Inefficient Policies in the Green Transition*

Juan Dodyk[†]

November 18, 2025

Abstract

Countries have employed a wide range of policy instruments to mitigate climate change. These policies share a common pattern: governments initially rely on subsidies, together with command-and-control regulations, and eventually adopt carbon pricing. I develop a dynamic model of climate policymaking that accounts for this pattern. Although the first-best policy is solely a carbon tax, a climate-concerned policymaker uses subsidies to induce investments in emissions-abatement technologies with the goal of building a coalition in support of efficient policies in the future. The model provides additional insights: First, a policy package that satisfies political constraints and passes a cost-benefit analysis only exists if the economic costs of decarbonization are not too large, and the social cost of carbon is intermediate. Second, soft commitments, such as net-zero targets, can have real consequences by shifting expectations, but only if initial political pressure is not too large and policymakers are sufficiently concerned about climate. Finally, a higher risk of electoral turnover that replaces a green proposer with a misaligned proposer can improve prospects for a green transition.

*Job Market Paper. Many thanks to my committee: Peter Buisseret, Jeffry Frieden, Alisha Holland and Dustin Tingley. Thanks also to Aleksandra Conevska, Todd Eisenstadt, Justin Fox, Bård Harstad, Emily Luxon, Jonas Meckling, Matto Mildenberger, Kris Ramsay and audiences at APSA, the Climate Pipeline Conference, Harvard, and WashU for comments.

[†]Washington University in St. Louis. Email: juandodyk@wustl.edu. Webpage: juandodyk.github.io.

1. Introduction

Countries representing 81% of global greenhouse gasses (GHG) emissions have communicated a target of net-zero emissions, most of them by 2050 (Climate Watch). There are two policies that, according to leading economists, can achieve these climate objectives at minimal cost: carbon pricing and R&D subsidies (e.g., Acemoglu *et al.* 2016; Metcalf 2019; Blanchard *et al.* 2023). However, in practice, a wide variety of policies are used, which differ substantially in how efficient they are in terms of the cost per ton of CO_2 abated (Mealy *et al.* 2024; Hahn *et al.* 2024). Moreover, different kinds of policies are sequenced in predictable patterns: less efficient policies such as command-and-control regulations, feed-in tariffs (FITs) and renewable portfolio standards (RTSs), which involve explicit or implicit subsidies for emissions abatement investments and renewable energy production, are eventually replaced by carbon pricing as the main policy tool (Linsenmeier *et al.* 2022).

The following examples illustrate this pattern. Germany relied on FITs to support the expansion of renewable energy production since the 1990s (von Hirschhausen *et al.* 2019), but replaced the FITs with auctions in 2014 (Clean Energy Wire). Despite initially being reluctant to participate in the EU Emissions Trading System (ETS) (Ellerman *et al.* 2010), Germany implemented a carbon price for the heating and transport sectors in 2021 (IEA). The EU followed a similar path. It enacted in 2001 the Renewable Energy Directive mandating member states to set national targets for renewable energy production; in 2009, the targets became legally binding, and their scope was expanded (Leipprand *et al.* 2020).[1] The ETS was introduced in 2003, and it imposed a low carbon price until its 2018 reform, which took effect in 2021; prices have been above 80€ per ton of CO_2 most of the time since 2022 (van den Bergh and Botzen 2024). The Canadian federal government implemented a series of inefficient regulations and subsidies in its unsuccessful attempt to comply with the Kyoto Protocol (Jaccard *et al.* 2006; Samson and Stamler 2009; Harrison 2010b); in 2018, it enacted an ambitious national carbon price (Harrison 2023).

Meckling *et al.* (2015, 2017) and Pahle *et al.* (2018) propose an explanation for the sequencing of policy instruments based on the idea of policy feedback: renewable energy support policies and technology mandates can be used to build a coalition in support of more efficient and stringent policies in the future. Policies that create concentrated benefits, while protecting powerful opponents from immediate costs, can induce economic agents to make investments tied to the long-term decarbonization of the economy, which disrupts the power of incumbent carbon-intensive industries in the future without leading to an immediate veto. World Bank researchers have adopted this logic in their policy recommendations (Hallegatte *et al.* 2024).

This argument raises a number of theoretical questions. First, how large is the distortion from

[1] In the top six European countries, "[t]he cost to society implied by the deployment of wind and solar technologies [in 2010] represented €48,300 million" (Dechezleprêtre and Popp 2017).

first-best policies required to pursue this strategy? Second, under what conditions is it possible for policymakers to design a policy package that attracts pivotal opponents without alienating existing members of the coalition? Third, how is the strategy affected by the possibility that future policymakers may not be willing to continue the intended policy sequence?

To address these questions, I develop a dynamic model of climate policymaking. The main ingredients are the following. First, policies emerge from legislative bargaining, and the legislature is heterogeneous—some legislators represent districts that are more invested in high emissions intensity technologies, and some districts face lower costs of decarbonization than others. Both present climate policy and the expectation of future policy affect the investments that economic actors make, which, in turn, affects their preferences in the future. Today's policy outcome is therefore constrained by today's legislature, but also shapes the legislative constraints a policymaker faces in the future.

Second, policymakers face the threat of turnover. The current policymaker may be replaced by another whose preferences differ in how they prioritize environmental concerns versus aggregate economic welfare.[2] This creates uncertainty, which affects incentives to invest in mitigation technologies.

In the baseline model I consider two policy instruments: a carbon tax, and a subsidy for investments in green capital (e.g., renewable energy, clean manufacturing technologies, carbon capture, energy efficiency, or electric vehicles). The carbon tax is a form of carbon pricing and is equivalent to an emissions trading scheme with auctioned allowances. I assume that the revenues are recycled as uniform lump-sum transfers. Later, I extend the model to consider other types of policies, including targeted transfers (e.g., free allowances in the context of cap-and-trade systems), output subsidies, tradable standards, and feed-in tariffs.

The first-best policy in the model is simply a Pigouvian carbon tax that equates the price of carbon emissions to the social cost of carbon. The investment subsidy is not needed, because the expectation of an optimal carbon price is enough to induce socially optimal investment decisions.

Results. The main findings of the paper concern the scenario in which a climate-concerned policymaker has agenda-setting power, the status quo is business-as-usual (BAU), and a majority of the members in the legislature represent carbon-intensive constituencies. In this scenario polluting interests can veto climate policy, and in a static model they would block any change over the status quo. However, under some conditions the policymaker is able to implement a climate policy package that eventually leads to first-best policy. This holds even if legislators care exclusively

[2] Most economic models express environmental damages in terms of a decrease in production (e.g., Golosov *et al.* 2014), so it is not obvious how to separate conceptually concerns for climate and for the economy. We may assume that the two parties differ in their beliefs about the causal link between greenhouse gas emissions and climate change. As an alternative, we may assume that the concern for GHG emissions expresses an attitude about the country's responsibility, possibly induced by international cooperation.

about the economic interests of their constituents, and do not internalize the damages from carbon emissions.

In this case, the equilibrium climate policy in the first period consists of an investment subsidy and a carbon tax that is below the Pigouvian level, and which can be zero. The subsidy is set at the level that leaves the median legislator (ordered by the productivity of the green alternative technology in their district) indifferent between, on the one hand, paying the proposed carbon tax, taking the subsidy, and investing in abatement, and, on the other hand, keeping their polluting capital under business-as-usual for both periods. All polluting districts with better decarbonization opportunities than the median strictly prefer the policy package over the status quo, and decide to invest in abatement. Therefore, in the second period the green districts form a majority, and thus the green policymaker (if still in power) implements the first-best policy. If, instead, the opposition party is in power, they are forced to keep the carbon tax at the new status quo level. Thus, the green proposer uses policy in the first period to reduce the power of polluting interests, and to build a coalition in support of efficient policy in the future.

The need to build a coalition in the legislature that contains both green and polluting interests imposes two political constraints on policy. Polluting districts demand a low carbon tax today and a generous investment subsidy. This creates a trade-off: an ambitious carbon tax in the first period requires a large subsidy to compensate polluting districts. Green districts support a carbon tax because they enjoy the increased government revenue, but oppose a raise in taxes to cover the green subsidies, so they impose a dynamic budget surplus constraint: the subsidy must pay itself with the current plus the future carbon tax revenue. If both constraints can be satisfied simultaneously, the proposer can start the climate policy sequence.

I show that these constraints cannot always be jointly satisfied: the alternative clean technology must be sufficiently productive, investment costs sufficiently low, and the social cost of carbon cannot be too large. The last condition is perhaps counterintuitive. The reason is that if the green policymaker weighs reductions in carbon emissions too heavily, polluting producers expect large losses if they do not invest in abatement. This, in turn, leads to a large demand for the subsidy in the first period, which creates a large fiscal cost that clashes with the budget surplus constraint and costs the support of green districts.

The fact that a climate policy sequence can be started, i.e., the political constraints can be jointly satisfied, does not imply that a climate-concerned policymaker will pursue it. Satisfying the demands of the pivotal polluting industries increases the social costs of climate policy due to inefficient investments and low levels of abatement. If these costs are large enough, the green policymaker will prefer to keep the status quo, unless their climate concerns are sufficiently serious. Thus, the model can explain not only how the policy sequencing strategy works, but also why in many cases it fails or is not pursued, even if politicians in power are concerned about the climate.

The dynamic linkage of policies raises a novel implication: under some conditions, the model has multiple equilibria.[3] There is always an equilibrium in which polluting interests expect that failure of climate policy in the present preserves their political power in the future. But, if the policymaker is sufficiently concerned about the climate, a self-fulfilling prophecy is possible in which, expecting a carbon price in the future, enough economic agents make abatement investments, which reduces the political influence of polluters in the future, leaving the policymaker room to enact the carbon pricing policy. The fact that polluters are expected to lose their political power in the future regardless of the policies implemented in the present reduces their bargaining power, which enables the policymaker to enact more ambitious policies than in the baseline equilibrium. However, the equilibrium features the same policy sequencing pattern: a green technology subsidy combined with a low carbon price, eventually replaced with a Pigouvian carbon tax.

The possibility that the policymaker is not in power in the future to continue the policy sequence may paradoxically help start it. If the future policymaker is not willing to sacrifice present economic consumption to mitigate climate change, they will implement a low carbon tax, breaking the policy sequence. Anticipating this, economic actors will be more reluctant to make investments in green capital. This reluctance has a benefit: it reduces the fiscal cost of the subsidy as fewer districts utilize it, leaving the current policymaker with more fiscal resources for transfers. If the initial share of brown districts is large enough this effect dominates, and increasing the risk of turnover relaxes the political constraints, enabling the implementation of more ambitious policies.

This finding challenges the intuitive notion that low polarization on climate issues is necessary for policy progress. Instead, a polarized party system—where one major party is more committed to climate change mitigation than the median voter, while another opposes stringent climate policy—may be more conducive to initial policy progress when polluting interests still wield significant political power. Such polarization can create windows of opportunity for initiating ambitious climate policy sequences, even though it increases the likelihood of these sequences being interrupted. The passage of the Inflation Reduction Act of 2022 in the US suggests the plausibility of this result.

Finally, the model can be extended to study other policy instruments such as renewable energy production subsidies, clean technology standards, and feed-in tariffs. The qualitative results are robust: if green technologies are sufficiently advanced, capital costs low, and discount factors high, these policies are used in equilibrium in a first stage to disrupt the power of polluting interests, build a green coalition, and create the ground for optimal carbon taxes in the future.

Literature. The paper contributes to the literature on the domestic political economy of climate policy (besides the work already cited, relevant papers include Harrison 2010a; Breetz *et al.* 2018; Dolphin *et al.* 2020; Battaglini and Harstad 2020; Besley and Persson 2023). The first contribution

[3] Biais and Landier (2022) and Smulders and Zhou (2024) also study models where climate policy creates multiple equilibria.

is methodological—I provide a new way to model policymaking under political constraints that are dynamic and microfounded. There are two main approaches in the literature to study policy distortions by special interests: ad-hoc constraints, and common agency models. Two examples of the first approach are Tornell (1991), who models the political pressure by a protected industry as a constraint that policy must keep the employment level of the industry above a certain baseline, and Rozenberg *et al.* (2020), who model the political pressure by fossil fuel energy producers as a "no stranded assets" constraint on policy. Two examples of the second approach are Grossman and Helpman (1994) and Gerlagh and Liski (2023), who assume that a subset of producers can offer transfers to the policymaker contingent on policies.[4] The drawback of both approaches is that they cannot account for the ability of policy to change the power of special interests, because that power is assumed to be exogenous.

Second, the paper contributes specifically to the literature on policy feedback in climate policy (Aklin and Urpelainen 2013; Meckling *et al.* 2015, 2017; Pahle *et al.* 2018; Stokes 2020) by analyzing a microfounded model of the feedback mechanism. More broadly, the paper contributes to the theoretical literature that studies how the dynamic political effects of policies impact their choice (Alesina and Tabellini 1990; Persson and Svensson 1989; Besley and Coate 1998; Prato 2017) by showing how dynamic strategic considerations can explain puzzling patterns of climate policy. Baldursson and von der Fehr (2007) follow a similar approach, but they study a different question, viz., why "brown" governments may sell long-lived allowances instead of pursuing a carbon tax, given that the latter is more efficient. This paper also provides further implications of dynamic policymaking with an endogenous status quo (Buisseret and Bernhardt 2017; Dziuda and Loeper 2018; Austen-Smith *et al.* 2019).

Third, I contribute to a related literature that studies the effects of partisan turnover on climate policy (Ulph and Ulph 2013; Schmitt 2014; Harstad 2020; Hochman and Zilberman 2021; Behmer 2023). A common result is that the possibility of an "anti-climate" policymaker in the future distorts policy in the present by increasing carbon taxes and clean technology subsidies relative to the first-best. Although the same forces are present in this paper, I reconcile the effect of turnover with the empirical observation that green subsidies coexist with low or zero carbon prices.

Fourth, the analysis in this paper contributes to the literature on the politics of instrument choice in climate policy (Buchanan and Tullock 1975; Aidt and Dutta 2004; Hughes and Urpelainen 2015; Meckling and Jenner 2016; Cullenward and Victor 2020; Konisky 2024) by showing how policies that offer benefits conditional on investments in emissions-cutting technologies can be used in

[4] Other papers using the ad-hoc constraints approach to study climate policy include Bovenberg *et al.* (2005, 2008); Kalkuhl *et al.* (2013); Kalk and Sorger (2023); Acharya *et al.* (2024). Others using the common agency approach include Fredriksson (1997); Damania (2001); Fredriksson and Svensson (2003); Damania and Fredriksson (2003); Fredriksson and Sterner (2005); Fredriksson and Wollscheid (2008); Habla and Winkler (2013); Aidt (1998); Aidt and Dutta (2004); Aidt (2010); Lai (2007, 2008); Hanoteau (2014); Grey (2018); Kalkuhl *et al.* (2020); Winkler (2022).

political equilibrium even though they have larger economic costs relative to other available policies. Aidt and Dutta (2004) explain the transition from command-and-control policies to carbon pricing, but they need to assume that there is an exogenous tightening in emissions-reduction objectives; in contrast, the increase in policy ambition emerges endogenously in my model. In addition, the model provides a political economy rationale for the use of green industrial policy (Rodrik 2014; Meckling 2021; Allan and Nahm 2024; Juhász and Lane 2024).

The paper also speaks to the research on the political acceptability of climate policy (Gaikwad *et al.* 2022; Meckling and Nahm 2022; Meckling and Strecker 2023; Bolet *et al.* 2023; Gazmararian and Tingley 2023) and lobbying (Grumbach 2015; Kim *et al.* 2016, 2021; Brulle 2018; Meng and Rode 2019; Goldberg *et al.* 2020; Kennard 2020) by showing how they shape both the ambition of policy objectives and the instruments used to achieve them.

Structure of the Paper. In Section 2 I introduce the model and discuss its assumptions. In Section 3 I determine optimal policy without political constraints, with and without political turnover. In Section 4 I analyze the full model and provide the main results. In Section 5 I show that the main qualitative findings are not affected by considering other commonly used policy instruments nor by changes in selected assumptions of the model. Section 6 discusses the possibility that there is political turnover before the investments in clean technology create the political coalition required to lock in the changes in climate policy, and concludes.

2. The Model

The Economy. There is a set of districts indexed by $i \in I = [0, 1]$. Each district has a unit of specific capital that cannot be traded.[5] Capital can be "green" or "polluting," and we denote $\chi_{it} = 1$ if the capital of i is polluting at time t, and $\chi_{it} = 0$ if it is green. Polluting districts can "upgrade" or "transition" their capital to the green kind by paying a cost $c > 0$. The decision to upgrade at time t is denoted $\iota_{it} = 1$; absent a transition, $\iota_{it} = 0$.

There is one good in the economy, which is used for production and consumption, and will serve as the numéraire. A polluting district produces y units of the good at cost $\frac{1}{2}y^2$, and emits y units of carbon by doing so. A green district produces y units at cost $\frac{1}{2A_i}y^2$, where $A_i > 0$ is the productivity of green technology in district i.[6] For simplicity, I will assume that $A_i = Ai$ for all $i \in I$, where $A > 1$ is a parameter. This captures in reduced form the assumption that decarbonization entails different changes in productivity in some districts than others—green technology is more

[5] This is a composite of the district's human capital and investments that are location-specific in the short run (e.g., energy production capacity and infrastructure).

[6] This is equivalent to assuming that there are firms with production function $f_i(x) = x_B^{\frac{1}{2}} k_B^{\frac{1}{2}} + A_i x_G^{\frac{1}{2}} k_G^{\frac{1}{2}}$, where x_B, x_G denote units of the numéraire used for production, k_B denotes units of i's polluting specific capital, and k_G denotes units of i's green specific capital.

productive than polluting technology in some districts, but less productive in others.

Policy. There are three policies. First, a carbon tax $\tau \in [0, 1]$. Second, there is a green investment subsidy $s \geqslant 0$ that is transferred to any polluting district that decides to upgrade its capital. Finally, there is a uniform lump-sum transfer (or tax) $T \in \mathbb{R}$. A climate policy bundle is defined a tuple (τ, s, T) of a carbon tax, a green investment subsidy, and a uniform transfer.

Policy Process. The districts are represented in a legislature. There are two proposers, G and B, and an initial policy bundle $p_0 = 0$. In each period $t \in \{1, 2\}$ the timing of events is:

1. The proposer $P_t \in \{G, B\}$ is drawn, with $P_0 = G$ and $\Pr(P_t = P_{t-1}|P_{t-1}) = \rho \in (0, 1]$.
2. P_t chooses a policy proposal $p'_t = (\tau'_t, s'_t, T'_t)$.
3. If a majority of districts i prefer p'_t over p_{t-1}, then $p_t = p'_t$, and otherwise $p_t = p_{t-1}$.
4. Districts make production and investment decisions, $y_i \geqslant 0$ and $\iota_{it} \in \{0, 1\}$.

In the last period T is not a choice, and is set so that the budget is intertemporally balanced:

$$\sum_{t=1}^{2} \delta^{t-1} \left[\tau_t e_t - T_t - \int_0^1 \iota_{it} s_t \, di \right] = 0,$$

where $e_t = \int_0^1 e_{it} \, di$ is the aggregate quantity of emissions. I assume that agents can take on debt for free subject to the same budget constraint, viz, $B_1 + \delta B_2 = 0$ if B_t is the net debt taken or paid in period t.

Preferences. The agents maximize expected discounted payoffs, with a common discount factor $\delta \in (0, 1]$. In each period t, districts' payoff is given by their income

$$\pi_{it} = \begin{cases} (1 - \tau_t) y_{it} - \frac{1}{2} y_{it}^2 - \iota_{it}(c - s_t) + T_t & \text{if } i \text{ is polluting,} \\ y_{it} - \frac{1}{2A i} y_{it}^2 + T_t & \text{if } i \text{ is green,} \end{cases}$$

which is given by the returns of their capital endowment net of taxes and transfers, and investment costs net of subsidies if they decide to transition.

The payoff of proposer $P \in \{G, B\}$ is given by

$$W_P = \sum_{t=1}^{2} \delta^{t-1} \left[\int_0^1 \pi_{it} \, di - \alpha_P D_t(E_t) \right],$$

where $E_t = e_1 + \cdots + e_t$ is the stock of local carbon emissions, D_t measures environmental damage at time t, with $D'_t > 0$ and $D''_t \geqslant 0$, and $\alpha_P \in [0, 1]$ measures how each proposer trades off consumption for environmental damage. I will assume that $\alpha_G = 1$ and $\alpha_B = 0$, and that environmental damage is linear in emissions, $\sum_{s \geqslant t} \delta^{s-t}(D_s(E) - D_s(0)) = \lambda E$ for all $t \geqslant 1$ and $E \geqslant 0$, where $\lambda \in (0, \frac{1}{2})$ measures the social cost of carbon.

Equilibrium Concept. Subgame perfect equilibrium.

Comments on the Assumptions. The model of the economy is stylized in order to focus on the political mechanism, but the simplifications are not uncommon in the literature. For example, Acharya *et al.* (2024) also assume that there is only one good in the economy and there is no market power, and Coate and Morris (1999) also assume that firms upgrade their technology by making a binary investment decision. Colmer *et al.* (2024) provide empirical justification for the latter assumption: in their study of the EU ETS the conclude that "[o]ur findings are consistent with firms paying an up-front fixed cost to invest in alternative 'clean' production technologies that reduce marginal variable costs." Ramadorai and Zeni (2024) show that firms react to beliefs about future carbon prices by investing in carbon abatement technology, which justifies the assumption that firms anticipate future climate policies. The assumption that the productivity of clean technology is heterogeneous across constituencies can be defended in two ways: the cost-competitiveness of renewable energy depends on location (Davis *et al.* 2023), and the advances in decarbonization technology depend on industry (Victor *et al.* 2019).

The assumption that different proposers may differ in their concerns for climate may be explained by differences in partisanship. Cadoret and Padovano (2016) find that left-wing governments promote the development of renewable energy more than right-wing parties in Europe. Lundquist (2024) finds that the degree of environmentalism expressed in parties' manifestos predicts the level of policy stringency they implement when in power. Knill *et al.* (2010), Jensen and Spoon (2011) and Jahn (2022) obtain similar results. Fankhauser *et al.* (2015) and Dolphin *et al.* (2020), however, do not find an effect of party ideology on climate legislation or the carbon price. In any case, the main results of this paper do not depend on this assumption. What is crucial is that carbon-intensive interests hold political power (because of their representation in the legislature) regardless of which party controls the agenda. Mildenberger (2020)'s logic of the double representation of these interests (with labor being represented via left-wing parties and business via right-wing parties) provides an empirically grounded justification for this assumption.

An important feature of the model is that investments in emissions-substituting capital in the present change the policy preferences of the constituencies where those investments take place. I offer three pieces of evidence to support this assumption. First, Alberdi (2024) finds that subsidized investments in rooftop solar panels increased support for ambitious climate policies and for the Green party in Germany. Second, Urpelainen and Zhang (2022) show that wind turbine installations increased vote shares of climate-concerned Democratic candidates in US House elections, and led to an increase in pro-climate votes in Congress, even though they may have created an electoral backlash among voters located close to the turbines. Third, Vormedal and Meckling (2023) provide evidence that the shale gas revolution led oil and gas industries to sincerely support carbon pricing (during the Trump administration, Exxon lobbied against withdrawal from the Paris Agreement, and a Republican-backed coalition involving Exxon and other oil companies promoted legislation for a federal carbon tax starting at $40 per ton), and the fuel-efficiency regulations on the car industry imposed by the Obama administration led some car manufacturers to resist Trump's decision to roll back those regulations, due to their investments in clean technologies.

There are many important issues involved in climate policy from which the model abstracts, such as innovation, learning-by-doing and network externalities (Stock 2020; Fischer *et al.* 2021; Bistline *et al.* 2023; Hahn *et al.* 2024), imperfect competition (Kennard 2020), regulation of energy markets (Reguant 2019; Davis *et al.* 2023), land use regulation (Sud *et al.* 2023), other tax distortions (Barrage 2020), international trade (Clausing and Wolfram 2023; Kotchen and Maggi 2024), conservation (Harstad 2023a), international cooperation (Battaglini and Harstad 2016; Harstad 2023b), private politics (Egorov and Harstad 2017), consumer preferences (Besley and Persson 2023), behavioral distortions of energy-efficiency investments (Allcott *et al.* 2014), and the possibility of fiscal illusion (Abbott and Jones 2023).

3. Benchmarks

Optimal Policy. Suppose that a green policymaker unilaterally chooses policy at both dates. What is the optimal policy choice? To answer this question, I will characterize first how policies affect production and investment decisions.

Given a tax τ_t, firms in polluting districts i choose the level of production y_{it} to maximize profits $(1-\tau_t)y_{it} - \frac{1}{2}y_{it}^2$. Thus, $y_{it} = 1-\tau_t$, and profits are $\frac{1}{2}(1-\tau_t)^2$. Similarly, firms in green districts i choose $y_{it} = A_i$, and their profit is $\frac{1}{2}A_i$.

Let $b_t = \int_0^1 \chi_{it}\, di$ be the share of polluting firms at time t. If a firm in district i decides to invest in green capital in the first period, they pay the cost of capital, $-c$, and receive the investment subsidy, s. Their discounted second-period profit is $\delta\frac{1}{2}A_i$. If the firm does not invest, its expected profit in the second period is $\delta\frac{1}{2}(1-\tau_2)^2$. Therefore, the firm invests in green capital if and only if $s - c + \frac{\delta A_i}{2} \geqslant \frac{\delta}{2}(1-\tau_2)^2$. Using the assumption that $A_i = Ai$, we obtain that the set of polluting firms that transition is $[b_2, b_1)$, where b_2 is given by

$$s - c + \frac{\delta A}{2}b_2 = \frac{\delta}{2}(1-\tau_2)^2.$$

Thus, the policy instruments have the following effects on the economy. Carbon taxes reduce emissions by reducing production in the polluting districts. The second-period carbon tax also induces investment in green capital, since, if correctly anticipated, it reduces the expected returns from using polluting capital. The subsidy s increases investment in the first period directly.

Proposition 0. *The optimal policy consists of Pigouvian carbon taxes and no subsidies. Moreover, it implements the optimal allocation.*

Proof. All proofs are in Appendix B. ■

There are two kinds of economic decisions, production and investment, and both have externalities—production creates carbon emissions, and investment reduces carbon emissions in the future. The carbon tax induces producers to internalize the first externality; thus the optimal tax is equal to the marginal environmental damage (this is the Pigouvian level). The subsidy can induce producers to internalize the second externality, viz, to give them incentives to invest in green capital and thus reduce future emissions. However, in equilibrium the policymaker does not use the subsidy, because once the first externality is corrected, the second one disappears. If the subsidy was used along with Pigouvian carbon taxes, it would lead to inefficient investments, i.e., investments in technologies that reduce emissions at an economic cost larger than the environmental cost of pollution.

Political Turnover. Suppose now that there is a probability $1-\rho$ that the green proposer is replaced with an opposition proposer between periods one and two. An opposition policymaker chooses zero taxes. Thus, from a first-period perspective, the expected future profit in the polluting sector is greater than if turnover was not possible, because with probability $1-\rho$ the future carbon tax is zero. As a result, fewer firms decide to transition for any first-period policy.

How should the green policymaker respond? As I show in Appendix B.2, the Pigouvian carbon tax is still optimal, but the optimal subsidy is now $s = \delta(1-\rho)\lambda$. The reason why the subsidy is required in this case is that political turnover reduces investment below the optimal level, and the subsidy is the appropriate instrument to correct this distortion. It is noteworthy and intuitive that the optimal subsidy increases with the probability of turnover and with the social cost of carbon.

The result that the equilibrium carbon tax is not affected by political turnover depends on the assumption that the environmental damage is linear. If the environmental damage is a strictly convex function of the stock of carbon emissions, the effect of an increase in the probability of turnover on the first-period carbon tax is ambiguous. There are two effects. First, given that the second-period carbon tax is likely to be repealed, emissions are likely to be larger than optimal; with a convex cost, this implies that the cost of additional first-period emissions is larger, which pushes first-period carbon taxes above the first-best level. However, a second effect is that turnover leads to the use of the subsidy, which increases investment, and hence brings emissions down, which decreases the cost of additional first-period emissions. See Appendix B.2 for details.

In sum, government turnover can explain the use of subsidies, but under this mechanism subsidies arise as a complement to carbon pricing, not as a substitute. I now incorporate political constraints into the model, and show that these lead both to positive subsidies and taxes below the first-best.

4. Legislative Bargaining

I study now the full model, incorporating the legislative bargaining game to the analysis. An initial observation is that if green districts form a majority in the first period then they are at least as willing to implement ambitious climate policy as the green policymaker, since they benefit from the fiscal revenue and do not suffer the economic costs of the policies. Therefore, in this case the green policymaker is effectively unconstrained, and we are back to the scenario in the previous section. From now on, I will focus on the interesting case in which polluting districts form a majority initially.

Assumption 1. The initial set of polluting districts are exactly the districts that do not have an incentive to transition if no climate policy in either period is expected, and are a majority. Formally, districts $i \in [0, b_1)$ are polluting in the first period, and $i \in [b_1, 1]$ are green, where b_1 is given by $-c + \frac{\delta A}{2} b_1 = \frac{\delta}{2}$ and satisfies $\frac{1}{2} < b_1 < 1$.

If polluting districts are still a majority in the second period, the legislature will block any proposal to raise carbon taxes. In the first period, the legislature only accepts a carbon tax if it is bundled with a subsidy that is generous enough to compensate for the cost imposed by the tax. A crucial observation is that in order for a district to benefit from the subsidy, it needs to invest in the clean technology, which turns them into a green district in the second period. Therefore, a legislative victory in the first period occurs only if there is a green majority in the second period. In that case, the future green majority accepts any increase in carbon taxes and blocks any proposal to lower them, because they do not suffer the economic costs but enjoy the fiscal benefits. Thus, if the green policymaker stays in power in the second period, they set the carbon tax at the Pigouvian level, and, if an opposition policymaker takes over, they keep the carbon tax at the level inherited from the first period.

Legislators anticipate that a first period subsidy can lead to higher carbon taxes in the future. They understand that even if the government runs a deficit in the present to pay for the subsidy, this may not lead to new taxes, because the future carbon tax is a source of revenue. Hence subsidies can be attractive even for green districts, who do not benefit directly from them. Polluting districts that do not plan to take the subsidy, on the other hand, realize that even if the first-period carbon tax is low or even zero, it can lead to a large tax in the future, and hence they oppose it.

To change the status quo policy, the policymaker needs to create a winning coalition that includes the green districts and enough polluting districts to form a majority. The pivotal polluting districts demand a subsidy that is large enough to compensate them for the losses from the carbon tax in the present and the costs from the investment in green capital. The green districts demand a subsidy that is not so large that exceeds the expected revenue from the carbon tax, because they are not willing to pay taxes themselves.

When casting their vote, legislators compare the policy that is proposed, in conjunction with the future policy they expect to follow, against the status quo plus the policy they expect to be enacted in the future if the status quo is preserved. There are two possibilities in the second case. If there is a polluting majority in the future, the future carbon tax is zero. However, if there is a green majority in the future and the green policymaker stays in power, they will enact a Pigouvian carbon tax. The first possibility is always an equilibrium of the subgame in which the proposal fails in the first period. I consider the second possibility in the next section. The following Lemma characterizes the two political constraints in the first case.

Lemma 1. *A carbon tax τ_1 and a subsidy s are accepted by the legislature if and only if*

$$\frac{1}{2}(1-\tau_1)^2 + s - c + \frac{\delta A}{4} + T \geqslant \frac{1}{2} + \frac{\delta}{2} \tag{PC_B}$$

and

$$T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2) \geqslant 0, \tag{PC_G}$$

hold, where $\tau_2 = \max\{\tau_1, \lambda\}$ with probability ρ and $\tau_2 = \tau_1$ with probability $1-\rho$, and b_2 is given by

$$s - c + \frac{\delta A}{2} b_2 = \frac{\delta}{2}\mathrm{E}\big[(1-\tau_2)^2\big]. \tag{1}$$

What do equilibrium policies look like? If the policymaker can satisfy the political constraints, the carbon tax is less than optimal, and it can be zero in equilibrium, while the subsidy is positive and can even be larger than the cost of green capital.

Proposition 1. *In equilibrium either the first-period carbon tax is below the social cost of carbon and the subsidy is positive, or no climate policy is enacted, i.e., $\tau_1 < \lambda$ and $s > 0$, or $\tau_1 = s = 0$.*

To understand this result, we can focus on the case in which there is no turnover for simplicity. Notice first that in equilibrium either the political constraint imposed by pivotal polluting districts, PC_B, binds, or the subsidy is zero. This is because, given the first-period tax, both the policymaker and the green districts prefer to reduce the subsidy, hence the policymaker chooses the minimum subsidy that is politically acceptable. Polluting districts are not willing to accept a carbon tax with no subsidies. Thus, if the policymaker decides to impose a carbon tax, the higher the tax, the greater the subsidy needs to be in order to compensate the pivotal polluting districts. There is a tradeoff: a larger carbon tax brings it closer to the Pigouvian level, but increases the size of the subsidy, whose efficient level is zero. The equilibrium carbon tax must be below the efficient level, because, starting from the Pigouvian level, a small decrease in the tax has a negative effect on the objective of the policymaker that is of second order, but makes possible a reduction in the subsidy that has a first order positive effect. For the same reason, the subsidy must be positive, because an increase from 0 is second order, but makes possible an increase in the carbon tax that has a positive first order effect.

This result can explain the phenomenon of policy sequences: the green policymaker initially obtains a partial political victory by enacting a low carbon tax and an inefficiently large subsidy, and this leads over time to efficient policies. The initial policies are designed to expand the green coalition by inducing pivotal polluting districts to transition. Germany's experience with the Renewable Energy Sources Act (EEG) of 2000 offers an illustration of this mechanism. The EEG, enacted by the Social Democrat–Green coalition, provided generous feed-in tariffs for renewable energy, which effectively acted as an investment subsidy. While these tariffs were inefficient as policies designed to reduce carbon emissions (Marcantonini and Ellerman 2015), they served to rapidly expand the renewable energy sector. This expansion created a growing constituency of firms, workers, and communities with vested interests in green policies. Over time, as the renewable energy industry matured and costs decreased, Germany was able to gradually reduce the subsidies and implement more market-based mechanisms: the 2017 revision of the EEG introduced competitive auctions for most renewable energy sources. Furthermore, despite initially being reluctant to participate in the EU Emissions Trading System (Ellerman *et al.* 2010), Germany implemented a carbon price for the heating and transport sectors in 2021 (IEA). This progression exemplifies how an initially inefficient policy paved the way for more comprehensive and efficient climate measures by altering the balance of power between green and polluting economic interests over time.

Proposition 1, however, raises the following question: under what conditions does the policy sequence start? Two key conditions must be met. First, there must be a policy package that is acceptable to both the green districts and the pivotal polluting districts. In other words, the two political constraints must be feasible. These constraints are in conflict: polluting districts close to the green frontier demand low carbon taxes and large subsidies, while green districts advocate for high carbon taxes and low subsidies. Second, there must exist a feasible policy that the policymaker prefers over the status quo. The policymaker, while concerned about environmental damages, also prioritizes aggregate economic performance. Political feasibility introduces distortions that conflict with the objective of improving economic outcomes. Consequently, for the policy sequence to initiate, a policy bundle must not only be politically feasible but also pass a cost-benefit analysis: the expected environmental benefits must outweigh the aggregate losses in consumption.

Political Constraints. There are two main forces that determine the stringency of the political constraints and, consequently, their feasibility and the extent to which they compel the proposer to deviate from optimal policies. On the one hand, pivotal polluting districts are more willing to accept climate policies if economic opportunities in the green sector improve, as accepting the subsidy requires them to transition. On the other hand, both polluting and green districts benefit from a larger fiscal balance. An improvement in the economic benefits of decarbonization implies that a larger share of the economy will have already transitioned without policy incentives. This, in turn, implies that the revenue from the carbon tax will be reduced, which shrinks the fiscal space for subsidies and transfers. Thus, lower costs of transition tighten the political constraint imposed by green districts and indirectly diminish the willingness of polluting districts to accept climate policy.

To formalize this argument, let us say that an intervention *relaxes* a political constraint if it expands the set of feasible policies that satisfy the constraint; the intervention *tightens* the constraint if the opposite happens. We say that the constraints are *feasible* if there exists a policy package that satisfies both constraints simultaneously. We have the following result.

Observation 1. *If economic agents are sufficiently patient, i.e.,* $\delta > \frac{1}{A-\frac{1}{2}}$,[7] *then an increase in the productivity of green technology* A *or a decrease in the cost of capital* c *relax the political constraint imposed by polluting districts* (PC_B) *and tighten the political constraint imposed by green districts* (PC_G). *If the constraints are feasible and* A *increases or* c *decreases, then the constraints continue to be feasible.*

A general improvement in green technology across all districts (represented by an increase in A) and a reduction in the cost of capital (represented by a decrease in c) enhance the economic opportunities associated with decarbonization. The first force described above relaxes the constraint imposed by pivotal polluting districts, but the second force tightens the constraint imposed by green districts. The second part of the result shows that the overall impact on feasibility is positive. This improvement in feasibility occurs despite the tightening of the green districts' constraint, implying that that the relaxation of the polluting districts' constraint is the dominant effect. As a result, we obtain that a political compromise between the two types of districts exists if and only if green technology is sufficiently advanced and the cost of capital is low.

An increase in the expected second-period carbon tax affects the constraints only through the second force: it increases the share of polluting districts that decide to transition, which impacts the fiscal balance but does not alter the benefits of transitioning for the pivotal polluting districts. Hence, it influences the willingness of the two types of districts to accept climate policy in the same direction. However, the direction of the effect is ambiguous. On the one hand, an increase in the expected carbon tax increases the expected fiscal balance mechanically. On the other hand, it induces more districts to transition, which reduces the fiscal surplus by shrinking the carbon tax base and increasing the cost of the subsidy.

[7]The condition $\delta > \frac{1}{A-\frac{1}{2}}$ is enough to provide a very simple proof of the result, but it is far from necessary.

The social cost of carbon λ and the probability that the green proposer stays in power ρ affect the constraints through their effect on the expected second-period carbon tax. Therefore, their effect is ambiguous. In particular, an increase in the social cost of carbon can tighten the political constraints. A greater λ indicates that the proposer is more committed to climate change mitigation. A naive intuition might suggest that a more committed proposer should be more likely to implement climate policy. However, this intuition overlooks the fact that a highly concerned proposer cannot credibly commit to not raising carbon taxes when political constraints eventually relax. This, in turn, implies that the effectiveness of subsidies as a "carrot" required to make progress when the political constraints still bind is reduced, because their fiscal cost increases with the expected future carbon tax. In fact, unless A is sufficiently large that to compensates for this effect, increasing λ beyond a certain point not only tightens the constraints but makes them impossible to satisfy jointly. This implies that, surprisingly, a highly committed proposer may find it impossible to make any climate policy progress, while a less committed policymaker may be able to start the policy transition. Formally, we have the following result.

Observation 2. *If the productivity of green technology is not too large, i.e., $A < \underline{A}$, where $\underline{A}(b_1, \rho)$ is an increasing function of b_1 and ρ, then the political constraints are feasible only if the social cost of carbon is not too large, i.e., $\lambda < \overline{\lambda}$, where $\overline{\lambda}(A, c, \delta, \rho) < \frac{1}{2}$.*

The ambiguous effect of the expected second-period carbon tax also has implications for political turnover. An increase in the probability of turnover $(1 - \rho)$ reduces the expected carbon tax, and consequently has an ambiguous effect on the willingness of legislators to accept climate policy. In fact, if the political opposition to climate policies is initially high, an increase in the probability of turnover can facilitate to initiate a policy sequence. In this case the possibility of turnover can paradoxically help a green policymaker start a policy sequence that leads to an efficient carbon tax in the future. This is despite the fact that turnover interrupts the sequence, and policy uncertainty reduces green investments. Formally, we have the following result.

Observation 3. *If the political opposition to climate policies is initially high, i.e., $b_1 > \frac{1-\frac{3}{4}\lambda}{1-\frac{1}{2}\lambda}$,[8] and the political constraints are feasible, then, if the probability that the second-period proposer is polluting increases, the constraints continue to be feasible.*

Proposer's Incentives. The fact that the political constraints are feasible, and there is a policy (τ_1, s) that the legislature is willing to approve, does not imply that the policymaker will choose the best feasible policy, because the distortions may be so large that inaction is preferable. In other words, the best politically feasible climate policy may not pass a cost-benefit analysis.

The proposer aims to maximize aggregate consumption while reducing carbon emissions, with the tradeoff determined by the social cost of carbon, λ. The model's primitives affect these two

[8] Again, this condition is enough to provide a very simple proof of the result, but it is far from necessary.

concerns in opposing ways. If the proposer cares more about the environment and the expected second-period policy is closer to its optimal level, the proposer becomes more willing to compromise in the first period. However, in this scenario, private actors' incentives to invest rather than consume increase, making the proposer more hesitant to implement a compromise policy.

In particular, the social cost of carbon λ has two opposing effects on the proposer's incentives to implement a compromise policy. First, a greater λ implies that the proposer values reductions in carbon emissions more highly. Since both the carbon tax and the investment subsidy reduce emissions, the proposer becomes more willing to implement them. However, a greater λ also implies that the the expected second-period carbon tax is higher, leading more districts to transition. Thus, if the legislature demands a subsidy so high that the level of investment exceeds the optimal level, increasing λ exacerbates the distortion and reduces the proposer's willingness to implement a compromise policy package. Unless the legislature demands an very large subsidy, the first effect dominates, and a more committed proposer will be more willing to enact a politically feasible policy. Conversely, if λ is low, the proposer will decide to maintain the status quo, even if there is a policy package acceptable to the legislature.

To formally state this result, let $\Delta W(\tau_1, s) = W_G(\tau_1, s) - W_G^{\text{BAU}}$, so that the green policymaker prefers (τ_1, s) over BAU if and only if $\Delta W(\tau_1, s) \geqslant 0$.

Observation 4. *There is a policy that satisfies the political constraints (PC_B and PC_G) and the proposer prefers over the status quo (i.e., $\Delta W \geqslant 0$) only if the social cost of carbon is not too low, i.e., $\lambda \geqslant \underline{\lambda}$, where $\underline{\lambda}(A, c, \delta, \rho) > 0$.*

An increase in ρ, the probability that the proposer stays in power in the second period, also has two opposing effects. First, if it is more likely that the climate-concerned proposer retains their agenda-setting power, the carbon tax in the second period is more likely to be set at the optimal level, increasing the willingness to make compromises in the first period. However, if the level of investment is excessive due to a subsidy above the optimal level, an increase in ρ raises the expected carbon tax in the second period, which in turn increases investment and further exacerbates the distortion. Interestingly, this second effect can dominate.

A key theoretical implication of this analysis is that the effect of political turnover differs significantly when political constraints are incorporated into the model. In Section 3, the possibility of turnover reduced the level of investment below the optimal level, requiring the policymaker to use subsidies to restore investment to its optimal level. However, the picture fundamentally changes with political constraints, as they force the policymaker to use subsidies even without the possibility of turnover. Thus, the policy resulting from political compromise leads to an inefficiently high level of investment. Introducing turnover attenuates this distortion and partially aligns the proposer's objective with that of the pivotal polluting districts, since the proposer now sees subsidies as a tool to mitigate the negative environmental effects of a future opposition policymaker.

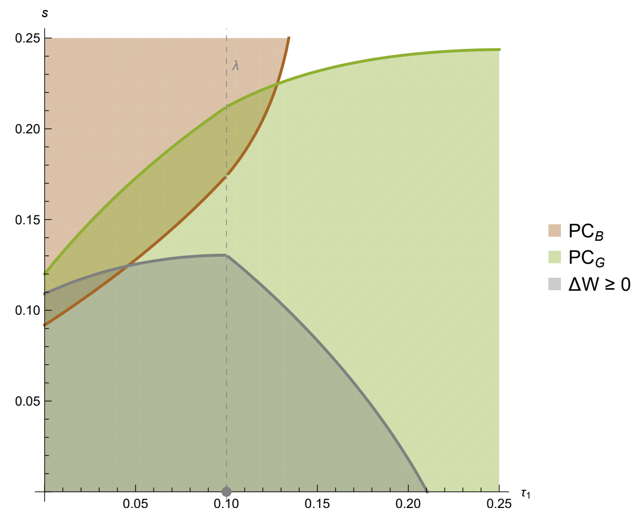


Figure 1: Political constraints when $A = 1.8$, $c = 0.06$, $\delta = 0.9$, $\lambda = 0.1$ and $\rho = 1$. The region PC_B is the set of policies that polluting districts prefer over the status quo (generous subsidies s and low carbon taxes τ_1). The region PC_G is the set of policies that polluting districts prefer over the status quo (low subsidies s and large carbon taxes τ_1). The region $\Delta W \geqslant 0$ is the set of policies that the green proposer prefers over the status quo. The gray dot is the ideal policy of the proposer, which is Pigouvian taxes $\tau_1 = \lambda$, the vertical dashed line, and no subsidies, $s = 0$.

Numerical Example. Figure 1 illustrates the political constraints. The pivotal polluting districts benefit from the subsidy and prefer low carbon taxes. Their ideal subsidy level is finite because a larger subsidy induces more districts to accept it, which increases its fiscal cost and reduces fiscal revenue; at some point this effect dominates the direct benefit of the subsidy. Green districts prefer the subsidy to be as small as possible and favor a large carbon tax due to the fiscal revenue it generates. The ideal level of the carbon tax is not confiscatory (the maximum of the Laffer curve is at $\tau_1 = \frac{1}{2}$), because, when $\tau_1 > \lambda$, the second-period carbon tax is expected to stay at τ_1; therefore, increasing τ_1 induces more districts to transition and reduces the fiscal surplus. The set of feasible policies is the intersection of the regions defined by PC_B and PC_G. The set of feasible policies that the proposer is willing to implement is the intersection of all three regions.

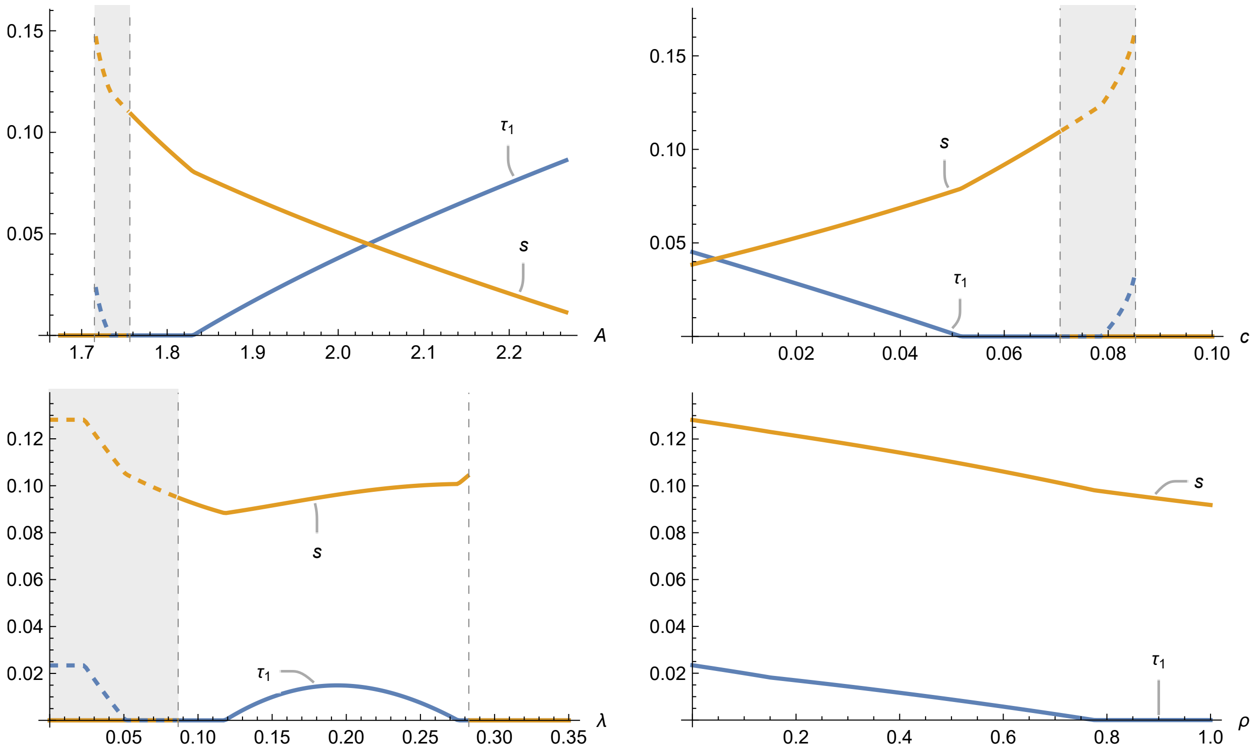


Figure 2: Equilibrium policy when $A = 1.8$, $c = 0.06$, $\delta = 0.9$, $\lambda = 0.1$ and $\rho = 1$, as each parameter changes. In the gray regions the proposer prefers the status quo ($\tau_1 = s = 0$) over the optimal feasible policy (the dashed curves).

Figure 2 shows how the equilibrium policy depends on parameters. Increasing the productivity of green technology A or decreasing the cost of capital c relaxes the political constraints, which enables the proposer to implement policy that is closer to the first-best: τ_1 increases, and s decreases. This holds as long as the green districts' political constraint does not bind. For low enough A, or large enough c, τ_1 is so small and s is so large that the fiscal constraint PC_G binds: the carbon tax is not enough to pay for the subsidies. This requires increasing taxes or decreasing spending in the future, which the green districts do not support. To compensate, the policymaker can increase τ_1, which requires an increase in s. This works until the constraints cannot be jointly satisfied. However, as we see in the Figure, the policy does no longer pass a cost-benefit analysis for the policymaker before PC_G binds.

The bottom-left panel of the Figure shows that, as discussed above, if the social cost of carbon λ is large then the political constraints cannot be satisfied jointly, and the equilibrium policy is laissez-faire. The Figure also shows that if the social cost of carbon is low then the political constraints are feasible, but the proposer is not willing to implement any policy that is acceptable to the legislature. As a result, the policy sequence starts in equilibrium if and only if λ is intermediate.

The bottom-right panel illustrates how the probability of turnover affects equilibrium policies.

In this example, more likely turnover (lower ρ) relaxes the political constraints, and leads to an increase in the carbon tax and the subsidy. Consistent with previous literature (Schmitt 2014; Harstad 2020), the probability of turnover increases the equilibrium subsidy, but in this model the carbon tax is significantly lower than the social cost of carbon, which is consistent with empirical observation.

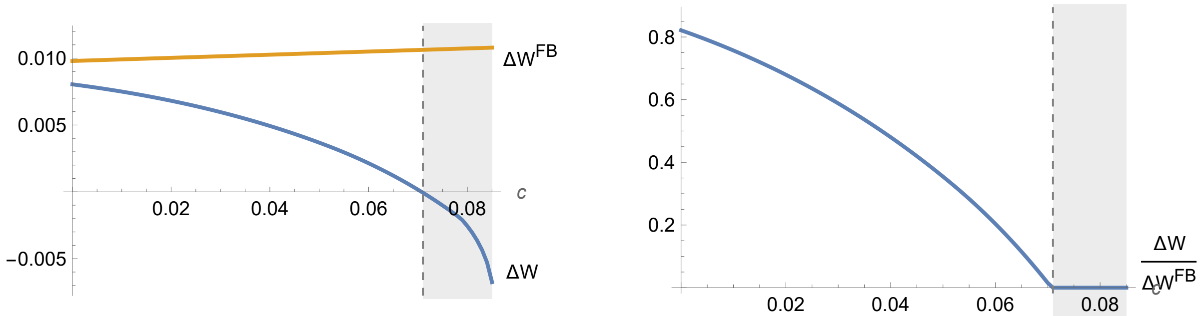


Figure 3: Increase in social welfare of the equilibrium policy (ΔW) and first-best policy (ΔW^{FB}) with respect to business-as-usual when $A = 1.8$, $\delta = 0.9$, $\lambda = 0.1$ and $\rho = 1$ as c changes. In the gray regions the optimal feasible policy is not better than BAU, and in equilibrium no policy is enacted.

Figure 3 shows how the objective of the policymaker compares to the first best, relative to business as usual. As c grows, and the political constraints become harder to satisfy, the improvement of equilibrium climate policy over the status quo BAU decreases, and diverges from the first best policy more, until it is no longer better than BAU. The welfare loss relative to the first best can vary enormously depending on parameters; for the parameters illustrated in the Figure, equilibrium policy goes from achieving more than 80% of the welfare gains produced by optimal policy if $c = 0$ to being no better than the status quo when $c \approx 0.07$.

Green Expectations and Soft Commitments. Under Assumption 1, the expectation of a second-period carbon tax induces some districts to transition even in the absence of an investment subsidy. This is because a future carbon tax decreases the value of the polluting capital in the future, and convinces a district that is indifferent between upgrading its capital or not to do it. In the previous section, I studied the equilibrium in which, absent policy in the first period, polluting districts expect no climate policy in the second period, which leads them to stay polluting. In this section, I consider an alternative equilibrium, the *green expectations* equilibrium, in which polluting districts expect a carbon tax in the second period, which leads some of them to decarbonize in the first period. This is possible only if the share of districts that transition is enough to create a green majority in the second period that allows the green policymaker to implement the carbon tax. In that case, the expectation of climate policy in the first period is confirmed, which makes it an equilibrium.

In terms of the model, this is an equilibrium if the second-period polluting districts do not form a majority, i.e., if $b_2 \leqslant \frac{1}{2}$, where the marginal district that transitions, b_2, is such that $-c + \frac{\delta A}{2} b_2 = \frac{\delta}{2}\left[\rho(1-\lambda)^2 + 1 - \rho\right]$. Using Assumption 1, this is equivalent to

$$b_1 \leqslant \frac{1}{2} + \frac{\rho\lambda(2-\lambda)}{A}. \tag{2}$$

In words, for the green expectations equilibrium to exist, the initial share of polluting districts cannot be too large. A greater Pigouvian carbon tax λ and a smaller probability of turnover $1-\rho$ increase the upper bound, because they lead more districts to transition if they expect a carbon tax in the future.

The fact that polluting districts expect a future carbon tax in the future absent any policy in the first period changes the political calculation for the green policymaker, because resisting a climate policy bundle (τ_1, s) is less attractive in this equilibrium than in the equilibrium considered in the previous section. The new political constraints are as follows:

$$\frac{1}{2}(1-\tau_1)^2 + s - c + \frac{\delta A}{4} + T \geqslant \frac{1}{2} \underbrace{- c + \frac{\delta A}{4}}_{\substack{\text{transition} \\ \text{without subsidy}}} + \tilde{T} \tag{PC$'_B$}$$

and

$$T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2) \geqslant \tilde{T} = \delta \tilde{b}_2 \rho\lambda(1-\lambda), \tag{PC$'_G$}$$

where $\tau_2 = \max\{\tau_1, \lambda\}$ with probability ρ and $\tau_2 = \tau_1$ with probability $1-\rho$, and $\tilde{b}_2$ is given by $-c + \frac{\delta A}{2}\tilde{b}_2 = \frac{\delta}{2}\left[\rho(1-\lambda)^2 + 1 - \rho\right]$.

We have the following.

Proposition 2. *In the green expectations equilibrium if $\lambda < \frac{1}{4}$ we have $\tau_1 < \lambda$ and $s > 0$.*

An active climate policy is always feasible in this equilibrium as long as λ is not too large, but it is still distorted relative to the first best. The reason is that the policymaker faces the same trade-off: a positive first-period carbon tax requires a positive subsidy if it is feasible. If the first-period carbon tax is set at the optimal level, the politically required subsidy induces too much investment, which creates an incentive to reduce the carbon tax. Thus, green expectations help implement climate policy, but they do not eliminate the policy distortion.

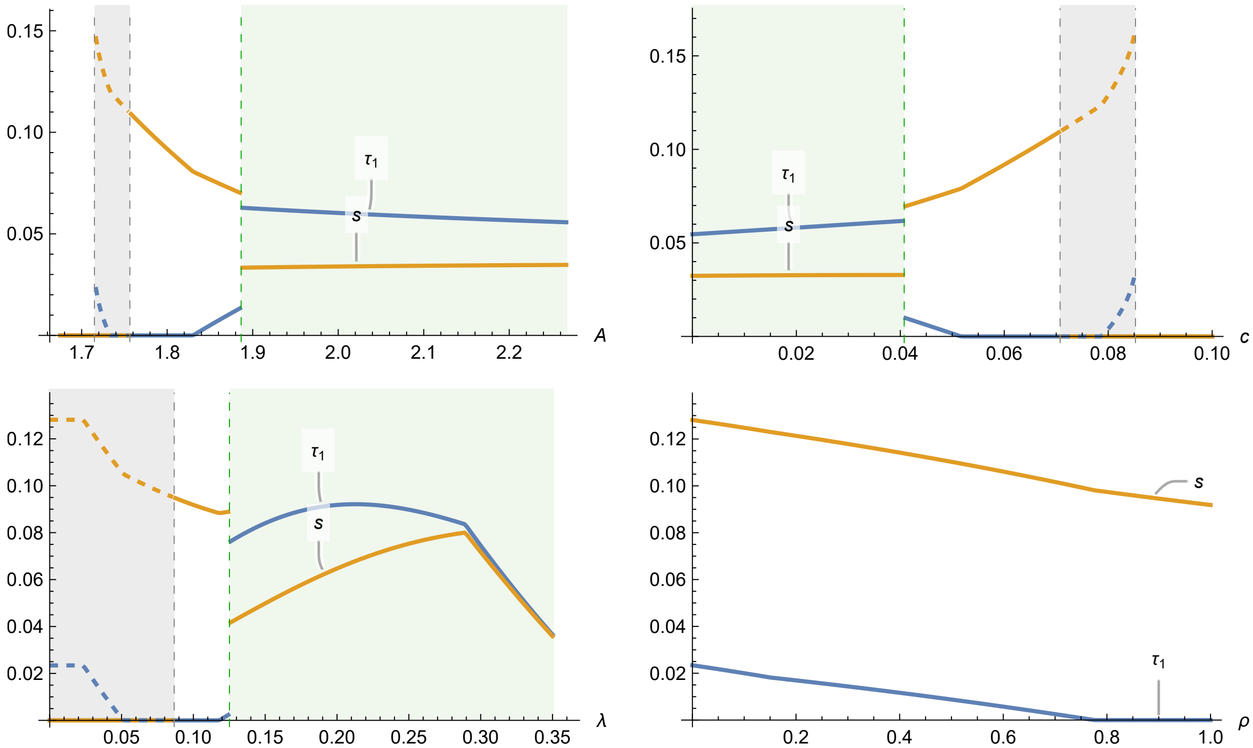


Figure 4: Equilibrium policy when $A = 1.8$, $c = 0.06$, $\delta = 0.9$, $\lambda = 0.1$ and $\rho = 1$, as each parameter changes. In the gray regions the proposer prefers the status quo over the optimal feasible policy (the dashed lines). In the green regions the green equilibrium exists and is displayed.

Figure 4 shows how green expectations affect equilibrium policy. When they are possible (for large A, small c and large λ), demarcated by the green regions, the first-period carbon tax is increased, and the subsidy decreased, relative to the baseline equilibrium.

The existence of the green expectations equilibrium suggests that *soft commitments* –non-binding declarations by governments that might otherwise be dismissed as mere "cheap talk"– can exert real effects on economic behavior and policy outcomes if they shift expectations of economic actors towards this equilibrium. Examples of such soft commitments include widespread net-zero targets and the Nationally Determined Contributions (NDCs) under the Paris Agreement. Stiglitz (2019) articulates this logic in the context of the latter: "[P]art of the rationale for the Paris strategy [was that] if enough firms believed that there was enough global commitment to climate change that there would be a high carbon price (implicit or explicit) going forward, they would have an incentive to make green investments; and to ensure that they were advantaged over firms that didn't make such investments and to ensure that they obtained the desired returns on those investments, they would then politically support, in coalition with other like-minded agents, a high carbon price." Ramadorai and Zeni (2024) show that, consistent with this idea, the announcement of the Paris Agreement led to a significant change in beliefs about future carbon taxation in a sample of North

American public firms.

The analysis of the model, however, reveals that soft commitments do not work in all circumstances. Carbon-abating technology needs to be sufficiently advanced, the social cost of carbon perceived to be sufficiently large, and polarization around climate policy (captured by the probability of turnover) sufficiently low. Equation (2) makes the conditions precise. The fact that many governments undershot their emissions targets (de Silva and Tenreyro 2021) suggests that policymakers may not always be able to trigger a change in expectations that paves the ground for ambitious policies.

5. Extensions

Targeted Subsidies and Transfers. In Appendix A.1 I consider the implications of allowing the policymaker to target transfers and subsidies, which so far I assumed to be uniform. I first argue that arbitrary targeting is implausible because of an information asymmetry: an optimal targeting strategy would require knowing exactly the identity of the minimal set of polluting districts that are closest to the decarbonization frontier and form a winning coalition. Which constituencies are willing to make investments in the transition given the right incentives is a difficult empirical question, as we can infer from historical cases. Moreover, to the extent that we observe targeting (for example, the allocation of free allowances in the initial phases of the EU ETS), it does not respond to an optimal coalition-building strategy, but can be explained as a result of lobbying by particularly powerful industries (Winkler 2022).

To take into account the information asymmetry I employ a mechanism design approach. I assume that legislators know the potential productivity of green capital investments in their district, but can choose to withhold that information. The policymaker proposes a menu of targeted taxes, subsidies and transfers. The legislature then votes on the entire menu, after which each district selects its preferred bundle. I find that the proposer can target subsidies and transfers to some extent, but the incentive compatibility constraints protect the polluting districts excluded from the winning coalition, and prevent the policymaker from restricting the subsidy to only the pivotal districts. Consequently, the same fundamental distortion that arises with uniform subsidies and transfers emerges with this richer set of policy instruments.

Green Preferences. In Appendix A.2 I assume legislators intrinsically care about the environment. In this case green districts no longer pose a constraint on the proposer, and the constraint imposed by pivotal districts is relaxed. Polluting districts prefer zero carbon taxes and, unless the social cost of carbon is very large, in equilibrium the proposer starts implementing a carbon tax that is lower than Pigouvian plus a positive investment subsidy, and, once the green capital is sunk, raises the carbon tax to the optimal level. In sum, even though green preferences relax political constraints, the equilibrium features the same distortions and dynamics. As in the baseline model, improvements in technology, and reductions in capital costs or the discount rate, make it more likely that the policy sequence is started.

Production Subsidies, Standards and FITs. I considered so far investment subsidies as the inefficient policy instrument that the policymaker can use to build a winning coalition, but in practice other instruments are used for this purpose. In Appendix A.3 I show how a production subsidy in the green sector can be used with the purpose of inducing polluting districts to transition. In Appendix A.4 I introduce an emissions standard, and show that it is equivalent to a revenue neutral combination of a production subsidy and a carbon tax. I also show that a feed-in tariff for goods produced using green capital is equivalent to a tradable emissions standard. Their equivalence arises because there is no volatility in output. A policymaker that can propose a production subsidy and a carbon tax would not choose to propose standards, because they can be replicated and, in general, improved upon. However, if a production subsidy was not available, the policymaker can (under some conditions) use the standard to start the policy sequence, because it implicitly subsidizes production in the green sector.

Imperfect Competition and Pass-through. The main model ignores market imperfections to focus exclusively on the climate externality. As a result, the incidence of the carbon tax falls solely on districts hosting polluting production. In Appendix A.5, I adapt the model from Besley and Persson (2023), which features monopolistic competition and thus allows for pass-through of carbon taxes to consumers. This extension shows that carbon taxes adversely affect all consumers—not only those in polluting districts as in the main model. Nevertheless, green districts still benefit on net after tax rebates, as long as the carbon price is not too large. While incorporating legislative bargaining into the Besley-Persson framework becomes analytically intractable, I demonstrate in the Appendix that the same political inefficiency I uncover in the main model emerges: polluting districts induce policymakers to use subsidies, and create incentives to set first-period carbon taxes below the optimal level.

Endogenous Reelection Probability. In Appendix A.6 I consider the implications of making the probability that the green proposer is reelected endogenous. Because the intensity of policy preferences matters in probabilistic voting models, the probability of reelection depends on the difference between the average expected payoff under each party, where the average is taken over the voting population. Thus, even though, once the policy sequence starts, a majority is in favor of increasing the stringency of climate policy in the second period, increasing the salience (i.e., the electoral implications) of climate policy decreases the reelection chances of the green party. This is because voters do not internalize the climate externality in the baseline model, and, therefore, the average voter is against carbon taxes. Counterintuitively, increasing the first-period carbon tax does not lead to a backlash against the green party in the context of a policy sequence, because in that context voters understand that the carbon tax has been locked in, and therefore the anti-climate party cannot credibly promise to eliminate it. In fact, increasing the first-period carbon tax helps the green incumbent get reelected because it reduces the difference in the second-period policies that they and the polluting party implement, since the polluting party is forced to keep the first-period carbon tax in place. Increasing the size of the green subsidy also helps the green policymaker, because it reduces the share of the population affected by the carbon tax in the second period. In Appendix A.6 I show that, if a policy sequence can be started, increasing the salience of climate policy leads to an increase in the levels of the carbon tax and the green subsidy implemented in the first period. This result, combined with Observation 3, implies that increasing the salience of climate policy can make it more likely that it is implemented, and can increase its ambition, even though it also decreases the electoral prospects of green parties.

6. Conclusion

Politics imposes constraints on climate policy that are dynamic and local. These constraints are dynamic because policy instruments transform current opponents into future supporters by shaping incentives for investments in decarbonization, thus linking policies intertemporally. The constraints are local because winning coalitions comprise interests that are tied to specific constituencies. Modeling climate policymaking in this way contrasts with existing approaches that take the power of carbon emitters as exogenous and unchanging, or only focus on political turnover as a source of political distortions. Thus, this paper aligns with the unifying framework for studying energy transitions proposed by Gazmararian and Tingley (2024), which is centered around credibility, both in terms of expectations of future policy and anticipated future economic welfare in local communities.

The model I developed in this paper can explain both the increase in stringency of climate policy over time and the shift from inefficient to efficient instruments. Several additional insights emerge from the analysis. First, if initial conditions are sufficiently adverse for the energy transition (e.g., the economy relies heavily on fossil-fuel production, or emissions abatement technology is not sufficiently advanced), then no climate policy is implemented in political equilibrium. This finding helps to understand why policy feedback has failed to take hold in several empirical cases. Second, even if there is a policy bundle that is time-consistent (i.e., credible) and acceptable to a winning coalition of interests, its social costs may be so large that a climate-concerned policymaker will decide not to implement it. This result suggests that a fruitful direction for future research is to estimate empirically the costs that political acceptability adds to climate policies. Recently, researchers have made important methodological advances in the estimation of the costs and benefits of climate policies (Hahn *et al.* 2024), as well as in the design of policies that achieve acceptability by key political actors (Gazmararian and Tingley 2023); future research should combine those efforts in order to determine the cost of political acceptability.

Third, the analysis reveals that managing economic actors' expectations of future policy is important, and can be achieved both by policy design and by soft commitments. The latter can influence beliefs because future policies are affected by investments that economic actors make in the present. This feature creates a coordination game that can have multiple equilibria. Equilibrium policy, in turn, responds to these beliefs—expectations of stringent policy in the future strengthen the bargaining power of policymakers to implement more stringent policies in the present. Thus, voluntary commitments such as the nationally determined contributions made in the context of the Paris Agreement can complement policy enactment domestically.

The analysis of the policy sequencing strategy in this paper shows that it is not always feasible, that it can be interrupted by political turnover, and that it can produce large welfare losses relative to the first best. These conclusions follow even under the favorable assumption that the political effects of green industrial policies arrive before political turnover. If an election takes place before there are enough investments in decarbonization to tip the balance of power between polluters and green constituencies, and an anti-climate policymaker is elected, we can expect the logic of the sequencing argument to work in reverse: a policymaker opposed to climate policy has dynamic incentives to repeal green subsidies and stall the deployment of green technologies to prevent further growth of the green coalition, even if this is inefficient or entails short-term political costs. The Trump Administration's 2025 rollback of key subsidies in the Inflation Reduction Act of 2022, despite opposition from Republican lawmakers and business lobbies, is consistent with this mechanism.[9] Studying the politics of climate-policy retrenchment is fertile ground for future work.

References

Abbott, A. and Jones, P. (2023). The cyclicality of government environmental expenditure: Political pressure in economic upturns and in recessions. *Journal of Environmental Economics and Policy*, **12** (2), 209–228.

Acemoglu, D., Akcigit, U., Hanley, D. and Kerr, W. (2016). Transition to Clean Technology. *Journal of Political Economy*, **124** (1), 52–104.

Acharya, V. V., Engle, R. and Wang, O. (2024). Strategic Commitments to Decarbonize: The Role of Large Firms, Common Ownership, and Governments.

Aidt, T. S. (1998). Political internalization of economic externalities and environmental policy. *Journal of Public Economics*, **69** (1), 1–16.

[9] In March 2025, 21 House Republicans urged Ways & Means Chair Jason Smith to spare the IRA's production and manufacturing credits (Garbarino *et al.* 2025). Their stance dovetailed with lobbying by major clean-energy trade groups and some fossil-fuel and GOP allies (see, e.g., American Clean Power Association 2025; Fuel Cell & Hydrogen Energy Association *et al.* 2025).

— (2010). Green taxes: Refunding rules and lobbying. *Journal of Environmental Economics and Management*, **60** (1), 31–43.

— and Dutta, J. (2004). Transitional politics: Emerging incentive-based instruments in environmental regulation. *Journal of Environmental Economics and Management*, **47** (3), 458–479.

Aklin, M. and Urpelainen, J. (2013). Political Competition, Path Dependence, and the Strategy of Sustainable Energy Transitions. *American Journal of Political Science*, **57** (3), 643–658.

Alberdi, M. (2024). Building Eco-Wealth: How Climate Subsidies Shape Support for the Greens.

Aldy, J. E., Gerarden, T. D. and Sweeney, R. L. (2023). Investment versus Output Subsidies: Implications of Alternative Incentives for Wind Energy. *Journal of the Association of Environmental and Resource Economists*, **10** (4), 981–1018.

Alesina, A. and Tabellini, G. (1990). A Positive Theory of Fiscal Deficits and Government Debt. *The Review of Economic Studies*, **57** (3), 403–414.

Allan, B. B. and Nahm, J. (2024). Strategies of Green Industrial Policy: How States Position Firms in Global Supply Chains. *American Political Science Review*, pp. 1–15.

Allcott, H., Mullainathan, S. and Taubinsky, D. (2014). Energy policy with externalities and internalities. *Journal of Public Economics*, **112**, 72–88.

American Clean Power Association (2025). ACP Statement on Reconciliation Bill Text. Press release. Available at https://cleanpower.org/news/acp-statement-on-reconciliation-bill-text/.

Austen-Smith, D., Dziuda, W., Harstad, B. and Loeper, A. (2019). Gridlock and inefficient policy instruments. *Theoretical Economics*, **14** (4), 1483–1534.

Baldursson, F. M. and von der Fehr, N.-H. M. (2007). A Whiter Shade of Pale: On the Political Economy of Regulatory Instruments. *Journal of the European Economic Association*, **5** (1), 37–65.

Baldwin, R. E. and Robert-Nicoud, F. (2007). Entry and Asymmetric Lobbying: Why Governments Pick Losers. *Journal of the European Economic Association*, **5** (5), 1064–1093.

Barrage, L. (2020). Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy. *The Review of Economic Studies*, **87** (1), 1–39.

Battaglini, M. and Harstad, B. (2016). Participation and Duration of Environmental Agreements. *Journal of Political Economy*, **124** (1), 160–204.

— and — (2020). The Political Economy of Weak Treaties. *Journal of Political Economy*, **128** (2), 544–590.

Behmer, S. (2023). Carrots vs Sticks: Climate Policy with Government Turnover.

Besley, T. and Coate, S. (1998). Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis. *The American Economic Review*, **88** (1), 139–156.

— and Persson, T. (2023). The Political Economics of Green Transitions. *The Quarterly Journal of Economics*, **138** (3), 1863–1906.

Biais, B. and Landier, A. (2022). Emission Caps and Investment in Green Technologies.

Bistline, J., Mehrotra, N. and Wolfram, C. (2023). Economic Implications of the Climate Provisions of the Inflation Reduction Act.

Blanchard, O., Gollier, C. and Tirole, J. (2023). The Portfolio of Economic Policies Needed to Fight Climate Change. *Annual Review of Economics*, **15** (1), 689–722.

Bolet, D., Green, F. and González-Eguino, M. (2023). How to Get Coal Country to Vote for Climate Policy: The Effect of a "Just Transition Agreement" on Spanish Election Results. *American Political Science Review*, pp. 1–16.

Bovenberg, A. L., Goulder, L. H. and Gurney, D. J. (2005). Efficiency Costs of Meeting Industry-Distributional Constraints Under Environmental Permits and Taxes. *The RAND Journal of Economics*, **36** (4), 951–971.

—, — and Jacobsen, M. R. (2008). Costs of alternative environmental policy instruments in the presence of industry compensation requirements. *Journal of Public Economics*, **92** (5), 1236–1253.

Breetz, H., Mildenberger, M. and Stokes, L. (2018). The political logics of clean energy transitions. *Business and Politics*, **20** (4), 492–522.

Brulle, R. J. (2018). The climate lobby: A sectoral analysis of lobbying spending on climate change in the USA, 2000 to 2016. *Climatic Change*, **149** (3), 289–303.

Buchanan, J. M. and Tullock, G. (1975). Polluters' Profits and Political Response: Direct Controls versus Taxes. *The American Economic Review*, **65** (1), 139–147.

Buisseret, P. and Bernhardt, D. (2017). Dynamics of Policymaking: Stepping Back to Leap Forward, Stepping Forward to Keep Back. *American Journal of Political Science*, **61** (4), 820–835.

Cadoret, I. and Padovano, F. (2016). The political drivers of renewable energies policies. *Energy Economics*, **56**, 261–269.

Clausing, K. A. and Wolfram, C. (2023). Carbon Border Adjustments, Climate Clubs, and Subsidy Races When Climate Policies Vary. *Journal of Economic Perspectives*, **37** (3), 137–162.

Coate, S. and Morris, S. (1995). On the Form of Transfers to Special Interests. *Journal of Political Economy*, **103** (6), 1210–1235.

— and — (1999). Policy Persistence. *American Economic Review*, **89** (5), 1327–1336.

Colmer, J., Martin, R., Muûls, M. and Wagner, U. J. (2024). Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence from the European Union Emissions Trading System. *The Review of Economic Studies*.

Cullenward, D. and Victor, D. G. (2020). *Making Climate Policy Work*. John Wiley & Sons.

Damania, R. (2001). When the Weak Win: The Role of Investment in Environmental Lobbying. *Journal of Environmental Economics and Management*, **42** (1), 1–22.

— and Fredriksson, P. G. (2003). Trade policy reform, endogenous lobby group formation, and environmental policy. *Journal of Economic Behavior & Organization*, **52** (1), 47–69.

Davis, L. W., Hausman, C. and Rose, N. L. (2023). Transmission Impossible? Prospects for Decarbonizing the US Grid. *Journal of Economic Perspectives*, **37** (4), 155–180.

de Silva, T. and Tenreyro, S. (2021). Climate-Change Pledges, Actions, and Outcomes. *Journal of the European Economic Association*, **19** (6), 2958–2991.

Dechezleprêtre, A. and Popp, D. (2017). Fiscal and Regulatory Instruments for Clean Technology Development in the European Union. In I. Parry, K. Pittel and H. Vollebergh (eds.), *Energy Tax and Regulatory Policy in Europe*, The MIT Press, pp. 167–214.

Dolphin, G., Pollitt, M. G. and Newbery, D. M. (2020). The political economy of carbon pricing: A panel analysis. *Oxford Economic Papers*, **72** (2), 472–500.

Dziuda, W. and Loeper, A. (2018). Dynamic Pivotal Politics. *American Political Science Review*, **112** (3), 580–601.

Egorov, G. and Harstad, B. (2017). Private Politics and Public Regulation. *The Review of Economic Studies*, **84** (4), 1652–1682.

Ellerman, A. D., Buchner, B. and Carraro, C. (eds.) (2007). *Allocation in the European Emissions Trading Scheme: Rights, Rents and Fairness*. Cambridge: Cambridge University Press.

—, Convery, F. J., de Perthuis, C. and Alberola, E. (eds.) (2010). *Pricing Carbon: The European Union Emissions Trading Scheme*. Cambridge: Cambridge Univ. Press, repr edn.

Fankhauser, S., Gennaioli, C. and Collins, M. (2015). The political economy of passing climate change legislation: Evidence from a survey. *Global Environmental Change*, **35**, 52–61.

Fischer, C., Hübler, M. and Schenker, O. (2021). More birds than stones – A framework for second-best energy and climate policy adjustments. *Journal of Public Economics*, **203**, 104515.

Fredriksson, P. G. (1997). The Political Economy of Pollution Taxes in a Small Open Economy. *Journal of Environmental Economics and Management*, **33** (1), 44–58.

— and Sterner, T. (2005). The political economy of refunded emissions payment programs. *Economics Letters*, **87** (1), 113–119.

— and Svensson, J. (2003). Political instability, corruption and policy formation: The case of environmental policy. *Journal of Public Economics*, **87** (7-8), 1383–1405.

— and Wollscheid, J. R. (2008). The political economy of investment: The case of pollution control technology. *European Journal of Political Economy*, **24** (1), 53–72.

Fuel Cell & Hydrogen Energy Association *et al.* (2025). Preserve the Section 45V Credit for Production of Clean Hydrogen. Coalition letter. Available at https://fchea.org/wp-content/uploads/2025/06/Senate-45V-Sign-On-Letter-Final-6-5-2025.pdf.

Gaikwad, N., Genovese, F. and Tingley, D. (2022). Creating Climate Coalitions: Mass Preferences for Compensating Vulnerability in the World's Two Largest Democracies. *American Political Science Review*, **116** (4), 1165–1183.

Garbarino, A. R. *et al.* (2025). Letter to Chairman Jason Smith Regarding Energy Tax Credits. Letter. Available at https://garbarino.house.gov/sites/evo-subsites/garbarino.house.gov/files/evo-media-document/2025.03.09-tax-credits-letter.pdf.

Gawel, E., Strunz, S. and Lehmann, P. (2017). Support Policies for Renewables: Instrument Choice and Instrument Change from a Public Choice Perspective. In D. Arent, C. Arndt, M. Miller, F. Tarp and O. Zinaman (eds.), *The Political Economy of Clean Energy Transitions*, Oxford University Press, p. 0.

Gazmararian, A. F. and Tingley, D. (2023). *Uncertain Futures: How to Unlock the Climate Impasse*. The Politics of Climate Change, Cambridge: Cambridge University Press.

— and — (2024). *Unifying Comparative and International Theories of Energy Transitions around Credibility*. Unpublished manuscript.

Gerlagh, R. and Liski, M. (2023). Special Interest Climate Policy. *Unpublished Manuscript*.

Goldberg, M. H., Marlon, J. R., Wang, X., van der Linden, S. and Leiserowitz, A. (2020). Oil and gas companies invest in legislators that vote against the environment. *Proceedings of the National Academy of Sciences*, **117** (10), 5111–5112.

Golosov, M., Hassler, J., Krusell, P. and Tsyvinski, A. (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica*, **82** (1), 41–88.

Grey, F. (2018). Corporate lobbying for environmental protection. *Journal of Environmental Economics and Management*, **90**, 23–40.

Grossman, G. M. and Helpman, E. (1994). Protection for Sale. *The American Economic Review*, **84** (4), 833–850.

— and — (1996). Rent dissipation, free riding, and trade policy. *European Economic Review*, **40** (3-5), 795–803.

Grumbach, J. M. (2015). Polluting industries as climate protagonists: Cap and trade and the problem of business preferences. *Business and Politics*, **17** (4), 633–659.

Habla, W. and Winkler, R. (2013). Political influence on non-cooperative international climate policy. *Journal of Environmental Economics and Management*, **66** (2), 219–234.

Hahn, R. W., Hendren, N., Metcalfe, R. D. and Sprung-Keyser, B. (2024). A Welfare Analysis of Policies Impacting Climate Change.

Hallegatte, S., Godinho, C., Rentschler, J., Avner, P., Dorband, I. I., Knudsen, C., Lemke, J. and Mealy, P. (2024). *Within Reach: Navigating the Political Economy of Decarbonization*. Climate Change and Development Series, Washington, DC: World Bank.

Hanoteau, J. (2014). Lobbying for carbon permits in Europe. *Recherches Économiques de Louvain / Louvain Economic Review*, **80** (1), 61–87.

Harrison, K. (2010a). The Comparative Politics of Carbon Taxation. *Annual Review of Law and Social Science*, **6** (1), 507–529.

— (2010b). The Struggle of Ideas and Self-Interest in Canadian Climate Policy. In K. Harrison and L. McIntosh Sundstrom (eds.), *Global Commons, Domestic Decisions: The Comparative Politics of Climate Change*, The MIT Press, p. 0.

— (2023). Climate Governance and Federalism in Canada. In A. Fenna, J. Setzer and S. Jodoin (eds.), *Climate Governance and Federalism: A Forum of Federations Comparative Policy Analysis*, Cambridge: Cambridge University Press, pp. 64–85.

Harstad, B. (2020). Technology and Time Inconsistency. *Journal of Political Economy*, **128** (7), 2653–2689.

— (2023a). The Conservation Multiplier. *Journal of Political Economy*, **131** (7), 1731–1771.

— (2023b). Pledge-and-Review Bargaining: From Kyoto to Paris. *The Economic Journal*, **133** (651), 1181–1216.

Helfand, G. E. (1991). Standards versus Standards: The Effects of Different Pollution Restrictions. *The American Economic Review*, **81** (3), 622–634.

Hochman, G. and Zilberman, D. (2021). Optimal environmental taxation in response to an environmentally-unfriendly political challenger. *Journal of Environmental Economics and Management*, **106**, 102407.

Holland, S. P. (2012). Emissions taxes versus intensity standards: Second-best environmental policies with incomplete regulation. *Journal of Environmental Economics and Management*, **63** (3), 375–387.

—, Hughes, J. E. and Knittel, C. R. (2009). Greenhouse Gas Reductions under Low Carbon Fuel Standards? *American Economic Journal: Economic Policy*, **1** (1), 106–146.

Hughes, L. and Urpelainen, J. (2015). Interests, institutions, and climate policy: Explaining the choice of policy instruments for the energy sector. *Environmental Science & Policy*, **54**, 52–63.

Jaccard, M., Rivers, N., Bataille, C., Murphy, R. and al, e. (2006). Burning Our Money to Warm the Planet. *Commentary - C.D. Howe Institute*, (234), 0_1,0_2,1–31.

Jahn, D. (2022). Party families and greenhouse gas emissions: A new perspective on an old concept. *Zeitschrift für Vergleichende Politikwissenschaft*, **15** (4), 477–496.

Jensen, C. B. and Spoon, J.-J. (2011). Testing the 'Party Matters' Thesis: Explaining Progress towards Kyoto Protocol Targets. *Political Studies*, **59** (1), 99–115.

Juhász, R. and Lane, N. J. (2024). The Political Economy of Industrial Policy.

Kalk, A. and Sorger, G. (2023). Climate policy under political pressure. *Journal of Environmental Economics and Management*, **122**, 102900.

Kalkuhl, M., Edenhofer, O. and Lessmann, K. (2013). Renewable energy subsidies: Second-best policy or fatal aberration for mitigation? *Resource and Energy Economics*, **35** (3), 217–234.

—, Steckel, J. C. and Edenhofer, O. (2020). All or nothing: Climate policy when assets can become stranded. *Journal of Environmental Economics and Management*, **100**, 102214.

Kennard, A. (2020). The Enemy of My Enemy: When Firms Support Climate Change Regulation. *International Organization*, **74** (2), 187–221.

Kim, S. E., Urpelainen, J. and Yang, J. (2016). Electric utilities and American climate policy: Lobbying by expected winners and losers. *Journal of Public Policy*, **36** (2), 251–275.

—, — and — (2021). State Policy and Lobbying in a Federal System: Evidence from the Production Tax Credit for Renewable Energy, 1998–2012. *State Politics & Policy Quarterly*, **21** (1), 1–30.

Knill, C., Debus, M. and Heichel, S. (2010). Do parties matter in internationalised policy areas? The impact of political parties on environmental policy outputs in 18 OECD countries, 1970–2000. *European Journal of Political Research*, **49** (3), 301–336.

Konisky, D. M. (2024). The Politics of Climate Policy Instruments. *PS: Political Science & Politics*, **57** (1), 21–24.

Kotchen, M. J. and Maggi, G. (2024). The Political Economy of Carbon Taxes and Green Subsidies in a World Economy.

Lai, Y.-B. (2007). The Optimal Distribution of Pollution Rights in the Presence of Political Distortions. *Environmental and Resource Economics*, **36** (3), 367–388.

— (2008). Auctions or grandfathering: The political economy of tradable emission permits. *Public Choice*, **136** (1-2), 181–200.

Leipprand, A., Flachsland, C. and Pahle, M. (2020). Starting low, reaching high? Sequencing in EU climate and energy policies. *Environmental Innovation and Societal Transitions*, **37**, 140–155.

Linsenmeier, M., Mohommad, A. and Schwerhoff, G. (2022). Policy sequencing towards carbon pricing among the world's largest emitters. *Nature Climate Change*, **12** (12), 1107–1110.

Lundquist, S. (2024). Do Parties Matter for Environmental Policy Stringency? Exploring the Program-to-Policy Link for Environmental Issues in 28 Countries 1990–2015. *Political Studies*, **72** (2), 612–633.

Marcantonini, C. and Ellerman, A. (2015). The Implicit Carbon Price of Renewable Energy Incentives in Germany. *The Energy Journal*, **36** (4), 205–240.

Mas-Colell, A. and Vives, X. (1993). Implementation in Economies with a Continuum of Agents. *The Review of Economic Studies*, **60** (3), 613.

Mealy, P., Ganslmeier, M., Godinho, C. and Hallegatte, S. (2024). Empirical identification of feasible and strategic climate policies.

Meckling, J. (2021). Making Industrial Policy Work for Decarbonization. *Global Environmental Politics*, **21** (4), 134–147.

— and Jenner, S. (2016). Varieties of market-based policy: Instrument choice in climate policy. *Environmental Politics*, **25** (5), 853–874.

—, Kelsey, N., Biber, E. and Zysman, J. (2015). Winning coalitions for climate policy. *Science*, **349** (6253), 1170–1171.

— and Nahm, J. (2022). Strategic State Capacity: How States Counter Opposition to Climate Policy. *Comparative Political Studies*, **55** (3), 493–523.

—, Sterner, T. and Wagner, G. (2017). Policy sequencing toward decarbonization. *Nature Energy*, **2** (12), 918–922.

— and Strecker, J. (2023). Green bargains: Leveraging public investment to advance climate regulation. *Climate Policy*, **23** (4), 418–429.

Meng, K. C. and Rode, A. (2019). The social cost of lobbying over climate policy. *Nature Climate Change*, **9** (6), 472–476.

Metcalf, G. E. (2019). *Paying for Pollution: Why a Carbon Tax Is Good for America*. New York: Oxford University Press.

Mildenberger, M. (2020). *Carbon Captured: How Business and Labor Control Climate Politics*. American and Comparative Environmental Policy, Cambridge, Massachusetts ; London, England: The MIT Press.

Pahle, M., Burtraw, D., Flachsland, C., Kelsey, N., Biber, E., Meckling, J., Edenhofer, O. and Zysman, J. (2018). Sequencing to ratchet up climate policy stringency. *Nature Climate Change*, **8** (10), 861–867.

Parish, R. M. and McLaren, K. R. (1982). Relative Cost-Effectiveness of Input and Output Subsidies*. *Australian Journal of Agricultural Economics*, **26** (1), 1–13.

Persson, T. and Svensson, L. E. O. (1989). Why a Stubborn Conservative would Run a Deficit: Policy with Time- Inconsistent Preferences. *The Quarterly Journal of Economics*, **104** (2), 325–345.

— and Tabellini, G. E. (2000). *Political Economics: Explaining Economic Policy*. Zeuthen Lecture Book Series, Cambridge, Mass: MIT Press.

Prato, C. (2017). Electoral Competition and Policy Feedback Effects. *The Journal of Politics*, **80** (1), 195–210.

Ramadorai, T. and Zeni, F. (2024). Climate Regulation and Emissions Abatement: Theory and Evidence from Firms' Disclosures. *Management Science*.

Reguant, M. (2019). The Efficiency and Sectoral Distributional Impacts of Large-Scale Renewable Energy Policies. *Journal of the Association of Environmental and Resource Economists*, **6** (S1), S129–S168.

Rodrik, D. (2014). Green industrial policy. *Oxford Review of Economic Policy*, **30** (3), 469–491.

Rozenberg, J., Vogt-Schilb, A. and Hallegatte, S. (2020). Instrument choice and stranded assets in the transition to clean capital. *Journal of Environmental Economics and Management*, **100**, 102183.

Samson, R. A. and Stamler, S. B. (2009). Going Green for Less: Cost-Effective Alternative Energy Sources. *Commentary - C.D. Howe Institute*, (282), 0_1,0_2,1–21.

Schmalensee, R. (2012). Evaluating Policies to Increase Electricity Generation from Renewable Energy. *Review of Environmental Economics and Policy*, **6** (1), 45–64.

Schmitt, A. (2014). *Beyond Pigou: Climate Change Mitigation, Policy Making and Distortions*. Ph.D. thesis, Department of Economics, Stockholm University.

Smulders, S. and Zhou, S. (2024). Self-fulfilling Prophecies in the Transition to Clean Technology.

Stiglitz, J. E. (2019). Addressing climate change through price and non-price interventions. *European Economic Review*, **119**, 594–612.

Stock, J. H. (2020). Climate Change, Climate Policy, and Economic Growth. *NBER Macroeconomics Annual*, **34**, 399–419.

Stokes, L. C. (2020). *Short Circuiting Policy: Interest Groups and the Battle Over Clean Energy and Climate Policy in the American States*. Oxford University Press.

Sud, R., Patnaik, S. and Glicksman, R. (2023). *How to Reform Federal Permitting to Accelerate Clean Energy Infrastructure*. Tech. rep., Brookings Institution.

Tornell, A. (1991). Time Inconsistency of Protectionist Programs*. *The Quarterly Journal of Economics*, **106** (3), 963–974.

ULPH, A. and ULPH, D. (2013). Optimal Climate Change Policies When Governments Cannot Commit. *Environmental and Resource Economics*, **56** (2), 161–176.

URPELAINEN, J. and ZHANG, A. T. (2022). Electoral Backlash or Positive Reinforcement? Wind Power and Congressional Elections in the United States. *The Journal of Politics*, **84** (3), 1306–1321.

VAN DEN BERGH, J. C. J. M. and BOTZEN, W. W. J. (2024). Assessing Criticisms of Carbon Pricing. *International Review of Environmental and Resource Economics*, **18** (3), 315–384.

VICTOR, D. G., GEELS, F. W. and SHARPE, S. (2019). *Accelerating the Low Carbon Transition: The Case for Stronger, More Targeted and Coordinated International Action*. Tech. rep., Commissioned by the UK Government Department for Business, Energy and Industrial Strategy.

VON HIRSCHHAUSEN, C., GERBAULET, C. and KEMFERT, C. (eds.) (2019). *Energiewende "Made in Germany": Low Carbon Electricity Sector Reform in the European Context*. Springer.

VORMEDAL, I. and MECKLING, J. (2023). How foes become allies: The shifting role of business in climate politics. *Policy Sciences*.

WINKLER, D. (2022). Pollution for Sale: Lobbying, Allowance Allocation and Firm Outcomes in the EU ETS.

Online Appendix

Contents

A. Extensions

A.1. *Targeted Transfers and Subsidies*

If policymakers can use arbitrarily targeted lump-sum transfers and have perfect information, optimal climate policy is simple: the government can offer each polluting firm in the economy a transfer contingent on technology adoption that makes the firm indifferent between upgrading or not. This requires no commitment, because lump-sum transfers are costless for the government. One reason why this is not feasible in reality is that the government lacks information: it does not know what the optimal abatement actions are, and cannot raise funds costlessly.

There are two facts that put this argument into question. First, governments do use targeted transfers to obtain political support for policies. This is clear in the case of free allowances in cap-and-trade systems, and is the basis of the idea of "just transition" strategies (Bolet *et al.* 2023), which bundle climate with redistributive policies to support affected communities, and "green bargains" (Meckling and Strecker 2023), which tie regulations to public investments. But, for example, in the first two phases of the EU ETS, even though there is evidence that lobbying by particular firms impacted the allocation of free allowances, policymakers used rigid rules (mainly based on historical emissions) and a large information asymmetry dominated the process (Ellerman *et al.* 2007). From a theoretical point of view, the fact that the process was transparent constrained the ability of politicians to be cynical, which may have created inefficiencies as policymakers had an incentive to try to look fair (Coate and Morris 1995). The fact that more than 90% of the allowances were allocated for free, plus the fact that there was substantial overallocation in several industries (i.e., they produced less emissions than the allowances they were given for free; see, e.g., Hanoteau 2014), indicates that the allocation was not targeted in a surgically precise manner to assuage opposition.

If we consider the possibility of quid-pro-quo, which is the implicit assumption of common agency or rent-seeking models, then we should expect a targeted (but distorted) allocation. This targeting, however, responds to the incentives and ability for lobbying of each firm or industry, which can be (and should be expected to be) very different to the targeting that results from a strategy of building a "green coalition"—the optimal green coalition is formed by the set of constituencies that have the lowest costs of decarbonization or the greatest expected opportunities in a decarbonized economy; in contrast, the firms or industries most willing to lobby may include declining industries (see Grossman and Helpman 1996 and Baldwin and Robert-Nicoud 2007) and industries with assets that will lose value with decarbonization. This argument helps to justify that a coalition-building strategy cannot plausibly rely on finely targeted transfers, despite the fact that we see quite targeted transfers as part of climate policies in practice.

The second counterargument is that in practice subsidies and regulations tend to be technology- and industry-specific. See, for example, Gawel *et al.* (2017) on the FITs in Germany, and Hahn *et al.* (2024) for a list of specific subsidies in the US; moreover, Cullenward and Victor (2020) argue that industry-specific regulations are not just prevalent but desirable, since broad policies, despite having the potential for being more efficient, tend to be watered down by the pressure of the most affected industries (they do not provide systematic evidence or a sound theoretical argument for this assertion, though; theoretically, targeted interventions may lead to less internalization, and, thus, more lobbying in opposition). However, the fact that in Germany the demand for FITs exploded beyond expectations, and the subsidies in the IRA are uncapped and there is considerable uncertainty about their fiscal cost (Bistline *et al.* 2023), suggest again that subsidies in practice are far from being surgically targeted and policymakers face substantial uncertainty when designing them.

To address these arguments in the context of our model we can relax the assumption that policymakers cannot propose targeted transfers, but assume that there is an information asymmetry between the policymaker and economic agents. To put structure on this assumption, suppose that the policymaker knows the distribution of A_i, but not the particular value of A_i for any i and, hence, can design the policies that I analyzed in the main model but cannot implement individually targeted rebates or subsidies. The policy choice decision given the information constraint becomes a mechanism design problem. Following Mas-Colell and Vives (1993), an (anonymous direct) *mechanism* is a function $f : [0, A] \to \mathbb{R}^3_+$ that takes the productivity of green technology of a district (which is the relevant characteristic) and assigns a specific carbon tax, a green subsidy and a tax rebate.[10] It is *strongly incentive compatible* (IC) if each district chooses to reveal their productivity type, i.e., district i communicates A_i and is subject to the policy $f(A_i)$, rather than any other productivity level A_j. The timing of the game is left unchanged: in each period, the agenda-setter proposes an IC mechanism to the legislature (instead of a uniform policy as in the main model); each legislator votes in favor of the proposal if and only if they prefer it over the status quo, and then firms submit their productivity levels to the mechanism, which assigns to them a carbon tax, a green subsidy and a tax rebate, and then decide their production and investment levels.

In the second period, green subsidies are irrelevant, IC mechanisms assign policies τ, r^b for polluters and r^g for green districts, and the IC constraint is that $\frac{1}{2}(1-\tau^b)^2 + r^b \geqslant r^g \geqslant r^b$. If there is a polluting majority, polluters must accept the mechanism for it to be politically implemented. If the alternative is no climate policy then the political constraint is $\frac{1}{2}(1-\tau)^2 + r^b + T \geqslant \frac{1}{2}$, where $T = b\tau(1-\tau) - br^b - (1-b)r^g \leqslant b\tau(1-\tau) - r^b$ (by IC), but $\frac{1}{2}(1-\tau)^2 + r^b + T \leqslant \frac{1}{2}(1-\tau)^2 - b\tau(1-\tau) = -[(1-b)\tau + (b-\frac{1}{2})\tau^2] \leqslant 0$, so the political constraint holds only if $\tau = 0$. In other words, if the alternative policy is void, it is impossible to implement a positive carbon tax. If greens are a majority, in contrast, any carbon tax is politically implementable.

Incentive-compatible mechanisms in the first period consist of three policies: a tax and rebate for polluting firms that do not transition, (τ_1^b, r_1^b), a tax, subsidy and rebate for firms that transition, $(\tau_1^{bg}, s, r_1^{bg})$, and a rebate for green firms r_1^g. Incentive compatibility imposes the following conditions: $\frac{1}{2}(1-\tau_1^b)^2 + r_1^b \geqslant \frac{1}{2}(1-\tau_1^{bg})^2 + r_1^{bg}$, i.e., polluting firms that do not plan to transition do not take the tax and rebate intended for firms that do transition, and $\frac{1}{2}(1-\tau_1^{bg})^2 + s + r_1^{bg} \geqslant \frac{1}{2}(1-\tau_1^b)^2 + r_1^b$, i.e., polluting firms that plan to transition do not take the tax and rebate intended for firms that do not. In addition, $\frac{1}{2}(1-\tau_1^b)^2 + r_1^b \geqslant r_1^g \geqslant r_1^b$ and $\frac{1}{2}(1-\tau_1^{bg})^2 + s + r_1^{bg} \geqslant r_1^g \geqslant r_1^{bg}$, i.e., polluters and greens do not impersonate each other.

To focus on the information asymmetry, I assume that there is no turnover, i.e., $\rho = 1$. The political constraints demand that polluting firms in $[\frac{1}{2}, b_1]$ and green firms accept the mechanism

[10] I focus on anonymous mechanisms in order to abstract from information leakage and ratcheting across periods.

that implements $(\tau_1^b, r_1^b, \tau_1^{bg}, s, r_1^{bg}, r_1^g)$ in the first period, and green firms accept the mechanism (τ_2, r_2^b, r_2^g) in the second period. The polluting firms that do not transition have no say in the legislature, but they are protected in the first period by the incentive compatibility constraint. The set of firms that transition is given by

$$\tilde{s} - c + \frac{\delta A}{2} b_2 = \frac{\delta}{2}(1-\tau_2)^2,$$

where $\tilde{s} = s - \Delta_1 + \Delta_2 \geqslant 0$, $\Delta_1 = \frac{1}{2}(1-\tau_1^b)^2 + r_1^b - \frac{1}{2}(1-\tau_1^{bg})^2 - r_1^{bg} \geqslant 0$, $\Delta_2 = \delta(r_2^g - r_2^b) \geqslant 0$. The political constraint for polluting districts in the first period can be written as

$$\frac{1}{2}(1-\tau_1^{bg})^2 + s - c + \frac{\delta A}{4} + r_1^{bg} + \delta r_2^g + T \geqslant \frac{1}{2} + \frac{\delta}{2},$$

where

$$\begin{aligned} T = {} & b_2[\tau_1^b(1-\tau_1^b) - r_1^b] + (b_1 - b_2)[\tau_1^{bg}(1-\tau_1^{bg}) - s - r_1^{bg}] - (1-b_1)r_1^g \\ & + \delta b_2[\tau_2^b(1-\tau_2^b) - r_2^b] - \delta(1-b_2)r_2^g. \end{aligned}$$

The political constraint for green districts is

$$r_1^g + \delta r_2^g + T \geqslant 0$$

in the first period and

$$r_2^g + b_2\tau_2^b(1-\tau_2^b) - b_2 r_2^b - (1-b_2)r_2^g \geqslant r_1^g + b_2\tau_1^b(1-\tau_1^b) - b_2 r_1^b - (1-b_2)r_1^g$$

in the second period.

The political constraints are relaxed to some extent relative to uniform rebates and subsidies, but the key inefficiency arises from the fact that the subsidy is used for political acceptability, but it spills over to wasteful investment decisions. The main driver of the results in the baseline model, viz., the fact that the political constraint creates a tradeoff between the first-period carbon tax and the subsidy (a larger carbon tax requires a larger subsidy to compensate the pivotal district, since rebates cannot be perfectly targeted to offset the loss in profits), persists in this variation of the model.

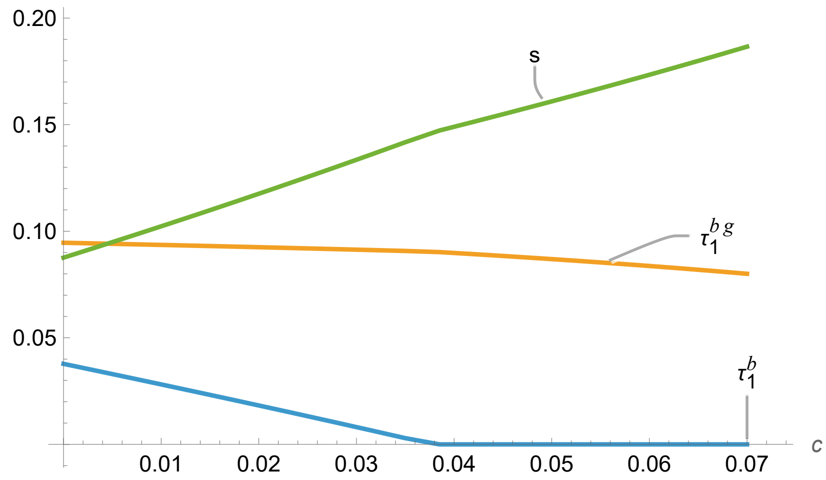


Figure 5: Equilibrium policy when $A = 1.8$, $\delta = 0.9$ and $\lambda = 0.1$, for different values of c.

Figure 5 shows a numerical example. The specific rebates are not used in equilibrium. As in the main model, the carbon tax in the first period is below the Pigouvian level, and the investment subsidy is used. The districts that transition pay a higher carbon tax rate than the ones that do not, even though the latter are left out of the political coalition. The reason is that, without specific rebates, the incentive compatibility constraint $\frac{1}{2}(1-\tau_1^b)^2 + r_1^b \geqslant \frac{1}{2}(1-\tau_1^{bg})^2 + r_1^{bg}$ implies $\tau_1^b \leqslant \tau_1^{bg}$. If we use the rebate r_1^b to obtain a larger τ_1^b, we must compensate the pivotal district (who bears part of the fiscal cost of the rebate); extracting from the green districts is not feasible because of their own political constraint; therefore increasing r_1^b requires increasing $\frac{1}{2}(1 - \tau_1^{bg})^2 + s$, which creates a welfare loss. In other words, the flexibility to provide specific policies is useful, but the incentive compatibility condition protects the polluters that are excluded from the political coalition. This, in turn, leads to the same tradeoff as in the main model: an optimal first-period carbon tax requires inefficient compensation to pivotal districts, and is, therefore, not used in equilibrium.

A.2. *Green Preferences*

In this section I assume that districts care about environmental damages as much as the green party, so payoffs are the same as before with an added term $-\lambda e_t$, i.e.,

$$\pi_{it} = \begin{cases} (1-\tau)y_{it} - \frac{1}{2}y_{it}^2 - \iota_{it}(c - s_t) + T_t - \lambda e_t & \text{if } i \text{ is polluting,} \\ y_{it} - \frac{1}{2Ai}y_i^2 + T_t - \lambda e_t & \text{if } i \text{ is green,} \end{cases}$$

where $e_t = \int_0^1 \chi_{it} y_{it}\, di$ are aggregate emissions at time t. For simplicity I assume that the first-period proposer stays in the second period, i.e., $\rho = 1$. In a one-period interaction, if the polluting districts are $[0, b)$ and the carbon tax is τ, the payoff of a polluting district is $\frac{1}{2}(1-\tau)^2 + b\tau(1-\tau) - \lambda b(1-\tau)$. If $(1 + \lambda)b > 1$, which requires $b > \frac{1}{2}$, then their ideal tax is $\tau = \frac{(1+\lambda)b-1}{2b-1} < \lambda$. Otherwise, their ideal tax is $\tau = 0$.

ASSUMPTION 2. $(1 + \lambda)b_1 < 1$.

Under Assumption 2, polluting districts always oppose a carbon tax. In this case, the equilibrium policy in the second period is the same as in the baseline model: $\tau_2 = \max\{\tau_1, \lambda\}$ if $b_2 \leqslant \frac{1}{2}$, and otherwise $\tau_2 = 0$.

As in the baseline model, the green districts prefer a high carbon tax. But, given that they care about carbon emissions, they are now willing to pay some taxes themselves in order to pay for investment subsidies if that is required to obtain a reduction in GHGs emissions. Moreover, given that, with these preferences, the payoff of the green policymaker is exactly aggregate welfare, $W_t = \int_0^1 \pi_{it}\, di$, because districts now internalize the environmental damage. Thus, if the green policymaker prefers a climate policy over BAU, then the green districts prefer it as well, because they value the benefits but do not internalize the costs. Therefore, green districts do not impose a constraint on policy in the first period.

Polluting districts impose a political constraint. The proposer needs to induce $b_2 \leqslant \frac{1}{2}$ and obtain the approval of the median district, i.e.,

$$\underbrace{\frac{1}{2}(1-\tau_1)^2 + s - c + \frac{\delta A}{4} + T}_{\text{economic payoff}} - \underbrace{\lambda b_1(1-\tau_1) - \delta\lambda b_2(1-\tau_2)}_{\text{environmental damage}} \geqslant \underbrace{(1+\delta)\left(\frac{1}{2} - \lambda b_1\right)}_{\text{BAU payoff}}. \qquad (\mathrm{PC}_B)$$

LEMMA 2. *Under the assumptions of this section, if PC_B holds then $b_2 \leqslant \frac{1}{2}$.*

Thus, the green policymaker has two options: implement the best policy subject to the constraint PC_B, or else keep BAU. Under what conditions is PC_B feasible? We have the following.

OBSERVATION 5. *If PC_B is feasible under (A, c, δ, λ) it is feasible under $(A', c', \delta', \lambda')$ if $A' \geqslant A$, $c' \leqslant c$, $\delta' \geqslant \delta$ and $\lambda' \geqslant \lambda$.*

In words, it is possible to implement climate policy in the first period if and only if A is large enough, c is small enough, δ is large enough, and λ is large enough. The first three conditions are also necessary in the baseline model, for essentially the same reason: they express that the economic cost of the transition is not too large. The final condition contrasts with the baseline model: if districts care about environmental damage as much as the green party, the higher the social cost of carbon is perceived to be, the easier it is for the green policymaker to convince economic actors to accept climate policy that leads to a carbon tax in the future.

PROPOSITION 3. *Under Assumptions 1 and 2, the equilibrium policy is either first-best ($\tau_1 = \lambda$, $s = 0$), satisfies $\tau_1 < \lambda$ and $s > 0$, or is business-as-usual.*

Qualitatively, the only difference with the baseline model is that optimal policy (Pigouvian carbon tax with no subsidies) can be feasible in the first period. Otherwise, in equilibrium the green policymaker faces the same tradeoff: a larger carbon tax requires a greater subsidy, which is costly, so in equilibrium the carbon tax is less than optimal and the subsidy is used, or else BAU is maintained.

To conclude, when districts care about environmental damages as much as the green policymaker, the set of politically feasible climate policies expands, and, moreover, in very restrictive conditions first-best policy becomes feasible. In general, though, the qualitative features of the equilibrium are the same: the green party enacts a low-ambition carbon tax plus a subsidy in the first period, and a carbon tax set at optimal level subsequently. Committed representatives make climate policy more likely and less distorted, but significant distortions can remain.

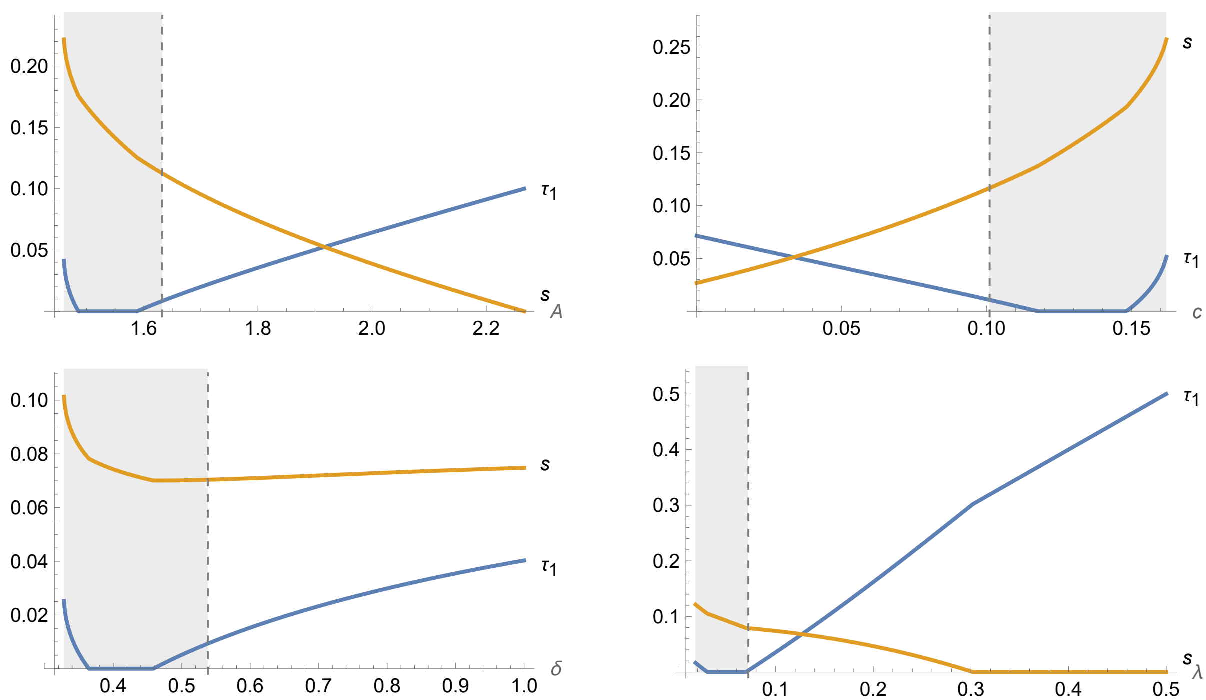


Figure 6: Equilibrium policy when $A = 1.8$, $c = 0.06$, $\delta = 0.9$ and $\lambda = 0.1$, as each parameter changes. The policies are always preferred over BAU.

Figure 6 shows the equilibrium policies as parameters change. The main differences with the baseline model (Figure 2) is that with green preferences for these parameters the best feasible policy is better than BAU, and $\tau_1 > 0$. Other than that, the comparative statics relative to A, c and δ are the same: the harder the political problem (small A, large c, small δ), the smaller the carbon tax and the greater the subsidy.

A.3. *Production Subsidies*

Consider a production subsidy for the green sector $\sigma \geqslant 0$. The way it works is that if a firm in district i produces y_i, they get a transfer σy_i. Given σ, the firm chooses y_i to maximize profits, $\pi_i = (1+\sigma)y_i - \frac{1}{2A_i}y_i^2$, so $y_i = (1+\sigma)A_i$. Thus, equilibrium profits are $\pi_i = (1+\sigma)^2\frac{1}{2}A_i$, and the fiscal cost is $\sigma(1+\sigma)A_i$, so the contribution to aggregate welfare is $(1+\sigma)^2\frac{1}{2}A_i - \sigma(1+\sigma)A_i = \frac{1}{2}(1-\sigma^2)A_i$.

To affect investment, the policymaker could use a production subsidy that starts acting in the second period. The political constraint for polluting districts $[\frac{1}{2}, b_1)$ is

$$\frac{1}{2}(1-\tau_1)^2 + s - c + \underbrace{(1+\sigma)^2\frac{\delta A}{4}}_{\substack{\text{period-2 profit} \\ \text{with production subsidy}}} + T \geqslant (1+\delta)\frac{1}{2}, \tag{3}$$

and the political constraint for the green districts $[b_1, 1]$ is

$$\underbrace{(1+\sigma)^2\frac{\delta A}{2}b_1}_{\substack{\text{period-2 profit} \\ \text{with production subsidy}}} + T \geqslant \frac{\delta A}{2}b_1, \tag{4}$$

where

$$T = \underbrace{b_1\tau_1(1-\tau_1) + \delta b_2\tau_2(1-\tau_2)}_{\text{carbon tax revenue}} - \underbrace{s(b_1-b_2)}_{\substack{\text{fiscal cost of the} \\ \text{investment subsidy}}} - \underbrace{\delta\int_{b_2}^{1}\sigma(1+\sigma)Ai\,di}_{\substack{\text{fiscal cost of the} \\ \text{production subsidy}}}.$$

The objective of the green policymaker is

$$\begin{aligned} W = {} & \frac{b_1}{2}(1-\tau_1^2) + \frac{A}{4}(1-b_1^2) - c(b_1-b_2) - \lambda b_1(1-\tau_1) \\ & + \delta\left[\frac{b_2}{2}(1-\tau_2^2) + \underbrace{(1-\sigma^2)}_{\substack{\text{distortion due to the} \\ \text{production subsidy}}}\frac{A}{4}(1-b_2^2) - \lambda b_2(1-\tau_2)\right], \end{aligned}$$

where b_2 is now given by $s - c + (1+\sigma)^2\frac{\delta A}{2}b_2 = \frac{\delta}{2}(1-\tau_2)^2$, and $\tau_2 = \max\{\tau_1, \lambda\}$.

What happens in numerical simulations is that in equilibrium σ is never used. The investment subsidy s is always preferred to σ. If s wasn't available then σ would be used, but it's worse.

There are theoretical reasons for subsidizing output rather than inputs (capital), but in general it depends. If the objective is to increase output, then subsidizing inputs will distort productive efficiency, so subsidizing output is prima facie better. However, even taking into account this distortion, it may be cheaper to subsidize one input if it's a lot more elastic than another one. See Parish and McLaren (1982) for an exposition. Now, even though this is theoretically ambiguous, empirically Aldy *et al.* (2023) show that production subsidies for wind energy were more cost effective than investment subsidies, which goes against the prediction of the model. The reasons have to do with aspects of the production of renewables (the almost zero marginal cost) that my model doesn't capture, so what I can say is that a more realistic model may (perhaps should) make a different prediction regarding investment versus production subsidies. So this is not something to take as a firm prediction. However, I should point out that maximizing social welfare is not the same as maximizing the output of renewable energy production. Increasing production of wind energy is not the goal of climate policy. The goal is to decarbonize the economy with the least social cost.

A.4. *Standards and Feed-in Tariffs*

Consider a clean production standard $\mu \in [0, 1]$, such as a RPS (see Helfand 1991; Holland *et al.* 2009; Holland 2012; Schmalensee 2012). It forces firms to emit no more than μ units of GHGs per unit of the good produced. Firms that produce more emissions than allowed can buy permits from firms that produce less emissions than allowed for a market price p. In the model, firms that use green capital do not emit GHGs, so if they produce y, they are allowed to emit μy, which they can sell, earning $p\mu y$. Firms with polluting capital emit one unit of GHGs per unit produced, so they need to buy $(1-\mu)y$ allowances, paying a cost $p(1-\mu)y$. Profit maximization implies that $y_i = (1+p\mu)Ai$ in green districts i, and $y_i = 1-(1-\mu)p$ in polluting districts. The price of allowances p is set by the market clearing condition:

$$\underbrace{\int_0^b (1-\mu) y_i \, di}_{\substack{\text{permits bought} \\ \text{by polluting firms}}} = \underbrace{\int_b^1 \mu y_i \, di}_{\substack{\text{permits sold} \\ \text{by green firms}}},$$

which implies $\int_0^b y_i \, di = \mu \int_0^1 y_i \, di$, so, as intended, emissions are capped at a fraction μ of aggregate production. The equilibrium price of the permits is

$$p = \frac{(1-\mu)b - \mu \frac{A}{2}(1-b^2)}{(1-\mu)^2 b + \mu^2 \frac{A}{2}(1-b^2)}$$

if $b \geqslant \mu\Big(b + \frac{A}{2}(1-b^2)\Big)$, i.e., the standard is binding: emissions per unit produced are more than μ under BAU. Otherwise, the price is 0. The clean production standard is thus equivalent to a carbon tax $\tau = p(1-\mu)$ combined with a green production subsidy $p\mu$.

Let $\tau = p(1 - \mu)$ and $\sigma = p\mu$ be the equivalent carbon tax and green production subsidy given a standard μ. The market clearing condition can be written in terms of τ and σ as

$$b\tau(1 - \tau) = \frac{A}{2}(1 - b^2)\sigma(1 + \sigma).$$

Notice that the LHS is the fiscal revenue from the equivalent carbon tax τ, and the RHS is the fiscal cost of the equivalent production subsidy σ, so the market clearing condition is effectively a joint revenue neutrality condition for the tax and the subsidy. Thus, choosing a standard is equivalent to choosing a carbon tax τ and using the revenue to finance a green production subsidy σ. The greater the tax τ, the greater the subsidy σ, as long as $\tau \leqslant \frac{1}{2}$, i.e., the carbon tax is on the left side of the Laffer curve.

We can introduce a feed-in tariff (FIT) into the model as follows. A retailer is forced to buy the numéraire good at price $p_G \geqslant 1$ from green producers, and sets the price $p_B \geqslant 0$ it buys from polluting producers to maximize profits obtained from selling the good at price 1. Profits are given by $\pi = (1 - p_B)\int_0^b y_i\, di - (p_G - 1)\int_b^1 y_i\, di = (1 - p_B)\int_0^b A_i p_B\, di - (p_G - 1)\int_b^1 p_G\, di = -\frac{A}{2}(1 - b^2)p_G(p_G - 1) + bp_B(1 - p_B)$. Perfect competition or free entry in the retail market brings profits to zero, so the price paid to polluting producers is given implicitly by $bp_B(1 - p_B) = \frac{A}{2}(1 - b^2)p_G(p_G - 1)$. This is again equivalent to a revenue-neutral combination of a carbon tax $\tau = 1 - p_B$ and an output subsidy for green producers $\sigma = p_G - 1$. Thus, in this simple model a standard and a FIT are equivalent. (The instruments differ in reality if, for example, there is uncertainty about demand or productivity. A standard fixes the quantity of green production, creating price risk for clean producers; a FIT fixes the price of renewable energy, which reduces risk for producers and can stimulate investment, but increases risk for consumers. See Schmalensee 2012.)

If a carbon tax and a production subsidy are available, a green policymaker would not use the standard or FIT, because its effects can be replicated, and in general improved, by using the other two instruments. However, if the carbon tax and the subsidy are not available, the standard can be used in equilibrium. Notice that polluting districts will oppose any binding standard in a static environment. However, if the median district transitions, they expect to receive a subsidy in the second period, which acts as an investment incentive, and, thus, as a carrot for accepting a standard today. The policymaker cannot commit in the first period to a generous standard in the future, though, except by setting the standard high in the first period and relying on the fact that, once a majority of districts is green, they will not accept a reduction in the standard, because they profit from the allowances. However, increasing the standard in the first period creates a cost for the median district that exceeds the benefit it provides by increasing the implicit subsidy in the future.

The green policymaker chooses the second period standard (τ_2, σ_2) to maximize W given b_2.

Anticipating this, the share of polluting districts that decide to transition in the first period is given by

$$-c + (1+\sigma_2)^2 \frac{\delta A}{2} b_2 = \frac{\delta}{2}(1-\tau_2)^2.$$

The political constraint in the first period is

$$\frac{1}{2}(1-\tau_1)^2 - c + (1+\sigma_2)^2 \frac{\delta A}{4} - (1+\delta)\frac{1}{2} \geqslant 0. \tag{PC_G}$$

Notice that in the first period the policymaker cannot affect σ_2 in order for PC_G to hold, except by making $\sigma_1 > \sigma_2$, but, as mentioned before, this is counterproductive as it increases τ_1.[11] Therefore, the policymaker can only implement a standard in the first period if PC_G holds for $\tau_1 = 0$. In that case, they can increase τ_1 until PC_G binds or the efficiency cost of the subsidy σ_1 is greater than the benefit of the tax. In the first case, the first-period standard is too low relative to the optimum choice. (I don't think that the second case can happen. It would be nice to prove this.)

Thus, climate policy in the first period relies on the second-period standard being large enough so that it provides an incentive for the median district to upgrade its capital. This, in turn, requires that the social cost of carbon λ is large enough to justify a stringent standard. In fact, for large λ the standard may be feasible when a carbon tax plus an investment subsidy is not, because the subsidy has to be paid by green districts as well as polluting districts, which makes it less appealing to the pivotal polluting district, to the point where the political constraints become infeasible. This seems to contradict the assertion I made before that the standard can be improved upon by a combination of a tax and a subsidy. What's happening is that if the other instruments are available the policymaker will use them instead of the standard, but, given the time-inconsistency problem, the fact that the policymaker has that power may make it harder to satisfy the political constraints.

A.5. *Imperfect Competition and Pass-through*

In this section I adapt the model in Besley and Persson (2023) to incorporate an investment subsidy and legislative bargaining. As in the main model, there is a set of districts indexed by $i \in I = [0,1]$. Each district produces a specific variety with a clean or dirty production technology, using labor as an input. Producing x_i units of variety i using the dirty technology costs x_i units of labor, and releases x_i units of carbon. Producing x_i units using the clean technology costs $c(i)x_i$ units of labor, with c decreasing, and has a fixed cost f. There is a numéraire good produced competitively using labor with unit cost. The wage is therefore equal to 1. Each district is endowed with L units of labor.

[11] Suppose $\sigma_2 = \sigma_1$. Differentiating the fiscal neutrality condition yields $b_1(1-2\tau_1)\frac{d\tau_1}{d\sigma_1} = \frac{A}{2}(1-b_1^2)(1+2\sigma_1)$. Differentiating the LHS of PC_G we obtain $-(1-\tau_1)\frac{d\tau_1}{d\sigma_1} + \frac{\delta A}{2}(1+\sigma_2) = \frac{A}{2}\left[-\frac{1-b_1^2}{b_1}\frac{1-\tau_1}{1-2\tau_1}(1+2\sigma_1) + \delta(1+\sigma_2)\right] \leqslant 0$ if $1-b_1^2 \geqslant b_1$, which is true if $b_1 \leqslant 0.618$. In that case, as claimed, increasing σ_1 beyond σ_2 is counterproductive.

Consumer preferences in each period are given by

$$U = \frac{1}{1-\sigma}\int_0^1 x_i^{1-\sigma}\, di + n,$$

where x_i is the quantity of variety i, n is the quantity of the numéraire consumed, and $\sigma \in (0,1)$ is an inverse measure of the elasticity of substitution between varieties. Let p_i be the price of variety i. The representative consumer in district i chooses $x_i^{-\sigma} = p_i$, assuming that L is large enough so that the solution is interior.

Producers are monopolistic and maximize profits $\pi_i = (p_i - \tau_i - c_i)x_i - f_i$, where τ_i is a production tax, c_i is the marginal labor cost (i.e., $c_i = 1$ if i is polluting, $c_i = c(i)$ if i is green, employing a slight abuse of notation), and f_i is the fixed cost of production (i.e., $f_i = 0$ if i is polluting, $f_i = f$ if i is green). Therefore, they choose prices $p_i = \frac{1}{1-\sigma}(c_i + \tau_i)$. Profits are thus $\pi_i = \sigma p_i x_i - f_i$.

Given that districts are ordered by the productivity of the green technology, varieties $i \in [0, b)$ are produced with the dirty technology, and $i \in (b, 1]$ are produced with the clean technology, for some $b \in [0, 1]$.

The planner's problem is to choose $(x_i)_{i\in I}$ and b to maximize

$$\frac{1}{1-\sigma}\int_0^1 x_i^{1-\sigma}\, di - \int_0^1 (c_i x_i + f_i)\, di + L - \lambda\int_0^b x_i\, di,$$

where the last term measures climate damages. The solution is given by $x_i^{-\sigma} = c_i + 1_{i\leqslant b}\lambda$ and $\frac{\sigma}{1-\sigma}((1+\lambda)^{1-\frac{1}{\sigma}} - c(b)^{1-\frac{1}{\sigma}}) + f = 0$. We can implement the social optimum using production taxes τ_i and a subsidy for the fixed cost of green varieties s. To find the optimal taxes we use that in the decentralized equilibrium $x_i^{-\sigma} = p_i = \frac{1}{1-\sigma}(c_i + \tau_i)$, and therefore we can implement the optimal quantities if $\frac{1}{1-\sigma}(c_i + \tau_i) = c_i + 1_{i\leqslant b}\lambda$, i.e.,

$$\tau_i = 1_{i\leqslant b}(1-\sigma)\lambda - \sigma c_i.$$

The first term is the usual carbon tax, which is equal to the social cost of carbon λ multiplied by $1-\sigma$ in order to correct for the market imperfection. The second term corrects the underproduction due to market power. The marginal green producer is indifferent between profits using each technology, taking into account the fixed cost and the green subsidy, and thus b is given in the decentralized equilibrium by $\sigma c(b)^{1-\frac{1}{\sigma}} - f + s = \sigma(1+\lambda)^{1-\frac{1}{\sigma}}$. Therefore, to implement the optimal b we need $s = \sigma f$, which is due to the fact that profits do not fully cover fixed costs.

Putting everything together, to implement the optimal allocation we need (1) production subsidies $\sigma(c_i x_i + f_i)$ given variable cost c_i, production quantity x_i and fixed cost f_i, and (2) a carbon tax $(1-\sigma)\lambda x_i$ given a quantity x_i produced with dirty technology. The first instrument corrects the distortion due to the market imperfection, and the second instrument corrects the distortion due to the climate externality. In order to focus on the climate issue I will assume that the instrument (1) is already in place, and, as always, the tax revenue is rebated as a uniform lump-sum to every district. For notational convenience I will write the carbon tax as $(1-\sigma)\tau$, noting that the optimal tax is $\tau = \lambda$.

Thus, as in the main model, there are two climate policies: a carbon tax $\tau \geqslant 0$ that collects $(1-\sigma)\tau x_i$ from polluting districts, and a green subsidy $s \geqslant 0$ that pays $\delta\sigma s$ to polluting districts that switch to green technology ($\delta\sigma$ is a normalizing factor). The policy process is the same as in the main model.

Given these policy instruments, statically we have $p_i = x_i^{-\sigma} = c_i + 1_{i \leqslant b}\tau$ and the variable component of profits (taking production subsidies and carbon taxes into account) is $\pi_i = \sigma p_i x_i$. Utility from consumption of varieties is $\frac{1}{1-\sigma}\int_0^1 x_i^{1-\sigma}\,di - \int_0^1 p_i x_i\,di = \frac{\sigma}{1-\sigma}\int_0^1 x_i^{1-\sigma}\,di$. The tax rebate, ignoring the green subsidy, is $T = -\sigma\int_0^1 c_i x_i\,di + (1-\sigma)\int_0^b \tau x_i\,di = -\sigma\int_0^1 p_i x_i\,di + \tau\int_0^b x_i\,di$. Putting everything together, we define

$$\begin{aligned} V(b,\tau) &= \frac{1}{1-\sigma}\int_0^1 x_i^{1-\sigma}\,di - \int_0^1 p_i x_i\,di + T \\ &= \frac{\sigma^2}{1-\sigma}\int_0^1 x_i^{1-\sigma}\,di + \tau\int_0^b x_i\,di \\ &= \frac{\sigma^2}{1-\sigma}\left(b(1+\tau)^{1-\frac{1}{\sigma}} + \int_b^1 c(i)^{1-\frac{1}{\sigma}}\,di\right) + b\tau(1+\tau)^{-\frac{1}{\sigma}}. \end{aligned}$$

We can write the payoff of district i as

$$U_i = V(\tau_1, b_1) + \delta V(\tau_2, b_2) + \pi_{i1} + \delta\pi_{i2} - \delta\sigma((1-\sigma)f - s)\iota_i - \delta\sigma(s + \sigma f)(b_1 - b_2),$$

with $\pi_{it} = \sigma[1_{i \leqslant b_t}(1+\tau_t)^{1-\frac{1}{\sigma}} + 1_{i > b_t} c(i)^{1-\frac{1}{\sigma}}]$, and $\iota_i = 1_{i \in (b_2, b_1)}$, i.e., an indicator of technology switching.

I assume as in the main model that the initial marginal polluting district equalizes profits using each technology assuming no climate policies, and so it's given by $c(b_1)^{1-\frac{1}{\sigma}} - (1-\sigma)f = 1$. The share of second-period polluting districts is determined endogenously in the model, and is given by the equation

$$c(b_2)^{1-\frac{1}{\sigma}} - (1-\sigma)f + s = \mathrm{E}[(1+\tau_2)^{1-\frac{1}{\sigma}}].$$

As before, b_2 decreases if the subsidy increases or if the expectation of the second-period carbon tax increases.

The objective of the policymaker is

$$
\begin{aligned}
W &= \int_0^1 U_i\,di - \lambda\left(\int_0^{b_1} x_{i1}\,di + \delta\int_0^{b_2} x_{i2}\,di\right) \\
&= \frac{\sigma}{1-\sigma}\left(b_1(1+\tau_1)^{1-\frac{1}{\sigma}} + \int_{b_1}^1 c(i)^{1-\frac{1}{\sigma}}\,di\right) + (\tau_1-\lambda)b_1(1+\tau_1)^{-\frac{1}{\sigma}} - \delta\sigma f(b_1-b_2) \\
&\quad + \delta\left[\frac{\sigma}{1-\sigma}\left(b_2(1+\tau_2)^{1-\frac{1}{\sigma}} + \int_{b_2}^1 c(i)^{1-\frac{1}{\sigma}}\,di\right) + (\tau_2-\lambda)b_2(1+\tau_2)^{-\frac{1}{\sigma}}\right].
\end{aligned}
$$

As in the main model, this is maximized with $\tau_1 = \tau_2 = \lambda$ and $s = 0$.

Notice that

$$
\frac{dV}{d\tau} = b\frac{(1+\tau)^{-1-\frac{1}{\sigma}}}{\sigma}(\sigma(1-\sigma)(1+\tau)-\tau),
$$

so $\frac{dV}{d\tau} > 0$ iff $\tau < \frac{\sigma(1-\sigma)}{1-\sigma(1-\sigma)}$. This means that after tax rebates every consumer benefits from a carbon tax as long as it is not too large. Of course, polluting districts dislike carbon taxes because the negative effect on their profits dominates. But carbon taxes only affect green districts as consumers (i.e., through V), and therefore those districts support increases in the carbon tax (after rebates). Therefore, the distributive tension is exactly as in the main model. In the second period, if a majority of districts is green then a green proposer can implement an optimal carbon tax (as long as λ is not too large), but if a majority is still polluting then polluters block any proposal that increases the carbon taxes.

To implement climate policy, the policymaker needs to build a winning coalition in the first period, which requires approval by green districts and by the median (polluting) district. The political constraints are thus

$$
V(\tau_1, b_1) + \delta \mathrm{E} V(\tau_2, b_2) - \delta\sigma(s+\sigma f)(b_1-b_2) \geqslant (1+\delta)V(0, b_1)
$$

for green districts, and

$$
\begin{aligned}
&V(\tau_1, b_1) + \delta \mathrm{E} V(\tau_2, b_2) + \sigma(1+\tau_1)^{1-\frac{1}{\sigma}} \\
&\quad + \delta\sigma\left[-(1-\sigma)f + s + c\left(\frac{1}{2}\right)^{1-\frac{1}{\sigma}} - (s+\sigma f)(b_1-b_2)\right] \geqslant (1+\delta)(V(0,b_1)+\sigma)
\end{aligned}
$$

for polluting districts. As in the main model, the optimal policy is not feasible. Polluters demand a subsidy (up to some level) and resist first-period carbon taxes unless they are compensated. Therefore, accommodating the pivotal polluters creates an incentive to raise s and reduce τ_1 from the Pigouvian level, as in Proposition 1.

A.6. *Endogenous Reelection Probability*

In this section I make the probability ρ that the green proposer is reelected in the second period endogenous. I follow the standard approach to probabilistic voting covered in, e.g., Persson and Tabellini (2000, Ch. 3). A continuum of voters j in each district $i \in I$ votes for the green proposer if $V_i^G \geqslant V_i^B + \sigma_{ij} + \nu$, where V_i^P is the expected payoff if P has agenda-setting power in the second period, and σ_{ij}, ν are individual- and population-level preference biases (capturing differences on dimensions orthogonal to climate policy). We assume that σ_{ij} are independent and uniformly distributed on $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$, and ν is independent of $(\sigma_i)_{i\in I}$ and uniformly distributed on $\left[\nu_0 - \frac{1}{2\phi}, \nu_0 + \frac{1}{2\phi}\right]$. Therefore, conditional on ν, the vote share for G is $\int_I \int 1_{\sigma_{ij} \leqslant V_i^G - V_i^B - \nu}\, dj\, di = \int_I \left(\frac{1}{2} + \psi(V_i^G - V_i^B - \nu)\right) di = \frac{1}{2} + \psi(V^G - V^B - \nu)$, where $V^P = \int_I V_i^P\, di$. Here I ignore measurability issues and assume that $\psi > 0$ is small enough. Hence, G wins iff this number is greater than $\frac{1}{2}$, which happens iff $V^G - V^B - \nu \geqslant 0$, and this is the case with probability $\phi(V^G - V^B - \nu_0 + \frac{1}{2\psi}) = \rho_0 + \phi(V^G - V^B)$, where $\rho_0 = \frac{1}{2} - \psi\nu_0$, and $\phi > 0$ and $|\nu_0|$ are assumed to be small enough. In sum, the probability that G is reelected is

$$\rho = \rho_0 + \phi(V^G - V^B) = \rho_0 - \phi\frac{\delta b_2}{2}\max\{\lambda^2 - \tau_1^2, 0\}.$$

In the main model I studied the case $\phi = 0$. Here I am interested in determining the effect of increasing ϕ, which can be understood as the salience of the climate issue, by a small amount. Notice that, since voters are assumed to not internalize the climate externality, the average voter is against climate policy, and, therefore, increasing the salience of climate policy reduces the chances of victory of the green party as long as $\tau_1 < \lambda$, i.e., if there is a difference in second-period policy between the parties. The green proposer can increase their chances of reelection by reducing the share of polluting districts in the second period (b_2), which can be achieved by providing a more generous subsidy s, or by increasing the first-period carbon tax τ_1, thus reducing the difference in expected policies between the parties (due to the fact that in the second period a majority is pro-climate and can impose the costs of climate policy on the polluting minority).

Notice that a large first-period carbon tax would intuitively lead to backlash against the green party. In this model, however, a carbon tax in the first period is only possible in the context of a policy sequence. And, in that context, voters realize that the carbon tax has been locked in, and therefore the anti-climate party cannot credibly promise to eliminate it. Two strong assumptions drive this conclusion: first, the first-period carbon tax becomes the status quo in the second period, and, second, the second-period proposer cannot replace the carbon tax with a more efficient form of redistribution from the polluting minority to the green majority (more efficient from their point of view, i.e., ignoring climate damages).

If we ignore the political constraints, how does ϕ (the electoral consequences of climate policy) affect the optimal τ_1 and s? We have

$$\frac{\partial \Delta W}{\partial \rho} = \delta b_2 \left[\frac{\lambda^2}{2} - \tau_1 \left(\lambda - \frac{\tau_1}{2} \right) \right] - \frac{1}{A}(s - s^*)(\lambda - \tau_1)(2 - \lambda - \tau_1)$$

and

$$\frac{\partial^2 \Delta W}{\partial \tau_1 \partial s} = -\frac{2}{A}(1-\rho)(1+\lambda-2\tau_1),$$

where the partial derivatives with respect to τ_1, s ignore the effect through ρ. We have

$$\begin{aligned} 0 = \frac{d\Delta W}{d\tau_1} &= \frac{\partial \Delta W}{\partial \tau_1} + \left(\frac{\partial \Delta W}{\partial \rho} + \frac{\partial \Delta W}{\partial b_2}\frac{\partial b_2}{\partial \rho} \right) \frac{\partial \rho}{\partial \tau_1} \\ &= \frac{\partial \Delta W}{\partial \tau_1} + \left(\frac{\partial \Delta W}{\partial \rho} + \frac{\partial \Delta W}{\partial b_2}\frac{\partial b_2}{\partial \rho} \right) \delta\phi \left(-b_2\tau_1 - \frac{1}{2}(\lambda^2 - \tau_1^2)\frac{\partial b_2}{\partial \tau_1} \right), \end{aligned}$$

so, differentiating with respect to ϕ, we obtain

$$\begin{aligned} 0 = \frac{\partial^2 \Delta W}{\partial^2 \tau_1}\frac{d\tau_1}{d\phi} + \frac{\partial^2 \Delta W}{\partial \tau_1 \partial s}\frac{ds}{d\phi} + \delta \left(\frac{\partial \Delta W}{\partial \rho} + \frac{\partial \Delta W}{\partial b_2}\frac{\partial b_2}{\partial \rho} \right) \left(-b_2\tau_1 - \frac{1}{2}(\lambda^2 - \tau_1^2)\frac{\partial b_2}{\partial \tau_1} \right) \\ + \delta\phi \frac{d}{d\phi} \left[\left(\frac{\partial \Delta W}{\partial \rho} + \frac{\partial \Delta W}{\partial b_2}\frac{\partial b_2}{\partial \rho} \right) \left(-b_2\tau_1 - \frac{1}{2}(\lambda^2 - \tau_1^2)\frac{\partial b_2}{\partial \tau_1} \right) \right]. \end{aligned}$$

When $\phi = 0$ we have $\tau_1 = \lambda$ and $s = s^*$, so $\frac{\partial \Delta W}{\partial \rho} = 0$ and $\frac{\partial \Delta W}{\partial b_2} = s - s^* = 0$, hence $\frac{\partial^2 \Delta W}{\partial^2 \tau_1}\frac{d\tau_1}{d\phi} + \frac{\partial^2 \Delta W}{\partial \tau_1 \partial s}\frac{ds}{d\phi} = 0$. Similarly, $\frac{\partial^2 \Delta W}{\partial^2 s}\frac{ds}{d\phi} + \frac{\partial^2 \Delta W}{\partial \tau_1 \partial s}\frac{d\tau_1}{d\phi} = 0$, which implies $\frac{d\tau_1}{d\phi} = \frac{ds}{d\phi} = 0$, because the determinant is not zero:

$$\begin{aligned} \frac{\partial^2 \Delta W}{\partial^2 \tau_1}\frac{\partial^2 \Delta W}{\partial^2 s} - \left(\frac{\partial^2 \Delta W}{\partial \tau_1 \partial s} \right)^2 &= \frac{2}{\delta A}\left[b_1 + (1-\rho)\delta b_2 + \frac{2\delta}{A}(1-\rho)^2(1-\lambda)^2 \right] \\ &\quad - \frac{4}{A^2}(1-\rho)^2(1-\lambda)^2 \\ &= \frac{2}{\delta A}[b_1 + (1-\rho)\delta b_2] > 0. \end{aligned}$$

This means that, without political constraints, a small increase in the electoral salience of climate policy has a null first-order effect on the optimal policy. The intuition is that the optimal policy is such that there is no difference between the second-period policy of the two parties. Therefore, increasing the salience at the margin doesn't significantly change the incentives.

With political constraints, if climate policy occurs in equilibrium we know from Proposition 1 that $\tau_1 < \lambda$, and therefore the parties present a meaningful difference to voters if $\phi > 0$. Figure 7 shows how an increase in salience affects equilibrium policies—we see that it increases both the first-period carbon tax τ_1 and the green subsidy s.

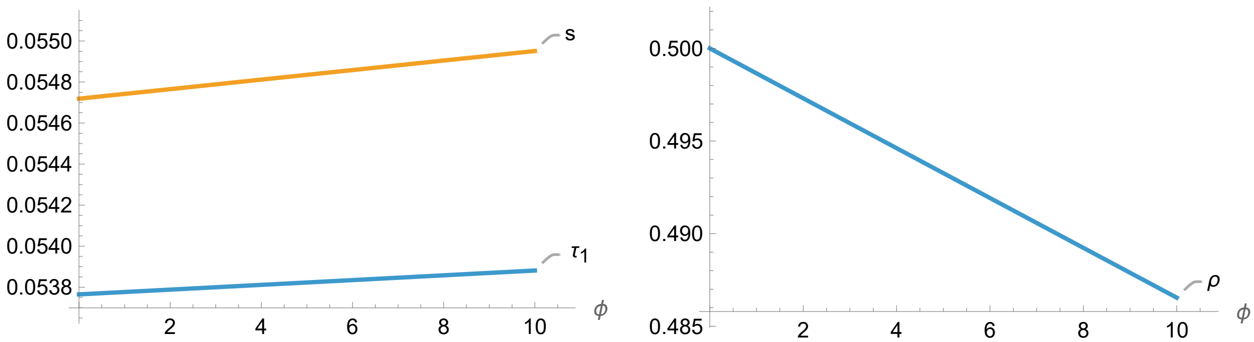


Figure 7: Equilibrium policy when $A = 1.8$, $c = 0.06$, $\delta = 0.9$, $\lambda = 0.1$ and $\rho_0 = \frac{1}{2}$, as ϕ goes from 0 to 10.

I will provide intuition for why $\phi > 0$ increases the equilibrium τ_1 in the numerical simulation. The reason is that if $\phi > 0$ then the probability that the green policymaker is reelected drops, and therefore the probability that the first-period carbon tax stays in the second period increases. This means that increasing the first-period carbon tax is now more valuable for the green policymaker. We can see this formally: when $\phi = 0$ we have

$$\begin{aligned}\frac{\partial^2 W}{\partial\phi\partial\tau_1} &= -\delta b_2(\lambda-\tau_1)\underbrace{\frac{d\rho}{d\phi}}_{<0} + \delta(1-\rho)(\lambda-\tau_1)\underbrace{\frac{db_2}{d\phi}}_{>0} \\ &\quad + \frac{2}{A}(1-\tau_1)[s-2\delta(1-\rho)(1-\tau_1)(\lambda-\tau_1)]\underbrace{\frac{d\rho}{d\phi}}_{<0} > 0\end{aligned}$$

as long as s is not too large. The first term captures the fact that increasing ϕ makes increasing τ_1 more valuable, because τ_1 stays in the second period with greater probability; the second term captures the fact that increasing ϕ, by reducing the expected τ_2, raises the number of polluters in the second period (b_2), therefore making the increase in τ_1 even more valuable. The third term comes from the fact that a large subsidy s creates a reason to reduce τ_1 in order to reduce excessive investments.

Raising τ_1 is also more valuable (or less costly) for the districts, because ϕ only affects their welfare through the expected fiscal transfer T, and a smaller ρ means that increasing τ_1 raises the expected fiscal revenue of the carbon tax more, both directly, by raising the expected tax rate τ_2, and indirectly, by reducing the share of producers that transition (i.e., raising b_2), and therefore increasing the tax base. We can see this formally: when $\phi = 0$ we have

$$\frac{\partial^2 T}{\partial\phi\partial\tau_1} = -\delta b_2(1-2\tau_1)\underbrace{\frac{d\rho}{d\phi}}_{<0} + \delta(1-\rho)(1-2\tau_1)\underbrace{\frac{db_2}{d\phi}}_{>0} + \delta[\lambda(1-\lambda) - \tau_1(1-\tau_1)]\underbrace{\frac{db_2}{d\tau_1}}_{<0}\underbrace{\frac{d\rho}{d\phi}}_{<0} > 0.$$

The first term captures the increase in the expected tax rate, the second captures the increase in the tax base, and the third one captures the fact that by reducing ϕ we lose tax revenue in expectation, and this loss is proportional to b_2, and so we reduce it by decreasing b_2, which is something that increasing τ_1 does. Since an increase in τ_1 is more valuable for both the proposer and every legislator after increasing ϕ, we should expect τ_1 to increase when we increase ϕ.

Now, notice that $\phi > 0$ also increases the value of s for the policymaker, since ρ decreases, the expected second-period carbon tax decreases, the level of investment in green technology falls below the optimal level, and therefore the need for the subsidy s grows. We see this formally: when $\phi = 0$ we have

$$\frac{\partial^2 W}{\partial\phi\partial s} = -\frac{2}{A}(1-\tau_1)(\lambda-\tau_1)\frac{d\rho}{d\phi} > 0.$$

Increasing ϕ from 0 also increases the value of s for the districts (or, rather, it decreases its cost), because a lower ρ decreases the number of districts who transition, and therefore the cost of the subsidy. In addition, a lower ρ increases the cost of having $\tau_1 < \lambda$ for the expected tax revenue; since this loss is proportional to b_2, reducing b_2 becomes more valuable, which is what increasing s does. Formally: when $\phi = 0$ we have

$$\frac{\partial^2 T}{\partial\phi\partial s} = \underbrace{\frac{db_2}{d\phi}}_{>0} + \delta[\lambda(1-\lambda) - \tau_1(1-\tau_1)]\underbrace{\frac{db_2}{ds}}_{<0}\underbrace{\frac{d\rho}{d\phi}}_{<0} > 0.$$

We see that $\phi > 0$ raises the value of s for both the policymaker and the districts, which justifies that in equilibrium we will see an increase in s.

The previous argument is not a formal proof because $\frac{\partial^2 W}{\partial\tau_1\partial s} < 0$ and $\frac{\partial^2 T}{\partial\tau_1\partial s} < 0$, which means that the incentive to raise τ_1 creates an incentive to reduce s (and vice versa), and therefore we would have to show that the incentive to raise s dominates this indirect incentive to reduce it (and the same for τ_1). Numerical simulations lend credence to the conjecture that this is always the case.

B. Proofs

B.1. *Proof of Proposition 0*

We have

$$W = \int_0^1 \pi_{i1}\,di - D_1(e_1) + \delta\left[\int_0^1 \pi_{i1}\,di - D_1(e_1+e_2)\right]$$
$$= \int_0^{b_1}\left((1-\lambda)y_{i1} - \frac{1}{2}y_{i1}^2\right)di + \int_{b_1}^1\left(y_{i1} - \frac{1}{2A_i}y_{i1}^2\right)di - c(b_1-b_2)$$
$$+\delta\left[\int_0^{b_2}\left((1-\lambda)y_{i2} - \frac{1}{2}y_{i2}^2\right)di + \int_{b_2}^1\left(y_{i2} - \frac{1}{2A_i}y_{i2}^2\right)di\right].$$

Pointwise maximization of the integrals yields $y_{it} = 1-\lambda$ for polluting i and $y_{it} = A_i$ for green i, so

$$W = \frac{b_1}{2}(1-\lambda)^2 + \frac{A}{4}(1-b_1^2) - c(b_1-b_2) + \delta\left[\frac{b_2}{2}(1-\lambda)^2 + \frac{A}{4}(1-b_2^2)\right].$$

We have $\frac{dW}{db_2} = c + \frac{\delta}{2}(1-\lambda)^2 - \frac{\delta A}{2}b_2$, so the optimal b_2 is given by $-c + \frac{\delta A}{2}b_2 = \frac{\delta}{2}(1-\lambda)^2$. Pigouvian carbon taxes $\tau_1 = \tau_2 = \lambda$ implement these choices in equilibrium, as desired.

B.2. *Proof of Observation 6*

Observation 6. *With political turnover the optimal carbon taxes are Pigouvian, and the investment subsidy is $s = \delta(1-\rho)\lambda$.*

Proof. Let $E_1 = E_0 + b_1(1-\tau_1)$ be the stock of emissions in period 1, $E_2^G = E_1 + b_2(1-\tau_2)$ be the stock of emissions in period 2 if the green party is in power, where $\tau_2 = D_2'(E_2^G)$ is the carbon tax rate set by the green party in period 2, and $E_2^B = E_1 + b_2$ be the stock of emissions in period 2 if the opposition party is in power. The objective of the green party in period 1 is

$$W_G = \frac{b_1}{2}(1-\tau_1^2) + \frac{A}{4}(1-b_1^2) - c(b_1-b_2) - D_1(E_1)$$
$$+\delta\left[\frac{b_2}{2}(1-\rho\tau_2^2) + \frac{A}{4}(1-b_2^2) - \rho D_2(E_2^G) - (1-\rho)D_2(E_2^B)\right].$$

We have

$$\frac{\partial W_G}{\partial b_2} = c + \frac{\delta}{2}(1-\rho\tau_2^2) - \frac{\delta A}{2}b_2 - \rho\delta D'(E_2^G)(1-\tau_2) + (1-\rho)\delta D_2'(E_2^B)$$
$$= c + \frac{\delta}{2}(1-\rho\tau_2^2) - \left(c - s + \frac{\delta}{2}(1-\rho(2\tau-\tau_2^2))\right) - \rho\delta\tau_2(1-\tau_2) - (1-\rho)\delta D_2'(E_2^B)$$
$$= s - (1-\rho)\delta D_2'(E_2^B).$$

Therefore,

$$\begin{aligned}
\frac{\partial W_G}{\partial s} &= \frac{\partial W_G}{\partial b_2}\frac{\partial b_2}{\partial s} = \frac{2}{\delta A}\Big((1-\rho)\delta D'(E_2^B) - s\Big),\\
\frac{\partial W_G}{\partial \tau_1} &= \Big(D_1'(E_1) + \delta \mathrm{E}\big[D_2'(E_2^P)\big] - \tau_1\Big)b_1 + \frac{\partial W_G}{\partial b_2}\frac{\partial b_2}{\partial \tau_1}.
\end{aligned}$$

The first equation implies that the optimal s is $(1-\rho)\delta D_2'(E_2^B)$. This implies $\frac{\partial W_G}{\partial b_2} = 0$, so the optimal τ_1 is $D_1'(E_1) + \delta\mathrm{E}\big[D_2'(E_2^P)\big]$. Under the assumption of linear cumulative damages we obtain $\tau_1 = \lambda$ and $s = (1-\rho)\delta\lambda$, as desired. ■

In order to relax the assumption that environmental damages are linear, I will assume that $D_t''(E) = \kappa \geqslant 0$ for all E and $t \in \{1,2\}$. We have $D_2'(E_2^B) = D_2''(\xi)(E_2^B - E_2^G) + D_2'(E_2^G) = \kappa(E_2^B - E_2^G) + D_2'(E_2^G) = \tau_2(1+\kappa b_2)$, where ξ is between E_2^G and E_2^B. Therefore, $s = (1-\rho)\delta\tau_2(1+\kappa b_2)$ and $\tau_1 = D_1'(E_1) + \delta\tau_2(1 + (1-\rho)\kappa b_2)$. Differentiating, we obtain

$$\begin{aligned}
\frac{\partial \tau_2}{\partial \rho} &= D_2''(E_2^G)\frac{\partial E_2^G}{\partial \rho} = \kappa\left[-b_1\frac{\partial \tau_1}{\partial \rho} + \frac{\partial b_2}{\partial \rho}(1-\tau_2) - b_2\frac{\partial \tau_2}{\partial \rho}\right],\\
\frac{\partial s}{\partial \rho} &= -\delta\tau_2(1+\kappa b_2) + (1-\rho)\left[\frac{\partial \tau_2}{\partial \rho}(1+\kappa b_2) + \tau_2\kappa\frac{\partial b_2}{\partial \rho}\right],\\
\frac{\partial b_2}{\partial \rho} &= -\frac{2}{\delta A}\left[\frac{\partial s}{\partial \rho} + \delta\tau_2\Big(1 - \frac{\tau_2}{2}\Big) + \delta\rho(1-\tau_2)\frac{\partial \tau_2}{\partial \rho}\right],\\
\frac{\partial \tau_1}{\partial \rho} &= -\kappa b_1\frac{\partial \tau_1}{\partial \rho} + \delta\frac{\partial \tau_2}{\partial \rho}(1 + (1-\rho)\kappa b_2) - \delta\tau_2\kappa b_2 + \delta\tau_2(1-\rho)\kappa\frac{\partial b_2}{\partial \rho}.
\end{aligned}$$

Taking $\rho = 1$, we obtain $\frac{\partial s}{\partial \rho} = -\delta\tau_2(1+\kappa b_2)$, so

$$\frac{\partial b_2}{\partial \rho} = \frac{2}{A}\left[\tau_2(1+\kappa b_2) - \tau_2\Big(1-\frac{\tau_2}{2}\Big) - (1-\tau_2)\frac{\partial \tau_2}{\partial \rho}\right] = \frac{2}{A}\left[\tau_2\Big(\kappa b_2 + \frac{\tau_2}{2}\Big) - (1-\tau_2)\frac{\partial \tau_2}{\partial \rho}\right].$$

Using this,

$$\begin{aligned}
\frac{\partial \tau_2}{\partial \rho} &= \kappa\left\{-b_1\frac{\partial \tau_1}{\partial \rho} + \frac{2}{A}\left[\tau_2\Big(\kappa b_2 + \frac{\tau_2}{2}\Big) - (1-\tau_2)\frac{\partial \tau_2}{\partial \rho}\right](1-\tau_2) - b_2\frac{\partial \tau_2}{\partial \rho}\right\}\\
&= \frac{\kappa}{1 + \kappa b_2 + \kappa\frac{2}{A}(1-\tau_2)^2}\left[-b_1\frac{\partial \tau_1}{\partial \rho} + \frac{2}{A}\tau_2(1-\tau_2)\Big(\kappa b_2 + \frac{\tau_2}{2}\Big)\right].
\end{aligned}$$

Therefore, $\frac{\partial \tau_1}{\partial \rho} = -\kappa b_1\frac{\partial \tau_1}{\partial \rho} + \delta\frac{\partial \tau_2}{\partial \rho} - \delta\tau_2\kappa b_2$, so

$$\left[1 + \kappa b_1 + \frac{\delta\kappa b_1}{1+\kappa b_2 + \kappa\frac{2}{A}(1-\tau_2)^2}\right]\frac{\partial \tau_1}{\partial \rho} = -\delta\tau_2\kappa\left[b_2 - \frac{2}{A}\Big(\kappa b_2 + \frac{\tau_2}{2}\Big)\frac{1}{1+\kappa b_2 + \kappa\frac{2}{A}(1-\tau_2)^2}\right].$$

If b_2 is small enough, the RHS is positive, so $\frac{\partial \tau_1}{\partial \rho} > 0$, which implies that an increase in the probability of turnover, i.e., a drop in ρ, leads to a reduction in the first-period carbon tax. However, if τ_2 is small, the RHS is negative, so $\frac{\partial \tau_1}{\partial \rho} < 0$ and we obtain the opposite conclusion. Thus, the effect of turnover on the first-period carbon tax is ambiguous, as claimed in the text.

B.3. *Proof of Lemma 1*

We start with the second period. Suppose that first-period policy has been (τ_1, s, T_1, B), where B is public debt or savings, taken to balance the budget,

$$\underbrace{b_1\tau_1(1-\tau_1)}_{\text{carbon tax revenue}} - \underbrace{s(b_1-b_2)}_{\text{cost of subsidy}} - \underbrace{T_1}_{\substack{\text{taxes or}\\ \text{transfers}}} + \underbrace{B}_{\substack{\text{debt or}\\ \text{savings}}} = 0. \tag{5}$$

In the second period, the policymaker proposes a carbon tax τ_2, and a uniform lump-sum tax or transfer $T_2(\tau_2)$ (that includes rebates) is implemented automatically to balance the budget. Thus,

$$\underbrace{b_2\tau_2(1-\tau_2)}_{\text{carbon tax revenue}} - \underbrace{T_2(\tau_2)}_{\substack{\text{taxes or}\\ \text{transfers}}} - \underbrace{\delta^{-1}B}_{\substack{\text{debt plus}\\ \text{interest}}} = 0. \tag{6}$$

Districts accept the proposal if they prefer it to the status quo, τ_1. For green districts, this happens iff

$$\frac{1}{2}Ai + T_2(\tau_2) \geqslant \frac{1}{2}Ai + T_2(\tau_1),$$

i.e., iff $b_2\tau_2(1-\tau_2) - \delta^{-1}B \geqslant b_2\tau_1(1-\tau_1) - \delta^{-1}B$, i.e., iff $\tau_2 \geqslant \tau_1$ as long as $\tau_1 + \tau_2 \leqslant 1$. Polluting districts accept the proposal iff

$$\frac{1}{2}(1-\tau_2)^2 + T_2(\tau_2) \geqslant \frac{1}{2}(1-\tau_2)^2 + T_2(\tau_1),$$

i.e., iff $\tau_2 \leqslant \tau_1$. The ideal carbon tax of a green policymaker is λ. Therefore, if $b_1 \leqslant \frac{1}{2}$, i.e., green districts form a majority in the legislature, the equilibrium carbon tax will be $\tau_2 = \max\{\tau_1, \lambda\}$. Otherwise, polluting districts form a majority, and the equilibrium carbon tax will be $\tau_2 = \min\{\tau_1, \lambda\}$. The ideal carbon tax of a green policymaker is 0, so $\tau_2 = \tau_1$ if $b_2 \leqslant \frac{1}{2}$, and $\tau_2 = 0$ otherwise.

If the first period policy fails, $\tau_1 = s = T_1 = 0$, and $b_2 > \frac{1}{2}$, $\tau_2 = 0$. In this case, a polluting district $i \in [0, b_1)$ upgrades its capital in the first period iff the cost of the investment plus the expected profits in the green sector is greater than the expected profits in the polluting sector, i.e., iff $-c + \frac{\delta}{2}Ai \geqslant \frac{\delta}{2}(1-\tau_2)^2 = \frac{\delta}{2}$, which reduces to $i \geqslant b_1$. Thus, in this case no polluting district transitions, and $b_2 = b_1$. Since $b_1 > \frac{1}{2}$ by assumption, $b_2 > \frac{1}{2}$, so this is an equilibrium of the proposal failure subgame. We call $\tilde{\tau}_2 = 0$ and $\tilde{b}_2 = b_1$ the carbon tax rate and the share of polluting districts in this equilibrium.

A polluting district i accepts a first-period policy (τ_1, s, T_1) if they prefer it to business as usual in both periods, i.e., iff

$$\underbrace{\frac{1}{2}(1-\tau_1)^2}_{\text{period-1 profit}} + \max\Big\{\underbrace{s - c + \frac{\delta}{2}Ai}_{\substack{\text{period-2 profit in}\\ \text{the green sector plus}\\ \text{net cost of transition}}},\ \underbrace{\frac{\delta}{2}\mathrm{E}\big[(1-\tau_2)^2\big]}_{\substack{\text{expected}\\ \text{period-2 profit in}\\ \text{the polluting sector}}}\Big\} + \underbrace{T_1 + \delta T_2}_{\text{expected transfers}} \geqslant \underbrace{\frac{1}{2} + \frac{\delta}{2}}_{\substack{\text{business as}\\ \text{usual profits}}}. \tag{7}$$

Notice that the LHS is weakly increasing in i, so if i accepts, every district $j \geqslant i$ accepts as well. Therefore, to get a majority to approve the policy, the policymaker has two options. They can create a coalition with the polluting districts in $[\frac{1}{2}, b_1)$ and the green districts, $[b_1, 1]$, or a purely polluting coalition $[b_1 - \frac{1}{2}, b_1)$. In any case, the median district, $i = \frac{1}{2}$, has to approve the proposal for it to be implemented.

Suppose that in equilibrium the median district does not transition. In that case, (7) is

$$\frac{1}{2}(1-\tau_1)^2 + \frac{\delta}{2}\mathrm{E}\big[(1-\tau_2)^2\big] + T \geqslant \frac{1}{2} + \frac{\delta}{2},$$

where $T = T_1 + \delta T_2 = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2)$, obtained by summing (5) and (6). Now, it is straightforward to verify that this condition cannot hold unless $\tau_1 = \tau_2 = s = 0$. Therefore, if the equilibrium policy is not business as usual, then (7) holds for the median district, which implies that this district transitions.

A green-polluting coalition requires (7) to hold for $i = \frac{1}{2}$, i.e., that the median district approves, and that green districts prefer the policy proposal to business as usual, which happens iff $T_1 + \delta T_2 \geqslant 0$, i.e., if the expected transfers are nonnegative. In other words, green districts may tolerate a tax in the present as long as it is compensated by a transfer in the future. Thus, a green-polluting coalition implements a non-BAU policy (τ_1, s) iff

$$\frac{1}{2}(1-\tau_1)^2 + s - c + \delta\frac{A}{4} + T \geqslant \frac{1}{2} + \frac{\delta}{2} \tag{PC$_B$}$$

and

$$T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2) \geqslant 0, \tag{PC$_G$}$$

where $\tau_2 = \max\{\tau_1, \lambda\}$ if the green policymaker stays in power, which happens with probability ρ, and is $\tau_2 = \tau_1$ in case of turnover, which occurs with probability $1-\rho$. The conditions PC$_B$ and

PC_G are the *political constraints* for the polluting and green districts, as desired.[12]

B.4. *Proof of Proposition 1*

Let

$$P = \frac{1}{2}(1-\tau_1)^2 + s - c + \delta\frac{A}{4} + T - (1+\delta)\frac{1}{2},$$
$$T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2),$$

so PC_B is $P \geqslant 0$, and PC_G is $T \geqslant 0$. Recall that $-c + \frac{\delta A}{2}b_1 = \frac{\delta}{2}$, $s - c + \frac{\delta A}{2}b_2 = \frac{\delta}{2}\mathrm{E}\big[(1-\tau_2)^2\big]$, $\tau_2 = \max\{\tau_1, \lambda\}$ with probability ρ, $\tau_2 = \tau_1$ with probability $1-\rho$, and

$$\Delta W = b_1\tau_1\left(\lambda - \frac{1}{2}\tau_1\right) + \delta b_2 \mathrm{E}\left[\tau_2\left(\lambda - \frac{1}{2}\tau_2\right)\right] + \delta\lambda(b_1 - b_2) - \frac{\delta A}{4}(b_1 - b_2)^2. \tag{8}$$

Observation. *In equilibrium, $\tau_1 \leqslant \frac{1}{2}$.*

Proof. If $\tau_1 > \frac{1}{2}$ then $\tau_1 > \lambda$, hence $\tau_2 = \tau_1$. We have $\frac{\partial T}{\partial \tau_1} = (b_1 + \delta b_2)(1-2\tau_1) - \frac{1}{A}(s + \delta\tau_1(1-\tau_2))(1-\tau_1) < 0$, $\frac{\partial P}{\partial \tau_1} = -(1-\tau_1) + \frac{\partial T}{\partial \tau_1} < 0$ and $\frac{\partial W}{\partial \tau_1} = (b_1 + \delta b_2 + \frac{\delta}{A}(1-\tau_1)^2)(\lambda - \tau_1) - \frac{s}{A}(1-\tau_1) < 0$. So the proposer can reduce τ_1, improving their objective and still satisfying the constraints, a contradiction. ■

Let $\tilde{s}$ be given by $P|_{s=\tilde{s}} = T|_{s=\tilde{s}}$, i.e., $\tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}(b_1 - \frac{1}{2}) > 0$.

Claim 1. *If $P, T \geqslant 0$ at (τ_1, s) then $P|_{s=\tilde{s}} \geqslant 0$.*

Proof. We have $\tilde{s} > 0$, $\frac{\partial^2 T}{\partial^2 s} = \frac{\partial^2 P}{\partial^2 s} = -\frac{4}{\delta A} < 0$ and $\frac{\partial T}{\partial s}\big|_{s=0} = -\frac{1}{A}\mathrm{E}[(4-3\tau_2)\tau_2] \leqslant 0$, so $\frac{\partial T}{\partial s}\big|_{s=\tilde{s}} < 0$. If $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} \geqslant 0$ we are done, because if $P|_{s=\tilde{s}} < 0$ then $P < 0$ for $s \leqslant \tilde{s}$, and $T < 0$ for $s > \tilde{s}$, so $P \geqslant 0$ and $T \geqslant 0$ cannot happen simultaneously, a contradiction. So it's enough to show that $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} \geqslant 0$. Notice, first, that $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}}$ is linear in ρ, so it's enough to prove $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} \geqslant 0$ for $\tau_2 = \max\{\lambda, \tau_1\}$ and $\tau_2 = \tau_1$.

We have $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} = 2(1-b_1) - \frac{1}{\delta A}[2(2-\tau_1)\tau_1 + \delta(4-3\tau_2)\tau_2]$. If $\tau_2 = 0$ then $\tau_1 = 0$, and $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} \geqslant 0$. So we can assume $\tau_2 > 0$. Suppose that $\frac{\partial P}{\partial s}\big|_{s=\tilde{s}} < 0$. We must have

$$A < \tilde{A} = \frac{2(2-\tau_1)\tau_1 + \delta(4-3\tau_2)\tau_2}{2\delta(1-b_1)}.$$

[12] A purely polluting coalition requires approval by $i = b_1 - \frac{1}{2}$, and so it implements a non-BAU policy (τ_1, s) iff

$$\frac{1}{2}(1-\tau_1)^2 + s - c + \frac{\delta}{2}A\left(b_1 - \frac{1}{2}\right) + T \geqslant \frac{1}{2} + \frac{\delta}{2}. \tag{PC_{BB}}$$

Therefore, in equilibrium (τ_1, s) has to satisfy the political constraints PC_B and PC_G, or PC_{BB}. I will restrict attention to PC_B and PC_G, because in general PC_{BB} is much more restrictive, and analyzing the conditions under which it is not doesn't provide further insights.

Let $\overline{P} = \max_{s' \in \mathbb{R}} P|_{s=s'}$. I will show that if $1 < A < \tilde{A}$ then $\overline{P} < 0$, which contradicts that $P \geqslant 0$ for some s.

We have

$$\overline{P} = -(\tau_1 + \delta\tau_2)(1 - b_1) - \left(b_1 - \frac{1}{2}\right)\tau_1^2 + \frac{\delta}{8}(A + 2\tau_2^2)(3 - 4b_1) + \frac{\delta\tau_2^4}{8A}. \tag{$*$}$$

Notice that if $b_1 < \frac{3}{4}$ then $\overline{P}$ is strictly convex in A, in which case it's maximized at either $A = 1$ or $A = \tilde{A}$, and if $b_1 \geqslant \frac{3}{4}$ then it's decreasing, in which case it's maximized at $A = 1$. So we only have to verify that $\overline{P}|_{A=1} \geqslant 0$ and $\overline{P}|_{A=\tilde{A}} \geqslant 0$.

Let $\tilde{P} = \overline{P}|_{A=\tilde{A}}$. By inspecting ($*$) it's clear that $\tilde{P}$ is concave in b_1, since the only nonlinear term in b_1 is $\propto \frac{3-4b_1}{1-b_1}$, whose second derivative is $\frac{-2}{(1-b_1)^3} < 0$. Now,

$$\frac{d\tilde{P}}{db_1} = \tau_1 + \delta\tau_2 - \tau_1^2 - \frac{\delta}{2}(\tilde{A} + 2\tau_2^2) + \frac{\delta}{8}\frac{\tilde{A}}{1 - b_1}(3 - 4b_1) - \frac{\delta\tau_2^4}{8\tilde{A}(1 - b_1)},$$

so

$$\left.\frac{d\tilde{P}}{db_1}\right|_{b_1=\frac{1}{2}} = -\frac{1}{2}\tau_1^2 - \frac{\delta}{4}\tau_2^2 - \frac{\delta^2\tau_2^4}{4(2(2 - \tau_1)\tau_1 + \delta(4 - 3\tau_2)\tau_2)} < 0.$$

This shows that $\tilde{P}$ is decreasing in b_1. Now,

$$\tilde{P}|_{b_1=\frac{1}{2}} = -\frac{1}{4}\tau_1^2 - \frac{\delta}{8}\tau_2^2 + \frac{\delta^2\tau_2^4}{8(2(2 - \tau_1)\tau_1 + \delta(4 - 3\tau_2)\tau_2)} < 0,$$

as we can easily verify. This proves that $\tilde{P} < 0$, as desired.

We have to prove that $\overline{P}|_{A=1} < 0$. Now, $\overline{P}$ is linear in δ, so it's enough to check this for $\delta = 0$, which is clear, and $\delta = 1$, so assume $\delta = 1$. Since $\tilde{A} > 1$, we have that $b_1 > \underline{b}_1$, where $\underline{b}_1$ is such that $\tilde{A}|_{b=\underline{b}_1} = 1$. Since $\overline{P}|_{A=1}$ is linear in b_1, it's enough to check the inequality for $b_1 = \underline{b}_1$ and $b_1 = 1$. In the first case, $\overline{P}|_{A=1,b_1=\underline{b}_1} = \overline{P}|_{A=\tilde{A},b_1=\underline{b}_1} = \tilde{P}|_{b_1=\underline{b}_1} < 0$ by the above. In the second case, $\overline{P}|_{A=1,b_1=1} = -\frac{1}{2}\tau_1^2 - \frac{\delta}{8}(1 + 2\tau_2^2 - \tau_2^4) < 0$, as desired. ■

Claim 2. *If $\tau_1 > \lambda$ and $P \geqslant 0$ then $\frac{\partial P}{\partial \tau_1} < 0$.*

Proof. Suppose that $\frac{\partial P}{\partial \tau_1} \geqslant 0$. We have $\tau_2 = \tau_1$ and

$$0 \leqslant \frac{\partial P}{\partial \tau_1} = -(1 - \tau_1) + (b_1 + \delta b_2)(1 - 2\tau_1) - \frac{2}{A}(1 - \tau_1)[s + \delta\tau_1(1 - \tau_1)].$$

This implies that $b_1 \geqslant \underline{b}_1$, where $\underline{b}_1$ is given by $\frac{\partial P}{\partial \tau_1}\big|_{b_1=\underline{b}_1} = 0$. We have

$$\underline{b}_1 = \frac{A(1-\tau_1) + 2(2-3\tau_1)s + \delta(4-9\tau_1+4\tau_1^2)\tau_1}{A(1+\delta)(1-2\tau_1)},$$

and $\underline{b}_1 \leqslant b_1 \leqslant 1$, so $2(2-3\tau_1)s + \delta(4-9\tau_1+4\tau_1^2)\tau_1 \leqslant A[\delta(1-2\tau_1)-\tau_1]$, which implies $\delta(1-2\tau_1) \geqslant \tau_1$.

Suppose that $\frac{\partial P}{\partial b_1} = -\frac{\delta A}{2} + (1+\delta)\tau_1(1-\tau_1) > 0$. In this case $P \leqslant P|_{b_1=1} \leqslant \max_{s'\in\mathbb{R}} P|_{b_1=1,s=s'} = -\frac{1}{8A}(\delta(A^2-\tau_1^4) + 2(2+\delta)A\tau_1) < 0$, contradiction. Hence $\frac{\partial P}{\partial b_1} \leqslant 0$, and $P \leqslant P|_{b_1=\underline{b}_1}$, so it's enough to prove that $P|_{b_1=\underline{b}_1} < 0$. Let $\overline{P} = A \max_{s'\in\mathbb{R}} P|_{b_1=\underline{b}_1,s=s'}$. The plan is to show that $\delta(1-2\tau_1) \geqslant \tau_1$, i.e., $\tau_1 \leqslant \frac{\delta}{1+2\delta}$ implies $\frac{\partial^2\overline{P}}{\partial^2 A} \leqslant 0$, $\frac{\partial\overline{P}}{\partial A}|_{A=1} \leqslant 0$ and $\overline{P}|_{A=1} < 0$, which shows that $\overline{P} < 0$, and therefore $P \leqslant P|_{b_1=\underline{b}_1} \leqslant \overline{P} < 0$, contradiction.

We have $\frac{\partial^2\overline{P}}{\partial^2 A} \propto -1 - 2\delta + 3\delta^2 + \tau_1(2\delta - 10\delta^2) + \tau_1^2(4 + 4\delta + 9\delta^2)$, which is convex in τ_1, so it's maximized at either $\tau_1 = 0$ or $\tau_1 = \frac{\delta}{1+2\delta}$. In the first case $\frac{\partial^2\overline{P}}{\partial^2 A} \propto -(1+2\delta)(1-\delta^2) \leqslant 0$, and in the second case $\frac{\partial^2\overline{P}}{\partial^2 A} \propto -\left[1 + 6\delta + 2\delta^2(1-\delta) + \delta^2(1-\delta^2)\right] < 0$, so $\frac{\partial^2\overline{P}}{\partial^2 A} \leqslant 0$, as desired.

We have $\frac{\partial\overline{P}}{\partial A}|_{A=1} \propto f(\delta) \equiv -4\tau_1^3 + 2\tau_1^2 + \delta(-10\tau_1^3 + 9\tau_1^2 - 1) + \delta^2(16\tau_1^4 - 49\tau_1^3 + 40\tau_1^2 - 6\tau_1 - 2) + \delta^3(16\tau_1^4 - 43\tau_1^3 + 42\tau_1^2 - 18\tau_1 + 3)$. Now $16\tau_1^4 - 49\tau_1^3 + 40\tau_1^2 - 6\tau_1 - 2 = 16x^4 + \frac{65}{3}x^3 + \frac{29}{3}x^2 + \frac{53}{27}x + \frac{22}{81} > 0$ where $x = \frac{1}{3} - \tau_1$ since $\tau_1 \leqslant \frac{\delta}{1+2\delta} \leqslant \frac{1}{3}$. Therefore $f'(\delta)$ is a convex quadratic, and $f'(0) = -(1-2\tau_1)(10\tau_1^2 + 24\tau_1 x + 9x^2) < 0$. Hence, f' has exactly one nonnegative root δ^*, and $f''(\delta^*) > 0$, so f has a local minimum at δ^*, which implies that f is maximized at either $\delta = \frac{\tau_1}{1-2\tau_1}$ or $\delta = 1$. We have $f(\frac{\tau_1}{1-2\tau_1}) \propto -97\tau_1^6 - 999\tau_1^5 x - 3834\tau_1^4 x^2 - 7128\tau_1^3 x^3 - 6885\tau_1^2 x^4 - 3402\tau_1 x^5 - 729x^6 < 0$ and $f(1) = -\tau_1(1-2\tau_1)(97\tau_1^2 + 297\tau_1 x + 216x^2) < 0$. This shows that $f < 0$, and $\frac{\partial\overline{P}}{\partial A}|_{A=1} < 0$, as desired.

We have $\overline{P}|_{A=1} \propto f(\delta) \equiv 4\tau_1^2 - 8\tau_1^3 + (-1 + 30\tau_1^2 - 100\tau_1^3 + 145\tau_1^4 - 112\tau_1^5 + 32\tau_1^6)\delta + (-2 - 14\tau_1 + 108\tau_1^2 - 258\tau_1^3 + 322\tau_1^4 - 224\tau_1^5 + 64\tau_1^6)\delta^2 + (3 - 26\tau_1 + 91\tau_1^2 - 166\tau_1^3 + 177\tau_1^4 - 112\tau_1^5 + 32\tau_1^6)\delta^3$. Now, the leading coefficient is equal to $\frac{47}{729} + \frac{46}{81}x + \frac{67}{9}x^2 + \frac{830}{27}x^3 + \frac{131}{3}x^4 + 48x^5 + 32x^6 > 0$ where $x = \frac{1}{3} - \tau_1$. Therefore $f'(\delta)$ is a convex quadratic, and $f'(\frac{\tau_1}{1-2\tau_1}) \propto -1433\tau_1^8 - 17016\tau_1^7 x - 82062\tau_1^6 x^2 - 212814\tau_1^5 x^3 - 327645\tau_1^4 x^4 - 308124\tau_1^3 x^5 - 172773\tau_1^2 x^6 - 52488\tau_1 x^7 - 6561x^8 < 0$. Hence f' has exactly one nonnegative root $\delta^* \geqslant \frac{\tau_1}{1-2\tau_1}$, and $f''(\delta^*) > 0$, so f is maximized at either $\delta = \frac{\tau_1}{1-2\tau_1}$ or $\delta = 1$. Now, $f(\frac{\tau_1}{1-2\tau_1}) \propto -631\tau_1^8 - 7332\tau_1^7 x - 34992\tau_1^6 x^2 - 91422\tau_1^5 x^3 - 146691\tau_1^4 x^4 - 152118\tau_1^3 x^5 - 101331\tau_1^2 x^6 - 39366\tau_1 x^7 - 6561x^8 < 0$, and $f(1) \propto -631\tau_1^6 - 5952\tau_1^5 x - 21258\tau_1^4 x^2 - 36072\tau_1^3 x^3 - 29727\tau_1^2 x^4 - 9720\tau_1 x^5 < 0$, so $f < 0$, and $\overline{P}|_{A=1} < 0$, as desired. ∎

Claim 3. *If $T = 0$ in equilibrium then $\tau_1 \leqslant \lambda$.*

Proof. Suppose that $\tau_1 > \lambda$, so $\tau_2 = \tau_1$. Let $\tilde{s}$ be given by $P|_{s=\tilde{s}} = T|_{s=\tilde{s}}$, i.e., $\tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}(b_1 - \frac{1}{2})$, and $\tilde{P} = P|_{s=\tilde{s}}$. Notice that $0 \leqslant P = s - \tilde{s} + T = s - \tilde{s}$, so $\tilde{s} \leqslant s$. We have

$\frac{\partial \Delta W}{\partial s} = -\frac{2}{\delta A}[s + \delta(\tau_1 - \lambda)(1-\tau_1)] < 0$. Therefore, we can reduce s to $\tilde{s}$, keeping $P, T \geqslant 0$ by Claim 1, and improving W. If $\tilde{P} > 0$ or $\frac{\partial \tilde{P}}{\partial \tau_1} < 0$ then we can decrease τ_1 a little and take $s = \tilde{s}$, since $\tilde{P} \geqslant 0$ is still the case, which implies that (τ_1, s) is feasible, and the effect on $\widetilde{\Delta W} = \Delta W|_{s=\tilde{s}}$ is positive, since

$$\frac{\partial \widetilde{\Delta W}}{\partial \tau_1} = \frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s}\frac{d\tilde{s}}{d\tau_1} = \frac{\partial \Delta W}{\partial \tau_1} - \frac{2}{\delta A}[\tilde{s} + \delta(\tau_1 - \lambda)(1-\tau_1)](1-\tau_1) < 0,$$

so the resulting policy is better, a contradiction. The remaining case is $\tilde{P} = 0$ and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$. In this case I will show that $\widetilde{\Delta W} < 0$. This shows that the proposer would rather choose the status quo policy $\tau_1 = s = 0$, a contradiction.

Suppose that $\widetilde{\Delta W} \geqslant 0$. Using (8) we obtain $\frac{\partial \Delta W}{\partial \lambda} = (b_1 + \delta b_2)\tau_1 + \delta(b_1 - b_2) \geqslant 0$, so $\widetilde{\Delta W} \leqslant \overline{\Delta W} \equiv \widetilde{\Delta W}|_{\lambda=\tau_1}$, and thus $\overline{\Delta W} \geqslant 0$. We have

$$\overline{\Delta W} = (b_1 + \delta b_2)\frac{\tau_1^2}{2} + \delta\tau_1(b_1 - b_2) - \frac{\delta A}{4}(b_1 - b_2)^2 \tag{9}$$

where $b_2 = \frac{1}{2} - \frac{2}{\delta A}(1+\delta)(\tau_1 - \frac{1}{2}\tau_1^2)$.

I will prove now that if $\tilde{P} = 0$ and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$ then τ_1 is the first positive root of $\tilde{P}$. First, I'll show show that $\tilde{P}$ is single-peaked in τ_1. We have $\frac{\partial^4 \tilde{P}}{\partial^4 \tau_1} = -\frac{12}{\delta A}(1 + 3\delta + 2\delta^2) < 0$ so $\frac{\partial^3 \tilde{P}}{\partial^3 \tau_1} \geqslant \frac{\partial^3 \tilde{P}}{\partial^3 \tau_1}|_{\tau_1 = \frac{1}{2}} = \frac{6}{\delta A}(1+\delta)^2 > 0$. Now, $\frac{\partial^2 \tilde{P}}{\partial^2 \tau_1}|_{\tau_1=0} = -(1 - \delta b_1) - \frac{3}{2}\delta - \frac{4}{\delta A}(1+\delta)^2 < 0$, so $\tilde{P}(\tau_1)$ is concave when $\tau_1 \in (0, \overline{\tau}_1]$ and convex for $\tau_1 \in (\overline{\tau}_1, \frac{1}{2}]$ for some $\overline{\tau}_1 \leqslant \frac{1}{2}$. Now, $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\tau_1=\frac{1}{2}} = -\frac{1}{4}(2+\delta)(2b_1 - 1) - \frac{1}{4\delta A}(3 + 4\delta + \delta^2) < 0$, so τ_1 must be decreasing in $(\overline{\tau}_1, \frac{1}{2}]$, which implies that $\tilde{P}$ is single peaked, as desired. This implies that $\tilde{P}(\tau_1)$ has at most two roots. Now, $\tilde{P}|_{\tau_1=0} = -\frac{1}{8}\delta A(2b_1 - 1)^2 < 0$. So, if $\tilde{P} = 0$ and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$ then τ_1 is the first positive root of $\tilde{P}$, as desired.

Now,

$$\tilde{P} = (b_1 + \delta b_2)\tau_1(1-\tau_1) - \left(\tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right)\right)(b_1 - b_2), \tag{10}$$

so $\frac{\partial \tilde{P}}{\partial b_1} = -\frac{\delta A}{2}(b_1 - b_2) - \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right) - \tau_1^2 < 0$. Therefore, if we decrease b_1 keeping $b_1 > \frac{1}{2}$, $\tilde{P} > 0$, but $\tilde{P}|_{\tau_1=0} < 0$, so there is still $\tau_1(b_1)$ with $\tilde{P} = 0$ and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$, i.e., a first positive root, and $\tau_1(b_1)$ is a continuous function. Suppose that $\overline{\Delta W} > 0$. If we decrease b_1, and we take $\tau_1 = \tau_1(b_1)$, then by continuity we either find b_1 with $\overline{\Delta W} = 0$, or in the limit we get to $b_1 = \frac{1}{2}$ with $\tilde{P} = 0$, $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$ and $\overline{\Delta W} \geqslant 0$.

Suppose that $\overline{\Delta W} = 0$. Using (10) and (9) we obtain

$$\left[-\delta\tau_1(b_1 - b_2) + \frac{\delta A}{4}(b_1 - b_2)^2\right]\tau_1(1-\tau_1) = \frac{\tau_1^2}{2}\left(\tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right)\right)(b_1 - b_2),$$

which implies

$$\left[-\delta\tau_1 + \frac{\delta A}{4}(b_1 - b_2)\right]2(1-\tau_1) = \tau_1\left(\tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right)\right)$$

because $b_1 > b_2$ and $\tau_1 > 0$. Therefore

$$b_1 = \frac{1}{2} - \frac{\tau_1}{\delta A(1-2\tau_1)}\Big(2(1-\delta) - (5-\delta)\tau_1 + (2+\delta)\tau_1^2\Big).$$

Now, using this and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$, $\tilde{P} = 0$, we obtain

$$0 \leqslant 2A(1-2\tau_1)\left(\tau_1(1-\tau_1)\frac{\partial \tilde{P}}{\partial \tau_1} - (1-2\tau_1)\tilde{P}\right) = -\tau_1^2(1-\tau_1)\Big(16 - 14(3+\delta)\tau_1 + (34+24\delta)\tau_1^2 - (8+7\delta)\tau_1^3\Big).$$

But $16-14(3+\delta)\tau_1+(34+24\delta)\tau_1^2-(8+7\delta)\tau_1^3 = 128\left(\frac{1}{2}-\tau_1\right)^3 + 8(27-7\delta)\left(\frac{1}{2}-\tau_1\right)^2\tau_1 + 4(29-16\delta)\left(\frac{1}{2}-\tau_1\right)\tau_1^2 + 5(4-3\delta)\tau_1^3 > 0$, a contradiction.

Finally, suppose that $\overline{\Delta W}|_{b_1=\tilde{b}_1,\tau_1=\tau_1(\tilde{b}_1)} > 0$ for all $\tilde{b}_1 \in (\frac{1}{2}, b_1]$. We have $\tilde{P}|_{\tau_1=0} = -\frac{\delta A}{2}(b_1-\frac{1}{2})^2$, so if $b_1 \to \frac{1}{2}$ then $\tau_1(b_1) \to 0$, since $\tau_1(b_1)$ is defined as the first positive root of $\tilde{P}$. Now, $\frac{\partial \tilde{P}}{\partial \tau_1}|_{b_1=\frac{1}{2},\tau_1=0} = \frac{1}{2}(1+\delta) > 0$, so by the implicit function theorem $\frac{d\tau_1}{db_1}|_{b_1=\frac{1}{2}} = -\frac{\partial \tilde{P}/\partial b_1}{\partial \tilde{P}/\partial \tau_1}|_{b_1=\frac{1}{2},\tau_1=0} = 0$. Now, $\frac{d\overline{\Delta W}}{db_1} = \frac{\partial \overline{\Delta W}}{\partial b_1} + \frac{\partial \overline{\Delta W}}{\partial \tau_1}\frac{d\tau_1}{db_1}$ so $\frac{d\overline{\Delta W}}{db_1}|_{b_1=\frac{1}{2}} = 0$, and $\frac{d^2\overline{\Delta W}}{d^2 b_1} = \frac{\partial^2\overline{\Delta W}}{\partial^2 b_1} + 2\frac{\partial^2\overline{\Delta W}}{\partial b_1 \partial \tau_1}\frac{d\tau_1}{db_1} + \frac{\partial^2\overline{\Delta W}}{\partial^2 \tau_1}(\frac{d\tau_1}{db_1})^2 + \frac{\partial\overline{\Delta W}}{\partial \tau_1}\frac{d^2\tau_1}{d^2 b_1}$, so $\frac{d^2\overline{\Delta W}}{d^2 b_1}|_{b_1=\frac{1}{2}} = -\frac{\delta A}{2}$. Therefore $\overline{\Delta W} = -\frac{\delta A}{4}(b_1-\frac{1}{2})^2 + O((b_1-\frac{1}{2})^3)$ as $b_1 \to \frac{1}{2}$, so $\overline{\Delta W} < 0$ for $b_1 > \frac{1}{2}$ sufficiently close to $\frac{1}{2}$, which is a contradiction. This finishes the proof. ■

CLAIM 4. *If $T = 0$ in equilibrium then $\tau_1 < \lambda$.*

Proof. We just proved that $\tau_1 \leqslant \lambda$, so let's prove that $T = 0$ implies $\tau_1 \neq \lambda$. Suppose that $\tau_1 = \lambda$. We will repeat the argument in the previous proof. Let $\tilde{s}$ and $\tilde{P} = P|_{s=\tilde{s}}$ as before. Notice that $0 \leqslant P = s - \tilde{s} + T = s - \tilde{s}$, so $0 < \tilde{s} \leqslant s$. We have $\frac{\partial \Delta W}{\partial s} = -\frac{2}{\delta A}s < 0$. Therefore, we can reduce s to $\tilde{s}$, keeping $P, T \geqslant 0$ by Claim 1, and improving W. If $\tilde{P} > 0$ or the left derivative $\frac{\partial \tilde{P}}{\partial \tau_1} < 0$ then we can decrease τ_1 a little and take $s = \tilde{s}$, since $\tilde{P} \geqslant 0$ is still the case, which implies that (τ_1, s) is feasible, and the effect on $\widetilde{\Delta W} = \Delta W|_{s=\tilde{s}}$ is positive, since

$$\frac{\partial \widetilde{\Delta W}}{\partial \tau_1} = \frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s}\frac{d\tilde{s}}{d\tau_1} = -\frac{2}{A}(1-\rho)\tilde{s}(1-\tau_1) - \frac{2}{\delta A}\tilde{s}(1-\tau_1) < 0,$$

so the resulting policy is better, a contradiction. The remaining case is $\tilde{P} = 0$ and $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$, where again the latter is a left derivative. In this case I will show that $\widetilde{\Delta W} < 0$. This shows that the proposer would rather choose the status quo policy $\tau_1 = s = 0$, a contradiction.

Now, the left derivative of $\tilde{P}$ with respect to τ_1 at λ is a linear function of ρ. So, $\frac{\partial \tilde{P}}{\partial \tau_1} \geqslant 0$ implies that $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=0} \geqslant 0$ or $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=1} \geqslant 0$. When $\rho = 0$, $\tau_2 = \tau_1$ for sure, and we are back to the case we considered in the previous proof. So it remains to prove that $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=1} \geqslant 0$ and $\tilde{P} = 0$ imply $\widetilde{\Delta W} < 0$. Given the previous proof, it's enough to show that $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=1} \geqslant 0$ implies $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=0} \geqslant 0$.

We have

$$\begin{aligned}
\left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\rho=1} &= 1 - b_1 - \frac{\lambda}{\delta A}\left[4 + \delta(4 + A) - (6 + 7\delta)\lambda + (3 + 3\delta)\lambda^2\right], \\
\left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\rho=0} &= (1 + \delta)(1 - b_1) \\
&\quad - \frac{\lambda}{\delta A}\left[4 + \delta(8 + A) + \delta^2\left(4 + \left(\frac{3}{2} - b_1\right)A\right) - 3(2 + 5\delta + 3\delta^2)\lambda + (2 + 6\delta + 4\delta^2)\lambda^2\right].
\end{aligned}$$

Now, $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=1} \geqslant 0$ implies $b_1 \leqslant \overline{b}_1 \equiv 1 - \frac{\lambda}{\delta A}\left[4 + \delta(4 + A) - (6 + 7\delta)\lambda + (3 + 3\delta)\lambda^2\right]$. Notice that $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=0}$ is decreasing in b_1, so if $\frac{\partial \tilde{P}}{\partial \tau_1}|_{\rho=0} < 0$ then this must also be the case taking $b_1 = \overline{b}_1$. In this case, we get, after some algebra, $A < \overline{A} \equiv \frac{1}{\delta(1-2\lambda)}\left[(2 + 2\delta)\lambda - (5 + 6\delta)\lambda^2 + (2 + 3\delta)\lambda^3\right]$. In addition, $-c + \frac{\delta A}{2} = \frac{\delta}{2}$ by definition of b_1, which using $c \geqslant 0$ implies $Ab_1 \geqslant 1$, and in particular $A\overline{b}_1 \geqslant 1$, i.e., $\delta A - \lambda\left[4 + \delta(4 + A) - (6 + 7\delta)\lambda + (3 + 3\delta)\lambda^2\right] \geqslant \delta$, i.e., $A \geqslant \underline{A} \equiv \frac{1}{\delta(1-\lambda)}\left[\delta + (4 + 4\delta)\lambda - (6 + 7\delta)\lambda^2 + (2 + 3\delta)\lambda^3\right]$. We obtain $\underline{A} < \overline{A}$, but $\overline{A} - \underline{A} = -\frac{1}{1-\lambda} - \frac{\lambda^2}{1-2\lambda} < 0$, contradiction. This finishes the proof. ∎

Claim 5. *If $\tau_1 = \lambda$ and $P \geqslant 0$ then $s > 0$.*

Proof. We have $\tau_2 = \lambda$. Suppose that $s = 0$. We have

$$0 \leqslant P = -\left(\lambda - \frac{1}{2}\lambda^2 + \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right)\right) + \left((1 + \delta)b_1 - \frac{\delta}{A}(2 - \lambda)\lambda\right)\lambda(1 - \lambda),$$

i.e.

$$\lambda - \frac{1}{2}\lambda^2 + \frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right) + \frac{\delta}{A}(2 - \lambda)(1 - \lambda)\lambda^2 \leqslant (1 + \delta)b_1\lambda(1 - \lambda).$$

When $\delta = 0$ this is

$$\lambda - \frac{1}{2}\lambda^2 \leqslant b_1\lambda(1 - \lambda),$$

i.e., $1 - \frac{1}{2}\lambda \leqslant b_1(1 - \lambda)$, but $1 - \frac{1}{2}\lambda - b_1(1 - \lambda) = 1 - b_1 + (b_1 - \frac{1}{2})\lambda > 0$, a contradiction. When $\delta = 1$ this is

$$\lambda - \frac{1}{2}\lambda^2 + \frac{A}{2}\left(b_1 - \frac{1}{2}\right) + \frac{1}{A}(2 - \lambda)(1 - \lambda)\lambda^2 \leqslant 2b_1\lambda(1 - \lambda).$$

When $b_1 = \frac{1}{2}$ this is

$$\lambda - \frac{1}{2}\lambda^2 + \frac{1}{A}(2 - \lambda)(1 - \lambda)\lambda^2 \leqslant \lambda(1 - \lambda),$$

which can't happen because $\lambda - \frac{1}{2}\lambda^2 > \lambda(1-\lambda)$. When $b_1 = 1$ this is

$$\lambda - \frac{1}{2}\lambda^2 + \frac{A}{4} + \frac{1}{A}(2-\lambda)(1-\lambda)\lambda^2 \leqslant 2\lambda(1-\lambda).$$

By AM-GM we have

$$\frac{A}{4} + \frac{1}{A}(2-\lambda)(1-\lambda)\lambda^2 \geqslant \sqrt{(2-\lambda)(1-\lambda)\lambda^2} > \sqrt{(1-\lambda)^2\lambda^2} = \lambda(1-\lambda),$$

so, summing, we get a contradiction. Since the expressions are linear in b_1 and the inequality doesn't hold at the extremes, it cannot hold for any b_1. By the same argument, it cannot hold for any δ, and, hence, it cannot hold. This proves that $P < 0$, a contradiction. ■

Claim 6. *In equilibrium, $\tau_1 < \lambda$.*

Proof. Suppose that $\tau_1 > \lambda$, so $\tau_2 = \tau_1$. By Claim 2 we have $\frac{\partial P}{\partial \tau_1} < 0$. This implies that the proposer can reduce τ_1 slightly without violating $P \geqslant 0$. We clearly have that $\frac{\partial \Delta W}{\partial \tau_1} < 0$ if $\tau_1 > \lambda$, so the proposer prefers to reduce τ_1. By Claim 4 we have $T > 0$, hence we can reduce τ_1 slightly and keep $T > 0$ by continuity, which contradicts that τ_1 is an equilibrium.

Suppose that $\tau_1 = \lambda$, and again $\tau_2 = \tau_1$. We have $s > 0$ by Claim 5 and $T > 0$ by Claim 4. We have $\frac{\partial P}{\partial s} \geqslant 0$, because if $\frac{\partial P}{\partial s} < 0$ then $\frac{\partial T}{\partial s} = \frac{\partial P}{\partial s} - 1 < 0$, so we can reduce s and increase P, T and ΔW, since $\frac{\partial \Delta W}{\partial s} = -\frac{2}{\delta A}[s + \delta(\tau_1 - \lambda)(1-\tau_1)] < 0$. If $\frac{\partial P}{\partial s} > 0$, by the implicit function theorem, there is a differentiable function $s(\tau_1)$ defined in an interval $(\lambda - \epsilon, \lambda]$ for some $\epsilon > 0$ such that $P|_{s=s(\tau_1)} = 0$, and we have $\frac{ds}{d\tau_1} = -\frac{\frac{\partial P}{\partial \tau_1}}{\frac{\partial P}{\partial s}}$. Now, $\frac{\partial P}{\partial \tau_1} = -(1 - b_1 + (2b_1 - 1)\tau_1) < 0$, so $\frac{ds}{d\tau_1} > 0$. We have

$$\frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s}\frac{ds}{d\tau_1} = \underbrace{b_1(\lambda - \tau_1)}_{=0} - \frac{2}{\delta A}[s + \delta(\tau_2 - \lambda)(1-\tau_2)]\frac{ds}{d\tau_1} < 0,$$

so we can reduce τ_1 a little bit, keeping $P = 0$ and $T > 0$, and we increase W, which contradicts that (τ_1, s) is optimal. If $\frac{\partial P}{\partial s} = 0$, the same argument essentially carries over, except that $\frac{ds}{d\tau_1} = +\infty$. Formally: take $h > 0$ and consider $(\tau_1 - h, s - h)$. We have $\frac{dP}{dh} = -\frac{\partial P}{\partial \tau_1} - \frac{\partial P}{\partial s} = -\frac{\partial P}{\partial \tau_1} > 0$ and $\frac{d\Delta W}{dh} = -\frac{\partial \Delta W}{\partial \tau_1} - \frac{\partial \Delta W}{\partial s} = \frac{2}{\delta A}[s + \delta(\tau_2 - \lambda)(1-\tau_2)] > 0$, hence if h is small enough we can increase P and ΔW, and keep $T > 0$ by continuity, so (τ_1, s) is not an equilibrium. This completes the proof. ■

Comment. If $\tau_1 < \lambda$, $\tau_2 = \lambda$ with probability ρ, and $\tau_2 = \tau_1$ with probability $1 - \rho$, so

$$\begin{aligned}
\frac{d\Delta W}{d\tau_1} &= [b_1 + (1-\rho)\delta b_2](\lambda - \tau_1) - \frac{2}{A}(1-\rho)(1-\tau_1)(s - s^*), \\
\frac{d\Delta W}{ds} &= -\frac{2}{\delta A}(s - s^*),
\end{aligned}$$

where $s^* = \delta(1-\rho)(1-\tau_1)(\lambda-\tau_1)$. It's reasonable to conjecture that in equilibrium $s > s^*$, i.e., not only is the tax too low ($\tau_1 < \lambda$) but compensating polluting districts requires choosing a subsidy level that is larger than the optimal level given the tax ($s > s^*$). But the constraint imposed by green districts can bind, and the equilibrium subsidy be too low, i.e., $s < s^*$. A numerical example is $A = 1.8, b_1 = 0.6, \delta = 1, \lambda = 0.5, \rho = 0$, in which case in equilibrium $\tau_1 \approx 0.16$ and $s \approx 0.24$, but $s^* \approx 0.28$. In other words, not only is the tax too low, but also the subsidy.

Claim 7. *In equilibrium, $s > 0$.*

Proof. Suppose that $s = 0$. In that case, $T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] > 0$ since $\rho > 0$. We have $\frac{d\Delta W}{d\tau_1} = \left[b_1 + (1-\rho)\delta b_2 + \frac{2}{A}\delta(1-\rho)^2(1-\tau_1)^2\right](\lambda-\tau_1) > 0$. If $P > 0$ we can increase τ_1 a bit, improving the objective and keeping $P, T \geqslant 0$, contradiction. Therefore $P = 0$. We have $\frac{dP}{ds} = \frac{1}{A}\mathrm{E}[A - (4\tau_2 - 3\tau_2^2)]$. There are two cases: $\frac{dP}{ds} > 0$ and $\frac{dP}{ds} \leqslant 0$.

Suppose that $\frac{dP}{ds} > 0$. By the implicit function theorem we can define $s(\tau_1)$ such that $P|_{s=s(\tau_1)} = 0$ in an interval around τ_1, and $s(\tau_1)$ is differentiable. We have

$$\frac{d\Delta W|_{s=s(\tau_1)}}{d\tau_1} = \frac{d\Delta W}{d\tau_1} + \frac{d\Delta W}{ds}\frac{ds}{d\tau_1} = \frac{d\Delta W}{d\tau_1} - \frac{d\Delta W}{ds}\frac{dP/d\tau_1}{dP/ds}.$$

I will show that if $\frac{dP}{ds} > 0$ then $\frac{d}{d\tau_1}\Delta W|_{s=s(\tau_1)} > 0$. This shows that the proposer can increase τ_1 a little while keeping $P, T \geqslant 0$ (in fact, $P = 0$, and $T > 0$ by continuity), which is a contradiction.

We have $\frac{d}{d\tau_1}\Delta W|_{s=s(\tau_1)} > 0$ iff $\frac{d\Delta W}{d\tau_1}\frac{dP}{ds} > \frac{d\Delta W}{ds}\frac{dP}{d\tau_1}$. We have $\frac{d\Delta W}{ds} = \frac{2}{\delta A}s^* \geqslant 0$, so if $\frac{dP}{d\tau_1} \leqslant 0$ we are done. Suppose that $\frac{dP}{d\tau_1} > 0$.

After canceling positive factors, we want to prove

$$\begin{aligned}&\left[b_1 + (1-\rho)\delta b_2 + \frac{2}{A}\delta(1-\rho)^2(1-\tau_1)^2\right]\mathrm{E}[A - (4\tau_2 - 3\tau_2^2)] > 2(1-\rho)(1-\tau_1)\\ &\qquad\times\left[-(1-\tau_1) + b_1(1-2\tau_1) - \delta\frac{2}{A}(1-\rho)(1-\tau_1)\mathrm{E}[\tau_2(1-\tau_2)] + \delta b_2(1-\rho)(1-2\tau_1)\right].\end{aligned}$$

This is linear in δ, since $b_2 = b_1 - \frac{2}{A}\mathrm{E}(\tau_2 - \frac{1}{2}\tau_2^2)$. If $\delta = 0$ this is

$$b_1\mathrm{E}[A - (4\tau_2 - 3\tau_2^2)] > -2(1-\rho)(1-\tau_1)[1 - b_1 + (2b_1 - 1)\tau_1],$$

which is true using $\frac{dP}{ds} > 0$. It's enough to prove the inequality for $\delta = 1$. Enough to show

$$\begin{aligned}&b_1\mathrm{E}[A - (4\tau_2 - 3\tau_2^2)]\\ &\qquad > 2\left[-(1-\tau_1) + b_1(1-2\tau_1) - \frac{2}{A}(1-\rho)(1-\tau_1)\mathrm{E}[\tau_2(1-\tau_2)] + b_2(1-\rho)(1-2\tau_1)\right].\end{aligned}$$

This is linear in b_1. If $b_1 = \frac{1}{2}$ this is

$$\frac{1}{2}\mathrm{E}[A-(4\tau_2-3\tau_2^2)]+1+\frac{4}{A}(1-\rho)(1-\tau_1)\mathrm{E}[\tau_2(1-\tau_2)] > \left(1-\frac{4}{A}\mathrm{E}(\tau_2-\frac{1}{2}\tau_2^2)\right)(1-\rho)(1-2\tau_1),$$

which is clearly true. If $b_1 = 1$ this is

$$\begin{aligned}&\mathrm{E}[A-(4\tau_2-3\tau_2^2)] \\ &\qquad > 2\left[-\tau_1-\frac{2}{A}(1-\rho)(1-\tau_1)\mathrm{E}[\tau_2(1-\tau_2)]+\left(1-\frac{2}{A}\mathrm{E}(\tau_2-\frac{1}{2}\tau_2^2)\right)(1-\rho)(1-2\tau_1)\right].\end{aligned}$$

It's enough to show $A-\mathrm{E}(4\tau_2-3\tau_2^2)+2\tau_1 > (1-\rho)(1-2\tau_1)$. It's linear in ρ. If $\rho = 0$ this is $A+3\tau_1^2 > 1$, which is true, since $A > 1$. If $\rho = 1$ this is $A-\mathrm{E}(4\tau_2-3\tau_2^2)+2\tau_1 > 0$, also true since $\frac{dP}{ds} > 0$. This finishes the proof assuming $\frac{dP}{ds} > 0$.

Suppose that $\frac{dP}{ds} \leqslant 0$, i.e., $A \leqslant \mathrm{E}(4\tau_2-3\tau_2^2)$. I will show that $P < 0$ in this case, which is a contradiction. We have

$$P = -\left(\tau_1-\frac{1}{2}\tau_1^2\right)-\frac{\delta A}{2}\left(b_1-\frac{1}{2}\right)+b_1\tau_1(1-\tau_1)+\delta b_2\mathrm{E}[\tau_2(1-\tau_2)],$$

which is linear in b_1 using $b_2 = b_1-\frac{2}{A}\mathrm{E}(\tau_2-\frac{1}{2}\tau_2^2)$. If $b_1 = \frac{1}{2}$ we have

$$P = -\frac{1}{2}\tau_1+\delta\left(\frac{1}{2}-\frac{2}{A}\mathrm{E}\left(\tau_2-\frac{1}{2}\tau_2^2\right)\right)\mathrm{E}[\tau_2(1-\tau_2)],$$

so it's enough to prove $\frac{1}{2}-\frac{2}{A}\mathrm{E}\left(\tau_2-\frac{1}{2}\tau_2^2\right) < 0$, and using $\frac{dP}{ds} \leqslant 0$ it's enough to prove $\mathrm{E}(4\tau_2-3\tau_2^2) < 4\mathrm{E}\left(\tau_2-\frac{1}{2}\tau_2^2\right)$, which is clearly true. If $b_1 = 1$ we have

$$P = -\frac{1}{2}\tau_1^2-\frac{\delta A}{4}+\delta\left(1-\frac{2}{A}\mathrm{E}\left(\tau_2-\frac{1}{2}\tau_2^2\right)\right)\mathrm{E}[\tau_2(1-\tau_2)].$$

Now, $\frac{\delta A}{4}+\frac{2\delta}{A}\mathrm{E}\left(\tau_2-\frac{1}{2}\tau_2^2\right)\mathrm{E}[\tau_2(1-\tau_2)] > \frac{\delta A}{4}+\frac{\delta}{A}\mathrm{E}[\tau_2(1-\tau_2)]^2 \geqslant \delta\mathrm{E}[\tau_2(1-\tau_2)]$ by AM-GM, so $P < 0$. Therefore, $P < 0$ if $b_1 \in \{\frac{1}{2}, 1\}$, which implies that $P < 0$ for every $b_1 \in (\frac{1}{2}, 1)$ by linearity, as desired. ■

B.5. *Proof of Observation 1*

We say that $(A, c, \delta, \lambda, \rho)$ is *feasible* if $\frac{1}{2} < b_1 < 1$ and there is (τ_1, s) such that $P \geqslant 0$ and $T \geqslant 0$, where (b_1, b_2, τ_2) are given by $-c+\frac{\delta A}{2}b_1 = \frac{\delta}{2}$, $s-c+\frac{\delta A}{2}b_2 = \frac{\delta}{2}\mathrm{E}\left[(1-\tau_2)^2\right]$, $\tau_2 = \max\{\tau_1, \lambda\}$ with probability ρ, $\tau_2 = \tau_1$ with probability $1-\rho$, and P, T are defined in Appendix B.4.

The following proves the first part of Observation 1.

Claim 8. *If $\delta > \frac{1}{A-\frac{1}{2}}$ then an increase in A or a decrease in c relax the political constraint imposed by polluting districts ($P \geqslant 0$) and tighten the political constraint imposed by green districts ($T \geqslant 0$).*

Proof. I simply prove that $\frac{dP}{dA} > 0$ and $\frac{dP}{dc} < 0$ if $P \geqslant 0$, which implies that the political constraint imposed by polluting districts is relaxed, and $\frac{dP}{dA} < 0$ and $\frac{dP}{dc} > 0$ if $T \geqslant 0$, which implies that the political constraint imposed by green districts are tightened.

We have $\frac{dT}{dA} = -\frac{T}{A}$, so the sign of T does not depend on A, and $\frac{dP}{dA} = \frac{\delta}{4} - \frac{T}{A} > 0$, since $T = b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2) \leqslant (b_1 + \delta b_2)\frac{1}{4} < \frac{1}{4}(1+\frac{\delta}{2}) = \frac{1}{4}(1+\delta) < \frac{\delta A}{4}$ using $x(1-x) \leqslant \frac{1}{4}$ for any $x \in \mathbb{R}$ and the hypothesis $\delta > \frac{1}{A-\frac{1}{2}}$.

We have

$$\frac{dT}{dc} = \left(\frac{\partial T}{\partial b_1} + \frac{\partial T}{\partial b_2}\right)\frac{2}{\delta A} = \frac{2}{\delta A}\left(\tau_1 - \tau_1^2 + \delta \mathrm{E}(\tau_2 - \tau_2^2)\right) > 0.$$

Also, $\frac{dP}{dc} = -1 + \frac{dT}{dc} < 0$ because $\frac{dT}{dc} = \frac{2}{\delta A}\left(\tau_1 - \tau_1^2 + \delta\mathrm{E}(\tau_2 - \tau_2^2)\right) \leqslant \frac{2}{\delta A}(1+\delta)\frac{1}{4} < 1$, using the hypothesis $\delta > \frac{1}{A-\frac{1}{2}}$. ■

Claim 9. *If $(A, c, \delta, \lambda, \rho)$ is feasible then there is (τ_1, s) such that $P = T = 0$.*

Proof. We use Claim 1. Let $f(\tau_1) = P(\tau_1, \tilde{s}(\tau_1))$. If $(A, c, \delta, \lambda, \rho)$ is feasible then there is $\tau_1 \in [0, \frac{1}{2}]$ such that $f(\tau_1) \geqslant 0$. Now,

$$f(1) = -\frac{(2 + \delta A(2b_1 - 1))(2 + 2\delta + \delta A(2b_1 - 1))}{8\delta A} < 0,$$

so there must be $\tau_1' \in [\tau_1, 1)$ such that $f(\tau_1') = 0$ by continuity of f. ■

Claim 10. *If $(A, c, \delta, \lambda, \rho)$ is feasible and $A' \geqslant A$, $c' \leqslant c$ are such that $(A', c', \delta, \lambda, \rho)$ satisfies $b_1 \in (\frac{1}{2}, 1)$ then $(A', c', \delta, \lambda, \rho)$ is feasible.*

Proof. Let (τ_1, s) such that $P = T = 0$. We have $\frac{\partial T}{\partial A} = -\frac{T}{A} = 0$ if $T = 0$ and $\frac{\partial P}{\partial A} = \frac{\delta}{2} + \frac{\partial T}{\partial A} = \frac{\delta}{2} > 0$, so if we increase A to A' we keep $T = 0$ and P increases, so (τ_1, s) is still feasible.

Take (τ_1, s) such that $P = T = 0$ for $(A, c, \delta, \lambda, \rho)$. Let $\tilde{P} = P|_{s=\tilde{s}}$. We have

$$\frac{\partial \tilde{P}}{\partial c} = \frac{\delta(A-2) - 4c - 2\tau_1 - \delta\mathrm{E}[(2-\tau_2)\tau_2]}{\delta A}.$$

If $A \geqslant 2$ we have $c > \frac{\delta}{4}(A-2)$ since $b_1 > \frac{1}{2}$. Hence $\frac{\partial \tilde{P}}{\partial c} \leqslant -\frac{2\tau_1 + \delta\mathrm{E}[(2-\tau_2)\tau_2]}{\delta A} < 0$. If $A < 2$ then $c \geqslant 0$ so $\frac{\partial \tilde{P}}{\partial c} < -\frac{2\tau_1 + \delta\mathrm{E}[(2-\tau_2)\tau_2]}{\delta A} \leqslant 0$ as well. Therefore, $\frac{\partial \tilde{P}}{\partial c} < 0$, so by decreasing c, starting from (τ_1, s) such that $P = T = 0$, we obtain $(\tau_1, \tilde{s}(\tau_1))$ such that $P = T \geqslant 0$, hence $(A', c', \delta, \lambda, \rho)$ is still feasible. ■

This establishes the second part of Observation 1, and concludes the proof.

B.6. *Proof of Observation 2*

Let $\tilde{s}$ be given by $P|_{s=\tilde{s}} = T|_{s=\tilde{s}}$, i.e., $\tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}(b_1 - \frac{1}{2}) > 0$, and $\tilde{P} = P|_{s=\tilde{s}}$.

Claim 11. *$\tilde{P}$ is single peaked in τ_1 for $\tau_1 \in [0, \lambda]$.*

Proof. If $\tilde{P}$ is strictly concave we are done. Otherwise, since

$$\frac{\partial^3 \tilde{P}}{\partial^3 \tau_1} = \frac{6}{\delta A}(1 + (1-\rho)\delta)(2 + 3(1-\rho)\delta - (2 + 4(1-\rho)\delta)\tau_1) \geqslant \frac{6}{\delta A}(1 + (1-\rho)\delta)^2 > 0$$

it must be strictly convex for $\tau_1 \geqslant \underline{\tau}_1$ for some $\underline{\tau}_1$. But keeping $\tau_2 = \lambda$ (even if $\tau_1 > \lambda$) we have

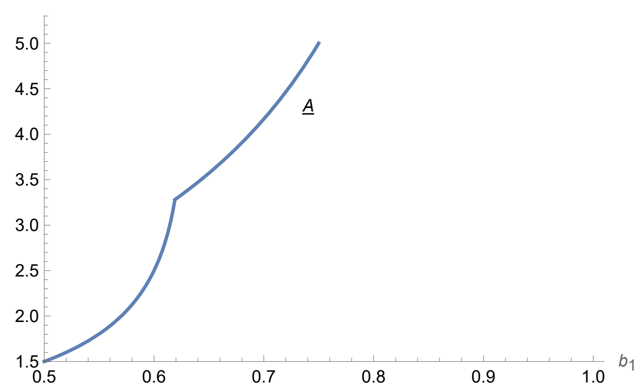
$$\left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\tau_1 = \frac{1}{2}} = -\frac{3 + 2\delta A(2b_1 - 1) + 2\rho\delta(4 - 3\lambda)\lambda + (1-\rho)\delta(4 + \delta + \delta A(2b_1 - 1))}{4\delta A} < 0.$$

Therefore, $\left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\tau_1 = \underline{\tau}_1} \leqslant \left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\tau_1 = \frac{1}{2}} < 0$, so $\tilde{P}$ is decreasing for $\tau_1 \geqslant \underline{\tau}_1$, and single-peaked for $\tau_1 < \underline{\tau}_1$ by strict concavity, so it's single-peaked, as desired. ■

Claim 12. *If $\lambda = \frac{1}{2}$ then $(A, c, \delta, \lambda, \rho)$ is feasible only if*

$$A \geqslant \underline{A} := \begin{cases} \frac{5 - 6b_1 - \sqrt{1 + 36b_1 - 60b_1^2}}{4(4b_1^2 - 4b_1 + 1)}\rho & \textit{if } b_1 \leqslant \frac{1}{64}(19 + 5\sqrt{17}) \approx 0.62, \\ \frac{5\rho}{4(1-b_1)} & \textit{otherwise,} \end{cases}$$

which is increasing in b_1:



Proof. Suppose that $\lambda = \frac{1}{2}$. We have $\left.\frac{\partial \tilde{P}}{\partial \tau_1}\right|_{\tau_1 = 0} \leqslant 0$ iff $A \leqslant \frac{5\rho}{4(1-b_1)}$ by a simple calculation. In this case, $\tilde{P}$ is maximized at $\tau_1 = 0$ by Claim 11. We have

$$\tilde{P}|_{\tau_1 = 0} = -\frac{\delta(2A^2(2b_1 - 1)^2 - A(5 - 6b_1)\rho + 3\rho^2)}{16A}.$$

The smallest root is

$$A_1 = \frac{5 - 6b_1 - \sqrt{1 + 36b_1 - 60b_1^2}}{4(4b_1^2 - 4b_1 + 1)}\rho,$$

which exists iff $1 + 36b_1 - 60b_1^2 \geqslant 0$, i.e., iff $b_1 \leqslant \frac{1}{30}(9 + 4\sqrt{6}) \approx 0.63$. Hence, if $A \leqslant \frac{5\rho}{4(1-b_1)}$ then the constraints are feasible only if $A \geqslant A_1$. So, either $A > \frac{5\rho}{4(1-b_1)}$ or $A_1 \leqslant A \leqslant \frac{5\rho}{4(1-b_1)}$. We have $A_1 \leqslant \frac{5\rho}{4(1-b_1)}$ iff $b_1 \leqslant \frac{1}{64}(19 + 5\sqrt{17}) \approx 0.62$, so A is feasible only if $A \geqslant A_1$ and $b_1 \leqslant \frac{1}{64}(19 + 5\sqrt{17})$ or, else, $A > \frac{5\rho}{4(1-b_1)}$, as desired. ∎

B.7. *Proof of Observation 3*

Using Claim 1 it's enough to show that if $\tilde{P} \geqslant 0$ for some τ_1 then, if we decrease ρ, $\tilde{P} \geqslant 0$ still holds. A sufficient condition is that $\frac{d\tilde{P}}{d\rho} < 0$. We have

$$\begin{aligned}
\frac{dT}{d\rho} &= \frac{\partial T}{\partial \rho} + \frac{\partial T}{\partial b_2}\frac{\partial b_2}{\partial \rho} \\
&= (\lambda - \tau_1)_+ \left\{ \delta b_2(1 - \lambda - \tau_1) - \frac{2}{A}(s + \delta \mathrm{E}[\tau_2(1 - \tau_2)])\left(1 - \frac{1}{2}(\lambda + \tau_1)\right)\right\},
\end{aligned}$$

where $x_+ = \max\{x, 0\}$, and $\tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2 + \frac{\delta A}{2}(b_1 - \frac{1}{2}) \geqslant \frac{\delta A}{2}(b_1 - \frac{1}{2})$. If $\tau_1 \geqslant \lambda$ then ρ doesn't affect T, so let's assume that $\tau_1 < \lambda$. We have

$$\begin{aligned}
\frac{d\tilde{P}}{d\rho} &= (\lambda - \tau_1)\left\{\delta b_2(1 - \lambda - \tau_1) - \frac{2}{A}(\tilde{s} + \delta \mathrm{E}[\tau_2(1 - \tau_2)])\left(1 - \frac{\lambda + \tau_1}{2}\right)\right\} \\
&\leqslant (\lambda - \tau_1)\left\{\frac{\delta}{2}(1 - \lambda - \tau_1) - \frac{2}{A}\frac{\delta A}{2}\left(b_1 - \frac{1}{2}\right)\left(1 - \frac{\lambda + \tau_1}{2}\right)\right\}.
\end{aligned}$$

Now the term in curly brackets is linear in τ_1, so we can bound it by the maximum of the value it takes for $\tau_1 = 0$ and $\tau_1 = \lambda$. In the first case it is $\frac{\delta}{2}(1 - \lambda) - \delta\left(b_1 - \frac{1}{2}\right)(1 - \frac{\lambda}{2}) < 0$ using the hypothesis $b_1 > \frac{1 - \frac{3}{4}\lambda}{1 - \frac{1}{2}\lambda}$. In the second case it is $\frac{\delta}{2}(1 - 2\lambda) - \delta\left(b_1 - \frac{1}{2}\right)(1 - \lambda) < 0$ using that $b_1 > \frac{1 - \frac{3}{4}\lambda}{1 - \frac{1}{2}\lambda} \geqslant \frac{1 - \frac{1}{2}\lambda}{1 - \lambda}$. This proves that $\frac{d\tilde{P}}{d\rho} < 0$, as desired.

B.8. *Proof of Observation 4*

We have

$$\Delta W = b_1\tau_1\left(\lambda - \frac{1}{2}\tau_1\right) + \delta b_2 \mathrm{E}\left[\tau_2\left(\lambda - \frac{1}{2}\tau_2\right)\right] + \delta\lambda(b_1 - b_2) - \frac{\delta A}{4}(b_1 - b_2)^2.$$

When $\lambda = 0$, $\Delta W = -\frac{1}{2}b_1\tau_1^2 - \frac{\delta}{2}\mathrm{E}\tau_2^2 - \frac{\delta A}{4}(b_1 - b_2)^2$, so $\Delta W \geqslant 0$ iff $\tau_1 = \tau_2 = s = 0$. However, in that case $P = -\frac{\delta A}{4}(2b_1 - 1) < 0$.

B.9. *Proof of Proposition 2*

In order to keep the notation consistent with the proof of the previous results, I will call P and T the quantities that need to be nonnegative for the political constraints PC_B and PC_G to hold, respectively. Thus, we have

$$\begin{aligned} T &= b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - s(b_1 - b_2) - \delta\rho\tilde{b}_2\lambda(1-\lambda), \\ P &= -\left(\tau_1 - \frac{1}{2}\tau_1^2\right) + s + T, \end{aligned}$$

where $\tilde{b}_2$ is given by $-c + \frac{\delta A}{2}\tilde{b}_2 = \frac{\delta}{2}\left[\rho(1-\lambda)^2 + 1 - \rho\right]$, i.e., $\tilde{b}_2 = b_1 - \frac{1}{A}\rho(2-\lambda)\lambda$.

Recall that τ_2 is $\max\{\tau_1, \lambda\}$ with probability ρ and τ_1 with probability $1-\rho$, b_1 and b_2 are given by $-c + \frac{\delta A}{2}b_1 = \frac{\delta}{2}$ and $s - c + \frac{\delta A}{2}b_2 = \frac{\delta}{2}\mathrm{E}[(1-\tau_2)^2]$,

$$\begin{aligned} W &= \frac{b_1}{2}(1-\tau_1^2) + \frac{A}{4}(1-b_1^2) - c(b_1 - b_2) - \lambda b_1(1-\tau_1), \\ &\quad + \delta\left[\frac{b_2}{2}(1-\mathrm{E}\tau_2^2) + \frac{A}{4}(1-b_2^2) - \lambda b_2(1-\mathrm{E}\tau_2)\right], \\ \tilde{W} &= \frac{b_2}{2} + \frac{A}{4}(1-b_1^2) - c(b_1 - \tilde{b}_2) - \lambda b_1, \\ &\quad + \delta\left[\frac{\tilde{b}_2}{2}(1-\rho\lambda^2) + \frac{A}{4}(1-\tilde{b}_2^2) - \lambda b_2(1-\rho\lambda)\right], \end{aligned}$$

and $\Delta W = W - \tilde{W}$. In all the calculations I am going to take $c = \frac{\delta}{2}(Ab_1 - 1)$, so $c \geqslant 0$ iff $Ab_1 \geqslant 1$. We have $b_2 = b_1 - \frac{2}{\delta A}s - \frac{1}{A}\mathrm{E}(2\tau_2 - \tau_2^2)$.

The policymaker chooses τ_1, s to maximize W subject to $P, T, \Delta W \geqslant 0$. If this is not feasible, $\tau_1 = s = 0$. We have to prove that if the problem is feasible, then the optimal policy satisfies $\tau_1 < \lambda$ and $s > 0$. Suppose that the problem is feasible and τ_1, s are optimal.

Strategy:

1. Suppose that $\tau_1 > \lambda$.
   - 1.1. Suppose that $T = 0$. Then $s \geqslant \tilde{s} \equiv \tau_1 - \frac{1}{2}\tau_1^2 > 0$. If $s > \tilde{s}$ then $P > 0$. I can reduce s improving W and keeping $P, T \geqslant 0$, contradiction. Therefore $s = \tilde{s}$, i.e., $P = 0$. Let $\tilde{P} = P|_{s=\tilde{s}}$.
     - 1.1.1. If $\frac{d\tilde{P}}{d\tau_1} < 0$ then I can reduce τ_1 improving P, T, W, contradiction.
     - 1.1.2. Suppose that $\frac{d\tilde{P}}{d\tau_1} \geqslant 0$. This contradicts $T = 0$.
   - 1.2. Suppose that $T > 0$. I prove that $\frac{dP}{d\tau_1} < 0$, so we can reduce τ_1 a little, contradiction.
2. Suppose that $\tau_1 = \lambda$.
   - 2.1. Suppose that $T = 0$. Same argument as before, but using the left derivative of $\frac{d\tilde{P}}{d\tau_1}$ at $\tau_1 = \lambda$. We obtain a contradiction.

2.2. Suppose that $T > 0$.

2.2.1. If $s = 0$ then I prove that $P < 0$, contradiction.

2.2.2. If $s > 0$ then $\frac{dP}{d\tau_1}|_{\tau_1=\lambda^-}, \frac{dW}{d\tau_1}|_{\tau_1=\lambda^-} < 0$. This means that by reducing τ_1 we can increase W keeping $P, T \geqslant 0$, contradiction.

3. Then $\tau_1 < \lambda$.

3.1. Suppose that $s = 0$ and $\tau_1 > 0$. We have $T > 0$ and $\frac{dP}{ds}, \frac{dW}{ds} > 0$, so it's better to increase s, contradiction.

3.2. Suppose that $s = 0$ but $\tau_1 = 0$. We have $P = T = 0$ and $\frac{dW}{ds}, \frac{dW}{d\tau_1} > 0$. I prove that we can increase both s and τ_1 a little so that both P and T increase, contradiction.

3.3. Therefore, $s > 0$. This finishes the proof.

Case 1. $\tau_1 > \lambda$.

We have $\tau_2 = \tau_1$, so

$$\frac{db_2}{d\tau_1} = -\frac{2}{A}(1-\tau_1), \qquad \frac{db_2}{ds} = -\frac{2}{\delta A}, \qquad \frac{\delta A}{2}(b_1 - b_2) = s + \delta\left(\tau_1 - \frac{1}{2}\tau_1^2\right)$$

and

$$\frac{dW}{db_2} = c + \frac{\delta}{2}\left[1 - \tau_1^2 - Ab_2 - 2\lambda(1-\tau_1)\right] = \frac{\delta}{2}\left[A(b_1 - b_2) - \tau_1^2 - 2\lambda(1-\tau_1)\right]$$
$$= s + \delta(\tau_1 - \lambda)(1-\tau_1).$$

Hence

$$\begin{aligned}\frac{dW}{d\tau_1} &= \frac{\partial W}{\partial \tau_1} + \frac{dW}{db_2}\frac{db_2}{d\tau_1} \\ &= (b_1 + \delta b_2)(\lambda - \tau_1) - \frac{2}{A}[s + \delta(\tau_1 - \lambda)(1-\tau_1)](1-\tau_1) < 0\end{aligned} \tag{11}$$

and

$$\frac{dW}{ds} = \frac{dW}{db_2}\frac{db_2}{ds} = -\frac{2}{\delta A}[s + \delta(\tau_1 - \lambda)(1-\tau_1)] < 0. \tag{12}$$

Also,

$$T = (b_1 + \delta b_2)\tau_1(1-\tau_1) - s(b_1 - b_2) - \delta\rho\tilde{b_2}\lambda(1-\lambda)$$

and

$$\frac{dT}{db_2} = s + \delta\tau_1(1-\tau_1),$$

so

$$\frac{dT}{ds} = -(b_1 - b_2) + \frac{dT}{db_2}\frac{db_2}{ds} = -(b_1 - b_2) - \frac{2}{\delta A}[s + \delta\tau_1(1-\tau_1)] < 0, \tag{13}$$

and

$$\begin{aligned}\frac{dT}{d\tau_1} &= (b_1+\delta b_2)(1-2\tau_1)+\frac{dT}{db_2}\frac{db_2}{d\tau_1}\\ &= (b_1+\delta b_2)(1-2\tau_1)-\frac{2}{A}[s+\delta\tau_1(1-\tau_1)](1-\tau_1).\end{aligned}\tag{14}$$

In addition,

$$\frac{dP}{d\tau_1} = -(1-\tau_1)+(b_1+\delta b_2)(1-2\tau_1)-\frac{2}{A}[s+\delta\tau_1(1-\tau_1)](1-\tau_1).\tag{15}$$

Case 1.1. Suppose that $T=0$.

Let $\tilde{s}=\tau_1-\frac{1}{2}\tau_1^2$. We have $0\leqslant P=s-\tilde{s}+T=s-\tilde{s}$, so $s\geqslant\tilde{s}$. Suppose that $s>\tilde{s}$. We have $P>0$, $\frac{dT}{ds}<0$ by (13) and $\frac{dW}{ds}<0$ (12). Therefore, we can decrease s a little bit making T and W larger and keeping $P>0$ by continuity. This contradicts that s is optimal. Therefore, $s=\tilde{s}$, and $P=0$. Let $\tilde{P}=P|_{s=\tilde{s}}$.

Case 1.1.1. $\frac{d\tilde{P}}{d\tau_1}<0$.

I can reduce τ_1, keeping $s=\tilde{s}$, and we obtain $P,T>0$. Now, if $\tilde{W}=W|_{s=\tilde{s}}$, using (11) and (12) we obtain

$$\frac{d\tilde{W}}{d\tau_1}=\frac{dW}{d\tau_1}+\frac{dW}{ds}\frac{d\tilde{s}}{d\tau_1}<0.$$

Therefore, we improve the objective keeping $P,T>0$, which contradicts that (τ_1,s) is optimal.

Case 1.1.2. $\frac{d\tilde{P}}{d\tau_1}\geqslant 0$.

Taking $s=\tilde{s}=\tau_1-\frac{1}{2}\tau_1^2$ we have $\frac{\delta A}{2}(b_1-b_2)=\tilde{s}+\delta\left(\tau_1-\frac{1}{2}\tau_1^2\right)=(1+\delta)\left(\tau_1-\frac{1}{2}\tau_1^2\right)$, and

$$\begin{aligned}\frac{d\tilde{P}}{d\tau_1} &= \frac{\partial\tilde{P}}{\partial\tau_1}+\frac{\partial\tilde{P}}{\partial s}\frac{d\tilde{s}}{d\tau_1}\\ &= -(1-\tau_1)+(b_1+\delta b_2)(1-2\tau_1)-\frac{2}{A}[\tilde{s}+\delta\tau_1(1-\tau_1)](1-\tau_1)\\ &\quad+\left[1-(b_1-b_2)-\frac{2}{\delta A}[\tilde{s}+\delta\tau_1(1-\tau_1)]\right](1-\tau_1)\\ &= \frac{1+\delta}{\delta A}\left[\delta Ab_1(1-2\tau_1)-\tau_1(4-6\tau_1+2\tau_1^2+\delta(4-9\tau_1+4\tau_1^2))\right].\end{aligned}$$

Now, using $0\leqslant\rho\leqslant 1$ and $\lambda<\tau_1$ we get

$$\begin{aligned}0=\tilde{P} &= (b_1+\delta b_2)\tau_1(1-\tau_1)-\tilde{s}(b_1-b_2)-\delta\rho\tilde{b_2}\lambda(1-\lambda)\\ &= \frac{1+\delta}{\delta A}\left[\delta Ab_1\tau_1(1-\tau_1)-\frac{1}{2}\tau_1^2(2-\tau_1)(2-\tau_1+\delta(2-2\tau_1))\right]\\ &\quad-\frac{\delta\rho}{A}\left[Ab_1-\rho(2\lambda-\lambda^2)\right]\lambda(1-\lambda)\end{aligned}$$

$$
\begin{aligned}
&> \frac{1+\delta}{\delta A}\left[\delta A b_1 \tau_1 (1-\tau_1) - \frac{1}{2}\tau_1^2 (2-\tau_1)(2-\tau_1+\delta(2-2\tau_1))\right] \\
&\quad - \frac{\delta}{A} A b_1 \tau_1 (1-\tau_1) \\
&= \frac{\tau_1}{\delta A}\left[\delta A b_1 (1-\tau_1) - (1+\delta)\frac{1}{2}\tau_1 (2-\tau_1)(2-\tau_1+\delta(2-2\tau_1))\right]
\end{aligned}
$$

so

$$
\delta A b_1 < \frac{1}{2}(1+\delta)\frac{\tau_1(2-\tau_1)}{1-\tau_1}(2-\tau_1+\delta(2-2\tau_1)) \leqslant \frac{\tau_1(2-\tau_1)}{1-\tau_1}(2-\tau_1+\delta(2-2\tau_1)).
$$

Therefore,

$$
\frac{d\tilde{P}}{d\tau_1} \leqslant \frac{1+\delta}{\delta A}\tau_1\left[\frac{2-\tau_1}{1-\tau_1}(2-\tau_1+\delta(2-2\tau_1))(1-2\tau_1) - (4-6\tau_1+2\tau_1^2+\delta(4-9\tau_1+4\tau_1^2))\right].
$$

The expression in brackets in the RHS is maximized at $\delta = 0$ or $\delta = 1$. For $\delta = 0$ it is

$$
\frac{2-\tau_1}{1-\tau_1}(2-\tau_1)(1-2\tau_1) - (4-6\tau_1+2\tau_1^2) = -\frac{\tau_1(2-\tau_1)}{1-\tau_1} < 0.
$$

For $\delta = 1$ it is

$$
\frac{2-\tau_1}{1-\tau_1}(4-3\tau_1)(1-2\tau_1) - (8-15\tau_1+6\tau_1^2) = -\frac{\tau_1(3-2\tau_1)}{1-\tau_1} < 0.
$$

Therefore, $\frac{d\tilde{P}}{d\tau_1} < 0$, a contradiction.

Case 1.2. Suppose that $T > 0$.

Claim 13. *If* $\tau_1 > \lambda$ *then* $\frac{dP}{d\tau_1} < 0$.

Proof. We want to prove that

$$
\frac{dP}{d\tau_1} = -(1-\tau_1) + (b_1+\delta b_2)(1-2\tau_1) - \frac{2}{A}[s+\delta\tau_1(1-\tau_1)](1-\tau_1) < 0.
$$

Now, using $b_2 = b_1 - \frac{2}{\delta A}s - \frac{1}{A}(2\tau_1 - \tau_1^2)$, we have

$$
\frac{dP}{d\tau_1} = -(1-\tau_1) - \frac{\delta}{A}\tau_1(4-9\tau_1+4\tau_1^2) + (1+\delta)b_1(1-2\tau_1) - \frac{2}{A}(2-3\tau_1)s.
$$

Using $s \geqslant 0$ and $Ab_1 \geqslant 1$, it's enough to show

$$
-(1-\tau_1) - \delta b_1 \tau_1(4-9\tau_1+4\tau_1^2) + (1+\delta)b_1(1-2\tau_1) < 0.
$$

By linearity, it's enough to show this for $\delta = 0$ and $\delta = 1$.

When $\delta = 0$ this is $-(1-\tau_1) + b_1(1-2\tau_1) < 0$. Now, $\tilde{b}_2 \leqslant \frac{1}{2}$, i.e.,

$$b_1 \leqslant \frac{1}{2} + \frac{\rho\lambda(2-\lambda)}{A} < \frac{1}{2} + \frac{\tau_1(2-\tau_1)}{A} \leqslant \frac{1}{2} + b_1\tau_1(2-\tau_1),$$

so $b_1 < \frac{1}{2(1-\tau_1)^2}$. Therefore, it's enough to show $1 - 2\tau_1 < 2(1-\tau_1)^3$. Expanding, this is $1 - 2\tau_1 < 2 - 6\tau_1 + 6\tau_1^2 - 2\tau_1^3$, i.e., $0 < 1 - 4\tau_1 + 6\tau_1^2 - 2\tau_1^3$. Now, $1 - 4\tau_1 + 6\tau_1^2 - 2\tau_1^3 = 1 - 2\tau_1(1-\tau_1)(2-\tau_1) \leqslant 1 - \frac{1}{2}(2-\tau_1) = \frac{1}{2}\tau_1 > 0$.

If $\delta = 1$ we want to prove

$$-(1-\tau_1) - b_1\tau_1(4 - 9\tau_1 + 4\tau_1^2) + 2b_1(1-2\tau_1) < 0,$$

i.e., $b_1(2 - 8\tau_1 + 9\tau_1^2 - 4\tau_1^3) < 1 - \tau_1$. We use $b_1 < \frac{1}{2(1-\tau_1)^2}$ again, and we have to show $2 - 8\tau_1 + 9\tau_1^2 - 4\tau_1^3 < 2(1-\tau_1)^3$, i.e., $0 < \tau_1(2 - 3\tau_1 + 2\tau_1^2) = \tau_1((2-\tau_1)(1-\tau_1) + \tau_1^2)$, which is true. ■

Therefore, we can reduce τ_1 a little keeping $P, T \geqslant 0$ and increasing W, which contradicts that τ_1 is optimal. This finished Case 1, showing that in equilibrium we must have $\tau_1 \leqslant \lambda$.

Case 2. $\tau_1 = \lambda$.

When $\tau_1 < \lambda$ we have that $\tau_2 = \lambda$ with probability ρ and τ_1 with probability $1-\rho$. In this case $b_2 = b_1 - \frac{2}{\delta A}s - \frac{1}{A}\mathrm{E}(2\tau_2 - \tau_2^2)$.

We have

$$\begin{aligned}
\frac{\partial W}{\partial b_2} &= c + \frac{\delta}{2}(1 - \mathrm{E}\tau_2^2) - \frac{\delta}{2}Ab_2 - \delta\lambda(1 - \mathrm{E}\tau_2) \\
&= \frac{\delta}{2}(Ab_1 - 1) + \frac{\delta}{2}(1 - \mathrm{E}\tau_2^2) - \frac{\delta}{2}Ab_2 - \delta\lambda(1 - \mathrm{E}\tau_2) \\
&= s + \delta\mathrm{E}[(\tau_2 - \lambda)(1 - \tau_2)] \\
&= s - \delta(1-\rho)(\lambda - \tau_1)(1 - \tau_1) = s - s^*
\end{aligned}$$

with $s^* \equiv \delta(1-\rho)(\lambda-\tau_1)(1-\tau_1)$,

$$\frac{dW}{d\tau_1} = \frac{\partial W}{\partial \tau_1} + \frac{\partial W}{\partial b_2}\frac{db_2}{d\tau_1} = [b_1 + \delta(1-\rho)b_2](\lambda - \tau_1) - \frac{2}{A}(s - s^*)(1-\rho)(1-\tau_1),$$

and

$$\frac{dW}{ds} = -\frac{2}{\delta A}(s - s^*).$$

The left derivatives at $\tau_1 = \lambda$ are

$$\left.\frac{dW}{d\tau_1}\right|_{\tau_1=\lambda^-} = -\frac{2}{A}(1-\rho)(1-\lambda)s \leqslant 0 \quad \text{and} \quad \left.\frac{dW}{ds}\right|_{\tau_1=\lambda^-} = -\frac{2}{\delta A}s \leqslant 0 \tag{16}$$

with equality iff $s = 0$.

When $\tau_1 < \lambda$ we have

$$\begin{aligned}
\frac{\partial T}{\partial b_2} &= s + \delta \mathrm{E}[\tau_2(1-\tau_2)], \\
\frac{dT}{d\tau_1} &= [b_1 + \delta(1-\rho)b_2](1-2\tau_1) - \frac{2}{A}(s + \delta\mathrm{E}[\tau_2(1-\tau_2)])(1-\rho)(1-\tau_1), \\
\frac{dT}{ds} &= -(b_1 - b_2) - \frac{2}{\delta A}(s + \delta\mathrm{E}[\tau_2(1-\tau_2)]) < 0, \\
\frac{dP}{d\tau_1} &= -(1-\tau_1) + [b_1 + \delta(1-\rho)b_2](1-2\tau_1) - \frac{2}{A}(s + \delta\mathrm{E}[\tau_2(1-\tau_2)])(1-\rho)(1-\tau_1), \\
\frac{dP}{ds} &= 1 - (b_1 - b_2) - \frac{2}{\delta A}(s + \delta\mathrm{E}[\tau_2(1-\tau_2)]) = \frac{4}{\delta A}\left[\frac{\delta}{4}\mathrm{E}(A - 4\tau_2 + 3\tau_2^2) - s\right]
\end{aligned}$$

Case 2.1. Suppose that $T = 0$.

Let $\tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2$. We have $0 \leqslant P = s - \tilde{s} + T = s - \tilde{s}$, so $s \geqslant \tilde{s}$. Suppose that $s > \tilde{s}$. We have $P > 0$, $\frac{dT}{ds} < 0$ and $\frac{dW}{ds} < 0$. Therefore, we can decrease s a little bit making T and W larger and keeping $P > 0$ by continuity. This contradicts that s is optimal. Therefore, $s = \tilde{s}$, and $P = 0$. Let $\tilde{P} = P|_{s=\tilde{s}}$.

Case 2.1.1. $\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} < 0$.

I can reduce τ_1, keeping $s = \tilde{s} > 0$, and we obtain $P, T > 0$. Now, if $\tilde{W} = W|_{s=\tilde{s}}$, using (16) we obtain

$$\left.\frac{d\tilde{W}}{d\tau_1}\right|_{\tau_1=\lambda^-} = \left.\frac{dW}{d\tau_1}\right|_{\tau_1=\lambda^-} + \frac{dW}{ds}\frac{d\tilde{s}}{d\tau_1} < 0.$$

Therefore, we improve the objective keeping $P, T > 0$, which contradicts that (τ_1, s) is optimal.

Case 2.1.2. $\frac{d\tilde{P}}{d\tau_1}|_{\tau_1=\lambda^-} \geqslant 0$.

Taking $s = \tilde{s} = \tau_1 - \frac{1}{2}\tau_1^2$ we have $b_2 = b_1 - \frac{1+\delta}{\delta A}(2\lambda - \lambda^2)$, and

$$\begin{aligned}
\frac{d\tilde{P}}{d\tau_1} &= \frac{\partial \tilde{P}}{\partial \tau_1} + \frac{\partial \tilde{P}}{\partial s}\frac{d\tilde{s}}{d\tau_1} \\
&= -(1-\tau_1) + [b_1 + \delta(1-\rho)b_2](1-2\tau_1) - \frac{2}{A}(\tilde{s} + \delta\mathrm{E}[\tau_2(1-\tau_2)])(1-\rho)(1-\tau_1) \\
&\quad + \left[1 - (b_1 - b_2) - \frac{2}{\delta A}(\tilde{s} + \delta\mathrm{E}[\tau_2(1-\tau_2)])\right](1-\tau_1),
\end{aligned}$$

so

$$\begin{aligned}
\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} &= -(1-\lambda) + [1+\delta(1-\rho)]b_1(1-2\lambda) - \delta(1-\rho)\frac{1+\delta}{\delta A}(2\lambda-\lambda^2)(1-2\lambda) \\
&\quad - \frac{1}{A}(1-\rho)(2\lambda-\lambda^2+2\delta\lambda(1-\lambda))(1-\lambda) \\
&\quad + \left[1 - \frac{1+\delta}{\delta A}(2\lambda-\lambda^2) - \frac{1}{\delta A}(2\lambda-\lambda^2+2\delta\lambda(1-\lambda))\right](1-\lambda) \\
&= \frac{1+\delta(1-\rho)}{\delta A}\left[\delta A b_1(1-2\lambda) - (2\lambda-\lambda^2+2\delta\lambda(1-\lambda))(1-\lambda)\right] \\
&\quad - \frac{1+\delta}{\delta A}\lambda(2-\lambda)(1-\lambda+\delta(1-\rho)(1-2\lambda)).
\end{aligned}$$

Now, we have

$$\begin{aligned}
T &= \frac{1+\delta}{A}(Ab_1-\lambda(2-\lambda))\lambda(1-\lambda) - \frac{1+\delta}{\delta A}\frac{1}{2}\lambda^2(2-\lambda)^2 - \frac{\delta}{A}\rho(Ab_1-\rho\lambda(2-\lambda))\lambda(1-\lambda) \\
&= \frac{\lambda}{\delta A}\left[(1+\delta(1-\rho))(1-\lambda)\delta A b_1 - \frac{1}{2}\lambda(2-\lambda)(2-\lambda+\delta(4-3\lambda)+2(1-\rho^2)\delta^2(1-\lambda))\right],
\end{aligned}$$

and $T = 0$, so

$$\begin{aligned}
\delta A b_1 &= \frac{1}{2}\frac{\lambda(2-\lambda)}{(1+\delta(1-\rho))(1-\lambda)}(2-\lambda+\delta(4-3\lambda)+2(1-\rho^2)\delta^2(1-\lambda)) \\
&\leqslant \frac{\lambda}{(1+\delta(1-\rho))(1-\lambda)}(2-\lambda+\delta(4-3\lambda)+2\delta(1-\lambda)) \\
&= \frac{\lambda}{(1+\delta(1-\rho))(1-\lambda)}(2-\lambda+\delta(6-5\lambda))
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} &\leqslant \frac{1}{\delta A}\frac{\lambda(1-2\lambda)}{1-\lambda}(2-\lambda+\delta(6-5\lambda)) \\
&\quad - \frac{1+\delta(1-\rho)}{\delta A}\lambda(2-\lambda+2\delta(1-\lambda))(1-\lambda) \\
&\quad - \frac{1+\delta}{\delta A}\lambda(2-\lambda)(1-\lambda+\delta(1-\rho)(1-2\lambda)).
\end{aligned}$$

The RHS is linear in ρ, so it's maximized at either $\rho = 0$ or $\rho = 1$. If it's maximized at $\rho = 0$ we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{1}{\delta A}\frac{\lambda(1-2\lambda)}{1-\lambda}(2-\lambda+\delta(6-5\lambda))$$

$$
\begin{aligned}
&-\frac{1+\delta}{\delta A}\lambda(2-\lambda+2\delta(1-\lambda))(1-\lambda)\\
&-\frac{1+\delta}{\delta A}\lambda(2-\lambda)(1-\lambda+\delta(1-2\lambda))\\
\leqslant\ &\frac{1}{\delta A}\frac{\lambda(1-2\lambda)}{1-\lambda}(2-\lambda+\delta(6-5\lambda))-2\frac{1+\delta}{\delta A}\lambda(2-\lambda)(1-\lambda)\\
=\ &\frac{\lambda}{\delta A(1-\lambda)}\left[(1-2\lambda)(2-\lambda+\delta(6-5\lambda))-2(1+\delta)(2-\lambda)(1-\lambda)^2\right].
\end{aligned}
$$

The expression in brackets is maximized at either $\delta = 0$ or $\delta = 1$. If it's maximized at $\delta = 0$, we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{\lambda}{\delta A(1-\lambda)}(2-\lambda)(-1+2\lambda(1-\lambda)) < 0.$$

If it's maximized at $\delta = 1$ we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{\lambda}{\delta A(1-\lambda)}\left[(1-2\lambda)(8-6\lambda)-4(2-\lambda)(1-\lambda)^2\right] = \frac{2\lambda^2(-1-2\lambda(1-\lambda))}{\delta A(1-\lambda)} < 0.$$

If the RHS above is maximized at $\rho = 1$ we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{1}{\delta A}\frac{\lambda}{1-\lambda}\left[(1-2\lambda)(2-\lambda+\delta(6-5\lambda))-((2+\delta)(2-\lambda)+2\delta(1-\lambda))(1-\lambda)^2\right].$$

The expression in brackets is maximized at either $\delta = 0$ or $\delta = 1$. If it's maximized at $\delta = 0$, we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{1}{\delta A}\frac{\lambda}{1-\lambda}(2-\lambda)(-1+2\lambda(1-\lambda)) < 0.$$

If it's maximized at $\delta = 1$ we have

$$\left.\frac{d\tilde{P}}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant \frac{1}{\delta A}\frac{\lambda}{1-\lambda}\lambda(-1-6\lambda(1-\lambda)-\lambda^2) < 0.$$

We obtained $\frac{d\tilde{P}}{d\tau_1}|_{\tau_1=\lambda^-} < 0$, a contradiction.

Case 2.2. Suppose that $T > 0$.

Case 2.2.1. Suppose that $s = 0$.

We have, using $\tilde{b}_2 \leqslant \frac{1}{2}$, i.e., $b_1 \leqslant \frac{1}{2} + \frac{\rho\lambda(2-\lambda)}{A}$,

$$
\begin{aligned}
P &= -\lambda+\frac{1}{2}\lambda^2+(b_1+\delta b_2)\lambda(1-\lambda)-\delta\rho\tilde{b}_2\lambda(1-\lambda)\\
&= -\lambda+\frac{1}{2}\lambda^2+(1+(1-\rho)\delta)b_1\lambda(1-\lambda)-\frac{\delta}{A}(1-\rho^2)\lambda^2(1-\lambda)(2-\lambda)
\end{aligned}
$$

$$
\begin{aligned}
&\leqslant -\lambda + \frac{1}{2}\lambda^2 + (1+(1-\rho)\delta)\left(\frac{1}{2} + \frac{\rho\lambda(2-\lambda)}{A}\right)\lambda(1-\lambda) - \frac{\delta}{A}(1-\rho^2)\lambda^2(1-\lambda)(2-\lambda) \\
&= -\lambda + \frac{1}{2}\lambda^2 + (1+(1-\rho)\delta)\frac{1}{2}\lambda(1-\lambda) + \frac{1}{A}[\rho - (1-\rho)\delta]\lambda^2(1-\lambda)(2-\lambda).
\end{aligned}
$$

This is linear in ρ, so it's maximized at either $\rho = 0$ or $\rho = 1$. In the first case,

$$
\begin{aligned}
P &\leqslant -\lambda + \frac{1}{2}\lambda^2 + (1+\delta)\frac{1}{2}\lambda(1-\lambda) - \frac{1}{A}\delta\lambda^2(1-\lambda)(2-\lambda) \\
&\leqslant -\lambda + \frac{1}{2}\lambda^2 + \lambda(1-\lambda) = -\frac{1}{2}\lambda^2 < 0.
\end{aligned}
$$

In the second case, using $A > 1$, we obtain

$$
\begin{aligned}
P &\leqslant -\lambda + \frac{1}{2}\lambda^2 + \frac{1}{2}\lambda(1-\lambda) + \frac{1}{A}\lambda^2(1-\lambda)(2-\lambda) \\
&< -\frac{1}{2}\lambda + \lambda^2(1-\lambda)(2-\lambda) = -\frac{1}{2}\lambda(1 - 2\lambda(1-\lambda)(2-\lambda)) \\
&\leqslant -\frac{1}{2}\lambda\left(1 - \frac{1}{2}(2-\lambda)\right) = -\frac{1}{4}\lambda^2 < 0.
\end{aligned}
$$

In both cases we obtain $P < 0$, a contradiction.

Case 2.2.1. Suppose that $s > 0$.

Claim. $\frac{dP}{d\tau_1}|_{\tau_1=\lambda^-} < 0$.

Proof. We have

$$
\left.\frac{dP}{d\tau_1}\right|_{\tau_1=\lambda^-} = -(1-\lambda) + [b_1 + \delta(1-\rho)b_2](1-2\lambda) - \frac{2}{A}(s + \delta\lambda(1-\lambda))(1-\rho)(1-\lambda)
$$

with $b_2 = b_1 - \frac{2}{\delta A}s - \frac{1}{A}(2\lambda - \lambda^2)$. This is maximized at $\rho = 0$ or $\rho = 1$. In the first case,

$$
\left.\frac{dP}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant -(1-\lambda) + (b_1 + \delta b_2)(1-2\lambda) - \frac{2}{A}(s + \delta\lambda(1-\lambda))(1-\lambda).
$$

The proof of Claim 13 works to show that this is negative, as desired. In the second case,

$$
\left.\frac{dP}{d\tau_1}\right|_{\tau_1=\lambda^-} \leqslant -(1-\lambda) + b_1(1-2\lambda) = -(1-b_2) + \lambda(1-2b_1) < 0
$$

since $b_1 \in (\frac{1}{2}, 1)$, as desired. ■

Now, $\frac{dW}{d\tau_1}|_{\tau_1=\lambda^-} = -\frac{2}{A}s(1-\rho)(1-\lambda) < 0$. Therefore, we can reduce τ_1 by a small amount increasing both P and W, and keeping $T > 0$ by continuity. This contradicts that τ_1 is optimal.

Case 3. Suppose that $\tau_1 < \lambda$.

Case 3.1. Suppose that $s = 0$ and $\tau_1 > 0$.

We have $T = P + \tau_1 - \frac{1}{2}\tau_1^2 > P \geqslant 0$ and $\frac{dW}{ds} = \frac{2}{\delta A}s^*$ with $s^* = \delta(1-\rho)(\lambda - \tau_1)(1-\tau_1)$. If $\rho = 1$ we have $\tau_2 = \lambda$ and $\tilde{b}_2 = b_2$, so

$$\begin{aligned} P &= -\tau_1 + \frac{1}{2}\tau_1^2 + b_1\tau_1(1-\tau_1) + \delta b_2 \mathrm{E}[\tau_2(1-\tau_2)] - \delta\rho\tilde{b}_2\lambda(1-\lambda) \\ &= -\tau_1 + \frac{1}{2}\tau_1^2 + b_1\tau_1(1-\tau_1) \leqslant -\tau_1 + \frac{1}{2}\tau_1^2 + \tau_1(1-\tau_1) = -\frac{1}{2}\tau_1^2 < 0, \end{aligned}$$

contradiction. If $\rho < 1$ then $s^* > 0$, so $\frac{dW}{ds} > 0$. We have $\frac{dP}{ds} = \frac{1}{A}\mathrm{E}(A - 4\tau_2 + 3\tau_2^2) > 0$, since $A > 1 > 4\lambda - 3\lambda^2 \geqslant \mathrm{E}(4\tau_2 - 3\tau_2^2)$ using the assumption that $\lambda < \frac{1}{4}$. We can increase s a little while increasing W and P and keeping $T > 0$ by continuity. This contradicts that $s = 0$ is optimal.

Case 3.2. Suppose that $s = 0$ and $\tau_1 = 0$.

We have $\frac{dW}{d\tau_1} = (b_1 + \delta(1-\rho)b_2)\lambda + \frac{2}{A}\delta(1-\rho)^2\lambda > 0$, $\frac{dW}{ds} = \frac{2}{A}(1-\rho)\lambda \geqslant 0$,

$$\begin{aligned} \frac{dT}{d\tau_1} &= b_1[1 + \delta(1-\rho)] - \frac{1}{A}\delta\rho(1-\rho)\lambda(4-3\lambda), \\ \frac{dT}{ds} &= -\frac{1}{A}(4\lambda - 3\lambda^2) < 0, \end{aligned}$$

$\frac{dP}{d\tau_1} = -1 + \frac{dT}{d\tau_1}$ and $\frac{dP}{ds} = 1 + \frac{dT}{ds} > 0$. I will show that $\frac{dT}{d\tau_1} > 0$. We have, using $Ab_1 \geqslant 1$ and $\lambda < \frac{1}{4}$,

$$\begin{aligned} \frac{dT}{d\tau_1} &= \frac{1}{A}[Ab_1[1 + \delta(1-\rho)] - \delta\rho(1-\rho)\lambda(4-3\lambda)] \\ &> \frac{1}{A}[1 + \delta(1-\rho) - \delta\rho(1-\rho)] \\ &= \frac{1}{A}\left[1 + \delta(1-\rho)^2\right] > 0, \end{aligned}$$

as desired.

Suppose that $\frac{dP}{d\tau_1} \leqslant 0$. I will show that there is $h > 0$ such that $\frac{dT}{d\tau_1} + \frac{dT}{ds}h > 0$ and $\frac{dP}{d\tau_1} + \frac{dP}{ds}h > 0$. We have $\frac{dT}{d\tau_1} + \frac{dT}{ds}h > 0$ iff $h < -\frac{dT}{d\tau_1}/\frac{dT}{ds}$, and $\frac{dP}{d\tau_1} + \frac{dP}{ds}h > 0$ iff $h > -\frac{dP}{d\tau_1}/\frac{dP}{ds}$. Then there is $h > 0$ such that $\frac{dT}{d\tau_1} + \frac{dT}{ds}h > 0$ and $\frac{dP}{d\tau_1} + \frac{dP}{ds}h > 0$ iff $-\frac{dP}{d\tau_1}/\frac{dP}{ds} < -\frac{dT}{d\tau_1}/\frac{dT}{ds}$, i.e., $\frac{dP}{d\tau_1}\frac{dT}{ds} < \frac{dP}{ds}\frac{dT}{d\tau_1}$. This simplifies to $\frac{dT}{d\tau_1} + \frac{dT}{ds} > 0$. Now, using $Ab_1 \geqslant 1$ and $\lambda < \frac{1}{4}$,

$$\begin{aligned} \frac{dT}{d\tau_1} + \frac{dT}{ds} &= \frac{1}{A}\left[Ab_1[1 + \delta(1-\rho)] - \delta\rho(1-\rho)\lambda(4-3\lambda) - (4\lambda - 3\lambda^2)\right] \\ &\geqslant \frac{1}{A}[1 + \delta(1-\rho) - (1 + \delta\rho(1-\rho))\lambda(4-3\lambda)] \end{aligned}$$

$$> \frac{1}{A}[1+\delta(1-\rho)-(1+\delta\rho(1-\rho))] = \frac{1}{A}\delta(1-\rho)^2 \geqslant 0,$$

as desired. So, taking $\tau_1 = t$ and $s = ht$, we have $\frac{dW}{dt} = \frac{dW}{d\tau_1} + \frac{dW}{ds}h > 0$, $\frac{dP}{dt} = \frac{dP}{d\tau_1} + \frac{dP}{ds}h > 0$, and $\frac{dT}{dt} = \frac{dT}{d\tau_1} + \frac{dT}{ds}h > 0$. Therefore, there is $t > 0$ such that τ_1, s increase W over $\tau_1 = s = 0$ keeping $P, T \geqslant 0$. This contradicts that $\tau_1 = s = 0$ is optimal.

To finish, suppose that $\frac{dP}{d\tau_1} > 0$. Then we can do what we did in the previous paragraph with any $h > 0$ such that $h < -\frac{dT}{d\tau_1}/\frac{dT}{ds}$.

We conclude that in equilibrium we have $s > 0$ and $\tau_1 < \lambda$, as desired.

B.10. *Proof of Lemma 2*

We have

$$P = \frac{1}{2}(1-\tau_1)^2 + s - c + \frac{\delta A}{4} + T - \lambda b_1(1-\tau_1) - \delta\lambda b_2(1-\tau_2) - (1+\delta)\left(\frac{1}{2} - \lambda b_1\right),$$
$$T = b_1\tau_1(1-\tau_1) + \delta b_2\tau_2(1-\tau_2) - s(b_1 - b_2).$$

We want to show that $P \geqslant 0$ implies that $b_2 \leqslant \frac{1}{2}$, i.e., $s - c + \frac{\delta A}{4} \geqslant \frac{\delta}{2}(1-\tau_2)^2$. It's enough to show that $s - c + \frac{\delta A}{4} - P \geqslant \frac{\delta}{2}(1-\tau_2)^2$, since I can sum $P \geqslant 0$ and obtain the result. Thus, it's enough to show that

$$J = \frac{1}{2}(1-\tau_1)^2 + \frac{\delta}{2}(1-\tau_2)^2 + T - \lambda b_1(1-\tau_1) - \delta\lambda b_2(1-\tau_2) - (1+\delta)\left(\frac{1}{2} - \lambda b_1\right) \leqslant 0.$$

If $\tau_1 \leqslant \lambda$ we have $\tau_2 = \lambda$, so

$$J = \frac{1}{2}(1-\tau_1)^2 + \frac{\delta}{2}(1-\lambda)^2 - b_1(\lambda - \tau_1)(1-\tau_1) - s(b_1 - b_2) - (1+\delta)\left(\frac{1}{2} - \lambda b_1\right).$$

We have $\frac{\partial J}{\partial \tau_1} = -(1-\tau_1) + b_1(1+\lambda-2\tau_1) = -(1-(1+\lambda)b_1 + (2b_1-1)\tau_1) \leqslant 0$, so

$$J \leqslant J|_{\tau_1=0} = -\delta\lambda\left(1 - b_1 - \frac{\lambda}{2}\right) - s(b_1 - b_2) \leqslant -\delta\lambda\left(1 - \frac{1}{1+\lambda} - \frac{\lambda}{2}\right) = -\delta\frac{\lambda^2(1-\lambda)}{2(1+\lambda)} < 0,$$

using Assumption 2.

If $\tau_1 > \lambda$ we have $\tau_2 = \tau_1$, so

$$J = -(1+\delta)\frac{1}{2}(2-\tau_1)\tau_1 + (b_1 + \delta b_2)(\tau_1 - \lambda)(1-\tau_1) + (1+\delta)\lambda b_1 - s(b_1 - b_2).$$

Noting that $b_1 - b_2 = \frac{2}{\delta A}\left[\frac{\delta}{2}(2\tau_1 - \tau_1^2) + s\right]$ it is clear that $\frac{\partial J}{\partial b_1} \geqslant 0$, so $J \leqslant \tilde{J} = J|_{b_1 = \frac{1}{1+\lambda}}$. Now,

$$\begin{aligned}\frac{\partial \tilde{J}}{\partial \lambda} &= \frac{\partial J}{\partial \lambda} - \frac{1}{(1+\lambda)^2}\frac{\partial J}{\partial b_1} = (b_1 + \delta b_2)\tau_1 + \delta(b_1 - b_2) - b_1^2[(\tau_1 - \lambda)(1-\tau_1) + (1+\delta)\lambda] \\ &\geqslant (1+\delta)(b_1 - b_1^2\lambda) - b_1^2(\tau_1 - \lambda)(1-\tau_1) \geqslant b_1^2\left[b_1^{-1} - \lambda - (\tau_1 - \lambda)(1-\tau_1)\right] \\ &= b_1^2[1 - (\tau_1 - \lambda)(1-\tau_1)] \geqslant 0,\end{aligned}$$

so $\tilde{J} \leqslant \tilde{J}|_{\lambda=\tau_1}$, but

$$\begin{aligned}\tilde{J}|_{\lambda=\tau_1} &= -(1+\delta)\tau_1\left(1 - b_1 - \frac{\tau_1}{2}\right) - s(b_1 - b_2) \leqslant -(1+\delta)\tau_1\left(1 - \frac{1}{1+\lambda} - \frac{\tau_1}{2}\right) \\ &= -(1+\delta)\tau_1\left(1 - \frac{1}{1+\tau_1} - \frac{\tau_1}{2}\right) = -\delta\frac{\tau_1^2(1-\tau_1)}{2(1+\tau_1)} < 0,\end{aligned}$$

hence $J < 0$, as desired.

B.11. *Proof of Observation 5*

PC_B is feasible given (A, c, δ, λ), iff there is (τ_1, s) such that $P \geqslant 0$. Now, P is a concave quadratic in s, maximized at $s = \hat{s} = \frac{\delta}{4}(A - (4\tau_2 - 3\tau_2^2) + 2\lambda(1-\tau_2)) > 0$. So $P \geqslant 0$ only if $\tilde{P} = P|_{s=\hat{s}} \geqslant 0$. Now, if $\tau_1 \leqslant \lambda$, $\frac{d\tilde{P}}{d\tau_1} = -(1 - (1+\lambda)b_1 + (2b_1 - 1)\tau_1) \leqslant 0$, so $\tilde{P} \geqslant 0$ implies $\tilde{P}|_{\tau_1=0} \geqslant 0$. If $\tau_1 > \lambda$, $\tau_2 = \tau_1$, so

$$\begin{aligned}\frac{d\tilde{P}}{d\tau_1} &= \frac{\partial P}{\partial \tau_1} + \frac{\partial P}{\partial \tau_2} + \frac{\partial P}{\partial b_2}\frac{db_2}{d\tau_2} + \frac{\partial P}{\partial s}\frac{\partial \hat{s}}{\partial \tau_2} \\ &= -(1 - (1+\lambda)b_1 + (2b_1 - 1)\tau_1) + \delta b_2(1 + \lambda - 2\tau_1) - \frac{2}{A}(1-\tau_1)[\hat{s} + \delta(\tau_1 - \lambda)(1-\tau_1)].\end{aligned}$$

Now, some algebra yields $\frac{\partial^2}{\partial^2\lambda}\frac{d\tilde{P}}{d\tau_1} = -\frac{2\delta}{A}(1-\tau_1) < 0$ and

$$\begin{aligned}\left.\frac{\partial}{\partial \lambda}\frac{d\tilde{P}}{d\tau_1}\right|_{\lambda=\tau_1} &= (1+\delta)b_1 - \frac{\delta}{2} - \frac{\delta}{A}\left(1 - \frac{1}{2}\tau_1\right)\tau_1 \geqslant (1+\delta)b_1 - \frac{\delta}{2} - \delta b_1\left(1 - \frac{1}{2}\tau_1\right)\tau_1 \\ &= \left(1 + \frac{\delta}{2}\left(1 + (1-\tau_1)^2\right)\right)b_1 - \frac{\delta}{2} \geqslant \frac{1}{2}\left(1 + \frac{\delta}{2}\left(1 + (1-\tau_1)^2\right) - \delta\right) > 0.\end{aligned}$$

Therefore, $\frac{\partial}{\partial \lambda}\frac{d\tilde{P}}{d\tau_1} > 0$, so $\frac{d\tilde{P}}{d\tau_1} \leqslant \frac{d\tilde{P}}{d\tau_1}\Big|_{\lambda=\tau_1} = -(1+\delta)(1-b_1)(1-\tau_1) \leqslant 0$. Therefore, $\tilde{P} \geqslant 0$ implies $\tilde{P}|_{\tau_1=\lambda} \geqslant 0$, which implies $\tilde{P}|_{\tau_1=0} \geqslant 0$ by the previous finding. In sum, (A, c, δ, λ) is feasible iff $\bar{P} = P|_{\tau_1=0, s=\hat{s}} \geqslant 0$, where $\hat{s} = \frac{\delta}{4}(A - 2\lambda + \lambda^2)$.

We have

$$\bar{P} = \frac{-8(A-2\lambda)c + \delta(3A^2 - 2(2+2\lambda-\lambda^2)A + \lambda(8+4\lambda-4\lambda^2+\lambda^3))}{8A}.$$

We have $b_1 \leqslant \frac{1}{1+\lambda}$, so $0 \leqslant c \leqslant \frac{\delta(A-1-\lambda)}{2(1+\lambda)}$, which implies $A \geqslant 1+\lambda$. Now,

$$\begin{aligned}\frac{\partial \bar{P}}{\partial A} &= \frac{3}{8}\delta - \frac{\lambda(16c + \delta(8+4\lambda-4\lambda^2+\lambda^3))}{8A^2} \geqslant \frac{3}{8}\delta - \frac{\lambda(8\frac{\delta(A-1-\lambda)}{1+\lambda} + \delta(8+4\lambda-4\lambda^2+\lambda^3))}{8A^2} \\ &= \frac{\delta}{8}\left(3 - \frac{\lambda^2(1-\lambda)^2}{A^2} - \frac{8\lambda}{A(1+\lambda)}\right) \geqslant \frac{\delta}{8}\left(3 - \frac{\lambda^2(1-\lambda)^2}{(1+\lambda)^2} - \frac{8\lambda}{(1+\lambda)^2}\right) > 0,\end{aligned}$$

since $3(1+\lambda)^2 - \lambda^2(1-\lambda)^2 - 8\lambda \geqslant 3(1+\lambda) - \frac{1}{16} - 8\lambda = 3 - \frac{1}{16} - 5\lambda \geqslant 3 - \frac{1}{16} - \frac{5}{2} = \frac{7}{16} > 0$. This shows that if $\bar{P} \geqslant 0$ for A, then $\bar{P} \geqslant 0$ for any $A' \geqslant A$. Also, $\frac{\partial \bar{P}}{\partial c} = -\frac{1}{A}(A-2\lambda) < 0$, hence if $\bar{P} \geqslant 0$ for c, then $\bar{P} \geqslant 0$ for any $c' \leqslant c$. Now, $\bar{P}$ is linear in δ. If $\frac{\partial \bar{P}}{\partial \delta} < 0$ then $\bar{P}$ is maximized at $\delta = 0$, but $\bar{P}|_{\delta=0} = -\frac{1}{A}(A-2\lambda)c \leqslant 0$, hence $\bar{P} < 0$ for any $\delta > 0$. Therefore, if δ is feasible, $\frac{\partial \bar{P}}{\partial \delta} \geqslant 0$; in that case we obtain the desired result, viz, that if $\bar{P} \geqslant 0$ for δ then it's also the case for $\delta' \geqslant \delta$.

Finally, $b_1 > \frac{1}{2}$ implies $c > \frac{\delta}{4}(A-2)$, so

$$\frac{\partial \bar{P}}{\partial \lambda} = \frac{4c + \delta(2 - (1-\lambda)A + 2\lambda - 3\lambda^2 + \lambda^3)}{2A} > \frac{\delta\lambda(A + 2 - 3\lambda + \lambda^2)}{2A} \geqslant \frac{\delta\lambda(A+\frac{1}{2})}{2A} \geqslant 0,$$

and $\bar{P} \geqslant 0$ for λ implies $\bar{P} \geqslant 0$ for any $\lambda' \in [\lambda, \frac{1}{2})$, as desired.

B.12. *Proof of Proposition 3*

We have to prove that $\tau_1 < \lambda$ implies $s > 0$, $\tau_1 = \lambda$ implies $s = 0$, and $\tau_1 > \lambda$ cannot happen.

Claim 14. *If (τ_1, s) is optimal and $\tau_1 < \lambda$ then $s > 0$.*

Proof. Suppose that $s = 0$. If $P > 0$ we can increase τ_1 a bit improving the objective, a contradiction, so $P = 0$. We have $\frac{\partial P}{\partial s} = \frac{1}{A}(A - (2-\lambda)\lambda) > 0$, so by the implicit function theorem there is $\tilde{s}(\tau_1)$ defined around τ_1 such that $P|_{s=\tilde{s}} = 0$, and it is differentiable. We have

$$\frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s}\frac{\partial \tilde{s}}{\partial \tau_1} = b_1(\lambda - \tau_1) - \frac{2}{\delta A}\underbrace{(s + \delta(\tau_2 - \lambda)(1-\tau_2))}_{=0}\frac{\partial \tilde{s}}{\partial \tau_1} = b_1(\lambda - \tau_1) > 0,$$

since $\tau_2 = \lambda$ and $s = 0$, so we should increase τ_1, a contradiction. ■

Claim 15. *If (τ_1, s) is optimal and $\tau_1 = \lambda$ then $s = 0$.*

Proof. Suppose that $s > 0$. If $P > 0$ we can decrease s and improve the objective, so $P = 0$. We have that the left derivative $\frac{\partial P}{\partial \tau_1}$ is $-(1-b_1)(1-\lambda) < 0$. If $\frac{\partial P}{\partial s} < 0$ we should decrease s, so

$\frac{\partial P}{\partial s} \geqslant 0$. If $\frac{\partial P}{\partial s} > 0$ then by the implicit function theorem there is a differentiable function $\tilde{s}(\tau_1)$ defined for τ_1 in an interval $(\lambda - \epsilon, \lambda]$ for some $\epsilon > 0$ such that $P|_{s=\tilde{s}} = 0$. We have

$$\frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s}\frac{\partial \tilde{s}}{\partial \tau_1} = b_1(\lambda - \tau_1) - \frac{2}{\delta A}(s + \delta(\tau_2 - \lambda)(1 - \tau_2))\frac{\partial \tilde{s}}{\partial \tau_1} = \frac{2s}{\delta A}\frac{\frac{\partial P}{\partial \tau_1}}{\frac{\partial P}{\partial s}} < 0,$$

so we should decrease τ_1, a contradiction. If $\frac{\partial P}{\partial s} = 0$ take $h \geqslant 0$ and consider $(\tau_1 - h, s - h)$. We have $\frac{dP}{dh}\big|_{h=0} = -\frac{\partial P}{\partial \tau_1} - \frac{\partial P}{\partial s} > 0$, and $\frac{d\Delta W}{dh}\big|_{h=0} = \frac{\partial \Delta W}{\partial \tau_1} + \frac{\partial \Delta W}{\partial s} = \frac{2s}{\delta A} > 0$, so by taking $h > 0$ small we improve the objective satisfying the constraint, a contradiction. ■

Claim 16. *If (τ_1, s) is optimal then $\tau_1 \leqslant \lambda$.*

Proof. Suppose that $\tau_1 > \lambda$, so $\tau_2 = \tau_1$. If $s > 0$ and $\frac{\partial P}{\partial s} < 0$ then we can decrease s, keeping $P \geqslant 0$ and improving the objective, a contradiction. Hence, either $s = 0$ or $\frac{\partial P}{\partial s} \geqslant 0$. Now, $\frac{\partial P}{\partial s} = \frac{4}{\delta A}(s - \bar{s})$ with $\bar{s} = \frac{\delta}{4}(A - 4\tau_1 + 3\tau_1^2 + 2(1 - \tau_1)\lambda) < \frac{\delta}{4}(A - (2 - \tau_1)\tau_1)$, so $\frac{\partial P}{\partial s} \geqslant 0$ implies $s \leqslant \frac{\delta}{4}(A - (2 - \tau_1)\tau_1)$.

Let $D = \frac{\partial \Delta W}{\partial \tau_1}\frac{\partial P}{\partial s} - \frac{\partial \Delta W}{\partial s}\frac{\partial P}{\partial \tau_1}$. I will show that $D < 0$. We have that D is linear in λ and b_1, so it's enough to show that $D < 0$ for $(\lambda, b_1) \in \{0, \tau_1\} \times \{\max\{\frac{1}{A}, \frac{1}{2}\}, \frac{1}{1+\lambda}\}$. We can verify this by brute-force algebra. ■