Political Dynamics, Public Goods and Private Spillovers*

Timothy Kam[†] Tina Kao[‡] Yingying Lu[§]

May 8, 2020

Abstract

Existing empirical work have established the quantitative importance of spillovers from private activities onto aggregate economic outcomes. These negative (or positive) spillovers detract from (or add to) economic outcomes through a public good effect. However, there is no empirical work connecting political mechanisms and the effects from private spillovers. We provide a theoretical framework that also suggests new avenues for future empirical work. If the negative spillover is sufficiently strong, the theoretical economy is trapped in a unique equilibrium with perpetually high tax rates and majority poor voters. Otherwise, there is multiple equilibria that can be used to interpret existing conundrums in empirical findings. In one of these equilibria, the economy with positive spillovers can sustain an equilibrium with perpetually low tax rates and majority rich voters. We use this to interpret an observed cross-country empirical regularity in terms of public funding for knowledge goods and income inequality.

Keywords: Environmental/Knowledge Public Good; Private Spillovers; Markovian Voting; Equilibrium Regimes; Redistribution

JEL Classification: D72; D78; E62; H21; H23

*We would like to thank Richard Cornes for advice and dedicate this work to his memory. We also thank Jaime Alonso-Carrera, Jürgen Meinecke, Solmaz Moslehi, Pietro Peretto, Rohan Pitchford, Maria Racionero, Yuji Saikai, Chung Tran, and Chris Waller for useful discussion, and also seminar participants at University of Tokyo, Kyoto University and GRIPS. We especially thank Elena del Rey, Simon Grant, and Ronald Stauber for detailed feedback and suggestions. The first author also thanks the University of Queensland and Monash University for their hospitality where most bits of this paper were developed. All errors and exceptions remain our own responsibility. This paper is a revised version of what was previously circulated as "Democracy, Public Goods and Private Activity Spillovers" and "Democracy and Impure Public Goods".

[†]Research School of Economics, Australian National University, ACT 2601, Australia. Email: `tcy.kam@gmail.com`

[‡]Research School of Economics, Australian National University, ACT 2601, Australia. Email: `tina.kao@anu.edu.au`

[§]Division of Land and Water, The Commonwealth Scientific and Industrial Research Organisation, ACT 2601, Australia. Email: `Yingying.Lu@csiro.au`

1 Introduction

The level of an economy's output or consumption for some goods are often affected by externalities or spillovers from private actions. These "private spillovers" are externalities in the sense that they arise in addition to, and are not internalized in, individual expenditure or investment decisions. Examples of positive spillovers include knowledge goods that benefit private productivity (see Romer, 1990) or consumption (see Stiglitz, 1974). In terms of negative spillovers, one would immediately think of carbon emissions from economic activity that damage aggregate economic outcomes through global temperature feedback effects (Hassler et al., 2016).

Motivation—some important empirical facts. How important are these spillover effects? Consider the case of negative spillover, by way of the carbon emissions problem. Empirical modelling consensus appears to suggest that for cumulative emissions below 2000 GtC, each one thousand Gigatons of carbon (GtC) emitted will lead to a 0.8° to 2.5° Celcius rise in global temperature (see IPCC, 2013; Matthews et al., 2012, 2009). In turn, global temperature rises can create heterogeneous economic damage at the micro level (see, e.g., Dell et al., 2014). At a more macro-global level, it has been estimated that a 3° Celcius increase in temperature would result in between 1.3% to 2.5% damage in terms of reduced global GDP (Nordhaus and Sztorc, 2013; Ciscar et al., 2011). These are non-trivial figures from well-accepted research on climate change and environmental economics (Hassler et al., 2016).

In the case of positive spillovers, the impure public goods literature shows that when a private activity jointly produces a public good and a private good, the private incentives, equilibrium configuration, and welfare considerations are very different from standard models without the spillovers (see Cornes and Sandler, 1994, 1984). For example, the difference between social and private returns of higher education has been studied in the education policy literature.[1] There is also convincing empirical evidence showing that private (knowledge) spillovers are large and important. Eberhardt et al. (2013) study their consequence for the measurement of private returns from investment in knowledge goods. To identify and measure the effect of private spillovers, Eberhardt et al. (2013) modify Griliches (1979)-style regressions with very flexible common-factor modelling that accommodates cross-sectional spillover heterogeneity. The authors show that much of what was previously measured as returns to private research and development (R&D) investment turns out to be largely accounted for by knowledge spillover channels. Also, by exploiting micro-level variations in US firm-level data (that cover the majority of private sector R&D spanning two decades), Bloom et al. (2013) were able to identify positive effects from knowledge spillovers, and separate these from negative, product-maket-rivalry effects. Moreover, the authors can identify a causal link from firm-level R&D to firm performance. The authors also showed that the positive knowledge spillover channel dominates quantitatively. They conclude that the social returns to private R&D investments are two to three times larger than private returns.

[1] In relation to positive knowledge spillovers, McMahon (2010, 2009), in his award-winning work on measuring private and social benefits of higher education, writes:

> Beyond the private benefits (market and non-market) of higher education are the external social benefits. There are ... [c]ontributions to social capital, to the generation of new ideas, and so forth. (*sic*)

Open empirics, our questions and approach. These empirical findings suggest that accounting for and measuring private investment spillover channels are very important for calculating private returns, and thus for informing the design of tax incentives. In turn, fiscal policies have implications for income inequality and economic growth.[2] Thus, one needs a model which considers the political economy and the taxation and redistributive policies while taking into account these spillovers.

However, present studies on private spillovers do not consider the mechanisms underlying such policies, in particular, their underlying politics. These are important connections to take into account: For example, one might think that a reason for why poorer countries (or even some wealthy countries) choose not to meet carbon-emissions targets is largely political. Likewise, corporate taxation with respect to R&D in reality may be influenced by a political majority that is rich. We have not been able to find any empirical work that looks at the connection between politics, private investment incentives and their spillovers onto aggregate activity, and the economic inequality outcomes.

These political-economic issues and empirical gaps lead us to take a first step by asking: How might politics interact with income inequality when private spillovers exist? How do they affect the design of tax policies from positive and normative perspectives? In this paper, we connect the empirically important insights on *private spillover effects* and their relation to policy via *politics*. We model the interactions of tax-dependent private spillovers with economic outcomes, along with politics, as equilibrium phenomena. As such, our theoretical exploration of political-economic equilibria in this paper will help to suggest possible channels and hypotheses for a next step using empirical or experimental methods.

In order to understand the additional channel of private spillovers—or what Eberhardt et al. (2013) refer to as causal misspecification in traditional empirical methods—a logical starting point is the very tractable political-economy model of Hassler et al. (2007). In this model, there is a simple source of intergenerational and cross-sectional heterogeneity that sufficiently motivates political disagreements in the determination of the tax rates.

In our context, the public good can also be interpreted as a knowledge consumption good that is primarily funded by the public, or, as an environmental good. We use the same voting model of Hassler et al. (2007), but incorporate externalities from private investments. Our contribution is in addressing how the presence of private spillovers onto the public good alters equilibrium voting outcomes. Furthermore, as in Hassler et al. (2007), dynamic beliefs matter. The agents in the model are forward-looking and current voters may attempt to manipulate agents' beliefs about future policies. We show that we can index our equilibria by the magnitude and sign of the externality.

Our insights. We show that a sufficiently large positive externality would give rise to a new class of manipulative equilibrium which is qualitatively different from the equilibria in Hassler et al. (2007).

We present two other insights in this paper. First, the positive or negative private spillovers underlying each politico-economic equilibrium can be used to interpret the long run prospects of the model economies. With a sufficiently large positive spillover, the economy can jump on to an equilibrium path in which there is a perpetual majority of rich agents and a double

[2] For example, Alesina and Rodrik (1994) present a political economy model with endogenous growth. They have empirical results which suggest negative relationship between income inequality and growth.

dividend on the public good. This new equilibrium outcome is absent in Hassler et al. (2007). This is attained through the equilibrium tax rates on private investments that encourage more private spillovers, and this reinforces future majority rich voters. This conclusion lends support to various policies which encourage private activities with large positive externalities. Examples of such policies include R&D tax credit and research degree scholarships.

In the setting with large negative spillovers, the economy may end up in a trap where there is a unique long-run majority of poor individuals. In such an equilibrium we say that agents' beliefs are strongly coordinated onto that unique path. This observation resonates with the reality of some economies that are democratic and yet persist in low states of economic development, and they have little concern for the environmental public good. However, if the negative externality is not too big, there is a possibility of an equilibrium with the previous property but it is no longer unique. There is also a continuum of equilibria with manipulative politics. In these equilibria, agents can be coordinated into varying investment efforts, and this weaker coordination effect is still sufficient to induce a long-run majority of the rich. However, the (environmental) public good is inefficiently provided, and the equilibria differ in terms of the negative private spillover and environmental good outcome. This weak coordination effect generates a continuum of equilibrium tax policies. Each policy configuration can coordinate agents' beliefs onto particular equilibria in a somewhat random, or sunspot manner. Thus, we can have the possibility of observing "similar" economies falling in different classes of politico-economic equilibria. In section 5, we suggest how this may help to rationalize the ambiguous empirical evidence on the environmental Kuznets curve hypothesis. (This hypothesis suggests that there is a U-shape in terms of the relation between economic progress and the environmental good outcome.)

Second, we also looks at the composition of the public goods for the case of positive (e.g., knowledge) spillovers. With numerical plots, we show that the share of public spending in the public good is higher in two classes of political equilibrium, compared to its share in a corresponding benchmark Ramsey planner's equilibrium. In what we term as a non-Machiavellian regime (or NM equilibrium), the lack of room for voters to dynamically manipulate agent's expectations results in a high tax rate on the old and also a nonzero tax rate on the young. As a result this limits the size of the positive spillover effect onto the public good, so that there is greater reliance on tax-funded public expenditure in the NM equilibrium. In another regime (called the Machiavellian $M2$ equilibrium), manipulative old voters induce zero taxes for themselves and nonzero tax on the young. This distorts the investment decisions of the young. Due to this under-investment, there is an over-reliance on public expenditure on the public good, relative to the economy's benchmark Ramsey planner's allocation. Furthermore, in this case, there is excessive public good in the economy compared to the Ramsey benchmark. Apart from the numerical examples, we use R&D as a proxy for the knowledge good and plotted some cross country comparison of the composition of R&D (public versus private) against various measures of income inequality. In cross-country data (for OECD members with explicit democratic institutions), there is an apparent positive statistical association between income inequality and the share of government-financed expenditure on R&D. This observation is consistent with a prediction of our theory.

Insights for future measurements. These theoretical predictions take us back to existing empirical considerations. In one of the most influential theories in political economy, Meltzer and

Richard (1981) suggest that high inequality ought to induce the median voter (viz. the politically dominant in reality) to vote for higher levels of taxes and redistribution, which would partially offset rising inequality. Our analyses suggest that when politics influence private incentives in the presence of private spillovers, the answer is: "It depends".

Our insights resonate with some recent empirical studies which suggest that democracy can result in changes in fiscal redistribution and in economic channels that have ambiguous consequences on inequality.[3] In our framework, this ambiguity is rationalized through multiple political-economic equilibrium regimes, given existing democratic institutions. While we stop short of contributing to more precise empirical methods for disentangling the effects of tax policies and private investment incentives on income inequality, we believe our theoretical insights will provide some guidance for a more nuanced approach to empirical/experimental hypothesis building, identification and measurement. In particular, empirical work that consider the connection between politics, redistributive taxation and unequal economic outcomes may have to consider identification issues with regard to multiple equilibrium possibilities.[4]

The remainder of the paper is as follows. In section 2, we describe the economic model. In section 3, we characterize the benchmark Ramsey planner's optimal policy. In section 4 we study the same economy in which policies are determined by state-dependent majority voting, and characterize the various politico-economic equilibrium regimes and their corresponding outcomes. (Readers have the option to skim across the formal results (or characterizations) in section 4 by reading the intuition of these results summarized in segments labelled "Discussion". These segments follow each major result.) In section 5, we discuss the implications of these politico-economic outcomes under two examples of private spillovers. Section 6 then focuses on the role of private spillovers in the composition of public good. Finally we conclude in section 7.

2 Model

Our model builds upon Hassler et al. (2007) with the additional feature that the public good is determined both by the government's public expenditure and the spillover from private investment (indexed by a parameter $\epsilon \in (-1, 1)$). Time is indexed by $t \in \mathbb{N} := \{0, 1, 2, ...\}$.

2.1 Agent types

The agents exist on a continuum and the constant population of young agents is normalized to unity. Each agent lives for two periods. Let a and z_a denote an agent's current age and current-age contingent productivity status respectively. We index them by "Y" for when they

[3] Acemoglu et al. (2015) give a summary of the literature and also present some new empirical evidence.

[4] There are some novel existing work that combine theoretical predictions with empirical/experimental study (see e.g., Battaglini et al., 2012; Bowen et al., 2014). However, attention in these papers is restricted to a unique political-economic equilibrium (e.g., legislative bargaining games) in public investments. Moreover, these papers do not focus on investments that have private externalities. There is also experimental work on intergenerational (environmental) public goods in the context of dynamic cooperation or common-pool games (see, e.g., Hauser et al., 2014). However, they do not explicitly map the experimental game to more intricate details of political-economic mechanisms nor do they consider private spillover effects. In contrast, the structural-econometric literature in Industrial Organization (see, e.g., Aguirregabiria and Jeon, 2020; Ciliberto and Tamer, 2009) often deal with multiple equilibria and partial identification of these for empirical inference. Likewise, in the Monetary Economics (and social-norms) literature, micro-founded theories predict multiple equilibria. A clever experimental study would allow one to infer what is a likely equilibrium (see, e.g., Duffy and Puzzello, 2014). Our theoretical analyses here may suggest a first step toward similar new empirical explorations in political economy.

are young, and "O", for when they are old. At the beginning of each period t, each Y agent realizes an idiosyncratic productivity shock, z_Y, $z_Y \in \{H, L\}$, where H and L refer to "high productivity" and "low productivity" agent types respectively. With probability μ, $\mu \in (0,1)$, a Y agent is of high productivity and with probability $(1-\mu)$, the agent is of low productivity.

Given the realization $z_Y = H$, a (Y, H)-type agent makes an investment decision $i_t \in [0,1]$ and incurs an an effort cost $(i_t)^2$. In addition, the investment i_t contributes a positive or negative spillover to the public good. With probability i_t, the agent succeeds in attaining the high level of per-period income of $\overline{R} \equiv 1$. This agent, when old ($a = O$), attains productivity status S for being successful in her investment when young – i.e., the agent is of type (O, S). The same (Y, H)-type agent faces the probability of $(1 - i_t)$ that her investment is unsuccessful (type U), and thus, yielding the lower per-period income of $\underline{R} \equiv 0$. This agent, when old, is labelled as (O, U). If a Y agent draws $z_Y = L$, then his per-period income is also $\underline{R} \equiv 0$. When his age is $a = O$, his type is labelled (O, L).

To summarize, each agent, at the start of each t, has individual state (a, z_a), where $a \in \{Y, O\}$, and, $z_a \in \{H, L\}$ if $a = Y$, and $z_a \in \{U, S, L\}$ if $a = O$.

2.2 Public good with private activity spillover

We now describe the key spillover mechanism underlying the determination of the public good in the model. The public good, E_t, is determined by two sources: (i) public expenditure A_t; and (ii) externality effect from investment activities of the current high ability young.

The public good is determined by the stylized process:

$$E_t = A_t + \epsilon \mu i_t^*, \tag{2.1}$$

where A_t is public spending on E and $\epsilon \in (-1, 1)$ denotes the effect of the private spillovers. Note that if $\epsilon = 0$, we have a pure public good with zero private activity spillover. This is a special case and corresponds to the model of Hassler et al. (2007).

2.3 Agents' preferences

Denote the age-contingent income tax rates at each time t as τ_t^Y and τ_t^O.[5] Apart from the public good, the policy maker funds a uniform lump-sum transfer of $s_t > 0$ to all agent types. Let $\tau^a \in \{\tau^Y, \tau^O\}$ and $R \in \{\underline{R}, \overline{R}\}$. We assume that agents' common per-period utility function is linear in private consumption $C(\tau^a, R, s)$ and the public good E:

$$C(\tau_t^a, R, s_t) + \lambda E_t;$$

where $\lambda > 0$ is per-period marginal utility of the public good E_t; or equivalently in this model, the marginal rate of substitution between C and E. Let $V_t^{(a,z_a)}$ denote the lifetime payoff of an agent, after their idiosyncratic type (a, z_a) is realized at time t. Denote $\beta \in (0,1)$ as the

[5] The assumption of age-dependent taxes follows Hassler et al. (2007) and is a simplification for tractability. This can be interpreted as part of a simplified tax structure approximating more real-world policy instruments such as progressive income tax rates and pay-as-you-go social security systems. We can also allow for a one-off tax rebate $b_t > 0$ (or $b_t < 0$) at time t, to (Y, H) agents who succeed in attaining $\overline{R} \equiv 1$. All these will not change our analysis and are omitted in the paper. Analysis of the model incorporating b_t can be obtained from the authors upon request.

common subjective discount factor. The expected lifetime payoffs of each type of agent at time t are described below. First, consider the (O,S)-agent with budget constraint $C(\tau_t^O, \overline{R}) = (1-\tau_t^O)\overline{R} + s_t$. This says that the agent's private-good consumption is financed by his after-tax income and (potential) lump-sum transfer. The (O,S)-agent has the following payoff:

$$V_t^{O,S} := V^{O,S}(s_t, E_t, \tau_t^O; \overline{R}) = (1-\tau_t^O) + s_t + \lambda E_t. \tag{2.2a}$$

Second, the (O,U)-agent, whose budget constraint is $C(\tau_t^O, \underline{R}, s_t) = (1-\tau_t^O)\underline{R} + s_t$ has the following payoff,

$$V_t^{O,U} := V^{O,U}(s_t, E_t, \tau_t^O; \underline{R}) = s_t + \lambda E_t, \tag{2.2b}$$

since we normalized $\underline{R} = 0$. Third, the (O,L)-agent has identical payoff to the (O,U) agent:

$$V_t^{O,L} := V^{O,L}(s_t, E_t, \tau_t^O; \underline{R}) = V_t^{O,U}. \tag{2.2c}$$

Fourth, if a young agent realizes type (Y,H), this agent incurs effort cost $(i_t)^2$ to obtain immediate expected payoff $i_t(1-\tau_t^Y)\overline{R} + (1-i_t)(1-\tau_t^Y)\underline{R}$ and continuation payoff $\beta\left[i_t V_{t+1}^{O,S} + (1-i_t)V_{t+1}^{O,U}\right]$. The (Y,H)-agent's total payoff thus reduces to the expression:

$$\begin{aligned} V_t^{Y,H} &:= V^{Y,H}(s_t, s_{t+1}, E_t, E_{t+1}, \tau_t^Y, \tau_t^O; \underline{R}, \overline{R}) \\ &= -(i_t)^2 + i_t\left[(1-\tau_t^Y) + \beta(1-\tau_{t+1}^O)\right] + (s_t + \lambda E_t) + \beta\left(\lambda E_{t+1}\right). \end{aligned} \tag{2.2d}$$

Fifth, the current (Y,L)-agent, whose young-age budget constraint is $C(\tau_t^Y, \underline{R}) = (1-\tau_t^Y)\underline{R} + s_t$, knows that his continuation payoff will be $V_t^{O,L}$, so the total discounted payoff for him is:

$$V_t^{Y,L} := V^{Y,L}(s_t, s_{t+1}, E_t, E_{t+1}, \tau_t^Y, 0, \tau_t^O; \underline{R}, \overline{R}) = (s_t + \lambda E_t) + \beta\left(s_{t+1} + \lambda E_{t+1}\right). \tag{2.2e}$$

2.4 Optimal investment

The investment decision, i_t, is only made by (Y,H)-type agents. Given policies (τ_t^Y, τ_{t+1}^O), the μ-measure of (Y,H) agents are identical and thus choose the same optimal investment function:

$$i_t^* := i^*(\tau_t^Y, \tau_{t+1}^O) = \max\left\{0, \min\left[\frac{(1-\tau_t^Y) + \beta\left(1-\tau_{t+1}^O\right)}{2}, 1\right]\right\}. \tag{2.3}$$

Note that the optimal decision rule (2.3) is derived with E_t taken as parametric by the agents.

2.5 Government budget constraint

Given the proportion μ of (Y,H)-type agents at time t, the proportion of old agents who are of type (O,S) in $t+1$ will be $\mu\pi_{t+1}$, where

$$\pi_{t+1} = i_t^*, \qquad \forall t \geq 1; \text{ and } \pi_0 \text{ given.} \tag{2.4}$$

Thus the model has a single natural state variable given by π_t.

The government is assumed to balance its budget period by period, so that total expenditure on the public good and general transfers must equal the tax revenue accruing from the current population of (O, S) agents and (Y, L) agents:

$$A_t + 2s_t = \tau_t^O \mu \pi_t + \tau_t^Y \mu i_t^*. \tag{2.5}$$

Note that this government budget constraint has already taken into account our normalization of individual incomes as $\underline{R} \equiv 0$ and $\overline{R} \equiv 1$.

2.6 Feasible allocations

The following describes what constitutes a feasible allocation.

Definition 1. A sequence of outcomes $\{A_t, s_t, \tau_t^Y, \tau_t^O\}_{t=0}^{\infty}$ is a *feasible allocation* if it satisfies, for all $t \in \mathbb{N}$: (1) the public good production function (2.1); (2) Young agents best response (2.3); (3) the evolution of the aggregate state (2.4); (4) government fiscal solvency (2.5); and (5) public policy feasibility:

$$A_t + 2s_t \geq 0, \tau_{t+1}^O \in [0,1], \text{ and}, \tau_t^Y \in [0,1]. \tag{2.6}$$

Define $\mathcal{F}^\dagger := \left\{(\tau_t^Y, \tau_{t+1}^O) | (2.3) - (2.1), \text{ and } (2.6) \text{ hold } \forall t \in \mathbb{N}\right\}$, as the set of feasible allocations. Note that $\mathcal{F}^\dagger$ is non-empty.

2.7 Economic policy trade-offs

The model setup implies a key trade-off for a policy maker. This trade-off is decomposable into a "policy intensive margin" (the direct effect of the tax rates) and an "extensive margin" (the indirect effects of the policy on the proportion of high-income agents). The latter extensive margin acts through two channels: one is a typical Laffer curve effect, and the other is our new spillover channel.

The Laffer curve effect suggests that while raising the present value of taxes $(\tau_t^Y + \beta\tau_{t+1}^O)$ will directly increase the public contribution (A_t) toward E_t, it lowers the incentive for (Y, H)-type agents to invest (i_t^*). This results in a lower current tax base (via i_t^*) and future tax base (via π_{t+1}) for providing A_t and A_{t+1}, and thus E_t and E_{t+1}.[6]

Second, raising the present value of taxes, and therefore lowering i_t^*, will now also result in a smaller total positive or negative spillover effect $(\epsilon\mu i_t^*)$ on E_t as described in the public good equation (2.1). The sign of the spillover effect has different implications on the aggregate public good E_t. If the spillover from private investment is positive, then a higher tax rate will reduce the positive incentive of such spillover. If the spillover is negative, a higher tax rate helps to reduce the negative spillover on the public good. This externality effect is new in the model and will play a vital role in inducing the possibility of a new class of politico-economic equilibrium (see Section 4). Moreover, the magnitude of externality, ϵ, will also matter in determining the composition of the public good between public provision and private sources (see section 6).

The overall impact of (equilibrium) tax rates on welfare and the state of the public good will depend on how a particular policy maker (Ramsey planner or decisive voter) evaluates this "intensive-versus-extensive" margin underlying her policy trade-off. In the following section, we describe a Ramsey planner's problem as the benchmark. The resulting allocation is second-best

[6] Note that all the equilibria induced by the various policy mechanisms considered later will turn out to be stationary equilibria, so that $i_t^* = \pi_{t+1} = \pi_t$, resulting in $E_{t+1} = E_t$, for all $t \geq 1$.

given the limited tax instruments available to the planner. Then, we analyze the equilibrium where the policy is determined by majority rule. We consider the two cases: $\lambda > 1/2$ (where $s_t = 0$ and $A_t \geq 0$) and $\lambda < 1/2$ (where $s_t > 0$ and $A_t = 0$).[7] Furthermore, if $\lambda < 1/2$, agents prefers receiving transfers than public goods. If $\epsilon < 0$, we have the case that agents' investment generate negative externality, but none of them care about public goods. This is not an intereseting case to analyse. Therefore, for if $\lambda < 1/2$, we only consider $\epsilon > 0$.

3 Benchmark Ramsey Allocations

The Ramsey planner maximizes a weighted sum of agents' payoffs subject to the requirement that the planner's allocation is feasible. More generally, suppose the planner may attach weight $J \geq 1$ to the payoffs of agents of types (Y, L) and (O, L). A larger J measures a larger bias of the planner towards low-productivity agents; and $J = 1$ implies that the planner assigns the same weight to the payoffs of all types of agents. The detailed description of this problem is in Appendix (S.1.4).

Assumption 1. Let $B := [\mu + (1-\mu)J]\,\lambda$ and $B' := B/\lambda$. If $\lambda > 1/2$, assume that $B > (1+\beta)/2(1+\beta-\epsilon)$. If $\lambda < 1/2$, assume that $B' > 2(1+\beta)/[1+\beta-2\epsilon\lambda]$ and $B' > 2$.

This assumption simplifies the characterization of the Ramsey planning solution without altering the main insights. In particular, the planner's problem can be reduced to an infinite series of static and identical optimization problems. Without this assumption, we would have to keep track of sequences of state-dependent, lower-bound constraints on τ_t^Y. These make the Ramsey characterization and computation history-dependent without affording too much additional insights. (See Appendix S.1, Lemma 7 for a more detailed explanation.) These sufficient conditions admit most values of $\beta \in (0,1)$ and $\epsilon \in (-1,1)$.

Proposition 1 (Optimal Ramsey tax plan)**.** *An optimal Ramsey tax plan, denoted by σ_R^*, exists and is characterized by*

$$\tau_0^O = 1 \tag{3.1}$$

and

$$\tau_t^Y(\sigma_R^*) + \beta\tau_{t+1}^O(\sigma_R^*) = \begin{cases} \frac{(1+\beta)(2B-1)-2B\epsilon}{4B-1} \equiv K > 0, & \text{if } \lambda > 1/2, \\ \frac{(1+\beta)(B'-2)-2B'\epsilon\lambda}{2(B'-1)} \equiv K' > 0, & \text{if } \lambda < 1/2, \end{cases} \tag{3.2}$$

and $(\tau_t^Y(\sigma_R^*), \tau_{t+1}^O(\sigma_R^*)) \in \mathcal{F}^\dagger$, for all $t \in \mathbb{N}$.

Discussion. The intuition for this characterization of our benchmark Ramsey tax plan is as follows. With Assumption 1, we place restrictions on parameters such that the composite parameter B has to be large enough. A higher planner's weight on the poor population (J) and a lower measure of ex-post high productivity agents (μ) both contribute to a larger B. That is, this requires that there is more low-productivity agents and the planner cares more about

[7] From the government budget constraint (2.5), the social rate of transformation between s_t and A_t is $-1/2$. The social marginal rate of substitution between s_t and A_t is $-\lambda$. In the knife-edge case, where $\lambda = 1/2$, the planner is indifferent between alternative compositions of (s_t, A_t). Therefore, we assume either $\lambda > 1/2$ or $\lambda < 1/2$. If $\lambda > 1/2$ ($\lambda < 1/2$), voters would unanimously choose $s_t = 0$ and $A_t \geq 0$ ($s_t \geq 0$ and $A_t = 0$).

them. Given these parametric conditions, the planner would optimally choose to tax the initial (or date-0) old and successful types completely. (Note that this extreme outcome is an artefact of a linear preference specification.) Similarly, the continuation tax plan (for dates $t \geq 1$) is increasing in B. This is stated in terms of the present value of taxes for each generation (i.e., K and K' in the proposition). Furthermore, higher positive externality lowers the subsequent taxes and negative externality increases the taxes.

This benchmark will be useful later for comparisons with the various politico-economic equilibria. Note that given the parameters (embedded in K or K'), the optimal Ramsey tax plan is indeterminate in the tax rates (τ^Y, τ^O) for all dates $t \geq 1$. (We will return to this feature again in Section 6.) We discuss the existence and properties of politico-economic equilibria next.

4 Political Conflict of Interest

In this section, we study how the private spillover may affect the emergence of different equilibrium regimes when tax rates are determined by majority voting. We compare the politico-economic equilibrium outcomes between our model and the one without private spillover—i.e., the Hassler et al. (2007) model.

We assume the following timing as in Hassler et al. (2007): (i) At the beginning of period $t \in \mathbb{N}$, the aggregate state π_t is publicly observed; (ii) the old vote on their preferred policies (τ_t^Y, τ_t^O), taking into consideration of the young agents' expectations, $(\tau_{t+1}^O)^e$, about political choices of the tax on the future old;[8] (iii) young agents realize their idiosyncratic productivity type, $z_Y \in \{H, L\}$; and the (Y, H)-type agents make their investment decisions.[9] Then the next period aggregate state π_{t+1} is realized, along with the level of the public good E_t.

In each generation, decisive voters care only about maximizing their own finite-life total payoffs over policy. This gives rise to the problem of politicians' lack of commitment to implementing the Ramsey optimal policy.

We begin by defining voters' payoffs over policies. First, we define some indicator functions to facilitate the discussion of the two cases. If $\lambda > 1/2$, we write $\mathbb{1}_{\{\lambda>1/2\}} = 1$ and $\mathbb{1}_{\{\lambda<1/2\}} = 0$. Similarly, if $\lambda < 1/2$, we have $\mathbb{1}_{\{\lambda<1/2\}} = 1$ and $\mathbb{1}_{\{\lambda>1/2\}} = 0$.

The payoff term

$$W(\pi_t, \tau_t^O) = \left[\mathbb{1}_{\left\{\lambda>\frac{1}{2}\right\}} + \mathbb{1}_{\left\{\lambda<\frac{1}{2}\right\}}\left(\frac{1}{2\lambda}\right)\right]\tau_t^O \mu \pi_t, \tag{4.1a}$$

is the payoff resulting from the tax revenue accruing from the old rich since preferences are

[8] As will be apparent in the equilibrium concept later, expectations $(\tau_{t+1}^O)^e$ matter for current voters since it affect the (Y, H)-type agents' investment.

[9] As argued in Hassler et al. (2007), the assumption that the tax rates are determined in the beginning of the period and only the old vote is observationally equivalent to assuming that all agents vote in the end of each period and the elected politicians sets tax rates for the following period. In the latter case, the old have no interests at stake and are assumed to abstain. Therefore, in the following period, it is as if the policy is determined by majority voting of the old and implemented in the beginning of the period. In our Online Appendix S.2, we also show how a more general setting, where all young and old agents vote and determine the tax rates for the current period, also yields similar results albeit with slightly more mathematical detail. We present results with this simpler voting assumption in the main paper.

linear; and,

$$Z\left(\tau_t^Y,(\tau_{t+1}^O)^e\right)=\left[\mathbb{1}_{\left\{\lambda>\frac{1}{2}\right\}}(\tau_t^Y+\epsilon)+\mathbb{1}_{\left\{\lambda<\frac{1}{2}\right\}}\left(\frac{\tau_t^Y+2\lambda\epsilon}{2\lambda}\right)\right]\mu i^*(\tau_t^Y,(\tau_{t+1}^O)^e), \tag{4.1b}$$

is the payoff derived from the tax revenue accruing from the current (Y,H)-type agents plus their utility flow arising from a positive/negative spillover on the public good (ϵ) as a result of their private choices i_t^*. Therefore, at time t, given state π_t, (O,z_O)-type agents' utility over policy (τ_t^Y,τ_t^O), given expectations $(\tau_{t+1}^O)^e$, is

$$w^{O,z_O}(\pi_t;(\tau_t^Y,\tau_t^O))=\mathbb{1}_{\{z_O\}}(1-\tau_t^O)+\lambda W(\pi_t,\tau_t^O)+\lambda Z\left(\tau_t^Y,(\tau_{t+1}^O)^e\right), \tag{4.1c}$$

for all $z_O\in\{\{S\},\{L\}\cup\{U\}\}$. The function $\mathbb{1}_{\{z_O\}}=1$ if $z_O=S$, and $\mathbb{1}_{\{z_O\}}=0$ otherwise.

At each each $t\in\mathbb{N}$, the majority voter may either be from the rich class $\{(O,S)\}$ or the poor $\{(O,L)\}\cup\{(O,U)\}$. If the current state is such that $\mu\pi_t\geq 1/2$, then the majority is rich. Otherwise, it is poor. This makes the notion of a median voter state dependent. A median voter is a voting agent whose indirect utility at state π_t is defined as follows.

Definition 2 (Current median voter)**.** A *current median voter* is one whose indirect utility at state π_t is

$$w^m(\pi_t)=\max_{\tau_t^O,\tau_t^Y\in[0,1]}\left\{(1-\tau_t^O)\mathbb{1}_{\{\pi_t\geq\frac{1}{2\mu}\}}+\lambda\left[W(\pi_t,\tau_t^O)+Z(\tau_t^Y,(\tau_{t+1}^O)^e)\right]\right\}. \tag{4.2}$$

where $\mathbb{1}_{\{X\}}=1$ if event $\{X\}$ holds, and $\mathbb{1}_{\{X\}}=0$ otherwise.

Discussion. Observe that from (4.2), if $\lambda>1/\mu$, all types of agents are unanimous on policy outcomes. Intuitively, when λ is very large, the (O,S)-type agents do not mind being taxed since the tax revenue contributes to the public good E_t which yields them high payoffs at the margin. Thus, they would agree on policy outcomes with the low-income agents. Therefore, any interesting political conflict of interest would arise only if marginal utility over the public good is not too large, i.e., $\lambda\leq 1/\mu$. In the politico-economic equilibrium characterizations below, we will work with this latter case, $\lambda\leq 1/\mu$, where we have either $\lambda>1/2$, or $\lambda<1/2$.

Given $\lambda\leq 1/\mu$, the class $\{(O,S)\}$ would disagree with $\{(O,L)\}\cup\{(O,U)\}$ on τ_t^O. It can be deduced from (4.2), that the former would prefer $\tau_t^O=0$, and the latter would prefer $\tau_t^O=1$. However, all classes of voters are unanimous on the determination of τ_t^Y, as they have identical utility term $\lambda Z\left[\tau_t^Y,(\tau_{t+1}^O)^e\right]$.

4.1 Politico-economic equilibrium

We focus on the class of stationary Markov-perfect equilibria.

Definition 3. Given initial aggregate state π_0, a *stationary Markov-perfect politico-economic equilibrium (SMPE)* is a triple $\langle T^O,P,T^Y\rangle$, where:

1. $T^O:[0,1]\to[0,1]$ is a tax rule on the current (O,S) agents such that

$$T^O(\pi_t)=\arg\max\left\{(1-\tau_t^O)\mathbb{1}_{\{\pi_t\geq\frac{1}{2\mu}\}}+\lambda W(\pi_t,\tau_t^O)\right\}; \tag{4.3a}$$

2. $P : [0,1] \to [0,1]$, is the (Y,H)-agents' best response, such that

$$P(\tau_t^Y) = i^*[\tau_t^Y, (\tau_{t+1}^O)^e]; \tag{4.3b}$$

3. $T^Y : \emptyset \to [0,1]$ is a tax rule on the (Y,H) agents such that

$$T^Y = \arg\max \left\{ \lambda Z \left[\tau_t^Y, (\tau_{t+1}^O)^e \right] \right\}; \tag{4.3c}$$

and $(\tau_{t+1}^O)^e = (T^O \circ P)(\tau_t^Y)$.

Discussion. The first component of the equilibrium definition requires that the mapping T^O maximizes the utility term W of the current median voter with respect to the tax on the current old τ_t^O, for every aggregate state π_t.

The second component requires that aggregate state outcomes are consistent with individual level decisions in equilibrium, which produces the equilibrium mapping P. Note that the functional equation (4.3b) defines an operator that maps from the space of (multi)-functions P into itself—i.e. we can write (4.3b) as

$$P(\tau_t^Y) = i^*[\tau_t^Y, (T^O \circ P)(\tau_t^Y)], \tag{4.4}$$

where given functions T^O and i^*, P is a fixed point of the operator defined by the right-hand side of (4.4).

The third component requires that the mapping T^Y maximizes the utility term Z of the current median voter with respect to the policy choice τ_t^Y. Implicit in the second and third components of a SMPE is also the requirement that expectations are consistent with equilibrium conditions. That is, the equilibrium composite function $T^O \circ P$ is used to form rational beliefs about future taxes τ_{t+1}^O in order to calculate current SMPE best responses T^Y and P.

4.2 SMPE characterization

Following Hassler et al. (2007), we exploit the block recursivity of the SMPE definition. We characterize the SMPE mapping T^O, P, and T^Y in turn.[10]

4.2.1 SMPE map T^O

Focusing on the caes $\lambda \leq 1/\mu$, if the current voter population is majority rich, then the SMPE rule prescribes a zero tax on the current (O,S) agents. Otherwise, it says to tax them completely.

Lemma 1. $T^O : [0,1] \to [0,1]$ *is an injective map such that*

$$T^O(\pi_t) = \begin{cases} 0, & \textit{if } \pi_t \geq 1/(2\mu) \\ 1, & \textit{if } \pi_t < 1/(2\mu) \end{cases}. \tag{4.5}$$

Proof. The solution for the function T^O is obtained from solving (4.3a). □

[10] This section is an extended analysis of the politico-economic equilibria in Hassler et al. (2007). The private-choice spillover on the public good (via ϵ) is a new feature of our SMPE characterization and outcomes. The direction of the spillover effect on the public good plays an important role. In particular, there exists a new class of politico-economic equilibrium if the spillover is positive and significant.

4.2.2 SMPE map P

Next, we describe the equilibrium transition law P.

Lemma 2. $P : [0,1] \to [0,1]$ *is a correspondence such that*

$$P(\tau_t^Y) \begin{cases} \in \left\{ \frac{1-\tau_t^Y}{2}, \frac{1-\tau_t^Y+\beta}{2} \right\}, & \text{if } \tau_t^Y \in [0, 1+\beta-1/\mu] \\ = \frac{1-\tau_t^Y}{2}, & \text{if } \tau_t^Y \in (1+\beta-1/\mu, 1] \end{cases}. \tag{4.6}$$

Discussion. Correspondence P summarizes the equilibrium best response of (Y,H)-type agents, given τ_t^Y, and the agents' beliefs about the tax rate on the old in the next period, $(\tau_{t+1}^O)^e$. By Lemma 1, we have a tractable description of the agents' expectation of future old-age tax. The agents can expect either a zero or a one-hundred percent tax rate on the future old. However, the expectation has to be consistent with the equilibrium strategy as defined in Definition 3. The threshold $1+\beta-1/\mu$ is the largest tax rate consistent with agents' expectation $(\tau_{t+1}^O)^e = 0$. Lemma 2 shows that if τ_t^Y is high enough, given its effect on the current young's investment, the agents should not expect a majority of rich voters in the next period (i.e. a zero future tax on the old). Therefore, the only equilibrium-consistent expectation when τ_t^Y is high, is to expect majority poor voters in the next period. However, when τ_t^Y is sufficiently small, both beliefs about the future tax rate on the old can be consistent with equilibrium strategies.

4.2.3 SMPE map T^Y and overall characterization

Now we characterize the last component of a SMPE, along with the rest of the SMPE requirements. Note that T^Y depends on agents' beliefs about the tax on the future old, $(\tau_{t+1}^O)^e$. This is a consequence of current voters taking into account that the current (Y,H) agents will best respond to τ_t^Y in their investment decisions (i^*).

Discussion. The impact of τ_t^Y is twofold. First, there is a direct effect via their net income flow in the current period. Second, there is an indirect effect via their income flow when old next period, since in equilibrium $(\tau_{t+1}^O)^e = (T^O \circ P)(\tau_t^Y)$.

These two effects create trade-offs. Increasing τ_t^Y raises the payoff $\lambda Z[\tau_t^Y, (\tau_{t+1}^O)^e]$ directly. However, increasing τ_t^Y lowers the current young's investment effort, thereby inducing a possible majority poor next period. The majority poor next period will set $\tau_{t+1}^O = 1$. This will lower the payoff $\lambda Z(\tau_t^Y, (\tau_{t+1}^O)^e)$ since Z is a decreasing function of $(\tau_{t+1}^O)^e$.

Following Hassler et al. (2007), we distinguish between two possible classes of voting strategies according to current young agents' beliefs. We will call these two classes of voting strategies "Machiavellian" and "non-Machiavellian".[11] Loosely, the former is one in which it is possible for current voters to manipulate $(\tau_{t+1}^O)^e$ when setting their preferred τ_t^Y in order to generate a larger tax base of endogenous measure $\mu i^*[\tau_t^Y, (\tau_{t+1}^O)^e]$ of the (Y,H) agents in the current period. The latter non-Machiavellian voting is the opposite—one in which current voters expect the worst outcome, $(\tau_{t+1}^O)^e = 1$.

[11]Note that Hassler et al. (2007) used, respectively, the terminology of "strategic" and "sincere" voting. We use the respective terms, "Machiavellian" and "non-Machiavellian", as strategic and sincere voting already have reserved meanings in the political economy literature.

Non-Machiavellian voting.

We begin by first characterizing T^Y in a non-Machiavellian voting SMPE.

Definition 4. A voter is said to be playing a *non-Machiavellian voting strategy* if his voting action is consistent with the belief $(\tau_{t+1}^O)^e = 1$, such that

$$T^Y = \arg \max_{\tau_t^Y \in [0,1]} \lambda Z(\tau_t^Y, 1); \tag{4.7}$$

and, this induces the current (Y, H)-agent's choice of i_t which would enforce the actual realization of $\tau_{t+1}^o = (\tau_{t+1}^O)^e = 1$ next period.

In a Non-Machiavellian equilibrium, we say that there is no room for the voter to manipulate beliefs about the future tax on the old, since the equilibrium-consistent belief is already at the maximal rate, $(\tau_{t+1}^O)^e = 1$. Then a non-Machiavellian voting strategy for τ_t^Y is such that

$$T^Y = (1 - \Psi(\lambda, \epsilon))/2 \tag{4.8}$$

where $\Psi(\lambda, \epsilon) = \left[\mathbb{I}_{\{\lambda>1/2\}} + \mathbb{I}_{\{\lambda<1/2\}} 2\lambda\right] \epsilon$. Next we define a critical tax rate which would facilitate our discussion of different classes of equilibria.

Definition 5. Let $\tilde{\theta} := \tilde{\theta}(\beta, \epsilon, \lambda)$ be the tax rate τ_t^Y satisfying

$$Z\left(T^Y, 1\right) = Z\left(\tilde{\theta}, 0\right), \tag{4.9}$$

where T^Y is defined in (4.8). We have $\tilde{\theta}(\beta, \epsilon, \lambda) := [(1 + \beta - \Psi(\epsilon, \lambda)) - \sqrt{\beta(2 + 2\Psi(\epsilon, \lambda) + \beta)}]/2$.

Definition 5, together with Lemmas 1 and 2, imply that if $\tilde{\theta} > 1 + \beta - 1/\mu$, the class of non-Machiavellian equilibrium is the only one that emerges. Otherwise, there exists some $\tau_t^Y \in [\tilde{\theta}, 1 + \beta - 1/\mu]$ such that $Z\left(T^Y, 1\right) < Z\left(\tau_t^Y, 0\right)$, and this is consistent with agents expecting $(\tau_{t+1}^O)^e = 0$.

Proposition 2 (Non-Machiavellian equilibria)**.** *A non-Machiavellian (NM) SMPE always exists such that,*

$$\begin{aligned} T^O(\pi_0) &= \begin{cases} 0, & \text{if } \pi_0 \geq 1/(2\mu) \\ 1, & \text{if } \pi_0 < 1/(2\mu) \end{cases}; \\ T^O(\pi_t) &= 1, \qquad (\forall t \geq 1); \end{aligned} \tag{4.10a}$$

$$P(\tau_t^Y) \begin{cases} \in \left\{\frac{1-\tau_t^Y}{2}, \frac{1-\tau_t^Y+\beta}{2}\right\}, & \text{if } \tau_t^Y \in [0, \tilde{\theta}(\beta, \epsilon, \lambda)] \\ = \frac{1-\tau_t^Y}{2}, & \text{if } \tau_t^Y \in (\tilde{\theta}(\beta, \epsilon, \lambda), 1] \end{cases}; \tag{4.10b}$$

and,

$$T^Y = \frac{1 - \Psi(\epsilon, \lambda)}{2}. \tag{4.10c}$$

If $\tilde{\theta} > 1 + \beta - 1/\mu$, the non-Machiavellian (NM) SMPE is the only equilibrium.

Proposition 3 (Non-Machiavellian SMPE outcome)**.** *The resulting non-Machiavellian (NM) SMPE outcome is unique for $t \geq 1$.*

$$\tau_t^Y = (1 - \Psi(\epsilon, \lambda))/2; \tag{4.11a}$$

$$\tau_t^O = 1; \tag{4.11b}$$

$$\pi_t = (1 + \Psi(\epsilon, \lambda))/4; \tag{4.11c}$$

$$E_t = \begin{cases} \frac{\mu(1+\epsilon)(3+\epsilon)}{8}, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu\epsilon(1+\lambda\epsilon)}{4}, & \text{if } \lambda < \frac{1}{2} \end{cases}; \tag{4.11d}$$

$$s_t = \begin{cases} 0, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu(1+2\lambda\epsilon)(3-2\lambda\epsilon)}{16}, & \text{if } \lambda < \frac{1}{2} \end{cases}; \tag{4.11e}$$

Proof. The non-Machiavellian voting equilibrium outcomes (4.11a)-(4.11e) can be derived from the characterization just proven. □

Discussion. We first discuss the positive spillover case. When agents care enough about the public good ($\lambda > 1/2$), there will be no public transfers ($s_t = 0$) and the government spends all tax revenue on the public good. Even though the tax on the young is falling with $\epsilon > 0$, a larger ϵ induces a greater level of the public good E_t through the spillover effect. Since λ captures the marginal utility of the public good, E_t increases in λ. If $\lambda < 1/2$, there is no public provision toward E_t, and E_t increases in ϵ purely through the spillover effect. Furthermore, since the tax on the young is decreasing in ϵ, this induces the successful young agents to exert more investment effort, and thereby increasing the tax base toward providing a higher level of s_t.

The negative spillover case has the opposite intuition. Because of the negative spillover of young agents' investment activity onto the public good, the NM equilibrium tax on the young is increasing in the magnitude of this negative spillover margin. This lowers both transfers and the spillover effect onto the public good.

Machiavellian voting

We now turn to the other class of voting equilibria which we termed Machiavellian. In Machiavellian equilibria, voters have the desire and ability to manipulate the voting outcomes such that a majority of rich voters will emerge in the next period. We first derive some useful observations in Lemma 3 and Lemma 4, and then, we describe the Machiavellian SMPE and their outcomes in Proposition 4 and Proposition 5.

Lemma 3. *Assume $0 \leq 1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$. If $(T^O \circ P)(\theta) = 0$, then for all $t \in \mathbb{N}$, a current median voter prefers $\tau_t^Y = \theta$ to any other $(\tau_t^Y)' \in [0, \theta)$.*

In the case of $\epsilon < 0$, it can be verified that $1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$ holds for all feasible parameterizations, since $-\Psi(\epsilon, \lambda)$ is always positive valued. Therefore, Lemma 3 always holds as long as $[0, 1 + \beta - 1/\mu]$ is non-empty, and $\epsilon < 0$. However, for the positive spillover case ($\epsilon > 0$), the assumption, $0 \leq 1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$ only holds if ϵ is sufficiently large. The following lemma illustrates the case where the assumption in Lemma 3 fails. This also results in the emergence of a new class of equilibrium.

Lemma 4 (positive spillover). *Consider the positive spillover case, $\epsilon > 0$, and assume $0 \leq (1+\beta-\Psi(\epsilon,\lambda))/2 \leq 1+\beta-1/\mu$. If $(T^O \circ P)[(1+\beta-\Psi(\epsilon,\lambda))/2] = 0$, then for all $t \in \mathbb{N}$, a current median voter prefers $\tau_t^Y = (1+\beta-\Psi(\epsilon,\lambda))/2$ to any other $(\tau_t^Y)' \in [0,(1+\beta-\Psi(\epsilon,\lambda))/2) \cup ((1+\beta-\Psi(\epsilon,\lambda))/2, 1+\beta-1/\mu]$.*

Lemma 4 follows directly from Lemma 3, since at $\tau_t^Y = (1+\beta-\Psi(\epsilon,\lambda))/2$, the function $Z(\tau_t^Y, 0)$ attains a maximum, and $Z(\tau_t^Y, 0)$ is strictly increasing and continuous on $[0,(1+\beta-\Psi(\epsilon,\lambda))/2)$. This result is new in our extension of Hassler et al. (2007). This is a consequence of the positive spillover effect on the public good from private decisions. In particular, such an equilibrium configuration could occur if, all else given, $\epsilon > 0$ is sufficiently large.

Discussion. For the positive spillover case, Lemma 3 says that given regularity conditions, if the current majority voter is rich, then the dominant strategy of the decisive voter is to extract from the young the highest possible tax rate θ which is consistent with the expectation of a majority rich voters again next period. When the positive externality is sufficiently large, Lemma 4 says that the tax rate $\tau_t^Y = (1+\beta-\Psi(\epsilon,\lambda))/2$, which maximises the tax revenue collected from the young can support expectations of a majority of rich voters in the next period and thus should be the equilibrium tax rate.[12]

We are now ready to summarize all the previous results. The following proposition provides the complete characterization of a Machiavellian SMPE.

Proposition 4 (Machiavellian equilibria). *Assume $\beta \geq [2-\mu(1-\Psi(\epsilon,\lambda))]^2/(4\mu)$. There exists a set of Machiavellian SMPE such that, for all $t \in \mathbb{N}$, we have:*

$$T^O(\pi_t) = \begin{cases} 0, & \text{if } \pi_t \geq 1/(2\mu) \\ 1, & \text{if } \pi_t < 1/(2\mu) \end{cases}; \tag{4.12a}$$

$$P(\tau_t^Y) = \begin{cases} \frac{1-\tau_t^Y+\beta}{2}, & \text{if } \tau_t^Y \in [0,\theta] \\ = \frac{1-\tau_t^Y}{2}, & \text{if } \tau_t^Y \in (\theta,1] \end{cases}; \tag{4.12b}$$

and,

(M1). if $0 \leq 1+\beta-1/\mu < (1+\beta-\Psi(\epsilon,\lambda))/2$, then,

$$T^Y = \theta \in [\tilde{\theta}(\beta,\epsilon,\lambda), 1+\beta-1/\mu]; \tag{4.12c}$$

(M2). or if $0 \leq (1+\beta-\Psi(\epsilon,\lambda))/2 \leq 1+\beta-1/\mu$, then,

$$T^Y = \frac{1+\beta-\Psi(\epsilon,\lambda)}{2} \in [\tilde{\theta}(\beta,\epsilon,\lambda), 1+\beta-1/\mu]; \tag{4.12d}$$

Proof. The assumption $\beta \geq [2-\mu(1-\Psi(\epsilon,\lambda))]^2/(4\mu)$ suffices to ensure that the set $[\tilde{\theta}(\beta,\epsilon,\lambda), 1+\beta-1/\mu]$ is non-empty, and that $\tilde{\theta}(\beta,\epsilon,\lambda) \leq 1+\beta-1/\mu$. Then the existence of, and characterization in (4.12a)-(4.12d) follow from Lemma 3 and Lemma 4, respectively. □

[12] As in Hassler et al. (2007), despite the possible non-monotonicity in $P(\tau_t^Y)$, either Lemma 3 or Lemma 4—depending on parameters $(\beta,\epsilon,\mu,\lambda)$—tells us that in terms of Machiavellian SMPE, we need only focus on the equivalent class of Machiavellian SMPE in which $P(\tau_t^Y)$ is always monotone.

Discussion. Note that a "sunspot" variable extraneous to the model $\theta \in [\tilde{\theta}(\beta, \epsilon, \lambda), 1+\beta-1/\mu]$ is part of the description in the $M1$ class of equilibria. That is $M1$ has a continuum of equilibria. Intuitively, The $M1$ equilibria feature a lack of coordination among voters so that there still needs to be some extraneous factor θ that helps to pin down a particular selection or equilibrium outcome. As a result, any value of θ in that interval can constitute an $M1$-equilibrium tax rate on the young. The presence of positive (negative) spillovers, ϵ, tends to lower (increase) $\tilde{\theta}(\beta, \epsilon, \lambda)$ and thus widens (narrows) the feasible range of the sunspot policy variable θ sustaining a particular selection of an $M1$ equilbrium.

In our additional $M2$ class of Machiavellian equilibrium there is no indeterminacy. The $M2$ equilibrium only emerges in the positive exernality case when the positive spillover effect ($\epsilon > 0$) is sufficiently large. A large enough positive spillover effect can solve the coordination issue amongst voters and facilitate a unique equilibrium. That is, the positive spillover is sufficiently large such that all the voters would still be manipulative, but will choose exactly the same tax rate to optimize theirs payoffs, while expecting that the majority will be the rich in the next period. We will discuss this more in the examples in the following section. Before we do so, the following summarizes analytical outcomes for each type of Machiavellian equilibria described in Proposition 4.

Proposition 5 (Machiavellian SMPE outcomes)**.** *Assume $\beta \geq [2 - \mu(1 - \Psi(\epsilon, \lambda))]^2/(4\mu)$. For every $t \geq 1$ such that $\theta \in [\tilde{\theta}(\beta, \epsilon, \lambda), 1 + \beta - 1/\mu]$, a Machiavellian SMPE outcome is given by the following:*

(M1). If $0 \leq 1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$, then outcomes are indeterminate and depend on $\theta \in [\tilde{\theta}(\beta, \epsilon, \lambda), 1 + \beta - 1/\mu]$:

$$\tau_t^Y = \theta; \tag{4.13a}$$

$$\tau_t^O = 0; \tag{4.13b}$$

$$\pi_t = \frac{(1 - \theta + \beta)}{2}; \tag{4.13c}$$

$$E_t = \begin{cases} \frac{\mu(\epsilon+\theta)(1-\theta+\beta)}{2}, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu\epsilon(1-\theta+\beta)}{2}, & \text{if } \lambda < \frac{1}{2} \end{cases}, \tag{4.13d}$$

$$s_t = \begin{cases} 0, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu\theta(1-\theta+\beta)}{4}, & \text{if } \lambda < \frac{1}{2}. \end{cases} \tag{4.13e}$$

(M2). If $\epsilon > 0$ and $0 \leq (1+\beta-\Psi(\epsilon,\lambda))/2 \leq 1+\beta-1/\mu$, then outcomes are uniquely:

$$\tau_t^Y = \frac{1+\beta-\Psi(\epsilon,\lambda)}{2}; \tag{4.14a}$$

$$\tau_t^O = 0; \tag{4.14b}$$

$$\pi_t = \frac{(1+\beta+\epsilon)}{4}; \tag{4.14c}$$

$$E_t = \begin{cases} \frac{\mu(1+\beta+\epsilon)^2}{8}, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu\epsilon(1+\beta+2\lambda\epsilon)}{4}, & \text{if } \lambda < \frac{1}{2} \end{cases}, \tag{4.14d}$$

$$s_t = \begin{cases} 0, & \text{if } \lambda > \frac{1}{2} \\ \frac{\mu\left[(1+\beta)^2-4(\lambda\epsilon)^2\right]}{16}, & \text{if } \lambda < \frac{1}{2}, \end{cases} \tag{4.14e}$$

Finally, we collect the two possible SMPE notions in the overall result below.

Theorem 1 (Existence and types of SMPE)**.** *Suppose $\lambda \leq 1/\mu$.*

1. *If $\beta < [2-\mu(1-\Psi(\epsilon,\lambda))]^2/(4\mu)$, then a SMPE is uniquely a non-Machiavellian equilibrium (NM) as stated in Proposition 2.*

2. *If $\beta \geq [2-\mu(1-\Psi(\epsilon,\lambda))]^2/(4\mu)$, a non-Machiavellian equilibrium (NM) as stated in Proposition 2 always exists. But there also exists either:*

   (a) an infinite set of Machiavellian equilibria ($M1$ type) as stated in Proposition 4 (M1), or,

   (b) a unique Machiavellian equilibria ($M2$ type) as stated in Proposition 4 (M2) when ϵ is positive and sufficiently large.

Discussion. Consider Propositions 3 and 5. With a negative spillover, the tax rates on the young τ_t^Y in all the classes of political equilibria are higher than their counterparts in the same economies without the spillover (c.f. Hassler et al., 2007). In contrast, with a positive net spillover economy, the tax rates on the young are lower than the ones in Hassler et al. (2007).[13]

If there is positive spillover ($\epsilon > 0$) from private activity, the preferred tax on the young is lower than if $\epsilon = 0$ to encourage higher investment effort by the young. The majority old voter benefits from more public good (and pay zero tax in the $M1$ and $M2$ equilibria), and the current young, in expectation, benefit from lower tax, higher expected income (or consumption), and also more public good consumption. The positive spillover, therefore, provides a "double dividend" in such a context. However, in a negative spillover economy, there would be a trade-off for the old voters: A higher tax rate on the young will reduce their investment efforts. This, on the one hand, will reduce the negative spillover on the public good, but on the other hand, it will reduce the tax base and therefore the public expenditure on public good.

A notable observation from Theorem 1 is that the equilibrium outcome is not symmetric for the two kinds of economies with negative and positive spillover. In particular, a new class of equilibrium ($M2$) emerges in the positive spillover case, in which the tax rate in the Machivaellian equilibrium is unique.

[13]The above statements also tend to be true in the $M1$ equilibria in which $\tau^Y = \theta \in [\tilde{\theta}(\beta,\epsilon,\lambda), 1+\beta-1/\mu]$. Our economy with the negative (positive) spillover admits relatively higher (lower) realizations of θ, since $\tilde{\theta}(\beta,\epsilon,\lambda) \gtrless \tilde{\theta}(\beta,0,\lambda)$ when $\epsilon \lessgtr 0$.

5 Equilibrium Interpretations

In this section, we discuss the the implications of the equilibria that we obtain together with some examples.

For a negative private spillover example, the public good can be interpreted as some environmental good while private investments generate some flow of transient pollutants such as sulphur dioxide (SO_2) or carbon dioxide (CO_2).[14] This case applies to countries that tend to rely on technologies with negative spillover effects on the natural environment. The government also spends (A_t) on cleaning up the pollution and providing environmental goods, such as national parks.

On the flipside, we can also have countries that may have a positive net spillover from private activity to the environmental good. Some examples may be industries which produce with recycled materials and green energy. Another positive spillover interpretation is to consider the public good as public knowledge. Every private agent in such a model world derives utility directly from public knowledge or know-how (see e.g. Stiglitz, 1974). The private investment can be regarded as investment in human capital which is positively associated with higher probability of attaining high income. In total, young high-productivity agents induce a positive spillover effect on aggregate knowledge through their human capital investment. The government also spends on providing knowledge—e.g., financing the flow of public-school teachers.

From Propositions 3 and 5, we have analytically derived the various equilibrium outcomes under the possible SMPE regimes. We now elaborate on their implications, focusing on the spillover's implication for political power and average incomes. Consider economies that differ in terms of the magnitude and direction of the private spillover ϵ, but they are otherwise identical in taste (β, λ) and distribution of skills (μ). Then, as a subset of our SMPE characterizations, we have the following observations:

1. If $\epsilon \searrow -1$, then the equilibrium regimes $M1$ and $M2$ vanish, and the only SMPE that exists is the NM class. Under the NM class, the economy has a perpetual majority of poor voters.

2. If the spillover is "sufficiently small"—i.e., if $\epsilon < 0$ or $\epsilon > 0$ such that $\beta \geq [2 - \mu(1 - \Psi(\epsilon, \lambda))]^2/(4\mu)$ and the interval $[\tilde{\theta}(\beta, \epsilon, \lambda), 1 + \beta - 1/\mu]$ is non-empty—then there exist multiple equilibria. In addition to the NM equilibrium, there also exists the $M1$ class, where there is a long-run majority of rich voters.

3. If $\epsilon \nearrow +1$, then in addition to the NM equilibrium, there exists the $M2$ class equilibrium. Under the $M2$ equilibrium, there is a perpetual majority of rich voters. Moreover, the $M2$ allocation of E_t is increasing in the magnitude of the positive spillover, ϵ. But if $\lambda < 1/2$, s_t is decreasing in ϵ.

We can interpret these results as follows. Consider the limiting case (1): For economies that have quite extreme negative spillovers of private activity onto everyone's enjoyment of the public good, they will tend to get stuck in a long run political equilibrium (NM) that exhibits a majority of poor voters. These majority poor cannot influence the incentives of high ability

[14] For a more detailed modeling of the carbon cycle, see Hassler et al. (2016) for a summary of the quantitative literature. The purpose of this paper is to provide potential theoretically-inspired hypotheses linking politics, private investment spillovers and public good outcomes (e.g., environmental good). It is not yet a quantitative or empirical exercise.

young generations of agents in order to push the economy onto a majority-rich path, unlike the voters in the other "Machiavellian equilibria".[15] In these economies, there is too much reliance on polluting technologies and too little (or non-existent) reliance on new ideas or technologies that have positive spillover effects. Therefore, the resulting net negative spillover effect on the average citizen's enjoyment of the public good, for example, the environment, can be large. In our model, such an economy can be perpetually trapped in a political-economic equilibrium that has a majority of poor voters. Additionally, the greater the magnitude of this negative spillover, the worse off the poor are, since they enjoy less public good (spillover).

Our results also show that there can be intermediate and less obvious possibilities involving equilibrium manipulative politics ($M1$ SMPE). In case (2), the politico-economic equilibrium tax rate on the young will be indeterminate. There is still possibility of inducing a majority of rich in the long run. This echoes a similar property in Hassler et al. (2007). This holds when the absolute value of the spillover is sufficiently small. The key mechanism that ensures perpetual survival of the majority rich is the ability of each successive cohort of majority voters to induce current high-productivity young agents to exert high effort in their private activities. In the case of "small" positive spillovers, this reinforces the voters' incentives to induce a succession of majority rich voters, so that there is always a lesser reliance on taxing private activity, since $\tilde{\theta}(\beta, \epsilon, \lambda)$ is decreasing with $\epsilon > 0$. In the case of "small" negative spillovers, although there is a tendency for drawing higher taxes on the high-productivity young agents, since $\tilde{\theta}(\beta, \epsilon, \lambda)$ is increasing with $|\epsilon|$, the equilibrium trade-off is still not sufficiently large to deter those young agents from exerting enough effort to ensure a majority of rich voters in subsequent periods.

These intermediate cases can be used to interpret a variety of countries that seem to have similar average incomes and democratic institutions, but they may differ in the degree and direction of spillovers of private activity onto the public good outcome. In the context of E being an enviromental good, the empirical ambiguity of the "Environmental Kuznets Curve" hypothesis can also be interpreted in our model as multiple political equilibrium classes and the possible indeterminacy of equilibrium outcomes across countries. The well-known "Environmental Kuznets Curve" hypothesis, suggests that there ought to be an inverted U-shaped (or U-shaped) relationship between income per capita and pollution (or environmental quality). Various empirical studies have attempted to test this hypothesis (see e.g. Grossman and Krueger, 1995; Torras and Boyce, 1998; Stern and Common, 2001; Harbaugh, Levinson and Wilson, 2002; Dijkgraaf and Vollebergh, 2005). Also see Dinda (2004) for a survey of this literature. However, there is no conclusive evidence in favor of the hypothesis. Specifically, the literature has produced a wide variety of estimated polynomial relationships (see e.g. Harbaugh et al., 2002).

In case (3), we have economies that have high-enough positive spillovers of private activity onto the public good, e.g., national knowledge flow. They may jump onto the $M2$ equilibrium path. We can interpret such a possible case to be countries where a high degree of knowledge building or innovation benefits from private income-generating activity spillovers, and the latter are encouraged through equilibrium tax incentives. In such $M2$ type of equilibrium outcomes, there will be a perpetual majority of rich voters as these voters have the incentive to keep tapping into the positive spillover gains on their individual welfare, and not rely too much on taxation that discourages private activity with those spillover effects.

[15] The presence of negative spillover from private activity on the public good makes the interval of $[\tilde{\theta}(\beta, \epsilon), 1 + \beta - 1/\mu]$ smaller than the interval in the model without such spillover (i.e., the HSZ model), as $\tilde{\theta}_{-}(\beta, \epsilon) > \tilde{\theta}(\beta, 0)$. This implies that it is easier for an NM equilibrium to emerge in an economy with $0 > \epsilon \searrow -1$, all else equal.

6 Public Expenditure Share in Public Good

In this section, we discuss the composition of the public good between public spending and private spillovers when the spillover is positive.[16] We compare the composition across the different classes of politico-economic equilibria, as well as against their respective benchmark Ramsey allocation with numerical examples. In all these numerical examples, we fix $\beta = 0.85$ and $\lambda = 1$, and vary μ and ϵ to generate each of the NM, $M1$ or $M2$ types of equilibria. Each corresponding Ramsey equilibrium is derived from the characterization in Proposition 1.

It should be noted that the tax rates are indeterminate in the Ramsey equilibrium σ_R^*. For any Ramsey optimal policy $\langle \tau^Y(\sigma_R^*), \tau^O(\sigma_R^*) \rangle$ satisfying the condition $\tau^Y + \beta\tau^O = K$ in (3.2), there is a unique equilibrium investment outcome $i^*[\tau^Y(\sigma_R^*), \tau^O(\sigma_R^*)]$. Any Ramsey optimal policy $\langle \tau^Y(\sigma_R^*), \tau^O(\sigma_R^*) \rangle$ satisfying (3.2) yields the same optimal ex-ante social welfare for the planner. Given this observation, we pick the combination of optimal Ramsey tax rates $\langle \tau^Y(\sigma_R^*), \tau^O(\sigma_R^*) \rangle$ that yields the highest possible public good provision. We set $\tau^Y(\sigma_R^*) = 0$ and solve for $\tau^O(\sigma_R^*)$ from the condition $\tau^Y + \beta\tau^O = K$ in (3.2). Also, for the $M1$ equilibria, τ^Y is indeterminate. We pick $\tau^Y = 1 + \beta - 1/\mu$ such that the highest utility of old voters can be attained.

Figure 1 presents a panel of plots showing the composition of public good between public expenditure and private spillovers under the positive spillover ($\epsilon > 0$) case. The vertical axes in the charts measure the level of the public good E attained in each example equilibrium outcome. Each type of the SMPE classes—NM, $M1$ and $M2$—is labelled on the horizontal axes, respectively in panels (a), (b) and (c) of the figure. Given each panel of the figure, we consider different economies with increasing magnitudes of $\epsilon > 0$. Adjacent to each SMPE outcome (indexed by some $\epsilon > 0$), is that of its corresponding Ramsey planner allocation.

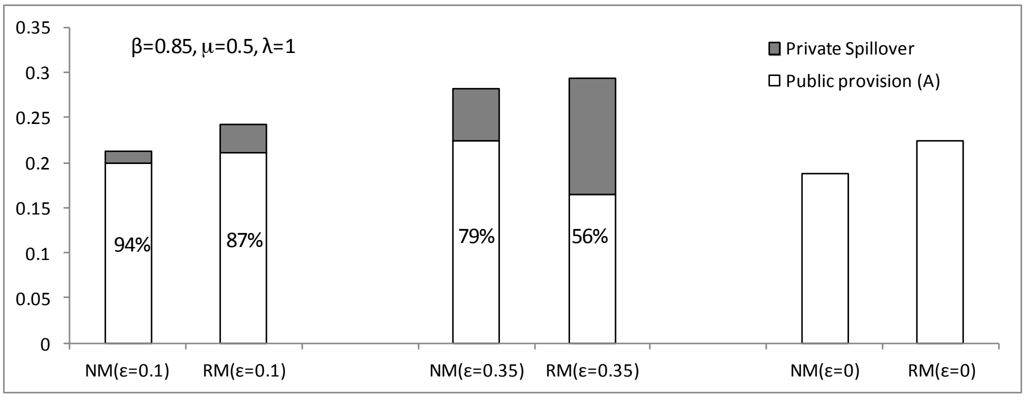
The share of public expenditure in the public good is larger in the NM and $M2$ equilibria than the corresponding share in the Ramsey allocation. Such a difference is more obvious as the spillover magnitude increases—see Figure 1 (a) and (c). Under our selection of the optimal Ramsey outcome, the public-expenditure channel is exploited as much as possible. The spillover effect $\epsilon\mu i^*(\sigma_R^*)$ in absolute terms is fixed, since $i^*(\sigma_R^*)$ is the same across any pair $\langle \tau^Y(\sigma_R^*), \tau^O(\sigma_R^*) \rangle$ consistent with the optimal Ramsey plan. However, in the NM equilibrium, the tax rate on the old is always 1 and the expectation that the next period will have a majority poor will restrict the tax rate on the young. This will limit the positive spillover that could be generated and the economy will rely more on public expenditure (A) to provide for the public good E.

Similarly, in the $M2$ equilibrium, voters' political manipulation will restrict the positive spillover channel from being fully exploited. In contrast, the Ramsey planner does not worry if the majority is poor or rich and therefore has the flexibility to take advantage of the private spillover onto the level of the public good.

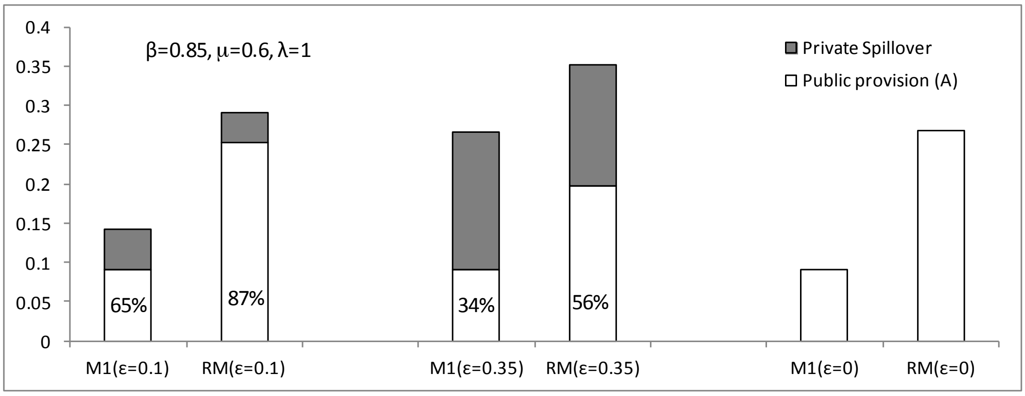
Interestingly, for the $M1$ equilibrium selection, the share from public expenditure in public good is smaller than that in the Ramsey allocation—see Figure 1 (b). In this example equilibrium, the old are not taxed whereas the young are set to be taxed at the rate $1 + \beta - 1/\mu$. However, for the Ramsey planner, since the absolute spillover effect $\epsilon\mu i^*(\sigma_R^*)$ is fixed, then to

[16] In the case of negative spillovers, the amount of public good would be the government funded public good minus the negative spillover. All the resulting public good in the economy would be government funded. Also, focusing on the positive spillover case allows us to put the emphasis on the additional $M2$ class of equilibrium.

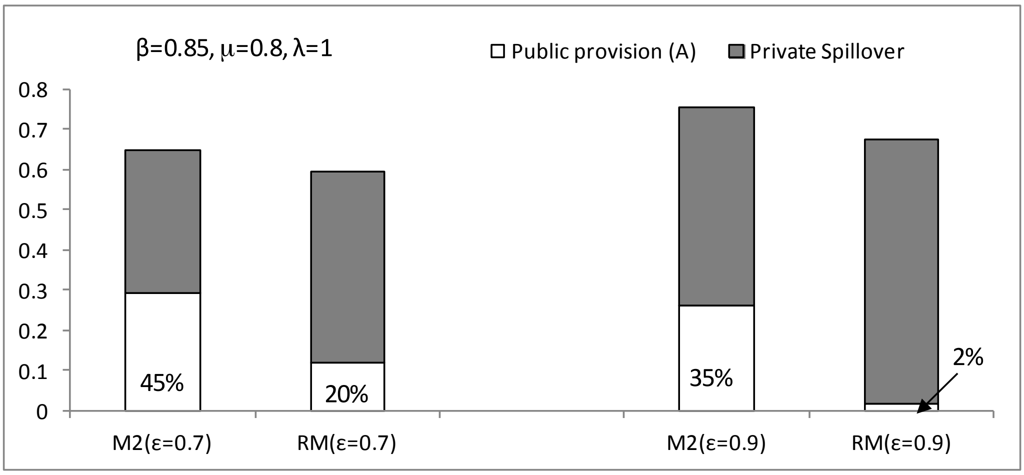
Figure 1: Public Good Composition (positive spillover economy, $\epsilon > 0$)



(a) NM Equilibrium



(b) $M1_+$ Equilibrium



(c) M2 Equilibrium

Note: "RM" refers to the Ramsey allocation in each corresponding economy.

attain the highest public good level, the planner can further exploit the maximum on the Laffer curve that is consistent with the equilibrium taxes. While not illustrated, these arguments also hold for the negative spillover case.[17]

To finish off, we would like to suggest some relevance of our theory's predictions to some a

[17] Additional numerical results are available from the authors.

cursory empirical regularity. We can potentially relate the various example NM, $M1$ and $M2$ type of economies (under the case of $\epsilon > 0$) to an empirical regularity. Note that with these possible equilibrium regimes, the prediction is that the public spending share of the public (e.g., knowledge) good is smaller with a larger proportion of the rich in the economy (i.e., the higher the income inequality). (This is also an implication deducible from from Proposition 5.)

In Figure 2 we plot cross-country public expenditure share in total R&D against various measures of income inequality for a selection of OECD countries democratic institutions.[18] These countries tend to have more R&D (knowledge goods production) than the rest of the world. There appears to be a positive correlation between income inequality and the share of government-financed expenditure on R&D (knowledge goods), as predicted in the theory. This relationship appears robust to whether one looks at the Gini coefficient [panel (a)], the poverty rate [panel (b)], or the income ratio of the top 10% to the bottom 10% of people in the income distribution (P90/P10) [panel (c)] as a measure of income inequality.

Nevertheless, we do need to be careful here in extrapolating these cursory insights since we have a very simple model: A more detailed, and quantitatively more flexible model version of our theory here will need be to used. Also, more careful econometric work will be to be done to control for other possible factors driving the correlation in the data. These are currently beyond the scale and scope of this paper.

7 Conclusion

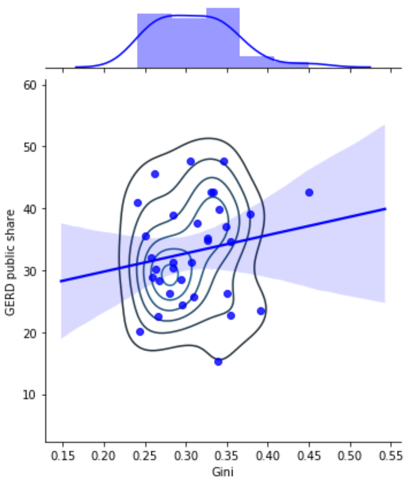
In this paper, we consider a public good that can be affected by spillovers from private activities. We analyze how this spillover would impact on the voting equilibrium when the voters can influence the expectations of future voters.

We present the following theoretical insights: First we show that the presence of the spillover generally means that equilibrium taxes on the young would encode the margin of the spillover effect. As a result, this encourages more (less) of the public good and/or general transfer payment when the spillover is positive (negative).

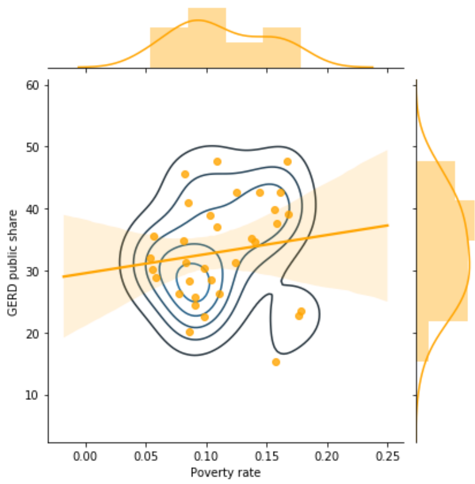
Second, with sufficiently large positive spillovers, a new class of equilibrium with manipulative voters emerges in our model. The presence of such a spillover results in manipulative voters ensuring that their majority-rich kind will benefit from the spillover and their types persist in dominating the policy outcome. Thus, in economies with sufficiently large positive spillovers, the long run path of the economy can be one with a majority of rich people. On the other hand, sufficiently large negative spillovers results in the unique outcome that has a perpetual majority of the poor.

Third, using numerical examples, we find that the voters tend to rely more on the public expenditure in public good provision under the NM and $M2$ politico-economic equilibrium classes than in their corresponding Ramsey planner outcomes. In the instances of the $M1$ or $M2$ classes of equilibria, there is a side effect from political manipulation of expectations. Private spillovers, under such circumstances, provide voters with another channel of manipulation. Not

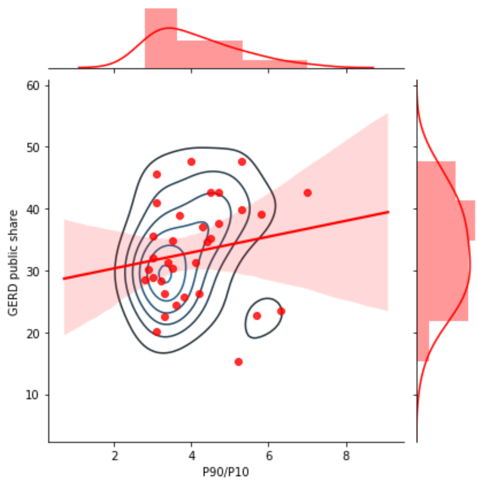
[18] The public expenditure and total expenditure data on R&D is available from the UNESCO.Stat at `http://data.uis.unesco.org/` and the inequality measures come from the OECD database at `https://data.oecd.org/inequality/income-inequality.htm`. These plots consider the most recent years in which the data is available. We have also considered the same annual observations pooled across multiple years: The evidence is the same. A Jupyter notebook performing the data plots is available on our public GitHub repository: `https://github.com/phantomachine/votespill`.



(a) Gini



(b) Poverty rate



(c) P90/P10

Figure 2: Inequality and percentage public share in R&D.

only is the overall allocation of public good inefficient from the Ramsey planner's perspective, the composition of the public good also gets distorted.

Using a stylized model, we show that spillover effects from private activities have interesting and important consequences for long-run welfare and distributional outcomes in a political economy. Assessing the magnitudes of these effects quantitatively would require more explicit and dynamic versions of the simple model considered here. Our work here may help provide a set of theoretical mechanisms that will inform future empirical work. In particular, given the empirically proven significance of private spillover effects, we argue for empirical work or experimental design to identify and account for the interplay of politics, private spillovers, and equilibrium incentives.

A Omitted Proofs

A.1 Proof of Lemma 2

Proof. $P : [0,1] \to [0,1]$ is given by

$$P(\tau_t^Y) := i^*[\tau_t^Y, (\tau_{t+1}^O)^e] = \frac{(1-\tau_t^Y) + \beta[1-(\tau_{t+1}^O)^e]}{2}. \tag{A.1}$$

Any rational beliefs system on $(\tau_{t+1}^O)^e$ must be such that $(\tau_{t+1}^O)^e \in [0,1]$. Therefore, any rational beliefs system would imply that the fixed point $P : [0,1] \to [0,1]$ to the equilibrium functional equation (4.4) has to be such that $P(\tau_t^Y) \in \left[(1-\tau_t^Y)/2, (1-\tau_t^Y+\beta)/2\right]$. However, in an SMPE, rational beliefs $(\tau_{t+1}^O)^e$ must also be consistent with equilibrium strategies, so then, by Lemma 1, $(\tau_{t+1}^O)^e = (T^O \circ P)(\tau_t^Y) \in \{0,1\} \subset [0,1]$. Therefore, we have the following possibilities. If the current tax on the young is too low, i.e. if $(1-\tau_t^Y+\beta)/2 \geq 1/(2\mu)$, then $P(\tau_t^Y) \in \left\{(1-\tau_t^Y)/2, (1-\tau_t^Y+\beta)/2\right\}$ since the majority rich next period will set $\tau_{t+1}^O = 0$ consistent with current voters' SMPE belief that $(T^O \circ P)(\tau_t^Y) = 0$; and if the strict equality holds, then there are equal proportions of rich and poor voters next period, so then $P(\tau_t^Y)$ can be either of the two points in $\left\{(1-\tau_t^Y)/2, (1-\tau_t^Y+\beta)/2\right\}$. Otherwise, if the current net tax on the young is too high, i.e. if $(1-\tau_t^Y+\beta)/2 < 1/(2\mu)$, then $P(\tau_t^Y) = (1-\tau_t^Y)/2$, since the majority is poor next period, and they will set $\tau_{t+1}^O = 1$. This is consistent with voters' SMPE beliefs that $(T^O \circ P)(\tau_t^Y) = 1$. □

A.2 Proof of Proposition 2

Proof. The existence proof is constructive and encompasses all possibilities for the economy in terms of $\lambda \gtrless 1/2$ and $\epsilon \gtreqless 0$. Assume at each t, agents know $T^Y = (1-\Psi(\epsilon,\lambda))/2$ and expect $(\tau_{t+1}^O)^e = 1$. We want to verify that this assumption is consistent with the requirements of a non-Machiavellian equilibrium. If $T^Y = (1-\Psi(\epsilon,\lambda))/2$ and $(\tau_{t+1}^O)^e = 1$, then, $P[T^Y] = i^*[((1-\Psi(\epsilon,\lambda))/2, 1] < 1/2 \leq 1/(2\mu)$. Therefore, for all $\mu \in (0,1)$, indeed the time-$(t+1)$ majority is poor, and they will prefer $\tau_{t+1}^O = 1$. Thus, current voters' expectations are self-fulfilling.

Given SMPE $P(\tau_t^Y)$ as summarized in Lemma 2—i.e. equation (4.6)—and if, the current median voter finds it optimal to set $T^Y = (1-\Psi(\epsilon,\lambda))/2$, then a non-Machiavellian equilibrium exists. Now we prove that setting $T^Y = (1-\Psi(\epsilon,\lambda))/2$ is indeed optimal for the current median voter. First, note that for $\tau_t^Y \in [0, \tilde{\theta}(\beta,\epsilon,\lambda)]$, $Z(\tau_t^Y, 0) = \mu(\tau_t^Y + \Psi(\epsilon,\lambda))\left[\frac{1-\tau_t^Y+\beta}{2}\right]$ is strictly increasing in τ_t^Y, by the definition of $\tilde{\theta}(\beta,\epsilon,\lambda)$. Second, observe that for all τ_t^Y,

$$Z(\tau_t^Y, 0) = \mu(\tau_t^Y + \Psi(\epsilon,\lambda))\left[\frac{1-\tau_t^Y+\beta}{2}\right] > \mu(\tau_t^Y + \Psi(\epsilon,\lambda))\left[\frac{1-\tau_t^Y}{2}\right] = Z(\tau_t^Y, 1).$$

By Definition 5, $Z[\tilde{\theta}(\beta,\epsilon,\lambda), 0]$ is the maximal tax revenue that can be extracted from the current (Y,H) agents by the current median voter, if the current median voter were to deviate from a status quo of a non-Machiavellian voting equilibrium, for some $\tau_t^Y \in [0, \tilde{\theta}(\beta,\epsilon,\lambda)]$. Thus given MPE $P(\tau_t^Y)$ (Lemma 2), if $\tilde{\theta} > 1 + \beta - 1/\mu$, the only equilibrium is non-Machiavellian. □

A.3 Proof of Lemma 3

Proof. Because the SMPE mapping $P : [0,1] \to [0,1]$ in (4.6) is a correspondence with its image possibly non-monotonic in τ_t^Y, there are two cases to consider.

First, consider the case that $P(\tau_t^Y)$ is strictly decreasing in τ_t^Y. Then $(T^O \circ P)(\theta) = 0$ implies that for all $(\tau_t^Y)' \leq \theta$, $(T^O \circ P)[(\tau_t^Y)'] = 0$ (self enforcing) and therefore, by (4.6), $P[(\tau_t^Y)'] = (1 - (\tau_t^Y)' + \beta)/2$. As a result the relevant payoff is $Z(\tau_t^Y, 0)$ on the set $[0, \theta]$. Now, $Z(\tau_t^Y, 0) = \lambda\mu(\tau_t^Y + \Psi(\epsilon, \lambda))(1 - \tau_t^Y + \beta)/2$ is a quadratic function in τ_t^Y and attains a maximum at $(1 + \beta - \Psi(\epsilon, \lambda))/2$. Since $\theta \leq 1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$, then $Z(\tau_t^Y, 0)$ is increasing on $[0, \theta]$. This implies that $Z(\theta, 0) > Z((\tau_t^Y)', 0)$ for all $(\tau_t^Y)' \in [0, \theta)$. That is, a current median voter can still raise higher revenue from taxing the current (Y, H) agents as long as $(\tau_t^Y)' < \theta$. Therefore, if $P(\tau_t^Y)$ is monotone decreasing in τ_t^Y, then the statement of the Lemma is proved.

Second, consider the possibility that $P(\tau_t^Y)$ is not strictly decreasing in τ_t^Y, as described in the upper branch of the SMPE map in (4.6). If $P(\tau_t^Y)$ is non-monotonic in τ_t^Y, then there exists some subset $S \subset [0, \theta)$ such that for all $\tau_t^Y \in S$, $(T^O \circ P)(\tau_t^Y) = 1$. However, we have proven (see the proof of Proposition 2) that $Z(\tau_t^Y, 0) > Z(\tau_t^Y, 1)$ for all $\tau_t^Y \in [0, 1]$. Since it is assumed that $\theta \leq 1 + \beta - 1/\mu < (1 + \beta - \Psi(\epsilon, \lambda))/2$, then we can deduce that $Z(\tau_t^Y, 0)$ is strictly increasing and continuous on $[0, \theta]$. Therefore, as in the previous case, setting $\tau_t^Y = \theta$ is still optimal—i.e. $Z(\theta, 0) > Z(\tau_t^Y, 0) > Z[(\tau_t^Y)'', 1]$, for all $\tau_t^Y \in [0, \theta)$ and all $(\tau_t^Y)'' \in S$. □

References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson, "Chapter 21 - Democracy, Redistribution, and Inequality," in Anthony B. Atkinson and cois Bourguignon Fran eds., *Handbook of Income Distribution*, Vol. 2 of *Handbook of Income Distribution*, Elsevier, 2015, pp. 1885 – 1966. Cited on page(s): [4]

Aguirregabiria, Victor and Jihye Jeon, "Firms'Beliefs and Learning: Models, Identification, and Empirical Evidence," *Review of Industrial Organization*, 2020, *56* (2), 203–235. Cited on page(s): [4]

Alesina, Alberto and Dani Rodrik, "Distributive Politics and Economic Growth," *The Quarterly Journal of Economics*, May 1994, *109* (2), 465–90. Cited on page(s): [2]

Austen-Smith, David and Jeffrey S. Banks, *Positive Political Theory II: Strategy and Structure*, The University of Michigan Press, 2005. Cited on page(s): [SA.7 — S.2]

Battaglini, Marco, Salvatore Nunnari, and Thomas R. Palfrey, "Legislative Bargaining and the Dynamics of Public Investment," *American Political Science Review*, 5 2012, *106*, 407–429. Cited on page(s): [4]

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393. Cited on page(s): [1]

Bowen, T. Renee, Ying Chen, and Hülya Eraslan, "Mandatory versus Discretionary Spending: The Status Quo Effect," *American Economic Review*, 2014, *104* (10), 2941–74. Cited on page(s): [4]

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 2009, *77* (6), 1791–1828. Cited on page(s): [4]

Ciscar, Juan-Carlos, Ana Iglesias, Luc Feyen, László Szabó, Denise Van Regemorter, Bas Amelung, Robert Nicholls, Paul Watkiss, Ole B. Christensen, Rutger Dankers, Luis Garrote, Clare M. Goodess, Alistair Hunt, Alvaro Moreno, Julie Richards, and Antonio Soria, "Physical and economic consequences of climate change in Europe," *Proceedings of the National Academy of Sciences*, 2011, *108* (7), 2678–2683. Cited on page(s): [1]

Cornes, Richard and Todd Sandler, "Easy Riders, Joint Production, and Public Goods," *Economic Journal*, September 1984, *94* (375), 580–98. Cited on page(s): [1]

— and — , "The Comparative Static Properties of the Impure Public Good Model," *Journal of Public Economics*, July 1994, *54* (3), 403–421. Cited on page(s): [1]

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken, "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, September 2014, *52* (3), 740–98. Cited on page(s): [1]

Dijkgraaf, Elbert and Herman Vollebergh, "A Test for Parameter Homogeneity in CO2 Panel EKC Estimations," *Environmental and Resource Economics*, 2005, *32*, 229–239. Cited on page(s): [19]

Dinda, Soumyananda, "Environmental Kuznets Curve Hypothesis: A Survey," *Ecological Economics*, 2004, *49* (4), 431 – 455. Cited on page(s): [19]

Duffy, John and Daniela Puzzello, "Gift Exchange versus Monetary Exchange: Theory and Evidence," *The American Economic Review*, 2014, *104* (6), 1735–1776. Cited on page(s): [4]

Eberhardt, Markus, Christian Helmers, and Hubert Strauss, "Do Spillovers Matter When Estimating Private Returns to R&D?," *The Review of Economics and Statistics*, May 2013, *95* (2), 436–448. Cited on page(s): [1], [2]

Griliches, Zvi, "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, Spring 1979, *10* (1), 92–116. Cited on page(s): [1]

Grossman, Gene and Alan B. Krueger, "Economic Growth and the Environment," *The Quarterly Journal of Economics*, 1995, *110* (2), 353–77. Cited on page(s): [19]

Harbaugh, William T., Arik Levinson, and David Molloy Wilson, "Reexamining The Empirical Evidence For An Environmental Kuznets Curve," *The Review of Economics and Statistics*, August 2002, *84* (3), 541–551. Cited on page(s): [19]

Hassler, J., P. Krusell, and A.A. Smith, "Chapter 24 - Environmental Macroeconomics," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2, Elsevier, 2016, pp. 1893 – 2008. Cited on page(s): [1], [18]

Hassler, John, Kjetil Storesletten, and Fabrizio Zilibotti, "Democratic public good provision," *Journal of Economic Theory*, March 2007, *133* (1), 127–151. Cited on page(s): [2], [3], [4], [5], [9], [11], [12], [15], [17], [19], [SA.4 — S.1]

Hauser, Oliver P., David G. Rand, Alexander Peysakhovich, and Martin A. Nowak, "Cooperating with the future," *Nature*, 2014, *511* (7508), 220–223. Cited on page(s): [4]

IPCC, "Climate change 2013 the physical science basis," in T.F. Stocker, D. Qin, G.-K. Plattner, M.M.B. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, eds., *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press 2013. Cited on page(s): [1]

Lindbeck, Assar and Jörgen Weibull, "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, January 1987, *52* (3), 273–297. Cited on page(s): [SA.6 — S.2]

Matthews, H. Damon, Nathan P. Gillett, Peter A. Stott, and Kirsten Zickfeld, "The proportionality of global warming to cumulative carbon emissions," *Nature*, 2009, *459* (7248), 829–832. Cited on page(s): [1]

— , Susan Solomon, and Raymond Pierrehumbert, "Cumulative carbon as a policy framework for achieving climate stabilization," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2012, *370* (1974), 4365–4379. Cited on page(s): [1]

McMahon, Walter W., *Higher Learning, Greater Good: The Private and Social Benefits of Higher Education*, Johns Hopkins University Press, 2009. Cited on page(s): [1]

— , "The Private and Social Benefits of Higher Education: The Evidence, Their Value, and Policy Implications," *Advancing Higher Education*, March 2010, pp. 1–12. TIAA-CREF Institute. Cited on page(s): [1]

Meltzer, Allan H. and Scott F. Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, *89* (5), 914–927. Cited on page(s): [3]

Nordhaus, W.P. and P. Sztorc, "DICE2013R: Introduction and Users Manual," mimeo, Yale University 2013. Cited on page(s): [1]

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102. Cited on page(s): [1]

Stern, David I. and Michael S. Common, "Is There an Environmental Kuznets Curve for Sulfur?," *Journal of Environmental Economics and Management*, 2001, *41* (2), 162 – 178. Cited on page(s): [19]

Stiglitz, Joseph E., "The Demand for Education in Public and Private School Systems," *Journal of Public Economics*, 1974, *3*, 349–385. Cited on page(s): [1], [18]

Torras, Mariano and James K Boyce, "Income, inequality, and pollution: a reassessment of the environmental Kuznets Curve," *Ecological Economics*, 1998, *25* (2), 147 – 160. Cited on page(s): [19]

NOT FOR PUBLICATION

— Supplementary (Online) Appendix —

S.1 Benchmark Ramsey Planner's Problem

In this appendix, we present our analytical derivation of a hypothetical Ramsey planner's policy or allocation. In the main paper, we presented a simplified version of the model where the tax rebates, b_t, are set to zero. In this appendix, we present the more general version where the rebates are allowed. The benchmark Ramsey solution (with the simplifying assumptions $b_t = 0$ is summarized in Proposition 1. The result is also referred to in Section 6 of the main paper.

Assume $b_t \geq 0$, so that the net tax on the (Y, H) agents, $\hat{\tau}_t^Y := \tau_t^Y - b_t$, may be unbounded below. The (Y, H)-agent's best response function is now:

$$i_t^* := i^*(\hat{\tau}_t^Y, \tau_{t+1}^O) = \min\left\{\frac{(1-\hat{\tau}_t^Y) + \beta\left(1-\tau_{t+1}^O\right)}{2}, 1\right\}. \tag{S.1.1}$$

The next section provides a model-consistent lower bound on $\hat{\tau}_t^Y$ in such a setting.

S.1.1. Natural lower bound on $\hat{\tau}_t^Y$

The following result provides a natural lower bound which is state-dependent.

Lemma 5. *The natural lower bound on $\hat{\tau}_t^Y$ is given by $\underline{\hat{\tau}}_t^Y$ satisfying*

$$\underline{\hat{\tau}}_t^Y = -\tau_t^O \pi_t, \tag{S.1.2}$$

and $\underline{\hat{\tau}}_t^Y$ is bounded in $[-1, 0]$ for all $t \in \mathbb{N}$.

Lemma 5 is necessary, along with other feasibility conditions defined next, for ensuring that a government's planning problem is well defined.

Proof. From the government budget constraint (2.5), we have

$$A_t + 2s_t = \tau_t^O \mu \pi_t + \hat{\tau}_t^Y \mu i_t^*.$$

Let $W(\pi_t, \tau_t^O) := \tau_t^O \mu \pi_t$ —i.e the tax revenue from the current population of (O, S)-type agents. Then we can write

$$\mu i_t^* \hat{\tau}_t^Y = A_t + 2s_t - W(\pi_t, \tau_t^O) \geq -W(\pi_t, \tau_t^O),$$

where the last weak inequality obtains from the fact that $A_t + 2s_t \geq 0$. This implies that $\hat{\tau}_t^Y \geq -\tau_t^O \pi_t / i_t^*$.

Since $i_t^* \equiv \min\left\{i^*(\underline{\hat{\tau}}_t^Y + \beta\tau_{t+1}^O), 1\right\} \in [0, 1]$, then in any equilibrium where the net tax on the (Y, H)-type agents is minimal, i.e. when $\hat{\tau}_t^Y = \inf\{\hat{\tau}_t^Y | \hat{\tau}_t^Y \in (-\infty, 1]\} =: \underline{\hat{\tau}}_t^Y$ (or when the rebate b_t is maximal) for any given τ_{t+1}^O, the induced effort level must be at its upper bound, $i^*(\underline{\hat{\tau}}_t^Y, \tau_{t+1}^O) = 1$. Therefore, $\underline{\hat{\tau}}_t^Y = -\tau_t^O \pi_t \geq -1$, where the last weak inequality is a result of the fact that $\tau_t^O \in [0,1]$ and $\pi_t \in [0,1]$, for all $t \in \mathbb{N}$. □

Definition 6. A sequence of outcomes $\{A_t, s_t, \hat{\tau}_t^Y, \tau_{t+1}^O\}_{t=0}^{\infty}$ is a *feasible allocation* if it satisfies, for all $t \in \mathbb{N}$:

1. young agents best response (S.1.1);
2. the evolution of the aggregate state (2.4);
3. government fiscal solvency (2.5); and
4. the environmental feedback law (2.1);
5. feasibility:

$$A_t \geq 0, s_t \geq 0, \tau_{t+1}^O \in [0,1], \text{ and}, \hat{\tau}_t^Y \in [\underline{\hat{\tau}}_t^Y, 1], \tag{S.1.3}$$

where $\underline{\hat{\tau}}_t^Y$ satisfies (S.1.2).

The time-t, state-π_t-contingent feasible set of government policies is denoted as

$$\mathcal{F}(\pi_t) := \left\{ (\hat{\tau}_t^Y, \tau_{t+1}^O) | (S.1.1), (2.4)\text{-}(2.1) \text{ and } (S.1.3) \text{ hold} \right\}.$$

S.1.2. Ramsey planning problem

The Ramsey planner maximizes a weighted sum of agents' payoffs subject to the requirement that the planner's allocation is feasible. More generally, suppose the planner may also attach a different (relative) weight $J \geq 1$ to the payoffs of agents of types (Y, L) and (O, L). That is, a larger J measures a larger bias of the planner towards low-productivity agents; and $J = 1$ implies that the planner assigns the same weight to the payoffs of all types of agents. The value to the planner beginning from a given initial state π_0 is

$$\begin{aligned} V(\pi_0) = \max_{\{\hat{\tau}_t^Y, \tau_t^O\}_{t=0}^{\infty}} \Bigg\{ & \mu \left[\pi_0 V_0^{O,S} + (1-\pi_0) V_0^{O,U} \right] + (1-\mu) J V_0^{O,L} \\ & + \sum_{t=0}^{\infty} \beta^t \left[\mu V_t^{Y,H} + (1-\mu) J V_t^{Y,L} \right] \Bigg| \tau_0^O \in [0,1] \text{ and} \\ & (\hat{\tau}_t^Y, \tau_{t+1}^O) \in \mathcal{F}(\pi_t) \Bigg\}, \end{aligned} \tag{S.1.4}$$

subject to $(V_0^{O,S}, V_0^{O,U}, V_0^{O,L}, V_t^{Y,H}, V_t^{Y,L})$ satisfying (2.2a)-(2.2e). That is the planner maximizes the weighted ($J \geq 1$) sum of all agents' lifetime payoffs, subject to the requirement of feasibility given in Definition 1.

Let $B := [\mu + (1-\mu)J]\lambda$ and $B' := B/\lambda$. After some algebra, it can be shown that the Ramsey (dual) problem is equivalently given by:

$$V(\pi_0) = \begin{cases} \mu \left\{ \max_{\tau_0^O} \left[\pi_0 \left(1 - \tau_0^O + 2B\tau_0^O\right) + \tilde{V}(\pi_0; \tau_0^O) \right] \right\}, & \text{if } \lambda > 1/2 \\ \mu \left\{ \max_{\tau_0^O} \left[\pi_0 \left(1 - \tau_0^O + \frac{B'}{2}\tau_0^O\right) + \tilde{V}(\pi_0; \tau_0^O) \right] \right\}, & \text{if } \lambda < 1/2 \end{cases} \tag{S.1.5a}$$

where

$$\tilde{V}(\pi_0; \tau_0^O) = \max_{\{\hat{\tau}_t^Y, \tau_{t+1}^O\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \beta^t U(\hat{\tau}_t^Y, \tau_{t+1}^O) \middle| (\hat{\tau}_t^Y, \tau_{t+1}^O) \in \mathcal{F}(\pi_t) \right\}, \tag{S.1.5b}$$

and,

$$U(\hat{\tau}_t^Y, \tau_{t+1}^O) = \begin{cases} \left[\left(1 - \hat{\tau}_t^Y\right) + \beta\left(1 - \hat{\tau}_{t+1}^O\right) - i_t^* + 2B\left(\hat{\tau}_t^Y + \epsilon + \beta\tau_{t+1}^O\right) \right] i_t^*, & \text{if } \lambda > 1/2 \\ \left[\left(1 - \hat{\tau}_t^Y\right) + \beta\left(1 - \hat{\tau}_{t+1}^O\right) - i_t^* + B'\left(\frac{\hat{\tau}_t^Y + \beta\tau_{t+1}^O}{2} + \lambda\epsilon\right) \right] i_t^*, & \text{if } \lambda < 1/2 \end{cases} . \tag{S.1.5c}$$

where i_t^* is defined in (2.3). Note that the only place where the state π_t may matter—i.e. making the problem dynamic—is through the natural lower bound on net tax on the young, as summarized in the set $\mathcal{F}(\pi_t)$.

Consider the case $\lambda > 1/2$. The Ramsey planning problem (S.1.5a)-(S.1.5c), written out explicitly, is:

$$\begin{aligned} \tilde{V}(\pi_0; \tau_0^O) = \max_{\{\hat{\tau}_t^Y, \tau_{t+1}^O\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \Big\{ & \left[(1+\beta) - (\hat{\tau}_t^Y + \beta\tau_{t+1}^O)\right] \\ & \times \left[\left((1+\beta) - (\hat{\tau}_t^Y + \beta\tau_{t+1}^O)\right) + 4B(\hat{\tau}_t^Y + \beta\tau_{t+1}^O + \epsilon)\right] \\ & + \overline{\eta}_t^Y(1 - \hat{\tau}_t^Y) + \underline{\eta}_t^Y\left[\hat{\tau}_t^Y + \tau_t^O\pi_t\right] \\ & + \overline{\eta}_t^O(1 - \tau_{t+1}^O) + \underline{\eta}_t^O\tau_{t+1}^O \Big\}, \end{aligned} \tag{S.1.6}$$

where

$$\pi_t = \min\left\{ \frac{(1+\beta) - (\hat{\tau}_{t-1}^Y + \beta\tau_t^O)}{2}, 1 \right\}.$$

There are two things to note about the Ramsey planning problem. First, the constrained optimization problem in (S.1.5a)-(S.1.5c), has two parts. The first term on the right of (S.1.5a) is a static problem involving τ_0^O. It is independent of any continuation strategy, $\sigma_R|_{(\pi_0;\tau_0^O)} := \{\hat{\tau}_t^Y(\pi_0;\tau_0^O), \tau_{t+1}^O(\pi_0;\tau_0^O)\}_{t=0}^{\infty}$. Second, we will show that the problem beginning from any given $(\pi_0;\tau_0^O)$ summarized by $\tilde{V}(\pi_0;\tau_0^O)$, i.e. the second term on the right of (S.1.5a), can be guaranteed to be analytically characterized (and independent of $(\pi_0;\tau_0^O)$) under a mild assumption on primitive parameters.

Now we consider the first observation regarding the Ramsey problem. The initial period problem can be solved for separately as stated in the following Lemma.

Lemma 6. *The Ramsey optimal tax on the initial old is*

$$\tau_0^O = \begin{cases} 1, & \text{if } \left(B > \frac{1}{2} \text{ and } \lambda > \frac{1}{2}\right) \text{ or } \left(B' > 2 \text{ and } \lambda < 1/2\right) \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

Proof. By inspection of the objective function $\pi_0\left(1 - \tau_0^O + 2B\tau_0^O\right)$ (when $\lambda > 1/2$) or $\pi_0\left(1 - \tau_0^O + B'\tau_0^O/2\right)$

(when $\lambda < 1/2$), in (S.1.5a), the optimal tax function in (3.1) obtains. $\square$

We first deal with the second observation, as summarized in Lemma 7 below, and then incorporate the first observation (see Lemma 6) to arrive at a complete characterization of the optimal Ramsey policy path, σ_R^*, in Proposition 1. The problem described in (S.1.5a)-(S.1.5c) is analogous to that found in Hassler et al. (2007). However, the approach to obtain the second observation is different here, due to the assumption that $b_t \geq 0$.

If $\lambda > 1/2$, then, since $J \geq 1$, we have $B \geq \lambda$. Then this implies that we will be working with $B > 1/2$. Alternatively, if $\lambda < 1/2$, then since $J \geq 1$, we have $B' \geq 1$. Given this, Assumption 1 admits most values of $\beta \in (0, 1)$ and $\epsilon \in (-1, 1)$.

Given Assumption 1, the next result in Lemma 7 says that the generic dynamic Ramsey planning problem (S.1.5a)-(S.1.5c) can be reduced to an infinite series of static and identical optimization problems, as the only state-dependent element of the generic dynamic problem—i.e. the endogenous lower bound $\underline{\hat{\tau}}_t^Y = -(\pi_t \tau_t^O)$ —is never binding.

Note that the minimal value of the lower bound on $\hat{\tau}_t^Y$ is $\inf\left(\{\underline{\hat{\tau}}_t^Y\}_{t\in\mathbb{N}}\right) = -1$. Therefore, given (2.3)-(2.5) and the following conditions:

$$A_t + 2s_t \geq 0, \tau_{t+1}^O \in [0, 1], \text{ and}, \hat{\tau}_t^Y \in [-1, 1], \tag{S.1.7}$$

we can alternatively define

$$\mathcal{F}^\dagger := \left(\bigcup_{t\in\mathbb{N}} \{(\hat{\tau}_t^Y, \tau_{t+1}^O) | (S.1.1), (2.1) - (2.5), \text{ and } (S.1.7) \text{ hold } \forall t \in \mathbb{N}\} \right),$$

which is a non-empty, *non-state-contingent* set of feasible allocations. In other words, the feasible set $\mathcal{F}^\dagger$ does not depend on the state π_t for any $t \in \mathbb{N}$.

Lemma 7. *The natural lower bound (S.1.2) is never binding, and, the optimal value to the Ramsey planner beginning from* any *given* $(\pi_0; \tau_0^O)$ *can be written as*

$$\tilde{V}(\pi_0; \tau_0^O) = \max_{\tau_t^Y, \tau_{t+1}^O} \left\{ \frac{U\left(\hat{\tau}_t^Y, \tau_{t+1}^O\right)}{(1-\beta)} : (\hat{\tau}_t^Y, \tau_{t+1}^O) \in \mathcal{F}^\dagger \right\}. \tag{S.1.8}$$

Proof. Suppose there exists some $t \in \mathbb{N}$ in which the lower bound $\underline{\hat{\tau}}_t^Y$ is binding, and therefore, the planner's problem is dynamically linked by the state π_t via $\mathcal{F}(\pi_t)$. We want to show that this results in a set of contradictions to all possible cases (describing an optimal planning outcome) induced by this assumption.

Consider the case $\lambda > 1/2$. Denote $(\overline{\eta}_t^Y, \underline{\eta}_t^Y, \overline{\eta}_t^O, \underline{\eta}_t^O)$ as the time-t Lagrange multipliers on the respective constraints: $\hat{\tau}_t^Y \leq 1$, $\hat{\tau}_t^Y \geq \underline{\hat{\tau}_t}^Y = -\tau_t^O \pi_t$, $\tau_{t+1}^O \leq 1$ and $\tau_{t+1}^O \geq 0$. The Karush-Kuhn-Tucker (KKT) optimality conditions with respect to $\hat{\tau}_t^Y$ and τ_{t+1}^O for all $t \in \mathbb{N}$ are, respectively,

$$(1-4B)[\hat{\tau}_t^Y + \beta\tau_{t+1}^O] = (1-2B)(1+\beta) + 2B\epsilon + 2(\overline{\eta}_t^Y - \underline{\eta}_t^Y) + \beta\underline{\eta}_{t+1}^Y \tau_{t+1}^O, \tag{S.1.9a}$$

and,

$$\begin{aligned}(1-4B)[\hat{\tau}_t^Y + \beta\tau_{t+1}^O] = (1-2B)(1+\beta) + 2B\epsilon + 2\beta^{-1}(\overline{\eta}_t^O - \underline{\eta}_t^O) \\ - [(1+\beta) - (\hat{\tau}_t^Y + 2\beta\tau_{t+1}^O)]\underline{\eta}_{t+1}^Y.\end{aligned} \tag{S.1.9b}$$

Suppose that $\underline{\hat{\tau}}_t^Y$ is binding for some $t \in \mathbb{N}$, then it must be that $\underline{\eta}_t^Y > 0$. Given this, there are six cases to consider:

C(1). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y = 0$, and $\tau_{t+1}^O = 1$;

C(2). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y = 0$, and $\tau_{t+1}^O = 0$;

C(3). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y = 0$, and $\tau_{t+1}^O \in (0, 1)$;

C(4). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y > 0$, and $\tau_{t+1}^O = 1$;

C(5). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y > 0$, and $\tau_{t+1}^O = 0$;

C(6). $\underline{\eta}_t^Y > 0$, with $\underline{\eta}_{t+1}^Y > 0$, and $\tau_{t+1}^O \in (0, 1)$;

Let $i \in \{1, 2, ..., 6\}$. Define an indicator function

$$\mathbb{1}_{\{i\}} = \begin{cases} 1 & \text{if C}(i) \text{ is true} \\ 0 & \text{otherwise} \end{cases}.$$

Note that the optimality conditions ($S.1.9a$) – ($S.1.9b$), under the assumptions of Cases C(1) to C(6), can then be respectively written as

$$\hat{\tau}_t^Y + \beta\tau_{t+1}^O = \frac{(1+\beta)(2B-1) - 2B\epsilon}{4B-1} + \sum_i \mathbb{1}_{\{i\}} \left(\zeta_{i,t} + \zeta_{i,t+1}\right), \tag{S.1.10a}$$

and

$$\hat{\tau}_t^Y + \beta\tau_{t+1}^O = \frac{(1+\beta)(2B-1) - 2B\epsilon}{4B-1} + \sum_i \mathbb{1}_{\{i\}} \left(\psi_{i,t} + \psi_{i,t+1}\right), \tag{S.1.10b}$$

The composite functions $(\zeta_{i,t}, \psi_{i,t})$ are such that, for example, if $i = 4$, then we have

$$\begin{aligned} \zeta_{4,t} &= 2(4B-1)^{-1}\underline{\eta}_t^Y, \\ \zeta_{4,t+1} &= -(4B-1)^{-1}\beta\underline{\eta}_{t+1}^Y, \\ \psi_{4,t} &= -2\beta^{-1}(4B-1)^{-1}\overline{\eta}_t^O, \\ \psi_{4,t+1} &= (4B-1)^{-1}\left[(1-\beta) - \underline{\hat{\tau}}_t^Y\right]\underline{\eta}_{t+1}^Y. \end{aligned}$$

Denote any Ramsey planner's continuation tax plan (not necessarily optimal) starting from a fixed $(\pi_0; \tau_0^O)$ as $\sigma_R(\pi_0; \tau_0^O) := \{\hat{\tau}_t^Y(\sigma_R), \tau_{t+1}^O(\sigma_R)\}_{t=0}^{\infty}$. We shall abbreviate this as σ_R. Denote an optimal strategy as σ_R^*. The induced total payoff under any σ_R is defined as

$$\begin{aligned} v(\pi_0; \tau_0^O|\sigma_R) &= \frac{1}{4}\sum_{t=0}^{\infty}\beta^t\left\{U(\hat{\tau}_t^Y, \tau_{t+1}^O)\right\} \\ &= \frac{1}{4}\sum_{t=0}^{\infty}\beta^t\Big\{\left[(1+\beta) - (\hat{\tau}_t^Y + \beta\tau_{t+1}^O)\right] \\ &\qquad \times \left[\left((1+\beta) - (\hat{\tau}_t^Y + \beta\tau_{t+1}^O)\right) + 4B(\hat{\tau}_t^Y + \beta\tau_{t+1}^O + \epsilon)\right]\Big\}, \end{aligned}$$

where it is understood that $\hat{\tau}_t^Y \equiv \hat{\tau}_t^Y(\sigma_R)$ and $\tau_{t+1}^O \equiv \tau_{t+1}^O(\sigma_R)$. Note that the per-period indirect utility U is quadratic in the present value of taxes, $\hat{\tau}_t^Y + \beta\tau_{t+1}^O$, and therefore, the value function

$v(\cdot|\sigma_R)$ is also quadratic. Thus, $v(\pi_0;\tau_0^O|\sigma_R)$ attains a global maximum when there is a strategy $\tilde{\sigma}_R$ such that

$$\hat{\tau}_t^Y(\tilde{\sigma}_R)+\beta\tau_{t+1}^O(\tilde{\sigma}_R)=\frac{(1+\beta)(2B-1)-2B\epsilon}{4B-1}\equiv K, \qquad \forall t\in\mathbb{N}. \tag{S.1.11}$$

However, we need to next check if a strategy such as $\tilde{\sigma}_R$ is feasible. Note that, since the upper bounds on $\hat{\tau}_t^Y$ and τ_{t+1}^O are both 1, then the upper bound for the present value of taxes $\hat{\tau}_t^Y+\beta\tau_{t+1}^O$ each period is $1+\beta$. Moreover, the lower bound on the present value of taxes per period is $\underline{\hat{\tau}}_t^Y\in[-1,0]$, by Lemma 5. Since $B>1/2$ and $\beta\in(0,1)$ and $\epsilon\in(0,1)$, then we have $K<1+\beta$. Moreover, since $B>1/2$, then Assumption 1 is sufficient and necessary for $K>0$. Therefore, $K\in(0,1+\beta)$, implying that $\tilde{\sigma}_R$ is a feasible *and* strictly interior strategy such that it is also optimal: $v(\pi_0;\tau_0^O|\tilde{\sigma}_R)=v(\pi_0;\tau_0^O|\sigma_R^*)$.

By assumption, a strategy $\sigma_R^{C(i)}$ satisfying any of Cases C(1)-C(6), and therefore (S.1.10a)-(S.1.10b), must yield the planner an indirect utility of $v(\pi_0,\tau_0^O|\sigma_R^{C(i)})=v(\pi_0,\tau_0^O|\sigma_R^*)$. However, note that (S.1.10a)-(S.1.10b) is always equivalent to (S.1.11) plus strictly non-zero terms: $\sum_i \mathbb{I}_{\{i\}}\left(\zeta_{i,t}+\zeta_{i,t+1}\right)$ or $\sum_i \mathbb{I}_{\{i\}}\left(\psi_{i,t}+\psi_{i,t+1}\right)$. Since the indirect utility function $v(\cdot|\sigma_R)$ is quadratic, any strategy $\sigma_R^{C(i)}$ generated by (S.1.10a)-(S.1.10b) for some $t\in\mathbb{N}$, must yield at least one per-period payoff that is strictly dominated:

$$U[\hat{\tau}_t^Y(\sigma_R^{C(i)}),\tau_{t+1}^O(\sigma_R^{C(i)})]<U[\hat{\tau}_t^Y(\tilde{\sigma}_R),\tau_{t+1}^O(\tilde{\sigma}_R)],$$

and all other per-period payoffs being at most, weakly dominated. This implies that

$$v(\pi_0,\tau_0^O|\sigma_R^{C(i)})<v(\pi_0,\tau_0^O|\sigma_R^*)=\tilde{V}(\pi_0,\tau_0^O).$$

We have a contradiction. The case $\lambda<1/2$ follows the same argument as above.

□

Combining this result with Lemma 6, we thus have a closed-form solution to the dynamic Ramsey planner's allocation problem as summarized in Proposition 1.

S.2 Probabilistic Voting

Here we consider an alternative probabilistic voting model (Lindbeck and Weibull, 1987) where all agents vote. The majority voting model considered in the paper coincides with this model if the old voters have all the influence on voting outcomes, relative to the young. In other words, the model in the paper can be thought of as the extreme setting where *probabilistically*, young voters have a measure zero impact on politician's vote-share objectives.

Consider any pairwise electoral competitive between political candidates, $c\in\{A,B\}$. Voters' voting decisions now depend on both announcements of policy platforms, $\mathbf{p}_c:=(s_t^c,E_t^c,(\tau_t^o)^c,(\hat{\tau}_t^Y)^c)$, and, on the ideologies (i.e., A or B) of the politicians. Without loss of generality, we assume that all young (Y) voters' policy-independent ideological bias toward candidate A, δ^Y, is drawn from the same uniform distribution, $\mathbf{U}\left([-1/(2\phi^Y),1/(2\phi^Y)]\right)$ where $\phi^Y>0$ is the (uniform) density measure. Likewise, all old voters' ideological bias toward candidate A, δ^O is drawn from $\mathbf{U}\left([-1/(2\phi^O),1/(2\phi^O)]\right)$.

From a politician's perspective, δ^a, where $a\in\{Y,O\}$, is a random variable, but it is known for sure by each voter himself. (This captures the idea that political candidates do not perfectly know everything about voters' preferences over policy.) We say that a voter group $a \in \{Y, O\}$ votes for candidate A, conditional on announced policy platforms $(\mathbf{p}_A, \mathbf{p}_B)$, if

$$V^{(a,z_a)}(\mathbf{p}_A) - V^{(a,z_a)}(\mathbf{p}_B) > \delta^a,$$

where $V^{(a,z_a)}(\mathbf{p}_c)$ is the policy preference of a voter of age-productivity type (a, z_a), previously defined in the paper. A swing voter in each group (a, z_a), is someone who, after the realization of δ^a, is indifferent between voting for A or B, and thus is someone with indirect utility

$$\sigma^{a,z_a}(\mathbf{p}_A, \mathbf{p}_B) = V^{(a,z_a)}(\mathbf{p}_A) - V^{(a,z_a)}(\mathbf{p}_B) - \delta^a.$$

Given δ^a, the vote share of candidate A is

$$\begin{aligned} \Pi_A(\mathbf{p}_A, \mathbf{p}_B) &= 1 - \Pi_B(\mathbf{p}_A, \mathbf{p}_B) \\ &= \phi^Y \left[(1-\mu) \left(\sigma^{Y,L}(\mathbf{p}_A, \mathbf{p}_B) + \frac{1}{2\phi^Y} \right) + \mu \left(\sigma^{Y,H}(\mathbf{p}_A, \mathbf{p}_B) + \frac{1}{2\phi^Y} \right) \right] \\ &\quad + \phi^O \left[(1-\mu\pi_t) \left(\sigma^{O,U}(\mathbf{p}_A, \mathbf{p}_B) + \frac{1}{2\phi^O} \right) + \mu\pi_t \left(\sigma^{O,S}(\mathbf{p}_A, \mathbf{p}_B) + \frac{1}{2\phi^O} \right) \right], \end{aligned}$$

noting that $\sigma^{O,U}(\mathbf{p}_A, \mathbf{p}_B) \equiv \sigma^{O,L}(\mathbf{p}_A, \mathbf{p}_B)$.

The probability that candidate A wins the election is

$$\begin{aligned} \Pr\left\{ \Pi_A(\mathbf{p}_A, \mathbf{p}_B) > \frac{1}{2} \right\} = \frac{1}{2} &+ (1-\omega') \left[(1-\mu) \left(V^{Y,L}(\mathbf{p}_A) - V^{Y,L}(\mathbf{p}_B) \right) \right. \\ &\qquad \left. + \mu \left(V^{Y,H}(\mathbf{p}_A) - V^{Y,H}(\mathbf{p}_B) \right) \right] \\ &+ \omega' \left[(1-\mu\pi_t) \left(V^{O,U}(\mathbf{p}_A) - V^{O,U}(\mathbf{p}_B) \right) \right. \\ &\qquad \left. + \mu\pi_t \left(V^{O,S}(\mathbf{p}_A) - V^{O,S}(\mathbf{p}_B) \right) \right], \end{aligned}$$

where $\omega' := \phi^O/(\phi^Y + \phi^O)$.

Let $\omega := (1-\omega')/\omega' \equiv \phi^Y/\phi^O$ denote the relative density of voter biases between the two age groups. Since both candidates aim to maximize their individual probability of winning the election, in any Nash equilibrium given the continuation state π_t, they neccessarily converge onto the same policy:

$$\mathbf{p}_A^\star = \mathbf{p}_B^\star = \arg\max_{\mathbf{p}} \{ W_b(\mathbf{p}; \pi_t) \},$$

where

$$W_b(\mathbf{p}; \pi_t) := \mu\pi_t V_t^{O,S} + (1-\mu\pi_t) V_t^{O,U} + \omega \left[\mu V_t^{Y,H} + (1-\mu) V_t^{Y,L} \right].$$

In other words, the equilibrium policy is one that maximizes the equivalent of a Benthamite aggregation function (see Austen-Smith and Banks, 2005), where the welfare weights are defined by individual voters' marginal vote-probabilities, which in turn depend on $(\phi^Y, \phi^O, \mu, \pi_t)$.

After some algebra (imposing the equilibrium value functions of the agents), the equivalent

Benthamite welfare function is:

$$
\begin{aligned}
W_b(\mathbf{p};\pi_t) \equiv & \mu\pi_t(1-\tau_t^O) + \frac{\mu\omega}{2}\left[\frac{1}{2}\left(1+\beta-\hat{\tau}_t^Y\right) - \beta\tau_{t+1}^O\left(1+\beta-\hat{\tau}_t^Y\right) + \frac{1}{2}\left(\beta\tau_{t+1}^O\right)^2\right] \\
& + \mathbb{I}_{\{\lambda>1/2\}}\left\{(1+\omega)\lambda\mu\left[\pi_t\tau_t^O + \left(\frac{1-\hat{\tau}_t^Y+\beta(1-\tau_{t+1}^O)}{2}\right)\left(\hat{\tau}_t^Y+\epsilon\right)\right]\right. \\
& \qquad + \beta\omega\lambda\mu\left[\left(\frac{1-\hat{\tau}_t^Y+\beta(1-\tau_{t+1}^O)}{2}\right)\tau_{t+1}^O\right. \\
& \qquad\qquad \left.\left. + \left(\frac{1-\hat{\tau}_{t+1}^Y+\beta(1-\tau_{t+2}^O)}{2}\right)\left(\hat{\tau}_{t+1}^Y+\epsilon\right)\right]\right\} \\
& + \mathbb{I}_{\{\lambda<1/2\}}\left\{\frac{(1+\omega)\mu}{2}\left[\pi_t\tau_t^O + \left(\frac{1-\hat{\tau}_t^Y+\beta(1-\tau_{t+1}^O)}{2}\right)\left(\hat{\tau}_t^Y+2\lambda\epsilon\right)\right]\right. \\
& \qquad + \frac{\beta\omega\mu}{2}\left[\left(\frac{1-\hat{\tau}_t^Y+\beta(1-\tau_{t+1}^O)}{2}\right)\tau_{t+1}^O\right. \\
& \qquad\qquad \left.\left. + \left(\frac{1-\hat{\tau}_{t+1}^Y+\beta(1-\tau_{t+2}^O)}{2}\right)\left(\hat{\tau}_{t+1}^Y+2\lambda\epsilon\right)\right]\right\}.
\end{aligned}
$$

Note that the limit economy, where $\omega = 0$, is equivalent to the majority voting setup we have in the paper. This is precisely the case that the old voters have all the influence in the probabilistic voting setup.

S.2.1. Characterizing SMPE policy

Observe that the marginal value of τ_t^O to the swing voter is

$$
\frac{\partial W_b(\mathbf{p};\pi_t)}{\partial \tau_t^O} = \begin{cases} \frac{-\mu\pi_t(1-\omega)}{2} \leq 0, & \text{if } \lambda < \frac{1}{2} \\ \mu\pi_t\left[(1+\omega)\lambda - 1\right] \gtreqless 0, & \text{if } \lambda > \frac{1}{2} \text{ and } \lambda \lesseqgtr \frac{\omega}{2(1+\omega)} \end{cases}.
$$

As a result, the equilibrium tax policy is time- and state-invariant: The SMPE (with probabilistic voting) has that

$$
T^O(\pi_t) = \begin{cases} 0, & \text{if } \lambda \in \left(\frac{1}{2}, \frac{1}{1+\omega}\right], \text{ or if } \lambda < \frac{1}{2} \\ 1, & \text{if } \lambda > \frac{1}{1+\omega} \geq \frac{1}{2} \end{cases} \tag{S.2.1}
$$

for any π_t and $t \geq 0$. This is unlike the majority voting setup in the paper, where the outcome of the SMPE tax function T^O is dependent of the current state π_t. Here, this will simplify the rest of the SMPE tax policy description.

Now consider the marginal valuation of $\hat{\tau}_t^Y$:

$$
\frac{\partial W_b(\mathbf{p};\pi_t)}{\partial \hat{\tau}_t^Y} = \begin{cases} \frac{\mu}{2}\left[\left((1+\omega)\left(\frac{1+\beta-2\lambda\epsilon}{2}\right) - (1+\beta)\omega\right) - \hat{\tau}_t^Y + \frac{\beta}{2}\tau_{t+1}^o\right], & \text{if } \lambda < \frac{1}{2} \\ \frac{\mu}{2}\begin{bmatrix} \left((1+\beta)\left(\lambda+\omega(\lambda-1)\right)\right) \\ -\lambda(1+\omega)\epsilon - \left(2\lambda(1+\omega)-\omega\right)\hat{\tau}_t^Y \\ -\beta\left(\lambda(1+\omega)+\omega(\lambda-1)\right)\tau_{t+1}^O \end{bmatrix}, & \text{if } \lambda > \frac{1}{2} \end{cases}.
$$

After taking into account (S.2.1) in a SMPE, this is:

$$
\frac{\partial W_b(\mathbf{p};\pi_t)}{\partial \hat{\tau}_t^Y} = \begin{cases}
\frac{\mu}{2}\left[\left((1+\omega)\left(\frac{1+\beta-2\lambda\epsilon}{2}\right) - (1+\beta)\omega\right) - \hat{\tau}_t^Y\right], & \text{if } \lambda < \frac{1}{2} \\
\frac{\mu}{2}\begin{bmatrix} ((1+\beta)(\lambda+\omega(\lambda-1))) \\ -\lambda(1+\omega)\epsilon - (2\lambda(1+\omega)-\omega)\,\hat{\tau}_t^Y \end{bmatrix}, & \text{if } \lambda \in \left(\frac{1}{2}\ \frac{1}{1+\omega}\right] \\
\frac{\mu}{2}\begin{bmatrix} ((1+\beta)(\lambda+\omega(\lambda-1))) \\ -\lambda(1+\omega)\epsilon - (2\lambda(1+\omega)-\omega)\,\hat{\tau}_t^Y \\ -\beta(\lambda(1+\omega)+\omega(\lambda-1)) \end{bmatrix}, & \text{if } \lambda > \frac{1}{1+\omega} \geq \frac{1}{2}
\end{cases} .
$$

However, at a SMPE, it must that $\partial W_b(\mathbf{p};\pi_t)/\partial \hat{\tau}_t^Y = 0$, so that the SMPE function for tax on the young is:

$$
T^Y = \begin{cases}
\frac{1}{2}\left[(1-\omega)(1+\beta) - (1+\omega)2\lambda\epsilon\right], & \text{if } \lambda < \frac{1}{2} \\
\frac{1+\beta[\lambda+\omega(\lambda-1)]-\lambda(1+\omega)\epsilon}{2\lambda(1+\omega)-\omega}, & \text{if } \lambda \in \left(\frac{1}{2}\ \frac{1}{1+\omega}\right] \\
\frac{-\omega(1+\beta\lambda)+\lambda(1+\omega)(1-\epsilon)}{2\lambda(1+\omega)-\omega}, & \text{if } \lambda > \frac{1}{1+\omega} \geq \frac{1}{2}
\end{cases} .
$$

S.2.2. Remarks

In short, if we allow all agents in the paper to vote, and assuming the probabilistic voting framework, then the resulting SMPE tax policies, subject to feasibility conditions are quite similar to the case of the Non-Machiavellian (NM) equilibrium or the M2 case of the Machiavellian family of equilibria. In particular, the tax on the young, $\hat{\tau}_t^Y$ is time invariant and is monotone decreasing with $\epsilon > 0$ or monotone increasing with $|\epsilon|$ if $\epsilon < 0$.

However, unlike NM or M2, here the tax on the old can either be 1 or 0 depending on agents marginal utlity over the public good E_t—i.e., the parameter λ. In NM for example, it is always the case that $\hat{\tau}_t^O = 1$, and in M2, this is always 0.