate to the model and findings are discussed in the analysis.

intervention and indoctrination instill in people the view that the state is essential for individual well-being., i.e., preferences are long lasting and individual social experience shapes economic beliefs - for instance toward redistribution.

⁵Importantly, the poor need to participate in the activism game to mitigate the low tax preferences of the rich.

⁶Prominent studies rely on exogenous asymmetries to avoid outcomes with full redistribution: Bénabou (2000) assumes exogenous differences in voting participation or Persson and Tabellini (2002) assume exogenous differential mass of swing voters across the population. Section 5.4 studies how asymmetries in activism technology influence our results.

⁷To be clear, without human capital accumulation, our model is static and activism has no social value: full redistribution is the equilibrium outcome.

⁸In an extension, we consider the possibility that candidates are interested in the policy implemented. This generates an outcome with activism in equilibrium. See Section 5.4.

There is a large literature relating political institutions and redistribution, see Persson and Tabellini (2002) and Alesina and Giuliano (2011) for reviews. In the classic analysis by Meltzer and Richard (1981), democratic enfranchisement leads to higher tax rates, higher redistribution and lower inequality. Alesina and Ferrara (2005) presents empirical evidences that income level explains individual preferences for redistribution. The puzzle remains: why don't we observe higher level of redistribution? The literature has considered four broad explanations.

First, democracy is imperfect. *De jure* or *de facto* elements break the ideal of *one person, one vote* and provide higher relative influence to richer agents. Examples include *one dollar, one vote* as in Karabarbounis (2011), implications of legislative bargaining as in Mattozzi and Snowberg (2018) or differential voting participation as in Bénabou (2000). Second, individuals might internalize either the benefits of low taxes from social mobility, as in Bénabou and Ok (2001), or various externalities from inequality and redistribution. Third, following Piketty (1995), some authors emphasize that voters are confronted with imperfect information on the origin of income inequality. This generates multiple systems of beliefs consistent with different levels of redistribution, as for instance in Alesina and Angeletos (2005), Bénabou and Tirole (2006) or Bénabou (2008).⁹ Finally, fully redistributive tax schedules are costly to implement under asymmetric information, because incentives must be provided for agents to report their income truthfully. Bierbrauer and Boyer (2013) and Bierbrauer and Boyer (2015) study the efficiency of political competition under this informational friction.

Our analysis introduces activism: a non-cooperative game across interest groups that compete to influence political opinion and redistributive outcomes. Becker (1983) presents a related theory, where competition among pressure groups to curry favor from politicians leads to beneficial economic outcomes. Political environments with strategic group choice also include the ethical voter's models by Coate and Conlin (2004) or Feddersen and Sandroni (2006), where voter's turnout is endogenous: individuals weigh voting participation cost against the benefits to their affiliated group. Bierbrauer, Tsyvinski, and Werquin (2017) develop this idea in a static game of redistribution. The participation margin induces left-leaning parties not to propose high tax rates, to demobilize rich agents from voting.¹⁰ Similarly, our analysis emphasizes the implications of competition among interest groups, but our focus lies on the dynamic benefits in an environment with a lack of commitment.

Several studies consider the implications of a lack of commitment on the political economy of dynamic taxation.¹¹ These include for instance Acemoglu, Golosov, and Tsyvinski (2010), where elections operate as a disciplining device on politicians to implement tax schedules similar to the one that would prevail under commitment. Our analysis rests on an original mechanism - *the power of persuasion* - to eliminate the socially undesirable outcome of excessive redistribution. In our model, democracy is imperfect in the sense that activism induces an equilibrium outcome aligned with the economic interests of the rich at the time of the vote. But the effect of group activism on voter's preferences leads to the electoral defeat of highly redistributive platforms and provides dynamic economic benefits to everyone.

⁹Alesina, Stantcheva, and Teso (2018) document how beliefs about intergenerational mobility affect preferences for redistribution using cross-country surveys.

¹⁰Levine and Mattozzi (2018) develop this framework to study social norm enforcement of voter turnout with costly peer punishment.

¹¹Farhi, Sleet, Werning, and Yeltekin (2012) or Scheuer and Wolitzky (2016) study nonlinear capital taxation under exogenous and partial commitment and various political protocols. To be clear, in the absence of commitment, full redistribution is the equilibrium outcome in their environment.

3 Environment

Consider a two period $t = 1, 2$ economy populated by a continuum of agents. Agents at $t = 1$ differ in ability $\theta \sim \log \mathcal{N}(m, \sigma^2)$, with cumulative distribution function noted $F(\theta)$. They decide on education which influences next period's income.¹² At $t = 2$, agents are subject to productive idiosyncratic shocks. Taxes and transfers are applied to period 2 gross income, then agents consume net income. The economic channels of fiscal policy are multiple: it *redistributes* income across the population, provides *insurance* against idiosyncratic shocks and *distorts dynamic choices*.

The specific functional forms adopted here are chosen to highlight transparently the interplay between political institutions, growth and redistribution. Appendix D recasts our game without these functional assumptions and establishes sufficient conditions that support our central result.

3.1 Individual dynamic choice

An agent with ability θ at $t = 1$ invests in education e to maximize lifetime utility:

$$\max_e \log(c) + \beta E_{z,\tau}(\log(c')), \quad (1)$$

subject to:

$$c = \theta - e \quad \text{budget constraint } t = 1 \quad (2)$$

$$\theta' = z\theta^\alpha e^\delta \quad \text{gross income } t = 2 \quad (3)$$

$$c' = \theta'^{1-\tau} \bar{\theta}'^\tau \quad \text{net income } t = 2 \quad (4)$$

An agent with ability θ is referred to as a type θ agent.

Equation (2) is the period 1 budget constraint: gross income is either consumed or spent on education. The wage is normalized at unity so that labor income is equated to initial ability θ . Equation (3) is the dynamic evolution of human capital and hence income. Both current ability θ and education choice e determine future income θ' up to an idiosyncratic shock $z \sim \log \mathcal{N}(-\frac{w^2}{2}, w^2)$. The parameters $\alpha > 0$ and $\delta \geq 0$ measure respectively depreciation of human capital and return to education.¹³

Fiscal policy in period 2 is summarized by (4). The isoelastic tax function is common in the literature on progressive labor taxes: the higher the redistribution rate $\tau \in [0, 1]$, the lower the dispersion of net income.¹⁴ From an *ex ante* perspective, the redistribution is a form of insurance against variations in z .

$\bar{\theta}'$ is a break-even income level which sorts the population in net beneficiaries and contributors. The expectation operator in (1) reflects uncertainty over the magnitude of taxes and transfers τ . The institutional structure that determines the period 2 tax rate will be a key element in the analysis.

¹²Throughout wages are set to unity so that human capital and income are the same.

¹³By design, individual education is not directly influenced by the human capital decisions of others. This allows to isolate the interaction of agents through fiscal policy.

¹⁴Bénabou (2000), Bénabou (2002) and Heathcote, Storesletten, and Violante (2017), among others, adopt a similar tax function. Under this specification, agents report their true income to the tax authority. *Ex post* incentives to misreport income are not included in this analysis, allowing us to focus on the *ex ante* adverse incentive effects of taxation on human capital accumulation.

The optimal education choice satisfies:

$$e(\theta, \bar{\tau}) = \epsilon(\bar{\tau})\theta, \quad (5)$$

where $\bar{\tau}$ is the expected average redistribution rate, and $\epsilon(\bar{\tau}) \equiv \frac{\beta\delta(1-\bar{\tau})}{1+\beta\delta(1-\bar{\tau})}$ is an education rate common to all agents. In the limit case of extreme redistribution, i.e. $\tau = \bar{\tau} = 1$, there is no private return to education, so that in period 2 agents have zero income and thus zero consumption. This highlights the incentive effects of the labor tax.

Evolution of the income distribution. Gross income at $t = 2$ is log-normally distributed, with mean m' and standard deviation σ' given by:

$$m' = (\alpha + \delta)m + \delta \log(\epsilon(\bar{\tau})) - \frac{w^2}{2}, \quad (6)$$

$$\sigma'^2 = (\alpha + \delta)^2 \sigma^2 + w^2. \quad (7)$$

Fiscal intervention. Through taxes and transfers, determined by τ , the fiscal intervention is purely redistributive. The critical income level $\bar{\theta}'$ sorts agents in net contributors $\theta' > \bar{\theta}'$ and net beneficiaries $\theta' \leq \bar{\theta}'$. It reads:

$$\log(\bar{\theta}') = m' + \frac{\sigma'^2}{2}(2 - \tau). \quad (8)$$

Note the dual influence of redistribution on $\bar{\theta}'$. First, average log-income m' is a function of education and the expected tax rate $\bar{\tau}$, as explicit in (6). Second, $\bar{\theta}'$ is directly decreasing in τ : the share of the population that pays more taxes than it receives transfers is increasing in the magnitude of the redistributive program.

This expression makes clear that $\bar{\theta}'$ also reflects the tax base for redistributive policies. Hence, through (8) the model contains a fiscal externality. If all agents in the economy increase their education level, then $\bar{\theta}'$ will increase, as well as the consumption of all agents. Though no single agent takes this into account in making its education choice, both *ex ante* bliss tax policies and group choices will internalize this effect, as developed below.

3.2 Value functions and bliss policies.

To highlight the influence of redistribution on individual preferences and welfare, we contrast individual bliss policies at two points in time: before agents form an education choice at $t = 1$, and after, at $t = 2$, when idiosyncratic uncertainty remains.¹⁵

At $t = 1$, before agents choose education. Let $V_1(\theta, \tau)$ be the value function of an agent of type θ evaluating a rate of fiscal redistribution τ :

$$V_1(\theta, \tau) = \log(\theta - \epsilon(\tau)\theta) + \beta \left(\underbrace{(1 - \tau) \left(\alpha \log(\theta) + \delta \log(\epsilon(\tau)\theta) - \frac{w^2}{2} \right)}_{=E_z(\log(\theta')|\theta)} + \tau \log(\bar{\theta}') \right). \quad (9)$$

¹⁵This timing allows the exogenous component of mobility to shape individual preferences. It is maintained throughout the analysis.

There are two ways in which τ influences $V_1(\theta, \tau)$. First, the individual education decision $e(\theta, \tau) = \epsilon(\tau)\theta$ is sensitive to τ , as explicit in (5). Second, the fiscal tax base, captured by the government break-even income level $\bar{\theta}'$ (8), responds to the tax rate as well.

The favorite redistribution rate $\tau^*(\theta)$ of a type θ agent is the tax rate $\tau \in [0, 1]$ that solves $\frac{dV_1(\cdot)}{d\tau} = 0$. From the first-order condition, it is implicitly given by:

$$\beta(\alpha + \delta)(m - \log(\theta)) + \beta \underbrace{((\alpha + \delta)^2 \sigma^2 + w^2)}_{=\sigma'^2} (1 - \tau) + \beta \delta \frac{\tau}{\epsilon(\tau)} \frac{d\epsilon(\tau)}{d\tau} = 0. \quad (10)$$

The first two terms in this expression capture the relative support for *redistribution*. At least all agents with income below (log) median level m benefits (on average) from redistributive policies. Also, an increase in individual risk w^2 generates higher desire for *insurance* via redistribution across the population. The third term is the elasticity of the education rate $\epsilon(\tau)$ to redistribution τ . It is negative and captures the willingness to preserve individual *dynamic incentives* against distortionary redistribution. It is straightforward to show that $\tau^*(\theta) < 1$ is unique and decreasing in θ : higher income agents prefer lower rates of redistribution.¹⁶

At $t = 2$, after individual choices. Let $V_2(\theta, \tau|\epsilon)$ be the value function of a type θ agent after $t = 1$ consumption and education choices, before a realization of idiosyncratic risk z .¹⁷

$$V_2(\theta, \tau|\epsilon) = \beta \left((1 - \tau) \underbrace{\left(\alpha \log(\theta) + \delta \log(\epsilon\theta) - \frac{w^2}{2} \right)}_{=E_z(\log(\theta')|\theta)} + \tau \log(\bar{\theta}') \right). \quad (11)$$

At that stage, education levels are no longer sensitive to tax policies τ : the education rate ϵ is a sufficient statistic to describe individual education $e = \epsilon\theta$ and the aggregate tax base. But, the break even income level $\bar{\theta}'$ is sensitive to τ insofar that it sorts agents between contributors and beneficiaries of the tax program, given average log income m' .

The favorite redistribution rate $\tau^d(\theta)$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau} = 0$:

$$\beta(\alpha + \delta)(m - \log(\theta)) + \beta \underbrace{((\alpha + \delta)^2 \sigma^2 + w^2)}_{=\sigma'^2} (1 - \tau) = 0, \quad (12)$$

or the corner solution $\tau^d(\theta) = 1$ for lower income agents.¹⁸ Further, as education is realized, there is no term capturing *dynamic incentives* effect of redistribution, in contrast to (10). This expression reflects otherwise the relative preferences for redistribution, decreasing with income, and desire of insurance against idiosyncratic shocks.

Figure 1 represents bliss policies before and after education, respectively $\tau^*(\theta)$ and $\tau^d(\theta)$. Favorite redistribution rates are decreasing in income θ . When agents internalize the effect of taxes on incentives, they favor lower rates

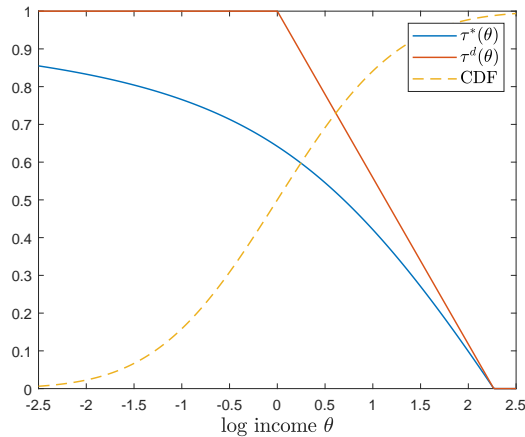
¹⁶For formal derivations, see Appendix A. This expression reflects in part our assumptions on preferences and distributions. But the generic trade-offs embedded in the interactions between dynamic choices and redistributive policies generalize to less restrictive preference and technology specifications, as shown in Appendix D.

¹⁷Given that the ranking of income before the realization of idiosyncratic uncertainty is maintained over time, we continue to order agents by initial income θ and CDF $F(\theta)$.

¹⁸Note that bliss policies $\tau^d(\theta)$ are only a function of income, not education ϵ : the size of the tax base is irrelevant and only redistributive conflicts determine bliss policies.

than they would after education is made: $\tau^*(\theta) \leq \tau^d(\theta)$ for all θ . Finally, after education, a majority supports complete redistribution.¹⁹

Figure 1: Individual Bliss Policies



This figure represents bliss policies before and after education, as a function of (log) income. Favorite rates are decreasing in income and agents support higher levels of redistribution once education is made. The yellow dashed line shows the underlying distribution of log-income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1.]

4 Outcomes under Commitment: τ^*

Policy maker with commitment. Before formally defining political protocols, we characterize a key normative benchmark. Assume the policy choice is made by a benevolent policy maker at $t = 1$ with commitment. That is, taxes are chosen prior to the education choice and *ex post* are not subject to change.²⁰

Formally, a benevolent policy maker with commitment chooses a tax rate τ^* to maximize expected utility over the population:

$$\max_{\tau} \int_{\theta} V_1(\theta, \tau) dF(\theta). \quad (13)$$

Using (10), $\tau^* \in (0, 1)$ is the solution to:

$$\underbrace{((\alpha + \delta)^2 \sigma^2 + w^2)}_{=\sigma'^2} (1 - \tau) + \delta \frac{\tau}{\epsilon(\tau)} \frac{d\epsilon(\tau)}{d\tau} = 0. \quad (14)$$

The terms in this expression highlight how $\tau^* \neq 1$ balances (average) preferences for *redistribution / insurance* and *incentives*. The first term captures the insurance over ability θ and human capital risk z . The optimal rate of redistribution is increasing in income inequality σ^2 and idiosyncratic risk w^2 .²¹

¹⁹The presence of positive skewness, as in the data of most if not all countries, is maintained throughout our analysis.

²⁰In contrast taxes set without commitment are chosen in period 2, after the education decision.

²¹If redistribution over initial income θ were allowed, this would decrease initial income inequality and decrease the optimal rate of redistribution. But the commitment tension at the heart of the mechanism would be maintained.

The second term captures the negative effect of higher taxes on human capital accumulation. The magnitude of this effect is parameterized by the return to education δ . But the overall effect of δ on τ^* is ambiguous, since it contributes both to higher return to education but also to income dispersion.

Political decentralization. This allocation can be decentralized through simple electoral competition with majority voting.²² Interestingly, this does not depend on the timing of the vote relative to the realization of ability, θ . But it is critical that the vote is taken prior to education decisions.

First, suppose majority voting were to take place before the realization of ability. In this case, all agents are identical and their preferred tax policy would coincide with τ^* defined in (14). There is no conflict behind the veil of ignorance. Second, suppose the vote on the tax rate takes place after the realization of ability. Now there are well defined rich and poor agents. But key to the outcome is whether the vote takes place before or after education choices.

Proposition 1. *The allocation under τ^* can be decentralized under electoral competition with majority voting if voting takes place **before** the education choice. If, instead, voting takes place **after** education, then full redistribution is the outcome of electoral competition.*

In the first case majority voting does yield the efficient outcome τ^* , without full redistribution. The channel though differs from that identified in the POUM argument. In our environment with human capital accumulation, the median voter, in determining a preferred tax, internalizes the effects of taxes on the education choice of all others, and accordingly the size of the tax base m' for fiscal interventions, as explicit in (9). The critical element for this result is the timing: taxes are set prior to education choices.

In contrast, if majority voting were to happen after education, individual agents would no longer internalize the effect of policies on education, as discussed in Section (3.2). This yields an outcome of full redistribution, as it coincides with the bliss policy of a majority of the population, see Figure 1. In these circumstances, the anticipation of full redistribution eliminates private incentives to invest in human capital.

5 Probabilistic voting and activism

A key result from the previous section is that majority voting can support the efficient tax rate τ^* **iff** the vote is taken **prior** to the education choice. To the extent that education decisions are made before the determination of taxes, the challenge is to identify conditions such that incentives for human capital accumulation are preserved.

Our focus is on the political system. In our setting, candidates propose tax rates to maximize vote share. We augment the environment with a political protocol that combines pre-election politics and probabilistic voting. Probabilistic voting follows Lindbeck and Weibull (1987): agents evaluate candidates based upon their preferred tax rates and another dimension that reflects political preferences. Political activism has an impact on the voting outcome through this second dimension.

Specifically, at the activism stage, agents with similar income levels organize as influence groups and decide cooperatively on non-pecuniary contributions to influence voters political preferences. Influence groups are self-

²²Simple electoral competition follows standard exposition in Persson and Tabellini (2002): two office seeking candidates propose competing policy platforms. An equilibrium policy survives pairwise evaluation to all possible policy alternative.

interested and large, in the sense that their actions can influence the voting outcome. Activism takes place prior to the education choice and after the candidates have chosen their tax platforms. This reflects the idea that activism is an ongoing process in contrast to electoral campaigns and elections that take place at discrete points in time. Of course, candidates appreciate that their choice of policy platforms can elicit a response through the level and direction of political activism.

In equilibrium, activism matters and can push the outcome away from complete redistribution despite lack of commitment. With this reduction in the probability of high taxes, agents retain an incentive to invest in human capital. As the influence of activism on voters' preferences grow, the outcome of the game converges to the *ex ante* efficient outcome τ^* . Interestingly, there is no activism along the equilibrium path. The credible prospect of activism is enough to discipline candidates.

These findings rest on two key properties of the model, established above. First, the POUM argument fails in our environment so that, absent commitment, tax rates are set at unity by all candidates. Second, with a positive return to education $\delta > 0$, the optimal tax rate is less than unity for all types: $\tau^*(\theta) < 1$ for all θ . In effect, all agents want to have taxes less than unity to preserve dynamic incentives but, without commitment, cannot achieve this outcome under simple electoral competition. Anticipating the activism response of interest groups, a candidate has an incentive to choose a tax platform lower than unit and increase the probability of winning the election. This argument is shown for our specific model in the analysis that follows and is presented in a more general setting in the Appendix D.

5.1 Timing of the game

The timing highlights both that human capital is determined prior to the vote on taxes and how activism influences the election outcome. Formally, the sequence of events is:

- i. Choice of platforms: two office seeking candidates from competing parties L and H propose redistributive platforms $\tau_l \leq \tau_h$.
- ii. Pre-election politics: activism.
- iii. Individual choice at $t = 1$: agents, given their type θ , chooses consumption and education e .
- iv. Political preferences: individuals are subject to idiosyncratic and aggregate political preference shocks.
- v. Vote: given policy platforms and political preferences, agents participate in a majority election and the winning candidate takes office.
- vi. Realizations of individual income shock z , tax and transfer and $t = 2$ consumption.

This timing calls for some comments. First, individual income uncertainty realizes after the vote, to give a chance, as in POUM, for *insurance* and *upward mobility* to influence the vote against the most redistributive policy platforms. Second, the vote takes place after the education choice, precisely to investigate whether highly redistributive platforms would emerge without commitment. Finally, activism takes place at a pre-election stage, allowing it to shape individual political preferences and in turn candidates policy platforms.

In order to isolate the effects of activism from the other elements of the political protocol, the first step is to characterize the political equilibrium of this game without activism. As we demonstrate, the outcome coincides with *ex post* simple majority rule: complete redistribution. Thus the results of Proposition 1 extend to a setting with probabilistic voting without activism.

5.2 No Activism

Absent pre-election politics, policy platforms (τ_l, τ_h) are decided by office seeking candidates anticipating the outcome of the probabilistic vote. The following exposition derives from the sequential nature of the game.

Voting outcome. Given policy platforms (τ_l, τ_h) and education rate ϵ , individuals trade off political and economic preferences and cast their vote sincerely. They evaluate policy platforms $\tau_l \leq \tau_h$ according to the value function $V_2(\theta, \tau|\epsilon)$ along with the realizations of idiosyncratic χ and aggregate ψ political preference shocks for candidate L. A type θ agent with education $e = \epsilon\theta$ votes for party H if and only if:

$$V_2(\theta, \tau_h|\epsilon) > V_2(\theta, \tau_l|\epsilon) + \chi + \psi. \quad (15)$$

As in Persson and Tabellini (2002), these shocks are distributed as:

$$\chi \sim U\left(-\frac{1}{2\phi}, \frac{1}{2\phi}\right) \quad \psi \sim U\left(-\frac{1}{2\Psi}, \frac{1}{2\Psi}\right). \quad (16)$$

They differ only because the average of the idiosyncratic shock χ across the population is zero, while ψ is common across agents.²³

Given a realization of aggregate preference ψ , let $\chi(\theta, \psi)$ be the *swing voter* for agents with income θ : type θ agents vote for party H if and only if $\chi \leq \chi(\theta, \psi)$. From (15),

$$\chi(\theta, \psi) = V_2(\theta, \tau_h|\epsilon) - V_2(\theta, \tau_l|\epsilon) - \psi = \Delta V_2(\theta) - \psi, \quad (17)$$

where $\Delta V_2(\theta)$ is the economic gain (or loss) to agents with initial income θ of τ_h over τ_l :

$$\Delta V_2(\theta) = \beta(\tau_h - \tau_l) \left[(\alpha + \delta)(m - \log(\theta)) + \frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) \right]. \quad (18)$$

This expression makes clear that $\chi(\theta, \psi)$ and $\Delta V_2(\theta)$ do not depend on the actual education rate ϵ , as individual preferences are only driven by distributional conflicts at this stage. The vote share for party H within group θ and across the population are:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi}}^{\chi(\theta,\psi)} \phi dj = \phi \left(\chi(\theta, \psi) + \frac{1}{2\phi} \right) \quad \pi_h(\psi) = \int_{\theta} \pi_{\theta,h}(\psi) dF(\theta). \quad (19)$$

²³Though these shocks determine political preferences for candidate L, they impact the voting outcome symmetrically. As shown in Appendix C, the results are not sensitive to the mean zero assumption, or to the strict majority requirement introduced below.

In a majority system, the probability p_h that the candidate from party H wins the election is:

$$p_h = P(\pi_h(\psi) \geq 1/2). \quad (20)$$

Combining previous expressions:²⁴

$$p_h = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) \geq \frac{1}{2}. \quad (21)$$

This expression highlights the tendency of the population to lean toward the most redistributive platform τ_h . When elections take place after the education choice, then a pure redistributive conflict drives economic preferences of agents. As illustrated in Figure 1, the positive skewness of the income distribution provides a majority mass of the population benefiting from high redistribution rates.

Choice of platforms. A candidate from party H seeking to maximize its probability of winning chooses to campaign on a redistributive program τ_h that solves:

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h), \quad (22)$$

where $p_h(\tau_l, \tau_h)$ is given by (21). The first order condition is:

$$\frac{dp_h(\cdot)}{d\tau_h} = \psi\beta\sigma'^2(1 - \tau_h) = 0. \quad (23)$$

$\tau_h = 1$ is a dominant strategy. The candidate from party L maximizes $p_l(\tau_l, \tau_h) = 1 - p_h(\tau_l, \tau_h)$, and again $\tau_l = 1$ is a dominant strategy.

Proposition 2. *The outcome of the game without pre-election politics is full redistribution $\tau^p = 1$.*

Intuitively, the election takes place after education choices have been formed. Hence despite redistributive conflicts, the median income agents with average political preferences support unconditionally a platform of full redistribution. In this case, dynamic incentives cannot be preserved and the economy collapses. Probabilistic voting does not alter the outcome under *ex post* simple majority voting characterized in Proposition 1.²⁵ This no longer holds once we introduce activism as an avenue to influence political preferences and electoral outcomes.

5.3 Electoral Competition with Activism

We now consider the equilibrium of the game when agents can actively support candidates and platforms.²⁶ By activism, we mean non-pecuniary contributions aimed to influence political preferences of voters. Which income groups promote the campaign of the high tax candidate? How does activism influence the probability of high taxes

²⁴Naturally $0 \leq p_h \leq 1$. We abstract from assumptions on parameters and restrictions on choices that are not relevant to characterize equilibrium outcomes.

²⁵Persson and Tabellini (2002) and Bénabou (2000) rely on exogenous asymmetry across income groups to generate outcomes of electoral competition that do not coincide with full redistribution, see discussion in Section 5.4. Our analysis moves beyond this exogenous source of asymmetry.

²⁶The equilibrium concept is now two stage Nash equilibrium: first candidates choice of platforms and then activism effort.

and equilibrium policy platforms? Eventually, does activism contribute to mitigate the high propensity of majority systems to lean toward redistributive policies?

We analyze activism intensity chosen by income groups and its impact on the voting outcome. The analysis takes as given group membership and focuses on the intensive margin of group specific contributions.²⁷ Groups are large and thus internalize the effects of their actions on the voting outcome.

Activism impact on the voting outcome. After the announcement of policy platforms (τ_l, τ_h) , at the pre-election stage, each income group θ decides on (non-pecuniary) activism contributions $A_\theta^h \geq 0$ and $A_\theta^l \geq 0$ to promote candidates and their economic platforms. Contributions influence political preferences of voters.

At the time of the vote, these choices are all given. The vote reflects the economic valuations of policy platforms along with political preferences. Note $A^i = \int_\theta A_\theta^i d\theta$ the aggregated influence of activism for each candidate $i \in \{l, h\}$. Given (A^l, A^h) , aggregate preference shock ψ , education rate ϵ and policy platforms (τ_l, τ_h) , an agent with initial income θ and preference shock χ votes for party H if and only if:

$$V_2(\theta, \tau_h | \epsilon) > V_2(\theta, \tau_l | \epsilon) + \chi + \psi + \gamma(A^l - A^h), \quad (24)$$

where $\gamma > 0$ measures how activism influences political preferences. The probability p_h that party H wins the election and τ_h is implemented reads:

$$p_h = \frac{1}{2} + \Psi \int_\theta \Delta V_2(\theta) dF(\theta) + \Psi \gamma (A^h - A^l). \quad (25)$$

Superficially, the term $\gamma(A^l - A^h)$ seems only to shift the distribution of the aggregate political preference shock. This misses an important dimension of the analysis: the choice of platforms is made anticipating the levels of activism. In this way, the prospect of political influence impacts the electoral platforms of candidates.

Activism choice. Every income group is active. The objective of each group is to influence political preferences of voters, with the goal of promoting their economic interests. In our setting, activism takes the form of effort, whose cost appears as a utility loss.

Given all other group activism decisions $\{A_{-\theta}^l, A_{-\theta}^h\}$ and competing platforms $\tau = (\tau_l, \tau_h)$, income group θ decides on total contributions (A_θ^l, A_θ^h) :

$$\max_{A_\theta^l, A_\theta^h \geq 0} f(\theta) V_1(\theta, \tau) - \frac{1}{2} \left((A_\theta^l)^2 + (A_\theta^h)^2 \right). \quad (26)$$

Here $V_1(\theta, \tau)$ is the expected value of a type θ household prior to the choice of education and to the election

²⁷For expositional convenience, it is easier to work with a continuum of agents where each ability type is a distinct group. This allows to highlight differential activism incentives across the income distribution. The same results would go through if agents would form two groups campaigning exclusively for one candidate.

outcome for given political platforms.²⁸ It is similar to (9) but captures uncertainty over the tax rate $\tau \in (\tau_l, \tau_h)$:

$$V_1(\theta, \tau) = E_\tau \left\{ \log(\theta - \epsilon(\bar{\tau})\theta) + \beta((1 - \tau)\alpha \log(\theta) + \delta \log(\epsilon(\bar{\tau})\theta) - \frac{w^2}{2} + \tau \log(\bar{\theta}')) \right\}, \quad (27)$$

Activism internalizes the effect of contributions on the education rate $\epsilon(\bar{\tau})$ and the outcome of the vote, captured by the probability p_h that party H defeats party L, see (25). Income break even level $\bar{\theta}'$ is a function of both $\bar{\tau}$ and τ , as explicit in (8). The group cost in (26) is assumed to be quadratic. One interpretation is there is a cost of organization associated with activism that depends on the contributions of all members of the group.²⁹

The first order condition for $A_\theta^i \geq 0$, $i \in (l, h)$, is

$$f(\theta) \frac{dV_1(\theta)}{dA_\theta^i} = A_\theta^i. \quad (28)$$

The sensitivity of group θ welfare to activism is:³⁰

$$\frac{dV_1(\cdot)}{dA_\theta^i} = \pm \Psi \gamma \beta (\tau_h - \tau_l) \left[(\alpha + \delta)(m_1 - \log(\theta)) + \frac{\sigma^2}{2}(2 - \tau_h - \tau_l) + \delta \bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} \right]. \quad (29)$$

The first two terms reflects the relative position of group θ in the income distribution and the preferences for redistribution, as in (10).

The last term reflects the effects of group contributions on the probability of high taxes and, through the expected tax rate, on the common accumulation of human capital $\epsilon(\bar{\tau})$.³¹ This term is central to understand the outcome of the activism game: regardless of an agent's position in the income distribution, this effect pushes in the direction of low taxes since all agents prefer a large tax base.

This last term is also the locus of strategic interactions, i.e. **conflicts** across the groups. Again, the decision to support a low or high tax candidate as well as the magnitude of group contributions depend on the probability p_h , which hinges on overall activism across the population, as explicit in (25). As long as $\tau_l < \tau_h$, all groups are active, and promote only one candidate: low income agents support only the champion of high taxes, while high income agents only the candidate from party L. Importantly, the split of the population is endogenous. All groups with initial income $\theta < \hat{\theta}$ contribute exclusively for τ_h , where the cut-off income level $\hat{\theta}$ is given by:

$$\log(\hat{\theta}) = m + \frac{1}{\alpha + \delta} \left[\frac{\sigma^2}{2}(2 - \tau_h - \tau_l) + \delta \bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} \right]. \quad (30)$$

Figure 2 report activism per capita $a_\theta^i = A_\theta^i/f(\theta)$ and total group activism A_θ^i for some policy platforms $\tau_l < \tau_h$. Individual activism is increasing in the income difference $|\hat{\theta} - \theta|$. Aggregate group effort is not monotonic, because of the relative size $f(\theta)$ of each income group. Finally, an increase in activism technology γ has two effects: it

²⁸The program (26) omits uncertainty regarding future political taste shocks. This does not change the analysis beyond overburdening notations.

²⁹If the individual cost to agent j with income θ of contributing $a_\theta^{i,j}$ is $\frac{a_\theta^{i,j}}{2} a_\theta^i f(\theta)$, where a_θ^i is the average group contribution, then total group costs is $\frac{A_\theta^i{}^2}{2} = \frac{(a_\theta^i f(\theta))^2}{2}$: individual cost is linear, but increases with group size and average contribution. Appendix D discusses our main result under general technology and preferences.

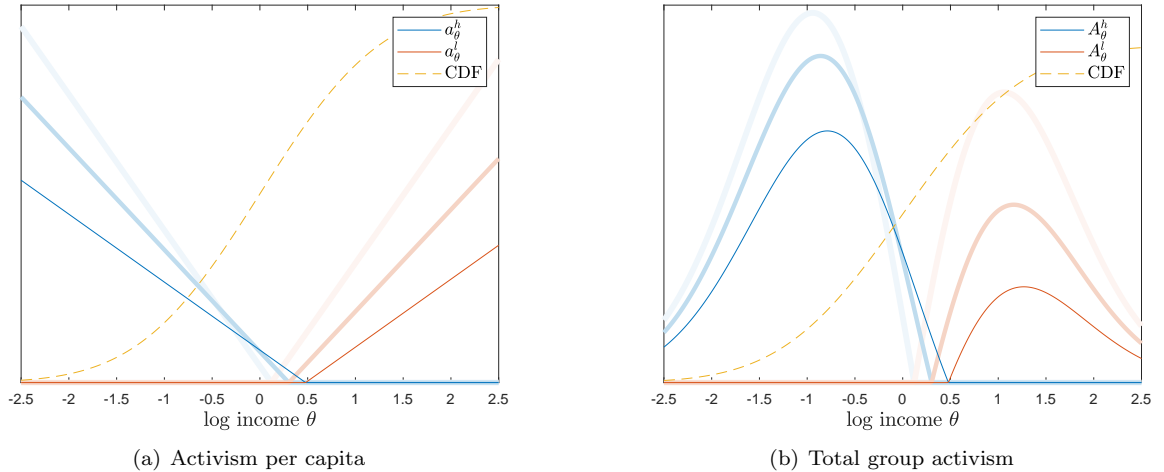
³⁰In this expression, $\pm = \mathbb{1}_{i=h} - \mathbb{1}_{i=l}$.

³¹Formally, $\bar{\tau} = p_h \tau_h + (1 - p_h) \tau_l$.

increases returns to activism but modifies the cut-off $\hat{\theta}$, i.e. it changes the composition of the population that promotes one candidate or the other. Aggregating group contributions (28), one gets:

$$A^h - A^l = \Psi\gamma\beta(\tau_h - \tau_l) \left[\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} \right]. \quad (31)$$

Figure 2: Activism Contribution



This figure represents activism per capita (left panel) and total group contributions (right panel) given two policy platforms $\tau_l = 1/4$ and $\tau_h = 3/4$. Lighter lines correspond to higher values of γ . Dashed yellow line reflects the distribution of (log) income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 1. *Given (τ_l, τ_h) , there is a unique Nash equilibrium of the activism subgame. The induced probability that the high tax candidate wins the election is given by:*

$$p_h(\tau_l, \tau_h) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l) \left[(1 + \Psi\gamma^2) \frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \Psi\gamma^2\delta\bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} \right]. \quad (32)$$

The probability of high taxes takes into account the aggregation of group activism (31) and reflects the resolution of conflicts over redistribution. Importantly, this expression highlights the interactions of activism and human capital accumulation in reducing the probability of high tax rates. In the absence of a return to education, i.e. $\delta = 0$, $p_h \geq \frac{1}{2}$ as in (21). Only when $\delta > 0$ and $\gamma > 0$ does the last term in (32) reduce the probability of high taxes.

Choice of platforms. How does activism influence political competition and equilibrium incentives? At the initial stage of the game, each candidate decides on its economic policy platform anticipating the effect of activism on voters' political preferences and the outcome of the vote. Formally, candidate from party H sets τ_h given τ_l to maximize (32):

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h). \quad (33)$$

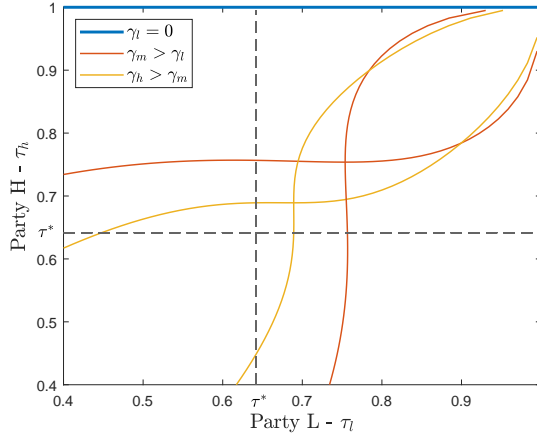
The first order condition leads to:

$$(1 + \Psi\gamma^2)\sigma'^2(1 - \tau_h) + \Psi\gamma^2\delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} + (\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}p_h(\tau_l, \tau_h) = 0, \quad (34)$$

where $\mathcal{E}(\bar{\tau}) = \bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} < 0$ is the elasticity of the education rate to the expected redistribution rate.

The platform choice internalizes the effect of the tax rate on the outcome of the activism subgame. Again, it highlights the interplay between activism and dynamic incentives. In the absence of dynamic choice, i.e. when $\delta = 0$, activism and associated conflicts are irrelevant since $\tau_h = 1$ is a dominant strategy, as in (23). In contrast, when $\delta > 0$, then activism induces strategic interactions across candidates. Figure (3) represents the best response functions of each candidate for different level of activism technology γ .

Figure 3: Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, without activism ($\gamma_l = 0$) and with varying degrees of activism ($\gamma_h > \gamma_m > 0$). [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 2. *There is a unique and symmetric Nash equilibrium $\tau^p = \tau_l = \tau_h$ that satisfies:*

$$\left(1 + \frac{1}{\Psi\gamma^2}\right)\sigma'^2(1 - \tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} = 0. \quad (35)$$

Activism does not imply that parties have an incentive to differentiate from each other in equilibrium, but as long as there are positive returns to education $\delta > 0$, the equilibrium no longer coincides with full redistribution.

Proposition 3. *The equilibrium rate of redistribution τ^p differs from $\tau = 1$ if and only if $\delta > 0$ and $\gamma > 0$. Further, it is decreasing in activism intensity γ , and in the limit case where $\gamma = +\infty$, the political game with activism implements the socially desirable level of redistribution τ^* .*

Along the equilibrium path of this game, there is no political activism since $\tau_l = \tau_h = \tau^p$. Still activist groups stand ready to influence voters political preferences if a candidate were to deviate from the equilibrium platform. Also, the equilibrium platform reflects the relative power of influence groups. If, for instance, the candidate from party L deviates from τ^p and runs on $\tau_l < \tau^p$, then relatively poor agents would engage into activism to convince voters away from this platform. And vice versa. This is the disciplining effect of activism on candidates.

The intensity of the activism technology γ influences the equilibrium rate of redistribution τ^p via the conflictual behavior across groups: for higher γ , the disciplining effect of activism is stronger. Again, the response of education to taxes, parameterized by δ is necessary for this channel.

Overall, activism is a complementary institution to majority voting to implement socially desirable policies in a dynamic environment with lack of commitment. It matters because agents, regardless of their position in the income distribution, appreciate the social benefit of human capital accumulation. So, as long as there is a response of education to activism's effect on expected taxes, $\delta > 0$, the bias towards redistribution under simple majority voting is, at least partially, redressed.

5.4 Extensions

This section studies two key extensions. What is the effect of asymmetry in activism technology on equilibrium outcomes? And, how do office-seeking candidates also interested to the policy outcome interact?

Asymmetric influence. The equilibrium outcome with activism characterized in Proposition 3 does not rely on exogenous asymmetry across groups to move away from full redistribution. In contrast, Persson and Tabellini (2002) consider asymmetry in the mass of swing voters across the population as a necessary condition to avoid full redistribution: when high income agents are more attractive targets to candidates, this reduces the likelihood of high taxes.³²

In this section, we show that the introduction of asymmetry in activism technology might further reduce the likelihood of high taxes. Formally, we consider a situation where income groups differ in their technology to influence political preferences. The cost of contributing is decreasing in α_θ , with $\int_\theta \alpha_\theta d\theta = 1$. Given policy platforms $\tau = (\tau_l, \tau_h)$ and other groups' contributions $\{A_{-\theta}^l, A_{-\theta}^h\}$, the activism choice of income group θ is the solution to:

$$\max_{A_\theta^l, A_\theta^h \geq 0} f(\theta)V_1(\theta, \tau) - \frac{1}{2\alpha_\theta} \left((A_\theta^l)^2 + (A_\theta^h)^2 \right). \quad (36)$$

Let $\rho = \text{cov}(\alpha_\theta, \log(\theta))$ be the covariance between income level and influence technology. The first order condition for $A_\theta^i \geq 0$ then reads:

$$\alpha_\theta f(\theta) \frac{dV_1(\theta, \tau)}{dA_\theta^i} = A_\theta^i. \quad (37)$$

Group contribution is increasing in influence technology α_θ . As $E(\alpha_\theta \log(\theta)) = m_1 + \rho$, the probability $p_h(\tau_l, \tau_h)$ of party H win is then implicitly defined by:

$$p_h(\tau_l, \tau_h) = \frac{1}{2} + (\Psi\gamma)^2 \beta (\tau_h - \tau_l) \left[\left(1 + \frac{1}{\Psi\gamma^2}\right) \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l) + \delta \bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha + \delta)\rho \right]. \quad (38)$$

³²Formally, in our set up assume $\chi_\theta \sim U\left(-\frac{1}{2\phi_\theta}, \frac{1}{2\phi_\theta}\right)$: if $\text{cov}(\phi_\theta, \log(\theta)) < 0$, then the outcome of the political game without activism yields $\tau^p < 1$. See Appendix C.1 for explicit derivations. The same results hold if we allow differential voting participation across income groups, as in Bénabou (2000).

A Nash equilibrium of the game across office seeking candidates $\tau^p = \tau_l = \tau_h$ is then the solution to:

$$\left(1 + \frac{1}{\Psi\gamma^2}\right)\sigma'^2(1 - \tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha + \delta)\rho = 0. \quad (39)$$

In that context, $\frac{d\tau^p}{d\rho} < 0$: the larger the influence of high income groups, the lower the equilibrium rate of redistribution. Accordingly, a positive covariance between income level and activism technology can either compensate a low aggregate activism technology γ and bring the equilibrium level of redistribution toward the optimal level or can tilt the policy rate towards partisan level of redistributions that are too low.

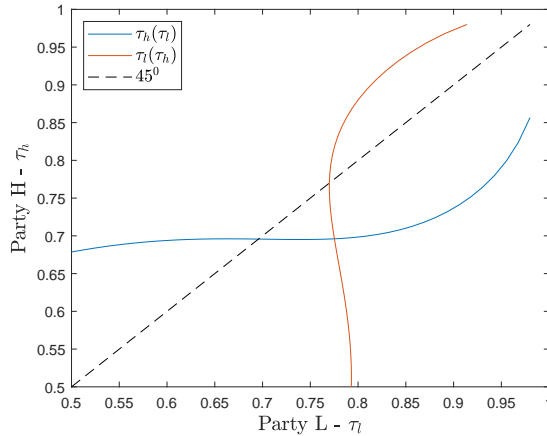
Citizen-office-seeking candidates. So far, candidates were simply interested in being elected. We now allow candidates to obtain utility from office but also to be interested in the policy outcome. Specifically, they suffer a loss that depends on the distance between their proposed tax and their bliss policy.³³ This has two effects on the equilibrium outcome. First, the proposed taxes by the two candidates are no longer equal. Second, there is activism in equilibrium.

Formally, the candidate from party H seeks to maximize the probability of getting elected, but would also like his bliss policy τ_h^* to be implemented in equilibrium. Given τ_l , this candidate solves the following program:

$$\max_{\tau_h} \mu p_h + (1 - \mu) E_\tau \left[\frac{(\tau - \tau_h^*)^2}{2} \right], \quad (40)$$

where p_h is given by (32). $0 < \mu < 1$ is a preference weight for being elected relative to the loss incurred as the actual policy rate deviates from the candidate's bliss policy. The candidate from party L solves a similar program with bliss policy τ_l^* .

Figure 4: Citizen-Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, with $\mu = 1/2$, $\tau_l^* = 0$, $\tau_h^* = 1$. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Figure 4 illustrates candidates' best responses and the equilibrium outcome: $\tau_l^p < \tau_h^p$ since candidates offer policy platforms that are now influenced by their individual preferences $\tau_l^* < \tau_h^*$. As explained in Section 5.3 and

³³Discussions with Annika Bacher led to the development of this case.

explicit in (31), different policy platforms are associated in equilibrium with conflictual activism.

6 Conclusion

Empirically, labor taxes are progressive but do not expropriate all the earning of the rich to compensate the poor. Instead, redistribution is limited. This paper studies the economic and political factors that limit redistribution in a democracy.

In the environment, the arguments put forth by Bénabou and Ok (2001) do not hold: absent commitment and despite the potential for upward mobility in the income distribution, the outcome under majority voting would be complete redistribution. In our model, this implies no capital accumulation and thus a massive reduction in the “economic pie”.

The analysis provides another mechanism: the power of persuasion. Coalitions of agents jointly decide on the level of activism which can influence the political preferences of agents and thus voting outcomes. Majority voting remains but the progress of political persuasion facilitates a redistribution of political power. Though there is no commitment, the outcome with activism is closer to the efficient allocation: redistribution is incomplete and incentives are retained for the accumulation of human capital.

References

- ACEMOGLU, D., M. GOLOSOV, AND A. TSYVINSKI (2010): “Dynamic Mirrlees Taxation under Political Economy Constraints,” The Review of Economic Studies, 77(3), 841–881.
- ALESINA, A., AND G.-M. ANGELETOS (2005): “Fairness and Redistribution,” American Economic Review, 95(4), 960–980.
- ALESINA, A., AND E. L. FERRARA (2005): “Preferences for Redistribution in the Land of Opportunities,” Journal of Public Economics, 89, 897–931.
- ALESINA, A., AND N. FUCHS-SCHÜNDELN (2007): “Goodbye Lenin (or Not?): The Effect of Communism on People’s Preferences,” American Economic Review, 97(4), 1507–1528.
- ALESINA, A., AND P. GIULIANO (2011): “Chapter 4 - Preferences for Redistribution,” vol. 1 of Handbook of Social Economics, pp. 93 – 131. North-Holland.
- ALESINA, A., S. STANTCHEVA, AND E. TESO (2018): “Intergenerational Mobility and Preferences for Redistribution,” American Economic Review, 108(2), 521–54.
- BECK, P. A., R. J. DALTON, S. GREENE, AND R. HUCKFELDT (2002): “The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices,” The American Political Science Review, 96(1), 57–73.
- BECKER, G. S. (1983): “A Theory of Competition Among Pressure Groups for Political Influence,” The Quarterly Journal of Economics, 98(3), 371–400.

- BIERBRAUER, F., A. TSYVINSKI, AND N. D. WERQUIN (2017): “Taxes and Turnout,” Working Paper 24123, National Bureau of Economic Research.
- BIERBRAUER, F. J., AND P. C. BOYER (2013): “Political competition and Mirrleesian income taxation: A first pass,” Journal of Public Economics, 103, 1 – 14.
- (2015): “Efficiency, Welfare, and Political Competition,” The Quarterly Journal of Economics, 131(1), 461–518.
- BÉNABOU, R. (2000): “Unequal Societies: Income Distribution and the Social Contract,” American Economic Review, 90(1), 96–129.
- (2002): “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?,” Econometrica, 70(2), 481–517.
- (2008): “Ideology,” Journal of the European Economic Association, 6(2-3), 321–352.
- BÉNABOU, R., AND E. A. OK (2001): “Social Mobility and the Demand for Redistribution: The Poul Hypothesis,” The Quarterly Journal of Economics, 116(2), 447–487.
- BÉNABOU, R., AND J. TIROLE (2006): “Belief in a Just World and Redistributive Politics,” The Quarterly Journal of Economics, 121(2), 699–746.
- COATE, S., AND M. CONLIN (2004): “A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence,” American Economic Review, 94(5), 1476–1504.
- FARHI, E., C. SLEET, I. WERNING, AND S. YELTEKIN (2012): “Nonlinear Capital Taxation without Commitment,” Review of Economic Studies, 79(4), 1469–1493.
- FEDDERSEN, T., AND A. SANDRONI (2006): “A Theory of Participation in Elections,” American Economic Review, 96(4), 1271–1282.
- HEATHCOTE, J., K. STORESLETTEN, AND G. L. VIOLANTE (2017): “Optimal Tax Progressivity: An Analytical Framework,” The Quarterly Journal of Economics, 132(4), 1693–1754.
- KARABARBOUNIS, L. (2011): “One Dollar, One Vote,” The Economic Journal, 121(553), 621–651.
- LEVINE, D. K., AND A. MATTOZZI (2018): “Voter Turnout with Peer Punishment,” Levine’s bibliography, UCLA Department of Economics.
- LINDBECK, A., AND J. W. WEIBULL (1987): “Balanced-budget redistribution as the outcome of political competition,” Public Choice, 52(3), 273–297.
- MAISEL, L. S., J. M. BERRY, G. C. EDWARDS, AND K. L. SCHLOZMAN (2010): Who Sings in the Heavenly Chorus? The Shape of the Organized Interest System. Oxford University Press.
- MATTOZZI, A., AND E. SNOWBERG (2018): “The right type of legislator: A theory of taxation and representation,” Journal of Public Economics, 159(C), 54–65.

- MELTZER, A. H., AND S. F. RICHARD (1981): “A Rational Theory of the Size of Government,” Journal of Political Economy, 89(5), 914–927.
- OLSON, M. (1965): The Logic of Collective Action: Public Goods and the Theory of Groups., vol. 1. Harvard University Press, 1 edn.
- PENNEC, C. L., AND V. PONS (2019): “Vote Choice Formation and Minimal Effects of TV Debates: Evidence from 61 Elections in 9 OECD Countries,” Working Paper 26572, National Bureau of Economic Research.
- PERSSON, T., AND G. TABELLINI (2002): Political Economics: Explaining Economic Policy, vol. 1. The MIT Press, 1 edn.
- PIKETTY, T. (1995): “Social Mobility and Redistributive Politics,” The Quarterly Journal of Economics, 110(3), 551–584.
- SCHEUER, F., AND A. WOLITZKY (2016): “Capital taxation under political constraints,” American Economic Review, 106(8), 2304–28.

Mathematical Appendix

A Section 3 - Environment

Individual choice. Consider household optimization (1), substitute the constraints into the objective function:

$$\max_e \log(\theta - e) + \beta \int_{\tau} \left\{ (1 - \tau) \left[\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2} \right] + \tau \log(\bar{\theta}') \right\} dG(\tau), \quad (41)$$

where $G(\cdot)$ captures uncertainty over rate τ . The first order condition reads:

$$-\frac{1}{\theta - e} + \frac{\beta \delta (1 - \bar{\tau})}{e} = 0, \quad (42)$$

where $\bar{\tau} = E(\tau)$. Reorganize and get (5).

Evolution of the income distribution. Start from (3), take the log and use (5):

$$\log(\theta') = \log(z) + (\alpha + \delta) \log(\theta) + \delta \log(\epsilon(\bar{\tau})) \quad (43)$$

The mean and variance of this expression yield (6) and (7).

Break-even income level. Given τ , income net of taxes and transfers satisfies (4), take the integral and then the log:

$$\log \int_{\theta} c' dF(\theta) = \tau \log(\bar{\theta}') + \log \int_{\theta} \theta'^{1-\tau} dF(\theta). \quad (44)$$

If $X \sim \log \mathcal{N}(\mu, \sigma^2)$, then $E(X^n) = e^{n\mu + n^2\sigma^2/2}$. Since $E(c') = E(\theta')$, it gives:

$$m' + \frac{\sigma'^2}{2} = \tau \log(\bar{\theta}') + (1 - \tau)m' + \frac{(1 - \tau)^2}{2} \sigma'^2. \quad (45)$$

Reorganize and get (8).

Value functions and bliss policies. Lifetime utility to an agent with initial income θ at $t = 1$ reads:

$$\begin{aligned} V_1(\theta, \tau) &= v_1(\theta, e, \tau, \bar{\theta}') \\ &= \log(\theta - e) + \beta \left[(1 - \tau) \underbrace{\left(\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2} \right)}_{=E_z(\log(\theta')|\theta)} + \tau \log(\bar{\theta}') \right], \end{aligned} \quad (46)$$

where $e = \epsilon(\tau)\theta$ as in (5) and $\bar{\theta}'$ is given by (8). The first order condition w.r.t. τ :

$$\frac{dV_1(\cdot)}{d\tau} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e} \frac{de}{d\tau}}_{=0} + \frac{\partial v_1(\cdot)}{\partial \tau} + \frac{\partial v_1(\cdot)}{\partial \log \bar{\theta}'} \frac{d \log \bar{\theta}'}{d\tau} = 0. \quad (47)$$

The first term is 0 from the envelope condition, the other terms are:

$$\frac{\partial v_1(\cdot)}{\partial \tau} = \log(\bar{\theta}') - (\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}), \quad (48)$$

$$\frac{d \log(\bar{\theta}')}{d\tau} = \frac{dm'}{d\tau} - \frac{\sigma'^2}{2} = \delta \frac{\epsilon'(\cdot)}{\epsilon(\cdot)} - \frac{\sigma'^2}{2}. \quad (49)$$

Reorganize and get (10):

$$\beta(\alpha + \delta)(m - \log(\theta)) + \beta \underbrace{((\alpha + \delta)^2 \sigma^2 + w^2)}_{=\sigma'^2} (1 - \tau) + \beta \delta \frac{\tau}{\epsilon(\tau)} \frac{d\epsilon(\tau)}{d\tau} = 0. \quad (50)$$

This expression implicitly defines $\tau^*(\theta)$. The first two terms form a decreasing linear function of τ , whose intercept is decreasing in θ . The third term, the elasticity of the saving rate to the redistribution rate τ , is decreasing in τ . Formally:

$$\frac{d}{d\tau} \left(\frac{\tau}{\epsilon(\tau)} \frac{d\epsilon(\tau)}{d\tau} \right) = - \left[\frac{\tau}{1 - \tau} \frac{\beta \delta}{[1 + \beta \delta (1 - \tau)]^2} + \frac{1}{(1 - \tau)^2} \frac{1}{1 + \beta \delta (1 - \tau)} \right] < 0, \quad (51)$$

which goes to $-\infty$ when τ goes to 1. Altogether there is a unique solution $\tau^*(\theta) < 1$ to (50), decreasing in θ .

The value function to an agent with initial income θ after $t = 1$ consumption and education choice reads:

$$\begin{aligned} V_2(\theta, \tau | \epsilon) &= v_2(\theta, e, \tau, \bar{\theta}' | \epsilon) \\ &= \beta \left[(1 - \tau) \left(\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2} \right) + \tau \log(\bar{\theta}') \right] \end{aligned} \quad (52)$$

The difference with (46) is that the education choice is no longer sensitive to the redistribution rate τ and $e = \epsilon\theta$. The sensitivity of individual value functions to τ satisfies:

$$\begin{aligned} \frac{dV_2(\cdot)}{d\tau} &= \frac{\partial v_2(\cdot)}{\partial \tau} + \frac{\partial v_2(\cdot)}{\partial \log \bar{\theta}'} \frac{d \log \bar{\theta}'}{d\tau} \\ &= \beta(\alpha + \delta)(m - \log(\theta)) + \beta(\alpha + \delta)^2 \underbrace{((\alpha + \delta)^2 \sigma^2 + w^2)}_{=\sigma'^2} (1 - \tau) \end{aligned} \quad (53)$$

Bliss policy $\tau^d(\theta) \in [0, 1]$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau} = 0$ or $\tau^d(\theta) = 1$. Given the linear nature of (53), there is a unique bliss policy, ordered by initial income θ .

B Section 4 - Outcome under Commitment

Sensitivity of optimal redistribution rate. To derive comparative statics for τ^* , first get

$$\frac{\tau}{\epsilon(\tau)} \frac{d\epsilon(\tau)}{d\tau} = -\frac{\tau}{1-\tau} \frac{1}{1+\beta\delta(1-\tau)}. \quad (54)$$

Then rewrite (14) as

$$((\alpha + \delta)^2 \sigma^2 + w^2)(1 - \tau)^2 - \frac{\delta \tau}{1 + \beta \delta (1 - \tau)} = 0 \quad (55)$$

The total derivative of this expression:

$$\begin{aligned} (\alpha + \delta)^2 (1 - \tau)^2 d\sigma^2 + (1 - \tau)^2 dw^2 + \left[2(\alpha + \delta)(1 - \tau)^2 - \tau \frac{1 + 2\beta\delta(1 - \tau)}{(1 + \beta\delta(1 - \tau))^2} \right] d\delta \\ = \left[2\sigma'^2 (1 - \tau) + \delta \frac{1 + \beta\delta}{(1 + \beta\delta(1 - \tau))^2} \right] d\tau \end{aligned} \quad (56)$$

Get immediately $\frac{d\tau^*}{d\sigma^2} > 0$ and $\frac{d\tau^*}{dw^2} > 0$ while the sign of $\frac{d\tau^*}{d\delta}$ is ambiguous.

Proof Proposition 1. Individual value functions are single peaked and bliss policies ordered by income level. Pairwise evaluation of policy alternative leads to a convergence of electoral platforms toward the bliss policy of the median agent, with (log) income $m = \log(\theta^m)$. If the vote takes place before the education choice, then from (10), $\tau^*(\theta^m)$ coincides with τ^* defined by (14). If the vote takes place after education, then from (12) the outcome of the vote is full redistribution $\tau^d(\theta^m) = 1$.

C Section 5 - Probabilistic Voting and Activism

C.1 No Activism.

Comparison of platforms (18). Start from (11):

$$V_2(\theta, \tau|\epsilon) = \beta \left((1 - \tau)(\alpha \log(\theta) + \delta \log(\epsilon\theta) - \frac{w^2}{2}) + \tau \log(\bar{\theta}') \right), \quad (57)$$

with $\log(\bar{\theta}')$ given by (8). Rearrange and get:

$$V_2(\theta, \tau|\epsilon) = \beta \left[(1 - \tau)(\alpha + \delta) \log(\theta) + \delta \log(\epsilon) - \frac{w^2}{2} + \tau(\alpha + \delta)m + \frac{\sigma'^2}{2}(2 - \tau)\tau \right]. \quad (58)$$

The difference of this expression with τ_h and τ_l :

$$\Delta V_2(\theta) = \beta(\tau_h - \tau_l)(\alpha + \delta)(m - \log(\theta)) + \beta \frac{\sigma'^2}{2} [(2 - \tau_h)\tau_h - (2 - \tau_l)\tau_l]. \quad (59)$$

Verify $(2 - \tau_h)\tau_h - (2 - \tau_l)\tau_l = (\tau_h - \tau_l)(2 - \tau_h - \tau_l)$ and get (18). Finally, note

$$\int_{\theta} \Delta V_2(\theta) dF(\theta) = \beta(\tau_h - \tau_l) \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l). \quad (60)$$

Probability p_h of high rate of redistribution τ_h . Consider the following distributions: $\chi_{\theta}^j \sim U\left(-\frac{1}{2\phi_{\theta}} + m, \frac{1}{2\phi_{\theta}} + m\right)$ and $\psi \sim U\left(-\frac{1}{2\Psi} + M, \frac{1}{2\Psi} + M\right)$. Note $\phi = E(\phi_{\theta})$. Given ψ and competing platforms (τ_l, τ_h) , agent j with income θ votes for party H if and only if $\chi_{\theta}^j \leq \chi(\theta, \psi) = \Delta V_2(\theta) - \psi$. Hence, the share of agents with income θ that vote for party H is:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi_{\theta}} + m}^{\chi(\theta, \psi)} \phi_{\theta} dj = \phi_{\theta} \left(\chi(\theta, \psi) + \frac{1}{2\phi_{\theta}} - m \right). \quad (61)$$

The share of votes across groups is then $\pi_h(\psi) = \int_{\theta} \pi_{\theta,h}(\psi) dF(\theta)$:

$$\pi_h(\psi) = \int_{\theta} \phi_{\theta} \left(\chi_{\theta} + \frac{1}{2\phi_{\theta}} - m \right) dF(\theta), \quad (62)$$

$$= \int_{\theta} \phi_{\theta} \Delta V_2(\theta) dF(\theta) - \phi(\psi + m) + \frac{1}{2}. \quad (63)$$

The probability that party H wins the election is $p_h = P(\pi_h(\psi) \geq \lambda)$, where $\lambda \in [0, 1]$ is a majority requirement for party H to win the election. The event $\pi_h(\psi) \geq \lambda$ is equivalent to the event

$$\psi \leq \bar{\psi} = \frac{1}{\phi} \int_{\theta} \phi_{\theta} \Delta V_2(\theta) dF(\theta) - m + \frac{1}{\phi} \left(\frac{1}{2} - \lambda \right). \quad (64)$$

Get then p_h as

$$p_h = P(\psi \leq \bar{\psi}) = \frac{1}{2} + \frac{\Psi}{\phi} \int_{\theta} \phi_{\theta} \Delta V_2(\theta) dF(\theta) - \Psi(m + M) + \frac{\Psi}{\phi} \left(\frac{1}{2} - \lambda \right). \quad (65)$$

Set $\phi_{\theta} = \phi$, $\lambda = \frac{1}{2}$, $m = M = 0$ and using (60) get (21).

Asymmetric political preferences. Persson and Tabellini (2002) study probabilistic voting allowing for a correlation between income and political preferences. To see how this mechanism works without activism, normalize average political preferences heterogeneity $\phi = E(\phi_{\theta}) = 1$ and assume that $\text{cov}(\phi_{\theta}, \log(\theta)) = a$: if $a > 0$ then high income agents are more responsive to economic factors. The probability $p_h(\tau_l, \tau_h)$ is then

$$p_h(\tau_h, \tau_l) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l) \left(-(\alpha + \delta)a + \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l) \right), \quad (66)$$

and a Nash equilibrium of economic platforms competition differs from full redistribution if and only if $a > 0$:

$$\tau_h = \tau_l = 1 - \frac{(\alpha + \delta)a}{\sigma'^2} < 1. \quad (67)$$

Overall, when high income agents are more sensitive to economic policy rather than political factors, they are attractive targets to candidates, which tilts equilibrium economic platform toward lower rate of redistribution.

C.2 Activism

Activism choice. To get (29), rewrite (27) as:

$$V_1(\theta, \tau) = v_1(\theta, \tau, e, p_h, \log(\bar{\theta}'_h), \log(\bar{\theta}'_l)) \quad (68)$$

where $e = \epsilon(\bar{\tau})\theta$. Further, $\frac{dV_1(\cdot)}{dA_\theta^i} = \frac{dV_1(\cdot)}{dp_h} \frac{dp_h}{dA_\theta^i}$. Consider the first term:

$$\frac{dV_1(\cdot)}{dp_h} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e}}_{=0} \frac{de}{dp_h} + \frac{\partial v_1(\cdot)}{\partial p_h} + \sum_{i=l,h} \frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)} \frac{d \log(\bar{\theta}'_i)}{dp_h}. \quad (69)$$

Term by term:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = -\beta(\tau_h - \tau_l) E_z(\log(\theta')|\theta) + \beta(\tau_h \log(\bar{\theta}'_h) - \tau_l \log(\bar{\theta}'_l)), \quad (70)$$

and

$$\tau_h \log(\bar{\theta}'_h) - \tau_l \log(\bar{\theta}'_l) = (\tau_h - \tau_l) \left[m' + \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l) \right]. \quad (71)$$

Using (6) and $E_z(\log(\theta')|\theta) = (\alpha + \delta) \log(\theta) + \delta \log(\epsilon(\bar{\tau})) - \frac{w^2}{2}$, rearrange and get:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = \beta(\tau_h - \tau_l) \left[(\alpha + \delta)(m - \log(\theta)) + \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l) \right]. \quad (72)$$

Then, since $\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)} = \beta p_i \tau_i$ and $\frac{d \log(\bar{\theta}'_i)}{dp_h} = \delta \frac{d\bar{\tau}}{dp_h} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ for $i \in (l, h)$,

$$\sum_{i=l,h} \frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)} \frac{d \log(\bar{\theta}'_i)}{dp_h} = \beta(\tau_h - \tau_l) \delta \underbrace{(p_h \tau_h + (1 - p_h) \tau_l)}_{=\bar{\tau}} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}. \quad (73)$$

Overall, with $\frac{dp_h}{dA_\theta^i} = \pm \Psi \gamma$, with $\pm = +$ for $x = h$ and $\pm = -$ for $x = l$, one gets (29):

$$\frac{dV_1(\cdot)}{dA_\theta^i} = \pm \Psi \gamma \beta(\tau_h - \tau_l) \left[(\alpha + \delta)(m - \log(\theta)) + \frac{\sigma'^2}{2} (2 - \tau_h - \tau_l) + \delta \bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} \right]. \quad (74)$$

Lemma 1. Given policy platforms (τ_l, τ_h) , an equilibrium of the activism subgame is a set of income group contributions $\{A_\theta^l, A_\theta^h\}$, aggregate contributions (A^l, A^h) defined as $A^i = \int_\theta A_\theta^i d\theta$, probability of high taxes p_h and expected tax rate $\bar{\tau} = E(\tau)$.

To establish existence, start from (28) characterizing optimal contribution of income groups. Sum this expression over income group and get the aggregate effects of activism on political preferences (31).

Using this expression with (25), get (32) which uniquely defines the probability p_h , since the right-hand side of this expression is decreasing in $\bar{\tau}$, hence in p_h . From this probability, one can recover aggregate and individual activism contributions. The probability p_h is unique, because the sensitivity of the education rate to the expected redistribution rate $\bar{\tau} \frac{\epsilon'(\cdot)}{\epsilon(\cdot)} = -\frac{\bar{\tau}}{1-\bar{\tau}} \frac{1}{1+\beta\delta(1-\bar{\tau})} < 0$ depends negatively on $p_h(\tau_l, \tau_h)$, since $\bar{\tau} = p_h\tau_h + (1-p_h)\tau_l$.

The argument for uniqueness goes as follow. Given (τ_h, τ_l) , the outcome of the contribution game yield a unique p_h (25). Is there another pair (A^h, A^l) and underlying group contributions that yield the same outcome? No, because the marginal return to contribution is a function of $A^h - A^l$, but the marginal cost depends on individual contribution only, see (28).

Endogenous choice of platforms. To derive (34), denote by $H(\tau_h, \tau_l, \bar{\tau})$ the right hand side of (32) and let $\mathcal{E}(\bar{\tau}) = \bar{\tau} \frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ be the elasticity of education rate to the expected tax rate. Totally differentiating (32) w.r.t. p_h and τ_h :

$$\left[1 - \frac{\partial H(\cdot)}{\partial \bar{\tau}} \frac{d\bar{\tau}}{dp_h}\right] dp_h = \left[\frac{\partial H(\cdot)}{\partial \tau_h} + \frac{\partial H(\cdot)}{\partial \bar{\tau}} \frac{d\bar{\tau}}{d\tau_h}\right] d\tau_h. \quad (75)$$

Term by term:

$$\frac{\partial H(\cdot)}{\partial \tau_h} = \Psi\beta[(1 + \Psi\gamma^2)\sigma'^2(1 - \tau_h) + \Psi\gamma^2\delta\mathcal{E}(\bar{\tau})], \quad (76)$$

$$\frac{\partial H(\cdot)}{\partial \bar{\tau}} = \Psi\beta(\tau_h - \tau_l)\Psi\gamma^2\delta \frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}, \quad (77)$$

$$\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} = -\frac{1 + \beta\delta - \beta\delta\tau^2}{(1 - \tau)^2[1 + \beta\delta(1 - \tau)]^2}. \quad (78)$$

$$(79)$$

Rearranging terms, get (34).

Proposition 3. If $\gamma = 0$ or $\delta = 0$, then the unique solution to (35) is $\tau = 1$. Otherwise, using (54), rewrite (35) as:

$$\left(1 + \frac{1}{\Psi\gamma^2}\right)\sigma'^2(1 - \tau) = \frac{\tau}{1 - \tau} \frac{\delta}{1 + \beta\delta(1 - \tau)}. \quad (80)$$

The left hand side is decreasing in τ , while the right hand side is increasing in τ , with limit when $\tau = 1$ is ∞ , which gives that the unique solution satisfies $0 < \tau^p < 1$. An increase in γ decreases the left hand side, which yield $\frac{d\tau^p}{d\gamma} < 0$. In the limit $\gamma = \infty$, (35) coincides with (14), hence $\tau^p = \tau^*$.

C.3 Citizen-office-seeking candidate

The program of the candidate from party H

$$\max_{\tau_h} \mu p_h + (1 - \mu) \left[p_h \frac{(\tau_h - \tau_h^*)^2}{2} + (1 - p_h) \frac{(\tau_l - \tau_h^*)^2}{2} \right] \quad (81)$$

$$\mu \frac{dp_h}{d\tau_h} + (1 - \mu) \left[\frac{dp_h}{d\tau_h} \frac{(\tau_h - \tau_h^*)^2}{2} + p_h(\tau_h - \tau_h^*) - \frac{dp_h}{d\tau_h} \frac{(\tau_l - \tau_h^*)^2}{2} \right] = 0 \quad (82)$$

Similarly, the first order condition characterizing $\tau_l(\tau_h)$:

$$-\mu \frac{dp_h}{d\tau_l} + (1 - \mu) \left[\frac{dp_h}{d\tau_l} \frac{(\tau_h - \tau_l^*)^2}{2} + (1 - p_h)(\tau_l - \tau_l^*) - \frac{dp_h}{d\tau_l} \frac{(\tau_l - \tau_l^*)^2}{2} \right] = 0 \quad (83)$$

To derive these expressions, denote $p_h = G(\tau_l, \tau_h, \bar{\tau})$ and get:

$$\left[1 - \frac{\partial G(\cdot)}{\partial \bar{\tau}} \frac{d\bar{\tau}}{dp_h} \right] dp_h = \left[\frac{\partial G(\cdot)}{\partial \tau_i} + \frac{\partial G(\cdot)}{\partial \bar{\tau}} \frac{d\bar{\tau}}{d\tau_i} \right] d\tau_i \quad (84)$$

Then derive

$$\frac{\partial G(\cdot)}{\partial \bar{\tau}} = \Psi \beta (\tau_h - \tau_l) \Psi \gamma^2 \delta \frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} \quad \frac{d\bar{\tau}}{dp_h} = \tau_h - \tau_l \quad (85)$$

$$\frac{d\bar{\tau}}{d\tau_h} = p_h \quad \frac{d\bar{\tau}}{d\tau_l} = (1 - p_h) \quad (86)$$

and

$$\frac{\partial G(\cdot)}{\partial \tau_h} = \Psi \beta [(1 + \Psi \gamma^2) \sigma'^2 (1 - \tau_h) + \Psi \gamma^2 \delta \mathcal{E}(\bar{\tau})] \quad (87)$$

$$\frac{\partial G(\cdot)}{\partial \tau_l} = -\Psi \beta [(1 + \Psi \gamma^2) \sigma'^2 (1 - \tau_l) + \Psi \gamma^2 \delta \mathcal{E}(\bar{\tau})] \quad (88)$$

Finally

$$\mathcal{E}(\tau) = -\frac{\tau}{1 - \tau} \frac{1}{1 + \beta \delta (1 - \tau)} \quad \frac{d\mathcal{E}(\tau)}{d\tau} = -\frac{1 + \beta \delta (1 - \tau^2)}{(1 - \tau)^2 [1 + \beta \delta (1 - \tau)]^2} \quad (89)$$

D General game

We consider a general formulation of preferences and technology and discuss minimal assumptions that support our main result: activism reduces the equilibrium tax rate. We maintain the demographic assumptions and the timing introduced in Section 5.1, and consider a version of the model with two types $\theta_l < \theta_h$.³⁴ We show that under two primitive properties, full redistribution cannot be a Nash equilibrium of the political game with activism. These properties are satisfied for the specific model presented in the exposition:

- i. POUM does not work. Absent activism and commitment, a pure redistributive conflict drives the tax choice, and because a majority of agent benefits from the expropriation of the rich, full redistribution $(1, 1)$ is the Nash equilibrium.
- ii. But all agents under commitment would refrain from such outcome because of the valuable dynamic incentives: $\tau^*(\theta) < 1$ for all θ .

³⁴The argument can easily be extended to a continuum of types since we consider the incentives of individuals / groups at the bounds of the distribution.

The stochastic outcome of electoral competition is captured by the probability of high taxes, $p_h = P(\pi_h(\cdot) > 1/2) = P_h(\tau_l, \tau_h, A)$, where $A = \mathcal{A}(A_l, A_h)$ aggregates the level of activism of each income group for the low tax candidate; this function is increasing in each argument and $\mathcal{A}(0, 0) = 0$. Given an aggregate level of activism A , the education choice of an agent with type θ is:

$$V_1(\theta, \tau_l, \tau_h, A) = \max_e E_\tau U(c(e, \tau, \theta, A), e). \quad (90)$$

This program implies $e = \phi(\theta, \tau, A)$. Here consumption $c(e, \tau, \theta, A) = \eta(\theta, \bar{\theta}', \tau)$ depends on productivity $\theta' = g(\theta, e)$, the tax base $\bar{\theta}'$ and the magnitude of redistribution τ . Note that taking the expectation of $\theta' = g(\theta, e(\cdot))$ over θ yields $\bar{\theta}' = G(\tau, A)$. Clearly a positive level of activism yields a higher tax base as it puts more weight on low taxes and everyone chooses more education: $G_A(\tau, A) > 0$. This increases consumption for all agents, i.e. $\frac{dc}{d\bar{\theta}'} > 0$.

The expectations operator in (90) incorporates $p_h = P_h(\tau_l, \tau_h, A)$, so there are two distinct influences of activism A . The first is directly on the probability of high taxes and thus on the amount of redistribution *ex post*. The second is on the education choices of all agents. This includes the fiscal externality through $G(\tau, A)$. The choice of activism for type θ A_θ , given $A_{-\theta}$, comes from:

$$A(\theta, \tau_l, \tau_h) = \operatorname{argmax}_a f(\theta) \left(V_1(\theta, \tau_l, \tau_h, A) \right) - k(a), \quad (91)$$

where $A = \mathcal{A}(A_\theta, A_{-\theta})$ is the vector of activism choice and $k(\cdot)$ is an activism cost function, symmetric: $k(a) = k(-a)$. The FOC for the choice of activism is then:

$$\frac{dV_1(\theta, \tau_l, \tau_h, A)}{da} = \frac{dp_h}{da} [U(c(e, \tau_h, \theta, A), e) - U(c(e, \tau_l, \theta, A), e)] + E_\tau U_1(c(e, \tau, \theta, A), e) \frac{dc}{da} = k'(a)/f(\theta) \quad (92)$$

The first term is redistribution. The second term captures the effect of group θ activism a on consumption through the aggregate tax base $\bar{\theta}'$, with

$$\frac{dc}{da} = \frac{dc}{d\bar{\theta}'} \frac{d\bar{\theta}'}{dA} \frac{dA(\cdot)}{da} = \frac{dc}{d\bar{\theta}'} G_A(\cdot) \frac{dA(\cdot)}{da} > 0 \quad (93)$$

The effect on the education choice drops from the envelope condition.

We now proceed to show that $(\tau_l, \tau_h) = (1, 1)$ cannot be an equilibrium of this game. We consider platforms $(\tau_l, \tau_h) = (1 - \varepsilon, 1)$ and show that the anticipation of activism would indeed induce a candidate to deviate from an outcome with full redistribution.³⁵ We review the incentives of each group in turn.

$A(\theta_h, 1 - \varepsilon, 1) > 0$ requires that the LHS of (92) be positive for $\theta = \theta_h$. The first term in (92) is positive as p_h is decreasing in A and the high ability types lose unambiguously from redistribution, i.e., $c(e, \tau, \theta, A)$ is decreasing in τ for θ_h type. The second term, using (93), is positive. Again, activism puts more weight on the low tax outcome and thus provides incentives for all to invest more in education. For any $\tau > 0$, all agents benefit from a higher tax base. This reflects the property noted earlier that the bliss tax rate is less than 1 for all θ , $\tau^*(\theta) < 1$.

If we only consider activism by one group, the argument effectively ends here: activism increases the probability

³⁵Since candidates are office seeking and play Nash over splitting the probability of win, the equilibrium in pure strategy is that they offer the same platform.

of $\tau_l = 1 - \varepsilon$ win over $\tau_h = 1$ which triggers an adjustment of candidates platform. Could activism by lower income groupe θ_l offset this effect? Their activism choice follows (92) as well.

If the θ_l group can only use activism to support higher taxes, then they will choose $A_l = 0$. To see why, note that their preferred tax rate at the time of activism is less than 1: $\tau^*(\theta_l) < 1$. At a tax rate of 1, the θ_l group would benefit from full redistribution, but without incentives, there is nothing to redistribute: they would not oppose an ε reduction in τ_l from 1. Thus $A(\theta_l, 1, 1 - \varepsilon) \geq 0$. So the overall response of activism to a deviation from full redistribution is positive $A = \mathcal{A}(A_l, A_h) \geq 0$.

Now consider the optimal choice of tax platform by the candidate from party L. The objective is to maximize $1 - P_h(\tau_l, \tau_h, A)$, given τ_h . Its best response function $\tau_l(\tau_h) \in [0, 1]$ is implied by:

$$\frac{\partial(1 - P_h(\cdot))}{\partial \tau_l} + \frac{d(1 - P_h(\cdot))}{dA} \left[\frac{d\mathcal{A}(\cdot)}{dA_l} \frac{dA_l}{d\tau_l} + \frac{d\mathcal{A}(\cdot)}{dA_h} \frac{dA_h}{d\tau_l} \right] = 0 \quad (94)$$

Note that if $\tau_h = \tau_l$ then $A = 0$, since no group has incentive to exert activism when there is no choice to influence. Also, since the Nash equilibrium is $(1, 1)$ in the absence of activism, we can see it cannot be a Nash equilibrium under activism if the sum of the second and third terms in (94) is negative at $(1 - \varepsilon, 1)$. As argued, θ_h unambiguously exert activism for $\tau_l = 1 - \varepsilon$, i.e. $\frac{dA}{d\tau_l} < 0$ for $(1 - \varepsilon, 1)$, while the θ_l group will not counter this influence as they benefit from the marginal improvement in incentives over redistribution.