THE UNIVERSITY OF CHICAGO

ESSAYS ON ECONOMIC POLICY DESIGN IN THE PRESENCE OF POLITICAL CONSTRAINTS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY

SCOTT FRANCIS BEHMER

CHICAGO, ILLINOIS

JUNE 2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank the many advisors who have helped me throughout my PhD, particularly Leonardo Bursztyn, Michael Dinerstein, Wiola Dziuda, Mikhail Golosov, and James Robinson.

On the teaching side, Kevin Murphy and Casey Mulligan's excellent price theory course inspired me to pursue a PhD. Chris Blattman and James Robinson's political economy courses were also very influential on the way that I think about the world.

I am grateful to all the PhD students at the University of Chicago for being so fun, smart, and kind. It was really an excellent group to be around. Interacting with great peers was the main highlight of my PhD experience.

I am also grateful to the undergraduate students at the University of Chicago who I TAed for. Teaching motivated and intellectually curious students was another highlight of the PhD.

Finally, I would like to thank my family and friends, especially Marissa Higdon, for so much encouragement and support throughout these past six years.

ABSTRACT

This dissertation contains two essays on optimal policy design in the presence of political constraints.

The first essay relates to an active debate among economists regarding clean energy subsidies. The models used to inform this debate have a common simplifying assumption: the preferences of the government are kept constant over time. In reality, control of the government often rotates between parties with very different policy preferences. This paper finds that adding turnover in party control of the government can have significant implications. Specifically, the party more concerned about the environment ("the green party") finds it optimal to subsidize irreversible investments in clean energy, even when carbon taxes are available and can be placed at any level. Quantitative evidence suggests that this mechanism can justify the use of relatively large clean energy investment subsidies.

The next essay relates to the design of international climate agreements in a context where countries must be incentivized to voluntarily join the agreement. Existing literature has focused on two possible ways to incentivize participation in a climate agreement: 1. participants can threaten to increase carbon emissions if other countries don't join or 2. participants can threaten to place trade sanctions on countries who don't join. This paper offers a novel justification for using trade sanctions to incentivize participation: if there's a chance that punishments will have to actually be carried out, using trade sanctions is welfare improving because trade sanctions are far more efficient punishments than carbon emissions increases. This argument is formalized using a game with incomplete information. A calibration exercise suggests that using trade sanctions to enforce a climate agreement could significantly increase global welfare.

CHAPTER 1

CARROTS VS STICKS: CLIMATE POLICY WITH GOVERNMENT TURNOVER

1.1 Introduction

Textbook economic models suggest that the optimal climate policy is a pigouvian tax (i.e. a carbon tax). While carbon taxes are fairly common in reality, many countries' climate plans have instead relied heavily on clean energy subsidies. Perhaps motivated by this disconnect, there is an active debate among economists on the relative merits of subsidies vs carbon taxes[1].

This paper contributes to that policy debate by relaxing a standard assumption: that the preferences of the government are constant over time. Instead, I use a model where control of the government rotates between two parties who may disagree on the size of the externality from carbon emissions. I find that if the parties' valuations are identical, then the optimal policy only involves a carbon tax, as in the textbook model. However, if the parties' valuations of the externality differ, then the more environmentally-conscious party (the green party) finds it optimal to subsidize clean energy investments rather than to rely exclusively on carbon taxes. The intuition is that clean investment today crowds out future fossil fuel production that would occur under a less-environmentally-conscious party (the brown party), and this crowd-out benefit is not internalized by the private sector.

We use a model with perfect competition, where the only market imperfection is from carbon emissions. This is what guarantees the textbook result that if one party has control forever, their optimal policy is just a carbon tax.

The model has two types of energy, green and brown, which are substitutes in consumption. Production of each type of energy requires a specialized type of capital (i.e. power

1. See the literature review section below.

plants). Furthermore, we assume that green energy is highly capital intensive and that investment is irreversible. These assumptions imply that once green capacity is built, it produces electricity at very low marginal cost, so it will continue to produce electricity throughout its lifetime regardless of the carbon tax level. This is what drives the result that the green party can reduce future carbon emissions by increasing green investment.

When considering cases with party turnover, we make election outcome probabilities exogenous, which allows us to isolate the effects of party turnover from any effects due to electoral competition. We consider two versions of the model. The first is the "single-election" case, where green party is in charge in the first period, but if they lose the second period election then they lose power in all future periods. This model is a useful starting point for building intuition, and it allows for more general functional forms. Next we look at the “multi-election" version of the model where there is an election in every period. While this model requires more functional form restrictions, it allows for a more realistic calibration exercise.

Both models provide the same sufficient statistic for the green party's optimal subsidy, which depends on a few objects: the discount factor, the extent of polarization between the two parties, the marginal cost of green energy, and how much increased green energy production crowds out emissions. These objects can be estimated from existing empirical studies. By plugging these estimates into the sufficient static formula, we find an optimal clean investment subsidy of 9.5%. This is a relatively large subsidy. For context, the base level of the clean investment subsidy in the Inflation Reduction Act of 2022 is 6%[2].

While the sufficient statistic approach has the advantage of being transparent and less

2. In the Inflation Reduction Act, the base level of the subsidy is 6%. It increases to 30% if the investment meets prevailing wage and apprenticeship requirements, and up to 50% if it relies on American made products and is located in an environmental justice community. Many of the apparent goals of this policy (i.e. protectionism and redistribution) are outside of our model. The key point is just that our model predicts an optimal subsidy level that is non-trivial relative to what is observed in practice. Source: https://www.energy.gov/eere/water/inflation-reduction-act-tax-credit-opportunities-hydropower-and-marine-energy

dependent on functional forms, it comes at the cost of relying on local estimates of endogenous objects. For example, we use an estimate for the marginal cost of green energy that is measured using US data from 2019; if a very large subsidy were put in place, the marginal cost of green energy would increase above that 2019 level, and thus our estimate may be inaccurate. To address this concern, we provide an alternative estimate of the optimal subsidy using a structural approach. This requires calibrating the full cost and demand functions for energy. In the baseline calibration, we find a subsidy of 7.7%, only slightly lower than the results from the sufficient statistic approach. This provides further evidence that party turnover can justify quantitatively significant subsidies.

The calibrated model also allows us check what would happen if the green party were to naively use their no-turnover optimal policy (that is, if they were to use only use a carbon tax). We find that this would lead to significant under-investment from the green party's perspective. Specifically, green investment would be about 34% lower than their optimal level.

What about the behavior of the brown party? First, note that our model is asymmetric: in the baseline case, we assume that capital does not depreciate and that the initial stock of brown capital is high. This makes it so that the brown party has no incentive to subsidize brown capital investment[3][4]. Next, one may think that the brown party would benefit from taxing green investment. We don't find that in any version of the model. Instead, we find that in some cases the brown party wants to *subsidize* green investment[5]. The intuition is

3. Intuitively, there is already an excess amount of brown capital in the economy, so further brown investment has no impact on future production decisions.

4. While we think these are reasonable approximations given that the primary focus of the paper is on the green party's optimal policy (see the model setup section for more discussion of these approximations), we think that a valuable opportunity for future work is to relax these assumptions and look more carefully at whether the brown party wants to subsidize fossil fuel investment.

5. This result does *not* say that the brown party wants more investment with turnover than they want without turnover. In our model, they want the same level of investment in either case. Without turnover, that level of investment can be implemented with just a carbon tax. With turnover, they need to use a subsidy.

that since firms expect large green investment subsidies under a future green government, they significantly reduce green investment under the brown party. If the brown party only uses a carbon tax instead of a subsidy, firms end up investing below the brown party's optimal level. In our calibration exercise, the brown party's optimal subsidy is 60% as large as the green party's.

1.1.1 Literature Review

As was mentioned above, there is an active debate among economists on the merits of clean energy subsidies vs carbon taxes[6]. A common argument against subsidies, which is present in my model, is that they do not provide proper incentives to conserve energy. By lowering the cost of energy, subsidies lead to increased energy consumption, which is the opposite of the efficient decrease in consumption caused by a carbon tax. Related arguments say that clean energy subsidies do not give the proper incentives to phase out relatively high emissions sources like coal instead of relatively low-emissions sources like natural gas (Borenstein and Kellogg 2023). Finally, subsidies must be financed with distortionary taxation, whereas the revenue from carbon taxes can be used to reduce distortionary taxes (see Jorgenson et al 2013). Economic arguments in favor of subsidies include the idea that the electricity market has large mark-ups (Kellogg and Borenstein 2023) and that there are positive externalities from clean energy due to learning-by-doing (see Rodrik 2014 for an example). All of these papers use models in which the preferences of the government are constant over time. Our paper's contribution is to relax that assumption, which we find provides an alternative consideration in favor of subsidies.

By introducing political economy considerations into public economics models, we follow

6. In light of all the existing arguments, there does not seem to be a consensus on this topic. In a 2017 IGM poll of expert economists, 18% of respondents believed that subsidies were more efficient policies than carbon taxes, 60% stated that carbon taxes were more efficient, and 22% were uncertain. In written responses to the poll, many economists question existing arguments in favor of subsidies. No one mentions the mechanism highlighted in this paper. Source: https://www.kentclarkcenter.org/surveys/energy-sources/

the broad suggestion made by Acemoglu and Robinson (2013). They suggest that in contrast to the standard economics approach, which “ignores politics", “sound economic policy should be based on a careful analysis of political economy". While we don't think our paper provides a complete analysis of the political economy concerns regarding clean energy subsidies, we think that it is a valuable step in the right direction, especially relative to the standard approach of assuming one-party rule.

In using a model with party turnover in control of the government, this paper relates to a literature going back to Persson and Svensson (1989) (see Persson and Tabellini (2002) for a more detailed review). Most of this early literature focused on debt accumulation, and it was primarily theoretical. In environmental economics specifically, there are a few papers looking at government turnover, but none of them focus on the choice of carbon taxes vs subsidies for clean energy investment. Hochman and Zilberman (2021) use a model where the government can only use taxes, whereas Watten (2021) uses a model where the government can only use subsidies. Schmitt (2014) works almost exclusively with social planner models that don't decentralize the optimal policy in terms of taxes and subsidies[7][8]. Finally, Ulph and Ulph (2013) focus on R&D subsidies instead of physical capital subsidies. They use a two period model where governments can subsidize a binary clean energy R&D project, and they show that a subsidy is optimal for the green party when there's turnover. While their model contains a similar mechanism, our more general setting allows us to clearly identify the theoretical mechanism[9] and allows for more realistic empirical exercises.

7. This is an important difference. In our baseline model the green party's optimal amount of investment is identical with or without government turnover. However, their way of implementing that level of investment with tax policy is significantly different with turnover. So, if we only worked with social planner models and didn't consider the decentralized implementation, we would be left with the misleading conclusion that government turnover is irrelevant for optimal policy.

8. In Schmitt (2014)'s two period model, they point out that subsidies are needed to decentralize the optimal policy. However, in their infinite horizon model and their quantification section, they only look at the planner solutions.

9. Specifically, Ulph and Ulph (2013) state that subsidies are optimal because turnover increases the benefit of R&D investment for the green party, while decreasing the benefit for the private sector. Our analysis shows that while their conditions are sufficient to make subsidies optimal, they aren't necessary. In

The rest of the paper proceeds as follows: first we present the setup of the model and the results for the case where there's no turnover. Next, we present theoretical results for the single-election case, where power only changes hands once. We then provide an estimate of the sufficient statistic for the optimal subsidy. Finally, we present theoretical and calibration results for the multi-election case, where elections occur in every period.

1.2 Model Setup

1.2.1 Technology

There are two types of energy, brown (E_b) and green (E_g), which are perfect substitutes in consumption. Green energy is produced with only capital[10] and has a strictly increasing and strictly concave production function $F(K_t)$. Brown energy is produced with only non-durable inputs (e.g. fuel)[11] and has a constant marginal cost of production mc.

We assume for tractability reasons that capital does not depreciate over time[12]. Investments x_t immediately add to the capital stock according to the following law of motion: $K_t = K_{t-1} + x_t$. Importantly, investment is fully irreversible ($x_t \geq 0$).

Carbon emissions in each period C_t are proportional to brown energy production $C_t = \gamma E_{bt}$. Green energy production is completely carbon-free.

---

our model, it's ambiguous whether turnover increases the marginal benefit of green investment for the green party, but it's unambiguous that the green party wants to use a subsidy. The only thing needed to make a subsidy optimal is that green investment crowds out future carbon emissions under the opposite party. See section 4.2 for details.

10. This assumption, while made for tractability reasons, is not too far from reality. In 2022, capital costs represent 73% and 79% of the levelized cost of new wind and solar plants, respectively (Nalley and LaRose 2022).

11. In appendix A.1.2, we show that this assumption can be micro-founded in a model where brown energy requires capital, but where the initial stock of brown capital is large enough that there is excess brown capacity available. Empirical evidence suggests that this is a reasonable approximation.

12. Empirically, power plants are long-lived assets, with a lifespans of between 25-50 years (Rhodes et al (2017)). This translates to a depreciation rate of 2% to 4% per year.

1.2.2 Preferences and Consumer Problem

Following Kellogg and Borenstein (2023), we use a partial equilibrium model of the electricity sector. In appendix A.1.1, we show that this partial equilibrium model is equivalent to a general equilibrium model with quasi-linear utility and production. The value of electricity consumption for a representative consumer at time t is given by $v(K_{gt} + E_{bt})$. v is increasing and strictly concave. $v'(\cdot)$ satisfies $lim_{x\to\infty}v'(x) = 0$ and $lim_{x\to 0^-}v'(x) \to \infty$.

There is a representative consumer who acts as a price taker. Prices at time t are conditional on the history of election outcomes h_t. The representative consumer problem is to choose energy consumption $E(h_t)$ for each time t and history of election outcomes h_t[13] to maximize consumer surplus:

$$\sum_{t=1}^{\infty} \sum_{h_t \in H_t} \Pi(h_t)\beta^{t-1}(v(E(h_t)) - p(h_t)E(h_t))$$

where $p(h_t)$ is the price of energy, H_t is the set of possible histories at time t, and $\Pi(h_t)$ is the probability of h_t being realized.

The solution to this problem is to set E_t to satisfy $v'(E(h_t)) = p(h_t)$ for all h_t. Thus, $v'(E)$ gives the inverse demand for energy and $D(p) \equiv v'^{-1}(p)$ is the demand curve for energy.

1.2.3 Firm Problem

A representative firm, taking prices $p(h_t)$, carbon taxes $\tau(h_t)$, and investment subsidies $s(h_t)$ as given, chooses paths for green investment $\{x(h_t)\}$, green capital $\{K(h_t)\}$, and brown

13. In the full control case, there's only one possible history at every period in time. In the single-election case, the history is simply a binary variable for any $t \geq 2$, specifying who won the election at the beginning of period 2.

production $\{E_b(h_t)\}$ to maximize expected profits:

$$\sum_{t=1}^{\infty} \sum_{h_t \in H_t} \Pi(h_t)\beta^{t-1}(\underbrace{(F(K(h_t)) + E_b(h_t))p(h_t)}_{\text{Revenue}} - \underbrace{(1 - s(h_t))x(h_t) - (mc + \tau(h_t)\gamma)E_b(h_t)}_{\text{Cost}})$$

Subject to the law of motion and irreversible investment constraints for all h_t:

$$K(h_t) = K(h_{t-1}) + x(h_t)$$

$$x(h_t) \geq 0$$

where the exogenous initial capital stock K_0 is zero.

1.2.4 Competitive Equilibrium

A competitive equilibrium is a set of quantities $(\{x(h_t)\}, \{K(h_t)\}, \{E_b(h_t)\}, \{E(h_t)\})$, prices $(\{p(h_t)\}$, and tax rates $(\{\tau(h_t)\}, \{s(h_t)\})$ which solve the firm problem, solve the consumer problem, and satisfy the market clearing conditions $E_b(h_t) + F(K(h_t)) = E(h_t)$ for all h_t.

1.2.5 Party Social Welfare Functions

The two parties have identical payoffs except for their valuation placed on the externality from emissions. Both are interested in maximizing total surplus, inclusive of climate damages from carbon emissions. The green party and the brown party value the externality from a unit of carbon emissions at d_g and d_b, respectively, with $d_g > d_b$. Party j's utility is given by:

$$\sum_{t=1}^{\infty} \sum_{h_t \in H_t} \Pi(h_t)\beta^{t-1}(\underbrace{v(F(K(h_t)) + E_b(h_t))}_{\text{Consumption Value}} - \underbrace{x(h_t) - mcE_b(h_t)}_{\text{Cost}} - \underbrace{d_j\gamma E_b(h_t)}_{\text{Externality}})$$

Note that taxes and subsidies do not appear in this formula, since we assume that lump sum taxes are available (see the general equilibrium model in appendix A.1.1 for details).

We do not take a stance on which party has the “correct" social welfare function. The purpose of the model is to analyze the (realistic) case where the two parties disagree about the size of the externality.

1.2.6 Tax Instruments

The party in power in period t can use two tax instruments: carbon taxes τ_t and green energy investment subsidies s_t. As we'll see, these instruments turn out to be sufficient to implement each party's optimal allocation as a competitive equilibrium.

1.3 No Turnover Case

1.3.1 Solution Concept

As a baseline, we'll first look at the textbook case, where the same party is always in power. Following standard practice, we'll solve for that party's optimal policy in two steps: first, solve the planner problem to find the optimal allocation, then find the tax policy which decentralizes that optimal allocation as a competitive equilibrium.

The planner problem for party j is to choose an allocation $(\{x_t\}, \{K_t\}, \{E_{bt}\})$ to maximize their social welfare function:

$$\left[\sum_t \beta^{t-1}\Big(\underbrace{v(F(K_t) + E_{bt})}_{\text{Consumption Value}} - \underbrace{x_t - mc\ E_{bt}}_{\text{Cost}} - \underbrace{d_j \gamma E_{bt}}_{\text{Externality}} \Big)\right]$$

Subject to the law of motion and irreversible investment constraints for all t:

$$K_t = K_{t-1} + x_t$$

$$x_t \geq 0$$

Note that there is no implementability constraint here. We are assuming that the solution to this planning problem can be implemented with carbon taxes and investment subsidies. This assumption is later verified in theorem 2.

We make the following assumptions on parameter values, which guarantee that both parties use a positive amount of both types of energy in their no-turnover planner solution:

Interior Assumptions:

1. $F'(0)(mc + \gamma d_b) > 1 - \beta$
2. $F'(F^{-1}(D(mc + \gamma d_g)))(mc + \gamma d_g) < 1 - \beta$

The first of these guarantees that both parties use some green energy in their no turnover solution. The second guarantees that both parties use some brown energy in their no turnover solution.

1.3.2 Results

First, we characterize the planner solution for each party.

Theorem 1: *Party j's no turnover solution has the following features:*

1. *The capital stock in each period is equal to a constant K_j^*, which is the unique solution to the following condition:* $(mc + \gamma d_j)F'(K_j^*) = 1 - \beta$
2. *Investment is positive in the first period and zero in all future periods*
3. *Brown energy production is constant in all periods and is equal to* $D(mc + \gamma d_j) - F(K_j^*) > 0$

So, the planner immediately makes enough investment to bring the capital stock up to its steady state level, K_j^*. Because there's no depreciation, investment in all future periods is zero. Brown energy production is constant and positive in each period.

The condition in the first part of theorem 1 for the steady state level can be rewritten as:

$$\sum_{t=0}^{\beta} {}^t F'(K_j^*)(mc + \gamma d_j) = 1 \tag{1.1}$$

The left side of this equation is the marginal benefit of a one-time increase in investment, while the right side is the marginal cost. The marginal benefit is the present value of the marginal product of capital $F'(K_j^*)$ times the marginal benefit of energy consumption, which in equilibrium is equal to the full marginal cost of brown energy $(mc + \gamma d_j)$.

The third part of theorem 1 pins down the level of brown energy production. Since brown energy has constant marginal costs, optimal brown energy production for party j is simply given by the difference between demand at the social marginal cost $D(mc + \gamma d_j)$ and the supply of green energy $F(K_t)$.

Theorem 2 below gives the familiar result that a pigouvian tax on carbon emissions is sufficient to implement the optimal allocation as a competitive equilibrium:

Theorem 2: *Party j's no-turnover planner solution can be implemented as a competitive equilibrium with a carbon tax equal to d_j and a subsidy equal to zero in every period.*

Theorem 2 establishes that subsidies aren't needed to implement the optimal allocation. One may still think that subsidies could be used instead of carbon taxes to implement the optimal allocation. The following theorem shows that that's not the case. It is impossible to implement the allocation with positive subsidies.

Theorem 3: *There is no tax policy $(\{\tau_t\}, \{s_t\})$ which implements party j's no turnover solution and has $s_t > 0$ for any t*

To see why this equation holds, consider a perturbation where the firm makes a one-time marginal increase in investment at time t. In any competitive equilibrium, this deviation can't increase profits, which gives the following condition:

$$\sum_{t'} \beta^{t'} (mc + \gamma d_j) F'(K_j^*) \leq 1 - s_t \tag{1.2}$$

The left side is the marginal revenue from the deviation[14], while the right side gives the marginal cost. Notice that this condition is identical to equation 1.1, except for the presence of the subsidy on the right side. So, the only way for both equation 1.1 and condition 1.2 to hold (which they must if the tax and subsidy policy implements the planner allocation) is for $s_t \leq 0$[15]. If positive subsidies were used, then firms would invest in green energy to the point where the social marginal benefit is greater than the social marginal cost (from the perspective of party j).

1.4 Single-election Model

Now we add turnover into the model. The green party is in power in period 1. At the beginning of period 2, there's an election which the green party wins with probability $1 - \theta$. Whoever wins that election is in charge for all $t \geq 2$.

1.4.1 Solution Concept with Turnover

Following the public finance literature[16], our solution concept in this case is analogous to the no turnover case. First, we find the equilibrium allocation by solving for the subgame

14. The marginal revenue is the discounted sum of the price of energy times the marginal product of capital. In equilibrium, the price of energy is equal to the marginal benefit of energy consumption, $mc + \gamma d_j$

15. For $t = 0$, the firm FOC must hold with equality since positive investment is made in that period, which means that $s_1 = 0$ in any tax policy which implements the planner allocation. For $t > 0$, no investment is made, so s_t could be strictly negative (an investment tax) and still implement the planner allocation.

16. See Farhi and Werning (2008), and the references cited therein, for examples.

perfect equilibrium of the "social planner game", where the party in power in each period directly chooses that period's allocation. Next, we find tax policies for each party which decentralize the equilibrium allocation as a competitive equilibrium.

The idea behind this solution concept is to analyze settings where governments have a large enough range of policy instruments available to achieve any allocation (this includes tax instruments, but also quantity setting instruments such as cap and trade policies). So, to pin down the equilibrium allocation, we can first analyze a model where the party in power in each period directly chooses that period's allocation. We then want to focus on the specific case where parties choose to use taxes and subsidies, so we solve for the tax and subsidy policies which implement the equilibrium allocation.

1.4.2 Turnover Results

The Equilibrium Allocation

First, we characterize the equilibrium strategy for the party who wins the second period election. This is easy to solve for, since from period 2 on the game collapses to the no-turnover planner case considered in the previous section. The difference is that instead of the capital stock starting at zero, it now starts at K_1, the capital stock inherited from the first period:

Theorem 4: *In any subgame perfect equilibrium, the party j who wins the period 2 election chooses the following allocation, conditional on the amount of capital inherited K_1:*

1. $K_t = max\{K_j^*, K_1\}$ *for all* $t \geq 2$
2. $E_{bt} = max\{0, D(mc + \gamma d_j) - F(K_t)\}$ *for all* $t \geq 2$

The first part of theorem 4 states that the party who wins the second period election sets capital to their no-turnover optimal level K_j^* if feasible. If they inherit $K_1 > K_j^*$, then K_j^*

isn't feasible due to the irreversible investment constraint, so they instead invest nothing in all future periods and keep the capital stock at K_1.

The second part of theorem 4 pins down the level of brown energy production. Like in theorem 1, optimal brown energy production for party j is simply given by the difference between total energy demand $D(mc + \gamma d_j)$ and the supply of green energy $F(K_t)$ (if that difference is negative, then brown energy production is zero).

Theorem 4 leads to an important follow up result:

Corollary 1: *Let $\bar{K}$ equal the unique solution to $D(mc + \gamma d_j) - F(\bar{K}) = 0$. If $K_1 \in [K_j^*, \bar{K})$,* then $\frac{dE_{bt}}{dK_1} = -F'(K_1) < 0$ for all $t \geq 2$

In other words, corollary 1 says that if K_1 is above party j's optimal level but not high enough to completely crowd out brown energy production, then a marginal increase in K_1 crowds out $F'(K_1)$ units of brown energy production in period 2 on. So, the party in charge in period 1 can reduce future emissions by increasing period 1 investment. This crowding out effect is what leads to the result that the green party wants to subsidize green investment in the first period.

The following result characterises the on-path equilibrium allocation:

Theorem 5: *There is a unique subgame perfect equilibrium allocation with the following characteristics on the equilibrium path:*

1. *In period 1, the green party sets K_1 to K_g^* and E_{b1} to $D(mc + \gamma d_g) - F(K_g^*)$*
2. *In period 2 on, the party in power j sets $K_t = K_g^*$ and $E_{bt} = D(mc + \gamma d_j) - F(K_1)$ for all t*

So, the green party immediately invests up to their full control optimal level K_g^*. If they win the election in the second period, they keep the capital stock at that optimal level. If the brown party wins the second period election, they also keep the capital stock at K_g^*, since they are bound by the irreversible investment constraint.

It may seem surprising that the green party's optimal level of investment is the same in the case with turnover as it is in the case where they have control forever. As we mentioned above, with turnover investment has the additional benefit of crowding out future emissions, so why wouldn't the green party want to invest more in that case? To see why, consider a one-shot deviation where the green party marginally changes investment in period 1. This deviation can't be profitable in equilibrium, so we get the following condition:

$$\underbrace{p_1 F'(K_g^*) + \frac{\beta}{1-\beta}[(1-\theta)p_2(g)F'(K_g^*) + \theta p_2(b)F'(K_g^*)]}_{\text{Consumption Benefit}} + \underbrace{\frac{1}{1-\beta}\theta\gamma(d_g - d_b)F'(K_g^*)}_{\text{Crowd-out benefit}} = 1 \tag{1.3}$$

where $p_1 = p_2(g) = mc + \gamma d_g$ (the marginal benefit of energy consumption under the green party) and $p_2(b) = mc + \gamma d_b$ (the marginal benefit of energy consumption under the brown party). The right side is the marginal cost of the investment, which is just equal to one. The left side is the marginal benefit. The first two terms on the left side correspond to the consumption benefit of increased green energy production. The third term on the left side is the benefit from crowding out future emissions by the brown party.

The analogous condition in the no turnover case (equation 1.1) can written in a very similar form:

$$\underbrace{p_1 F'(K_g^*) + \frac{\beta}{1-\beta}p_2(g)F'(K_g^*)}_{\text{Consumption Benefit}} = 1 \tag{1.4}$$

Comparing the conditions with turnover (equation 1.3) and without turnover (equation 1.4), we see that there are two differences:

1. Equation 1.3 has the crowd out term on the left side. This makes the marginal benefit of investment larger in the turnover case than in the no turnover case.

2. Equation 1.3 has a lower consumption benefit term than equation 1.4. This is because with turnover there's a chance that the brown party will have power in the future and will significantly increase brown energy production, which lowers the marginal benefit of energy consumption. This makes the marginal benefit of investment *lower* in the turnover case than in the no turnover case.

So, there are two countervailing forces. In this version of the model with constant marginal costs of brown energy production, these two forces perfectly cancel out, and the optimal level of investment for the green party is identical in the case with turnover vs without turnover. If we relax the assumption of constant marginal costs (see the model in appendix A.1.3), the two forces may not perfectly cancel out, and instead it's ambiguous whether optimal investment is larger or smaller with turnover than without turnover. The key conceptual point, however, goes through in both models: the green party doesn't necessarily want to invest more when there's turnover vs when there's no turnover. The presence of party turnover doesn't have robust implications for the level of investment. However, as we see in theorem 6, the presence of party turnover does have robust implications for investment subsidy policy.

Equilibrium Tax Policies

Theorem 6: *The unique subgame perfect equilibrium allocation can be implemented as a competitive equilibrium with the following tax policies:*

- *In period 1, the green party uses:*
  1. *A positive investment subsidy equal to* $\theta(\frac{1}{1-\beta})(d_g - d_b)\gamma F'(K_g^*)$
  2. *A carbon tax equal to* d_g
- *From period 2 on, whichever party* j *is in power uses:*

1. *No investment subsidies*

2. *A carbon tax equal to d_j*

The second part of theorem 6 is straightforward. From the second period on, whichever party is in power is back to the no turnover setting, where the optimal policy is a carbon tax and no subsidies.

The first part of theorem 6 establishes that, unlike in the case with no turnover, it can now be optimal for the green party to subsidize green investment in the first period. The result below strengthens this, by establishing that the only way for the green party to implement their allocation is to use a positive subsidy in the first period:

Theorem 7:*Any tax policy that implements the subgame perfect equilibrium allocation has the following feature: In period 1, the green party uses a positive investment subsidy equal to* $\theta(\frac{1}{1-\beta})(d_g - d_b)\gamma F'(K_g^*)$

To build intuition behind this result, suppose that the green party naively used just a carbon tax equal to d_g (their no turnover optimal policy). In the no turnover case, this is enough to generate the efficient amount of investment K_g^*. In that case, firms expect brown energy to be expensive now and forever, so they find it optimal to immediately make large investments in green energy. Things are different when we add turnover. In that case, firms expect that, with probability θ, carbon taxes and energy prices will be much lower in the future. So it becomes optimal for them to invest less than K_g^* in the first period. To get around this under-investment issue, the green party can use an investment subsidy, which gets firms to again invest K_g^* in the first period.

An alternative way the green party could try to stimulate first period green investment is to use a carbon tax above their valuation of the externality d_g. In some cases, this can be enough to implement their optimal level of investment K_g^*. However, this can never implement their full optimal period 1 allocation because it will induce underproduction of brown energy.

To get intuition behind the exact expression for the optimal subsidy, consider a perturbation where the firm marginally changes investment in the first period. This must not be profitable in equilibrium, which gives us the following condition:

$$\underbrace{p_1 F'(K_g^*) + \frac{\beta}{1-\beta}[(1-\theta)p_2(g)F'(K_g^*) + \theta p_2(b)F'(K_g^*)]}_{\text{Marginal Revenue}} = \underbrace{1 - s_1}_{\text{Marginal Cost}} \tag{1.5}$$

where $p_1 = p_2(g) = mc + d_g$ (the price of energy under the green party) and $p_2(b) = mc + d_b$ (the price of energy under the brown party). This condition is very similar to the condition from the green party's problem (1.3). The differences are:

1. Equation 1.5 includes the value of the subsidy on the right side.

2. Equation 1.3 includes an additional term on the left hand side which captures the marginal benefit of crowding out future emissions under the brown party. This crowding out term is not internalized by firms because carbon taxes in the event that the brown party gains power are equal to γd_b rather than γd_g.

To make the two equations hold, the subsidy must be equal to the size of the crowd-out externality: $s_1 = \frac{1}{1-\beta}\theta(d_g - d_b)\gamma F'(K_g^*)$.

The Sufficient Statistic

To get a more general and convenient expression for the optimal subsidy, first note that we can write equilibrium carbon emissions if the brown party gains control as a function of first period investment: $C_b(x_1) = \gamma(D(mc + \gamma d_b) - F(x_1))$. The optimal subsidy can then be rewritten as:

$$s_1 = -\frac{1}{1-\beta}\theta(d_g - d_b)\frac{dC_b(x_1)}{dx_1} \tag{1.6}$$

This expression turns out to also hold when the model is generalized to allow for upward sloping marginal costs of brown energy (see appendix A.1.3), and we find the same expression in the multi-election version of the model.

Equation 1.6 says the the optimal subsidy only depends on:

1. $\frac{dC_b(x_1)}{x_1}$, the extent to which increased investment today reduces future carbon emissions *in the event that the brown party gains power.* Notice that increased investment today also crowds out future emissions in the event that the green party keeps power, but (from the green party's perspective) there's no externality from that because emissions under future green governments are correctly priced.

2. The level of polarization $(d_g - d_b)$. If there is no polarization, green party expects that future brown governments will correctly price carbon emissions, and so there's no need to subsidize green investment. As polarization increases, subsidies increase.

3. The probability of losing power θ. If $\theta = 0$, then we are back to the case where the green party has full control, and the optimal subsidy is zero. As θ increases, the probability that investment today will crowd out under-priced future emissions increases, and so the subsidy increases.

4. The discount factor β. If the discount factor is higher, then the green party cares more about reducing future emissions, and thus the subsidy is higher.

The Importance of Irreversibility

Finally, we would like to point out that the irreversibility constraint plays a key role here, as the theorem below shows:

Theorem 8: *If investment is fully reversible, then there is a unique equilibrium allocation and any tax policy which implements that allocation involves:*

1. *Each party j uses a carbon tax equal to d_j whenever they're in power*
2. *Investment subsidies are always equal to zero*

So, if investment is fully reversible, then we're back to the textbook case where subsidies are suboptimal. In this case, equation 1.6 still holds, but marginal first period investment has no impact on future emissions ($\frac{dC_b(K_1)}{dx_1} = 0$), since future governments can simply reverse the investment. So, regardless of other parameter values, the optimal subsidy is always zero with fully reversible investment. Although fully reversible investment is commonly assumed in economic models for tractability purposes, we consider it to be very unrealistic, as it implies that an existing wind or solar farm can be taken apart and sold to recover the entire original cost of building it.

1.5 Estimating the sufficient statistic

This takes us to our first method for quantifying the optimal subsidy: estimating the quantities in the sufficient statistic formula (equation 1.6).

First, we'll rewrite equation 1.6 in a way that can be directly estimated from the data. As was noted earlier, we can write carbon emissions under the brown party from period 2 on as a function of x_1: $C_b(x_1) = \gamma(D(mc + \gamma d_b) - F(x_1))$. Since x_1 can be written as a function of green energy production from period 2 on ($x_1 = F^{-1}(E_g)$), we can also write future carbon emissions as a function of future green energy production: $\hat{C}_b(E_g) = \gamma(D(mc + \gamma d_b) - E_g)$. Taking derivatives, we find:

$$\frac{dC_b}{dx_1} = -F'(x_1)\frac{d\hat{C}_b}{dE_g}$$

Estimates of $\frac{d\hat{C}_b}{dE_g}$ (the change in emissions caused by increased green energy production) already exist in the literature.

Next, define the marginal levelized cost of electricity as:

$$MLCOE(K) = (1-\beta)/F'(K)$$

This corresponds to the marginal cost of producing a unit of green energy.

With these definitions, we can rewrite our expression for the optimal subsidy as:

$$s_1 = -\frac{\beta\theta(d_g - d_b)\frac{d\hat{C}_b}{dE_g}}{MLCOE(x_1)} \tag{1.7}$$

We estimate the quantities in equation 1.7 in the following way:

- The discount factor β is estimated from market interest rates. For the baseline estimates, we use 5% a year (sensitivity analyses are done for 7% and 3% as well). Since the period length is taken to be four years, this corresponds to $\beta = .95^4 = .81$
- θ is set to 1/2, based on the balance of power between the two US political parties in the past 30 years.
- d_g = \$51 per ton of CO2, based on the Biden and Obama administrations' official social cost of carbon estimate[17][18].
- d_b = \$1 per ton of CO2, based on the Trump administration's social cost of carbon estimate.
- The crowd out effect $\frac{d\hat{C}_b}{dE_g}$ is taken from Abrell et al (2018). Their average estimate is .3 tons of CO2 per MWh of wind and solar production.

17. Source: https://www.eenews.net/articles/federal-agencies-can-use-social-cost-of-carbon-for-now/

18. In the US, government agencies are required to do a cost benefit analysis when passing new regulations. The president sets an executive-branch wide number, called the social cost of carbon, which is to be used by agencies to value the externality from carbon emissions. So, these social cost of carbon numbers have real consequences for policy decisions.

- $MLCOE(x_1)$ is taken from Kellogg and Borenstein (2023), who estimate the full marginal levelized cost of solar and wind in 2019 to be equal to \$64 per MWh.

More details on these estimates are presented in the appendix A.1.4.

1.5.1 Results

Using these baseline estimates, we find the green party's optimal subsidy to be 9.5%. For reference, the clean investment subsidies included in the inflation reduction act have a base level of 6%, and rise to 30% if the investment meets prevailing wage and investment requirements. So, our estimates are roughly in line with the level of subsidies observed in practice.

A number of sensitivity analyses are shown in appendix A.1.4. These show that the plausible range for the optimal subsidy is very large. The lowest value found in the sensitivity analysis is an optimal subsidy of 4.3%. The highest value found is 66%. In light of this wide range, we interpret these results as only suggestive evidence that government turnover justifies large subsidies for the green party. We also think it highlights important opportunities for future empirical work. For example, more precise empirical estimates of the crowd out effect $\frac{d\hat{C}_b}{dE_g}$ would greatly increase the precision of our estimates.

One especially important sensitivity check to highlight involves the social cost of carbon estimates. The Biden administration is currently using a temporary value of \$51 per ton of CO2. They are considering an updated value of about \$191/ton of CO2[19]. If we use this new value in our estimation, then the optimal subsidy estimate rises to 35.6%, which is in line with the higher end of the subsidies in the Inflation Reduction Act.

Finally, a caveat to these estimates is that they are based on local empirical estimates of the crowd out effect and of the marginal cost of green energy. As policies change, these empirical objects change endogenously, which means that the local estimates used in this

19. Source: https://www.eenews.net/articles/epa-floats-sharply-increased-social-cost-of-carbon/

exercise may not be accurate. This is a common issue with the sufficient statistic approach, and it motivates the use of a more structural approach, such as our calibration exercise in section 7.

1.6 Multiple-election Model

The setup of the multiple-election model is the same, except now elections happen in every period. We assume that each party has a 50% chance of winning each election.

In this more realistic setting, we end up finding the same behavior for the green party as in the first period of the single-election model (Theorem 9 shows that they choose the same level of investment. Theorem 10 shows that they choose the same subsidy). The brown party's behavior is different from the single-election model, as they now want to subsidize green investment (see theorem 10 and the following discussion).

Before showing the results in detail, we first define our equilibrium refinement in this case.

1.6.1 Solution Concept

This version of the model has an issue that is common with many infinitely repeated games: if the discount factor is low enough, the set of subgame perfect equilibria is extremely large. For example, there's a subgame perfect equilibrium where both parties use no green energy, and if any party does use green energy, then the other party will punish them by producing a very high amount of brown energy in future periods[20]. To rule out these implausible-seeming punishment equilibria, we need a stronger equilibrium refinement.

To motivate our equilibrium refinement definition, recall that in the full-control case and the single-election case, the parties' optimal investment strategies had the following form:

20. This is in the spirit of standard folk theorem results that rely on minimax punishments (see Fudenberg and Maskin 1986). A full example of an equilibrium strategy profile is spelled out in appendix A.1.6

$x_t = max\{0, K_j^* - K_{t-1}\}$. That is, party j has some capital target K_j^*; if at any time t they inherit a capital stock below K_j^*, then they invest enough to bring the capital stock up to K_j^*; if they inherit a capital stock above K_j^*, then they invest nothing. Our refinement proposed below requires that strategies have this same form:

Definition 1: *A "History Independent Capital Target Equilibrium" is a Markov Perfect Equilibrium where strategies satisfy the following condition:*

- *There exists some (K_b, K_g) such that whenever party j is in power, they set x_t to* $\max\{0, K_j - K_{t-1}\}$

The Markov Perfection requirement means that equilibrium strategies can only depend on payoff relevant variables, which in this case is just the inherited level of capital K_{t-1}. This rules out strategies which use brown energy production as a punishment for past behavior. The requirement on investment behavior rules out strategies which use investment as a punishment for past behavior[21]. In what follows, we will refer to any History Independent Capital Target Equilibrium as simply "an equilibrium".

1.6.2 Results

Theorem 9: *There is a unique equilibrium allocation with the following characteristics:*

1. *Whenever party j is in power, they set investment to* $\max\{0, K_j^* - K_{t-1}\}$
2. *Whenever party j is in power, they set E_{bt} to $D(mc + \gamma d_j) - F(K_t)$*

So, if the brown party gains power first, they raise the capital stock to K_b^* and keep it there until the green party comes to power. Once the green party first gains power, they raise the capital stock to K_g^* and it stays there forever.

21. We haven't proven that there exist other Markov Perfect Equilibria other than the unique History Independent Capital Target equilibrium. There's a chance that any Markov Perfect equilibrium also has history independent capital targets.

Recall that K_j^* is party j's capital target when they have permanent control of the government. So, like in the single-election model, we find that each party sets a capital target equal to their target in the no-turnover case. And, similar to the single-election model, the theorem below says that to implement that capital target requires a different tax policy than in the no-turnover case:

Theorem 10: *Any tax policy which implements the equilibrium allocation as a competitive equilibrium has the following characteristics:*

- *Whenever party j is in control, they use a carbon tax equal to d_j.*
- *In the first period that the green party gains control, they use a subsidy s_g^* equal to $(1/2)(\frac{\beta}{1-\beta})(d_g - d_b)\gamma F'(K_g^*)$*
- *In the first period, if the brown party has control, they use a subsidy equal to $\frac{(1/2)\beta}{1-(1/2)\beta} s_g^*$*

There are a few things to notice about theorem 10. First, the green party's behavior is the same as their behavior in the first period of the single-election model. They use both a carbon tax and a subsidy on green investment. The expression for the subsidy is also identical, and the intuition is the same: Marginal investment today crowds out under-taxed brown energy production under future brown governments. From the green party's perspective, this is a positive externality and thus investment should be subsidized.

A new thing here is the result that the brown party wants to subsidize clean investment. To get intuition for this, consider first what happens when the brown party has permanent control. In that case, the brown party finds it optimal to only use a carbon tax equal to d_b, and firms respond by setting the capital stock to K_b^*. Now consider the case where there's turnover, and the brown party naively uses just a carbon tax equal to d_b in the first period. Since firms expect investment to be subsidized by a future green government, they find it optimal to invest less than K_b^* and wait until the green party comes to power to invest more.

According to theorem 9, the brown party still wants firms to invest K_b^*, so they get around this underinvestment issue by using a subsidy[22]. Note, however, that the subsidy is still smaller than the green party's (with a discount factor of 5%, s_b^* is about 60% as large as s_g^*)[23]. Also recall that we assumed that the initial stock of green capital is zero. If the initial stock is at or above K_b^*, then there's no need for the brown party to use any subsidy since their optimal level of investment is zero anyway.

Finally, note that theorem 8 only says that subsidies are positive for the green party in the first period when they have control. What about the periods after that? In the equilibrium allocation, after that period, no investment is made by either party. Since investment is irreversible, this means that firms are at a corner solution in those periods, and so any subsidy which is low enough to get them to invest zero can implement the equilibrium allocation. They could implement the optimal allocation by using the same subsidy that they used in the first period, but they could also implement it by using any subsidy level below that. The same reasoning applies to why theorem 10 only pins down the brown party's optimal subsidy in the first period[24].

The Sufficient Statistic

We can rewrite the optimal subsidy to get the same sufficient statistic formulas that we found in the single election case (equations 1.6 and 1.7). First, consider a deviation where

22. Another way of looking at this: from the brown party's perspective, investment under a future green party has a negative fiscal externality due to it being subsidized. Investment today crowds out future investment, which mitigates the negative fiscal externality, so investment today has a positive externality.

23. Preliminary evidence suggests that this result of a positive subsidy may not be robust to changes in the setup. For example, if investment doesn't immediately add to the capital stock, but instead increases it next period, then it's ambiguous whether the brown party wants to tax or subsidize green investment. However, even in that model, the mechanism driving the brown party to want to subsidize investment is still present. There's just another force pushing in the opposite direction which may turn out to be stronger.

24. More formally:
Proposition 1: *The equilibrium allocation can be implemented with party j using a carbon tax equal to d_j and an investment subsidy equal to s_j^* whenever they're in power.*

the green party increases investment in the first period to $x_1 \geq x_1^*$. Given the equilibrium behavior in theorem 9, we can write emissions in all future periods when the brown party has control as function of x_1:

$$C_b(x_1) = \gamma(D(mc + \gamma d_b) - F(x_1))$$

Using this expression, we can rewrite the optimal subsidy formula as:

$$s_g^* = -(1/2)\frac{\beta}{1-\beta}(d_g - d_b)\frac{dC_b(x_1)}{dx_1}$$

where the derivative $\frac{dC_b(x_1)}{dx_1}$ is evaluated at the equilibrium investment level from theorem 9 ($x_1^* = K_g^*$). This is the same as what we found in the single election case (equation 1.6), except now the probability of turnover θ is set equal to 1/2. Just as in the single-election case, the optimal subsidy depends on the level of polarization, the discount factor, and how much investment today crowds out future emissions under brown governments.

Next, following the same steps we used to derive equation 1.7, we find an identical expression in this version of the model:

$$s_g^* = -\frac{(1/2)\beta(d_g - d_b)\frac{d\hat{C})b}{dE_g}}{MLCOE(x_1^*)}$$

where $\frac{d\hat{C})b}{dE_g}$ is the impact of a marginal increase in green energy production on carbon emissions under the brown party[25] and $MLCOE(x_1^*)$ is the marginal levelized cost of green energy.

Since the sufficient statistic formula is identical, the empirical estimates of the optimal subsidy that we found in section 5 also apply in this version of the model. As was mentioned in section 5, this sufficient statistic approach has the benefit of being transparent and relying

25. This is evaluated at the equilibrium level of green energy production given in theorem 9 ($F(K_g^*)$).

on well-identified micro estimates, but it comes at the cost of relying on local estimates of endogenous objects. To address this concern, in the next section we provide alternative estimates of the green party's optimal subsidy using a more structural approach.

1.7 Calibration

In this section we calibrate the model to provide alternative estimates for the optimal subsidy and to get policy counterfactuals.

We assume a constant elasticity of demand function and a linear marginal cost function for green energy[26].

Values for parameters β, d_g, and d_b are the same as in the sufficient statistic exercise. Other parameters are calibrated in the following way:

- The demand elasticity is taken from the long run estimate in Deryugina et al (2020).
- The constant in the demand function is set to match 2022 levels of electricity demand.
- The parameters of the linear marginal cost function are taken from Borenstein and Kellogg (2023).
- The emissions content of brown energy γ is set to match the crowd out effect from Abrell et al (2018).

Details of these estimates are in the appendix A.1.5.

One key parameter estimate is the value of γ, the emissions content of brown energy. In the baseline, we report results for $\gamma = .18$ and $\gamma = .59$ tons of CO2 per MWh. This value was chosen so that the crowd out effect of increased green energy production (measured at the laissez faire equilibrium) agrees with the upper and lower bound from Abrell et al (2017),

26. See appendix A.1.5 for details of what these assumptions imply for the functional forms of $v(E)$ and $F(K)$

which is the same study that was used in the sufficient statistic estimation. An alternative way to set γ is to use the average emissions content of non wind and solar energy in the US economy. This leads to a value of .46 tons of CO2 per MWh, which is within our baseline range[27]. In either case, the crowd out effect in the calibrated model agrees with the micro evidence from Abrell et al (2017). This is important since the sufficient statistic formula tells us that the crowd out effect is a key determinant of the optimal subsidy.

1.7.1 Results

In the baseline calibration, we find that the green party's optimal subsidy is 7.7%. This is slightly lower than what we found in the sufficient statistic estimation. The reason for this discrepancy is that the sufficient statistic relied on local estimates of the marginal cost of green energy around the laissez faire equilibrium. With positive subsidies in place, the marginal cost of green energy rises, which reduces the optimal subsidy.

For the brown party, we find an optimal subsidy of about 5.2%, suggesting that government turnover can justify non-trivial subsidies even for the party who cares less about climate change.

Figure 1.1 shows a key comparative static: the green party's optimal subsidy as a function of the amount of polarization[28]. We see that for low levels of polarization, subsidies are near the full control level of zero. As polarization rises, subsidies rise to non-trivial levels.

Appendix A.1.5 shows a number of sensitivity analyses. One important scenario to

27. Average emissions content of non solar and wind energy is taken from Nalley and LaRose (2022)

28. For every point on the graph, the average of the two party's social costs of carbon $\frac{d_g+d_b}{2}$ is held constant at $\bar{d} = \$26$ per ton of CO2 (the average social cost of carbon in our baseline estimate). As polarization σ increases, the two parties' social costs of carbon become further apart:

$$d_g(\sigma) = \bar{d}(1+\sigma)$$

$$d_b(\sigma) = \bar{d}(1-\sigma)$$

$\sigma = 0$ gives the full control case, where the two parties have identical preferences. $\sigma = 1$ corresponds to the extreme case where the brown party places no value on the externality and the green party places $2\bar{d}$. In our baseline calibration, $\sigma = .96$.

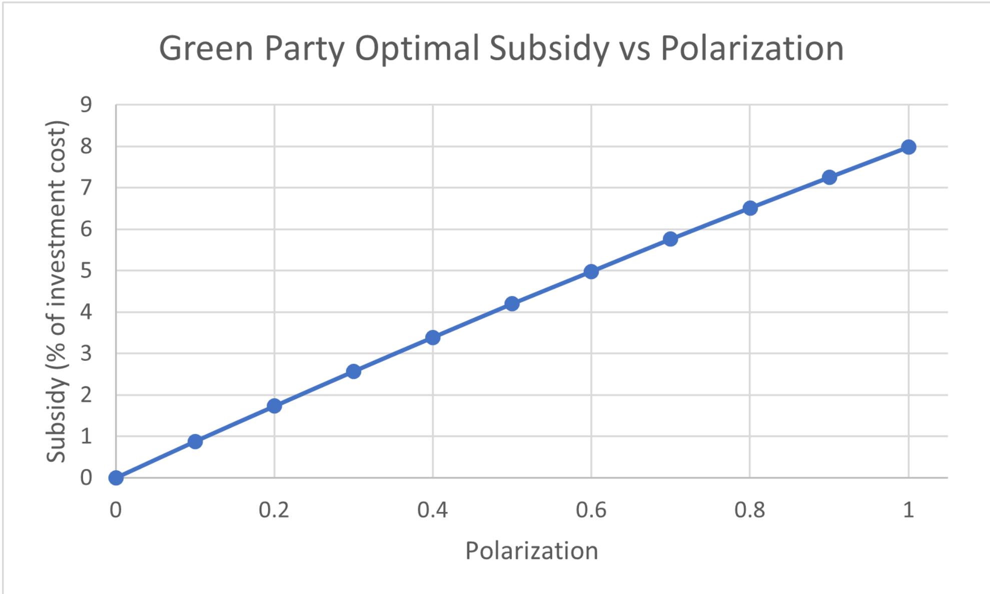

Figure 1.1: The green party's optimal subsidy as a function of polarization

Notes: With no polarization, we get the textbook result that subsidies are not optimal. As polarization rises, the optimal subsidy increases.

highlight is if we set d_g equal to the Biden Administration's suggested social cost of carbon number of \$195/ton of CO2. In this case, the green party's optimal subsidy rises significantly to 21.5%.

The calibration exercise also allows us to check what would happen in the counterfactual where the green party naively uses their no-turnover solution of just a carbon tax equal to d_g. For this exercise, we assume that the brown party also acts naively and uses their no-turnover solution of a carbon tax equal to d_b.

The results are shown in figure 1.2 for the baseline calibration. The first bar shows the steady state level if the green party had full control and just used a carbon tax. This permanent carbon tax is enough to get firms to invest up to the green party's optimal level. The next bar shows what happens if the green party naively follows that same policy when there's turnover. In this case, investment from firms is 29% lower, since they expect low carbon taxes under future brown governments. The final bar shows what happens if there's turnover and the green party acts optimally (that is, if the green party uses the optimal policy of a tax *and* subsidy). In that case they induce the same large amount of investment as in the full control case.

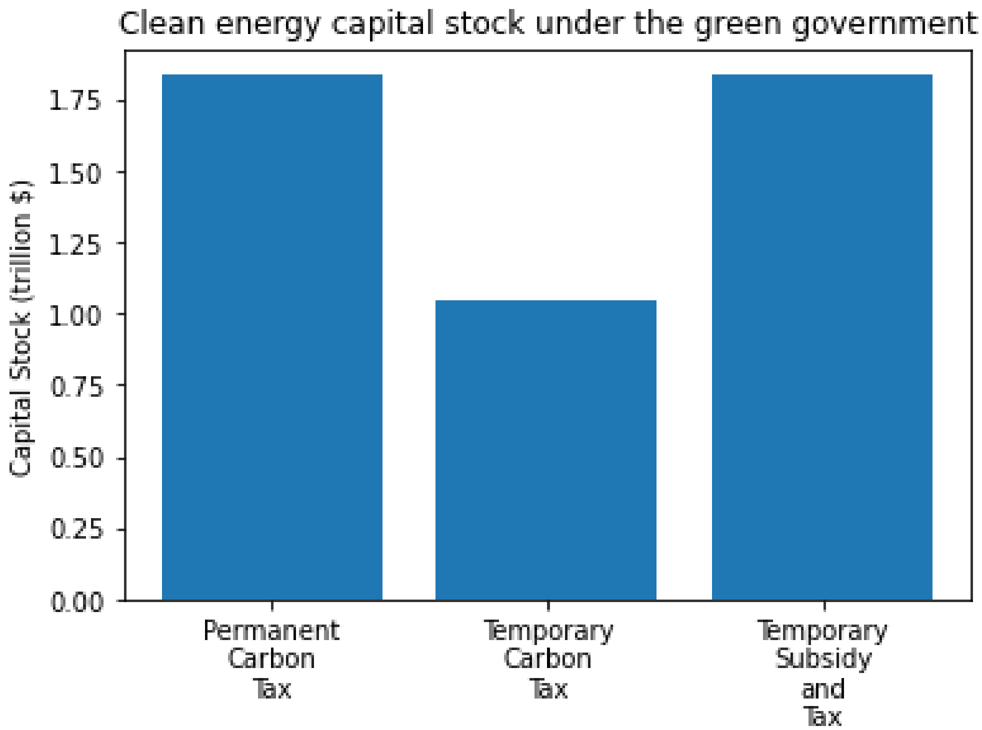

Figure 1.2: The steady state capital stock in three scenarios

Notes: "Permanent carbon tax" is the case where the green party has full control and acts optimally. "Temporary carbon tax" is the case where there's turnover and the green party naively uses just a carbon tax. "Temporary subsidy and tax" is the case where there's turnover and the green party uses their optimal policy, which is a carbon tax and an investment subsidy. The y axis is measured relative to the laissez faire level of capital.

1.8 Conclusion

This paper makes a novel contribution to the debate over the value of clean energy subsidies. Specifically, if control of the government rotates between parties with different levels of concern about climate change, then the more concerned party finds it optimal to use clean energy subsidies. This result relies on the assumption that investment is irreversible, so that renewable plants built under the green party will remain operational under future brown governments, and thus will crowd out future fossil fuel production.

The model gives a simple sufficient statistic formula for the optimal subsidy. The key empirical objects that it depends on are the discount factor, the difference between the two parties' social costs of carbon, and the extent to which increased clean investment reduces future carbon emissions. This both guides our quantification exercises, and (as is discussed more below) suggests opportunities for future empirical work.

We use two methods to investigate whether the green party's optimal subsidy is quantitatively large. First, we estimate the objects in the sufficient statistic formula using existing estimates in the literature. This has the advantage of being transparent and less dependent on functional forms, but comes at the cost of relying on local estimates of endogenous objects. To address some of these concerns, our second method is to calibrate the full model. In each case, the results suggest that the optimal clean energy investment subsidy is relatively large, between 5% and 17%. The calibration exercise also gives results for the counterfactual case where the green party naively uses just a carbon tax rather than a subsidy. In this case, we find that this leads to large underinvestment in green energy relative to the optimal level. In total, these results provide suggestive evidence that government turnover matters quantitatively for optimal climate policy.

There are many opportunities for future work on this topic. First, the models presented in this paper are still fairly stylized, and a number of assumptions could be relaxed to test the robustness of the results. A promising route to do this is to switch to a finitely repeated game, as in Schmitt (2014), since many of the strong assumptions in our model were needed to make the infinitely repeated game tractable[29]. With a finite time horizon, we could likely use a model with non-zero depreciation, more than two types of energy, and more than one

29. The specific issue was the multiplicity of subgame-perfect equilibria in the infinitely repeated game. Our strong assumptions were needed so that we could use a sensible refinement to shrink the set of equilibria. With a finitely repeated game, the set of subgame-perfect equilibria is typically much smaller.

input for each type of energy.

Next, the results in this paper suggest opportunities for future empirical work. Specifically, more precise estimates of the impact of green investment on future emissions would significantly reduce uncertainty about the green party's optimal subsidy.

An interesting extension of the model would be to look at possible pareto-improving compromises between the two parties. For example, if the brown party agrees to raise their carbon taxes conditional on the green party lowering their subsidies, then both parties could potentially be better off.

Finally, and more broadly, future work could use a similar model to examine whether government turnover matters in other policy contexts. For example, previous empirical work has found that trade policy reforms in developing countries are often reversed by future governments (Rodrik 1992). Perhaps in this context a reform-minded political party would find it optimal to subsidize irreversible investments in exporting sectors, rather than to just lower trade barriers. Similar mechanisms may be at play for a wide variety of policy issues on which political parties are polarized.

CHAPTER 2

ISSUE LINKAGE IN INTERNATIONAL AGREEMENTS WITH INCOMPLETE PARTICIPATION

2.1 Introduction

Global public goods problems include many of the most pressing issues of our time, such as climate change, global health, refugee resettlement, conflict prevention, and global poverty alleviation. In a noncooperative environment, standard free rider problems result in inadequate funding for global public goods, as each country fails to internalize the externalities that their actions have on others. This creates the potential for significant pareto improvements if countries can cooperate on policy, and, in fact, all of the above mentioned global public goods are subject to international agreements.

A major challenge with agreements on public goods is that participation is voluntary: countries must expect to be punished for not joining the agreement, and it must be rational for other countries to carry out the punishments. The simplest way to punish non-participants is to reduce provision of public goods by participants. For example, a climate agreement could specify that if any country doesn't join, participant countries will increase carbon emissions. However, some international agreements instead involve "issue linkage", where other policies are used as punishments for free riding on the public good. Most notably, the Montreal Agreement on ozone-depleting substances imposes trade sanctions on non-participants. Partially due to the success of that and other agreements, there have been recent calls to include trade sanctions in future climate agreements (Nordhaus 2015).

Existing literature offers a few justifications for issue linkage. The most relevant to this paper are Nordhaus (2015) and Barrett (1996)[1], who argue that countries cannot credibly threaten to cut funding for the public good because that punishment would be costly to carry

1. a full literature review is presented in section 2

out. In comparison, trade sanctions can be credibly threatened because they are significantly less costly to carry out. Thus, the threat of punishment needed to sustain cooperation on public goods is only possible with issue linkage. However, these papers rely on ad hoc restrictions on the strategies of countries (as was pointed out by Maggi (2016)). Also, these papers imply that issue linkage wouldn't be valuable if countries had some way to commit to threats.

This paper offers an alternative justification for issue linkage. In contrast to previous work, I impose no strong assumptions on credibility[2]; both public good cuts and trade sanctions can be credibly threatened. Gains from issue linkage are instead driven by the inclusion of incomplete information about countries' preferences. Specifically, each country in the model has a chance of being a "isolationist type" which won't participate in any international agreement. This creates the possibility that even if countries are credibly threatened with public good cuts as punishments, some countries still won't participate in the international agreement, and thus the punishments will have to be carried out. This, in turn, means that less costly punishments are preferred, which creates gains from issue linkage because trade sanctions are significantly less costly than public good cuts.

Why are trade sanctions less costly than public good cuts? The intuition for this is simple. The average country internalizes less than 1% of the benefits of a global public good. This means that to punish that country by $1 through public good funding cuts, the rest of the world must hurt itself by about $100. On the other hand, to punish that country by $1 with trade sanctions, the rest of the world may actually benefit through the terms of trade effect. Even without any terms of trade gains, most of the deadweight cost from sanctions will be internalized by the target of the sanctions, so the rest of the world will be hurt by less than $1.

My baseline model only has two periods. In the first period, countries choose whether

2. I use a standard equilibrium refinement, sequential equilibrium, and have no ad hoc restrictions on countries' public good or trade policies.

to sign on to a binding international agreement which determines their tariff and carbon emissions levels in the second period. This simple model is enough to generate the main result that issue linkage will be beneficial as long as tariffs are a more efficient punishment than carbon emissions increases. Next, I calibrate the model using data from Nordhaus (2015). The calibration exercise suggests that the quantitative gains from issue linkage are large in this context. Importantly, the calibration exercise finds significant welfare gains even when the probability of a country being isolationist is very small (i.e. 1 in 1000).

The rest of the paper is structured as follows: section 2 is the literature review, section 3 shows the setup of the baseline model, section 4 shows theoretical results for when issue linkage is beneficial, section 5 shows the infinitely repeated game version of the model, and section 5 presents the calibration exercise.

2.2 Literature Review

The existing literature has to overcome folk-theorem results which state that in a repeated game environment with perfect information and sufficiently patient countries, full cooperation can be achieved without issue linkage. One strand of the literature shows that when countries aren't patient enough, issue linkage can be beneficial because it can increase the severity of the maximium possible punishment (see examples in Maggi 2016). However, for many public goods, it is hard to see how impatience could be a major issue unless it takes many years to detect a deviation[3]. Also, it is not clear that the strength of the maximum possible punishment is a relevant constraint. As Maggi (2016) points out in his survey of the literature, it's often not optimal to use the maximum possible punishment because the

3. For example, in the climate change calibration presented below, the the per capita benefits from an efficient climate agreement are around $1000/year, while the per capita costs are around $250/year. If a country chooses to deviate and not contribute, they would save $250 per year until the deviation is detected, and then they would lose $750 per year until the end of time (assuming that nash reversion punishments are used). Standard discount rates are lower than 10% per year, so unless it takes decades to detect a deviation, the short run benefits cannot overcome the long term costs of deviating.

real world is complicated enough that punishments occur with positive probability. My paper captures one such "complication": there's a chance that some countries won't value the public good.

A few papers look at issue linkage in a context where punishments occur with positive probability (Ederington 2003, Bajona et al 2012, Chisik 2010). These papers each find cases in which issue linkage can actually be *worse* than using unlinked agreements, which is the opposite of what I find in my model. The key difference is that these papers assume that only nash reversion punishments can be used. Nash reversion punishments in unlinked agreements are often already excessively strong, so there's no need for issue linkage. Furthermore, issue linkage in that context can actually be harmful since it leads to even more excessively costly punishments being used. In my model, in contrast, international agreements are not limited to using only nash reversion punishments. In this case, issue linkage has benefits because it allows countries to substitute away from inefficient, costly punishments (like increasing carbon emissions) toward more efficient punishments (like trade sanctions).

Another strand of the literature, covered briefly in the introduction, uses models which place stronger restrictions on credibility than subgame perfect equilibrium. If more costly punishments, like public good reductions, can't be credibly threatened, then cooperation can only be achieved by using less costly punishments like trade sanctions. Nordhaus (2015) uses a repeated game and requires that the solution be a coalition-proof Nash equilibrium of the stage game, which rules out history dependent strategies. Barrett (1996) uses a two-period model in which countries form coalitions in the first period and coalitions set policies in the next period. Both papers have been criticized for relying on ad hoc restrictions on trade sanctions, which are needed to get the main results (Maggi 2016). My paper, in comparison, does not include ad hoc restrictions on strategies.

It's also worth noting that a number of authors have argued that, at least in some cases, countries have ways to commit to threats (Fearon 1994) (Schelling 1960). If credibility issues were the sole justification for issue linkage, then the presence of these commitment mechanisms would make issue linkage unnecessary. In my model, however, countries can commit to using powerful punishments and yet issue linkage is still valuable.

In view of the entire literature, the main contribution of this paper is to show that the case for issue linkage is robust. Even if countries are fairly patient and have some ability to commit to costly punishments, there are still significant gains from issue linkage.

2.3 Model

2.3.1 Actions and Timing

There are N identical countries and two periods. In the second period, each country i will set their level of carbon emissions cuts $c_i \geq 0$ and uniform tariffs $\tau_{i,j} \in [0, \bar{\tau}]$[4] on each other country j.

In the first period, countries will choose whether or not to join international agreements which will determine their carbon emissions levels and/or tariff levels in period 2. In the unlinked setting, there are two separate international agreements, a trade agreement and a climate agreement. The trade agreement specifies tariff levels on non-participating countries $\tau_{out}(n_{trade})$ as a function of the number of participating countries n_{trade}. The trade agreement always restricts tariffs on other participating countries to be equal to zero[5]. Countries who don't join the trade agreement are free to set whatever tariffs they want in period 2. Similarly, the climate agreement specifies the level of emissions cuts by participating countries $c(n_{climate})$ as a function of the number of participating countries $n_{climate}$. Countries who don't join the climate agreement are free to choose any emissions level in period 2.

4. The upper bound on tariffs $\bar{\tau}$ is taken to be larger than the nash tariff level, which is defined later in the section.

5. Assuming that tariffs between participating countries are equal to zero is only done for simplicity of exposition. If we relax this assumption, the results will be identical, since the optimal agreement will always end up specifying zero tariffs between participants.

In the linked setting, there is a single international agreement which specifies tariffs on nonparticipants $\tau_{out}(n)$ and emissions cuts for participants $c(n)$ as a function of the number of participants n. Countries who don't join the agreement are free to choose any emissions and tariff levels in the second period. Because there's only a single agreement, the "linked setting" allows for tariffs to be used as a punishment for non-participation on climate policy, and visa versa.

The design of international agreements ($c(n)$ and $\tau(n)$) is exogenous from the perspective of the countries. The goal of this paper will be to find the design of international agreements which maximizes global welfare in equilibrium.

The full timing of the game is as follows:

1. (period 0) Country types are drawn and privately observed.
2. (period 1) Countries simultaneously announce which international agreements (if any) they will participate in. These decisions are publicly observed.
3. (period 2) Countries whose policies aren't bound by international agreements simultaneously choose tariff levels $\{\tau_{ij}\}$ and/or carbon emission cuts $\{c_i\}$. Countries who have joined international agreements set tariffs and/or carbon emissions at the levels specified in the agreements.

The solution concept we will use is sequential equilibrium.

2.3.2 Preferences and Types

There are two types. Following Nordhaus (2015), non-isolationist types have utility functions which only depend on carbon emissions and tariff levels:

$$u_i(\{c_j\}, \{\tau_{ij}\}; noniso) = (1/N)SCC(\sum_j c_j) - ac_i^2 + (\sum_j W_h(\tau_{ij}) + W_f(\tau_{ji}))$$

Where τ_{ij} is the uniform tariff level that country i places on imports from country j, the functions $W_h(\tau)$ and $W_f(\tau)$ give the welfare impact of tariffs on the home country and the target country, respectively (This is discussed in more detail later in the section), SCC is the constant global social cost of carbon, and a an abatement cost parameter.

Isolationist type countries have the same utility function except they also get an arbitrarily large utility penalty from joining any international agreement. This guarantees that they will never join any international agreement.

$$u_i(\{c_j\}, \{\tau_{ij}\}, join_i; iso) = (1/N)SCC(\sum_j c_j) - ac_i^2 + (\sum_j W_h(\tau_{ij}) + W_f(\tau_{ji})) - Ajoin_i$$

where $join_i$ is an indicator variable for whether they joined any agreement, and the constant A is taken to be arbitrary large.

The component of the utility functions that depends on carbon emissions is taken from Nordhaus' (2015) influential paper on climate clubs. It assumes a linear global damage function from carbon emissions ($SCC(\sum_j c_j)$) and a quadratic abatement cost for each country (ac_i^2). Emissions cuts are a global public good since each country's utility is increasing in other countries' emissions cuts.

Also following Nordhaus (2015), I assume quadratic forms for the welfare losses from trade:

$$W_h(\tau) = \alpha(1 - \phi)\tau - b(1 - (1/2)\phi)\tau^2$$

$$W_f(\tau) = -\alpha(1 - \phi)\tau + (1/2)b(1 - \phi)\tau^2$$

Parameter α determines the terms of trade effect from tariffs, ϕ is the incidence of tariffs on the home country, and $(1/2)b\tau^2$ gives the deadweight cost of tariffs. All trade parameters are assumed to be positive.

I define nash tariffs and emissions in the following way:

$$\tau_{nash} \equiv \frac{a(1-\phi)}{2b(1-(1/2)\phi)}$$

$$c_{nash} \equiv (1/N)SCC/(2a)$$

In a one shot game without any international agreements, all countries would set tariffs and carbon taxes to these levels.

2.3.3 The optimal international agreement

I will restrict attention to international agreements which, in equilibrium, provide incentives for all non-isolationist type countries to join. So the resulting sequential equilibrium will always have the following form:

1. In period 1, all non-isolationist countries sign on to the agreement(s). All isolationist types do not sign any agreements.

2. In period 2, all non-isolationist countries will set trade and climate policies in accordance with the international agreement(s). All isolationist countries will set tariffs and carbon emissions to nash levels[6].

The optimal international agreements are those which lead to an equilibrium that maximizes global welfare, defined simply as the sum of all countries' utilities:

$$SCC(\sum_i (c_i)) - \sum_i ac_i^2 - \sum_{i,j} (1/2)b\tau_{ij}^2$$

where I substituted in our quadratic expression for $W_i(\tau)$ and cancelled out the zero-sum terms of trade effects.

6. Notice that in this model, nash tariffs and nash carbon emissions (defined in the previous section) are strictly dominant actions for any country who is not part of an international agreement.

Finally, we can define the first-best level of emissions cuts as:

$$c_{fb} = SCC/(2a)$$

This is the global welfare maximizing level of emissions cuts for each country, which would be chosen by a global social planner.

2.4 Conditions for strict improvements

The goal of this section is to find conditions under which global welfare can be strictly increased with a linked agreement.

I will start with the benchmark perfect-information case where no countries can be isolationist types. In this case there is no reason to use issue linkage. I will then move to a setting where only one country has a positive probability of being an isolationist type; this simple setting is useful for building intuition for why issue linkage is valuable. Finally, I will cover the setting where every country has a possibility of being isolationist; this more realistic version of the model will be used in the calibration section.

2.4.1 Benchmark: No Isolationist Countries

In this section, there is no uncertainty. All countries are always non-isolationist.

Let's first consider the case where international agreements must be "unlinked". Because utility is separable in terms of carbon emissions and trade, we can first solve for the optimal climate agreement and then the optimal trade agreement. To give countries an incentive to sign on, the agreement must satisfy the following IC constraint:

$$SCCc(N) - ac(N)^2 \geq (1/N)SCC((N-1)c(N-1) + c_{nash}) - ac_{nash}^2 \tag{2.1}$$

The left side gives a country's equilibrium utility level if they join the agreement, while the right side gives their utility level if they deviate and don't join.

The optimal climate agreement maximizes global welfare:

$$\max_{\{c(n)\}} SCCNc(N) - aNc(N^2)$$

subject to the IC constraint 2.1.

The following proposition establishes that the first best is always achievable in this environment:

Proposition 1: *With no chance of isolationist types, the optimal unlinked climate agreement always achieves the first best level of emissions cuts in equilibrium.*

Proof. $c(N-1)$ does not enter the objective function (since all countries sign up in equilibrium), but it enters the IC constraint. By making $c(N-1)$ sufficiently low, the IC constraint can be satisfied for any choice of $c(N)$. So, any solution to the optimization problem must involve the $c(N)$ which maximizes aggregate welfare (the first best). □

The intuition behind this result is simple. By threatening large enough increases in emissions for non-participation, the IC constraint can always be satisfied. As long as the IC constraint is satisfied, full participation will always occur in equilibrium, so the costly punishments never have to be used. The result is that the first best level of emissions cuts is always achieved along the equilibrium path. Notice that level of emissions cuts under incomplete participation $c(N-1)$ isn't pinned down in this setting. All we know is that it must be low enough to satisfy the IC constraint. Because punishments never occur in equilibrium, there is no cost to using an excessively costly punishment. That will change once we add in a positive probability of there being isolationist countries.

Similarly, the optimal unlinked trade agreement will minimize global deadweight costs from trade $\sum_{i,j}(d_1 + d_2)\tau_{ij}^2$ subject to the following IC constraint:

$$0 \geq (N-1)(W_h(\tau_{nash}) + W_f(\tau(N-1)))$$

The left side of this expression gives a country's payoff from joining the agreement. The right side gives the countries payoff from deviating and not joining. We see that by making the retaliation tariffs $\tau(N-1)$ sufficiently large, $W_f(\tau(N-1))$ will become negative enough to satisfy this FOC. This takes us to the following proposition:

Proposition 2: *With no chance of isolationist types, the optimal unlinked trade agreement always achieves zero tariffs in equilibrium.*

Proof. Zero tariffs maximize the objective function. The IC constraint can always be satisfied by making $\tau(N-1)$ sufficiently large. □

So, with perfect information, the full first best allocation (on both trade and carbon emissions) can be achieved without issue linkage. This echoes many of the folk theorem results from the existing literature (see Maggi, 2016), which state that issue linkage is unnecessary as long as sufficiently strong punishments are available and can be credibly threatened.

2.4.2 Only one potential isolationist country

In this section, country 1 has a probability $p > 0$ of being an isolationist type. All other countries are always non-isolationist.

First, let's consider the best unlinked climate agreement. The agreement must give all non-isolationist countries an incentive to join. To give country 1 an incentive to join when they're non-isolationist, the following IC constraint must be satisfied:

$$SCCc(N) - ac(N)^2 \geq (1/N)SCC((N-1)c(N-1) + c_{nash}) - ac_{nash}^2 \tag{2.2}$$

The left side gives country 1's payoff if they choose to sign the agreement. The right side gives their payoff if they deviate and don't sign the agreement.

For all other countries, we arrive at the following IC constraint:

$$\begin{aligned}(1-p)(SCCc(N) - ac(N)^2) + p(SCC(1/N)((N-1)c(N-1) + c_{nash}) - ac(N-1)^2 \\ \geq (1-p)(SCC(1/N)((N-1)c(N-1) + c_{nash}) - ac_{nash}^2) \\ + p(SCC(1/N)((N-2)c(N-2) + 2c_{nash}) - ac_{nash}^2)\end{aligned} \tag{2.3}$$

The optimal unlinked climate agreement maximizes expected global welfare:

$$\begin{aligned}\max_{\{c(n)\}} (1-p)(SCCNc(N) - Nac(N)^2) + p(SCC((N-1)c(N-1) + c_{nash}) \\ - (N-1)ac(N-1)^2 - ac_{nash}^2\end{aligned}$$

subject to IC constraints 2.2 and 2.3.

Similarly, the optimal unlinked trade agreement minimizes expected deadweight costs:

$$\min_{\{\tau(n)\}} p(N-1)(1/2)b\tau(N-1)^2$$

subject to the following IC constraints:

$$0 \geq (N-1)(W_h(\tau_{nash}) + W_f(\tau(N-1))) \tag{2.4}$$

$$p(N-1)(W_h(\tau(N-1))+W_f(\tau_{nash})) \geq (1-p)(N-1)(W_h(\tau_{nash})+W_f(\tau(N-1)))$$
$$+\,p[(N-1)W_h(\tau_{nash}+(N-2)W_f(\tau(N-2))+W_f(\tau_{nash})] \quad (2.5)$$

The first says that country 1 has an incentive to join. The second says that all of the other countries have an incentive to join.

The following propositions establishes that, unlike the case with perfect information, the first best outcome is no longer always achieved along the equilibrium path.

Proposition 3.1: *In unlinked agreements, the first best is not always achieved along the equilibrium path.*

Proof. Let n be the number of realized non-isolationist types. Along the equilibrium path, both $n = N$ and $n = N-1$ are realized. So, for the first best to always be realized, $c(N) = c(N-1) = c_{fb}$. But if we plug those values into IC constraint 2.2, we see that it is violated. So, the first best can never be incentive compatible. Similarly, for the trade agreement to achieve the first best means $\tau(N) = \tau(N-1) = 0$, but this would violate IC constraint 2.4. □

Since the first best is no longer achieved, there may be opportunities for improvements by moving to a linked agreement.

The following propositions establish that, unlike in the perfect information case, the punishment levels used in the best unlinked agreements (i.e. $c(N-1)$ and $\tau(N-1)$) can now be pinned down:

Proposition 3.2: *In the best unlinked climate agreeemnt, $c(N) = c_{fb}$ and $c(N-1)$ is equal to the unique level that satisfies IC constraint 2.2 with equality.*

Proof. See appendix. □

Proposition 3.3: *In the best unlinked trade agreeemnt, $\tau(N-1)$ is equal to the unique level that satisfies IC constraint 2.4 with equality.*

Proof. See appendix ☐

Propositions 3.2 and 3.3 are a consequence of the fact that incomplete participation now occurs along the equilibrium path. In the previous section, punishments are only used off the equilibrium path, so there was no cost in using an excessive punishment. In this section, since punishments occur with positive probability, the optimal agreement never involves using excessive punishments. Instead, it will always involve using punishments that are just strong enough to give countries an incentive to join.

Now, let's consider the best linked agreement. This agreement will maximize total expected global welfare:

$$\max_{\{c(n)\},\{\tau(n)\}} (1-p)[SCCNc(N) - Nac(N)^2] + p[SCC(N-1)c(N-1) - (N-1)ac(N-1)^2 \\ - (N-1)((1/2)b\tau_{out}(N-1)^2)]$$

subject to the following IC constraints:

$$SCCc(N) - ac(N)^2 \geq \\ (1/N)SCC((N-1)c(N-1) + c_{nash}) - ac_{nash}^2 + (N-1)(W_h(\tau_{nash}) + W_f(\tau(N-1))) \tag{2.6}$$

$$(1-p)[SCCc(N) - ac(N)^2]$$

$$+ p[SCC(1/N)((N-1)c(N-1) + c_{nash}) - ac(N-1)^2 + W_h(\tau(N-1)) + W_f(\tau_{nash})]$$

$$\geq (1-p)[SCC(1/N)((N-1)c(N-1)+c_{nash}) - ac_{nash}^2 + (N-1)(W_h(\tau_{nash}) + W_f(\tau(N-1)))]$$

$$+ p[SCC(1/N)((N-2)c(N-2) + 2c_{nash}) - ac_{nash}^2 + (N-1)W_h(\tau_{nash})$$

$$+ (N-2)W_f(\tau(N-2)) + W_f(\tau_{nash)}] \quad (2.7)$$

Where the first IC constraint says that country 1 has an incentive to join the agreement and the second IC constraint says that all other countries have an incentive to join the agreement.

The following proposition establishes a relationship between tariffs and emissions cuts in the event of incomplete participation in the optimal linked agreement:

Proposition 4: *As long as the optimal linked agreement is interior*[7] *($tau^*(N-1) < \bar{\tau}$ and $c^*(N-1) > 0$), the following relationship holds:*

$$\frac{SCC/N}{p(SCC - 2ac^*(N-1))} = \frac{-W_f(\tau^*(N-1))}{pb\tau^*(N-1)}$$

Proposition 4 can be interpreted as saying that the marginal efficiency of each punishment must be equal in the optimal agreement. The numerator on the left side says how much increasing carbon emissions in the event of incomplete participation harms the non-

7. If there's a corner solution where $\tau^*(N-1) = t\bar{a}u$, then the condition becomes:

$$\frac{SCC/N}{p(SCC - 2ac^*(N-1))} \leq \frac{-W_f(\tau^*(N-1))}{pb\tau^*(N-1)}$$

If there's a corner solution where $c^*(N-1) = 0$, then the condition becomes:

$$\frac{SCC/N}{p(SCC - 2ac^*(N-1))} \geq \frac{-W_f(\tau^*(N-1))}{pb\tau^*(N-1)}$$

participating country. The denominator on the left side says how much those increased carbon emissions harm global welfare. So the full ratio on left side is a measure of how efficient of a punishment carbon emissions increases are. Similarly, the numerator on the right side is how much marginal increases in tariffs on nonparticipants increases the net cost of non-participation, while the denominator is the global welfare cost of increasing tariffs. So, the ratio on the right hand side is a measure of how efficient of a punishment tariffs are. Proposition 4 thus says that, for the optimal agreement, both types of punishments are equally efficient at the margin.

Theorem 1 below gives conditions for when linked agreements will be strictly welfare improving. It says that there are strict improvements except in the knife edge case where the two types of punishments (carbon emissions cuts and tariffs) happen to be equally efficient in the optimal unlinked agreements.

Definition 1: *Let* $c^*_{ul}(N-1)$ *and* $\tau_u l^*(N-1)$ *be the unique values for* $c(N-1)$ *and* $\tau(N-1)$, *respectively, in the best unlinked climate and trade agreements.*

Theorem 1: Linked agreements are strictly welfare improving iff:

$$\frac{SCC/N}{p(SCC - 2ac^*_{ul}(N-1))} \neq \frac{-W_f(\tau^*_{ul}(N-1))}{pb\tau^*_{ul}(N-1)}$$

As long as the two ratios are not already equal in the unlinked agreements, then welfare can be improved by using a linked agreement to substitute the marginally less efficient punishment for the more efficient one. In our calibration exercise, we end up finding that carbon emissions cuts are far less efficient in the unlinked setting. This means that there are benefits from increasing tariffs on non-participants while increasing emissions cuts in the event of incomplete participation. In other words, there are benefits to relying more on tariffs as a way to incentivize participation rather than cutting contributions to the global public good.

2.4.3 Full Model

Now all countries could possibly be isolationist types. Every country has a probability p of being isolationist, with all types being independent. Let $P(n) = binomial(n; 1-p, N)$ be the probability that there are n non-isolationist types.

The best unlinked climate agreement will maximize global welfare:

$$\max_{\{c(n)\}} E\left[SCCnc(n) - nac(n)^2\right]$$

subject to the IC constraint:

$$E\left[(SCC/N)(nc(n) + (N-n)c_{nash}) - ac(n)^2\right] \geq$$
$$E\left[(SCC/N)((n-1)c(n-1) + (N-n+1)c_{nash}) - ac_{nash}^2\right] \quad (2.8)$$

The left side of the IC constraint gives a country's expected payoff from joining the agreement. The right side gives their expected payoff from deviating and not joining the agreement.

The best unlinked trade agreement will minimize deadweight costs:

$$\min_{\{\tau(n)\}} E\left[(1/2)(N-n)b\tau(n)^2\right]$$

subject to the IC constraint:

$$E\left[(N-n)W_h(\tau(n)) + (N-n)W_f(t_{nash})\right] \geq$$
$$E\left[(n-1)W_f(\tau(n-1)) + (N-1)W_h(\tau_{nash}) + (N-n)W_f(\tau_{nash})\right] \quad (2.9)$$

The best linked agreement maximizes total global welfare:

$$\max_{\{c(n)\},\{\tau(n)\}} E\left[SCCnc(n) - nac(n)^2 - (1/2)(N-n)b\tau(n)^2\right]$$

subject to the IC constraint:

$$\begin{aligned} E\Big[(SCC/N)(nc(n) + (N-n)c_{nash}) - ac(n)^2 + (N-n)W_h(\tau(n)) \\ +(N-n)W_f(t_{nash})\Big] \geq \\ E\Big[(SCC/N)((n-1)c(n-1) + (N-n+1)c_{nash}) - ac_{nash}^2 \\ +(n-1)W_f(\tau(n-1)) + (N-1)W_h(\tau_{nash}) + (N-n)W_f(\tau_{nash})\Big] \end{aligned} \quad (2.10)$$

The theorem below establishes when linked agreements yield welfare gains.

Theorem 2: *Let $\{c_{ul}(n)^*\}$ and $\{\tau_{ul}(n)^*\}$ be emissions cuts and tariff levels specified in the best unlinked climate and trade agreements, respectively. The best linked agreement will yield higher welfare than the best unlinked agreements if for any n, n' with $c(n) > 0$ and $\tau(n') < \bar{\tau}$, the following relationship holds:*

$$\frac{P(n)((SCC/N)n - 2ac_{ul}^*(n)) - P(n+1)(SCC/N)n}{P(n)(nSCC - 2nac_{ul}^*(n))} = \frac{P(n')(N-n')W_h'(\tau_{ul}^*(n')) - P(n'+1)W_f'(\tau_{ul}^*(n'))n'}{P(n')(N-n')b\tau_{ul}^*(n')}$$

Proof. See appendix □

This is very similar to Theorem 1[8]. The ratios in Theorem 2 are measures of the marginal

8. The numerator on the left side of the condition in theorem 2 looks slightly different from the numerator in theorem 1. In both cases, this numerator represents the net impact of decreasing $c(n)$ (that is, increasing carbon emissions for participating countries in the event that n countries participate) on countries' incentive to participate. In the case where only one country could be isolationist (theorem 1), the only relevant incentive problem was for that one isolationist country (IC constraint 2.2), and reducing $c(N-1)$ only impacted the right side of that IC constraint (it harms country 1 in the event that they deviated). However, in the case where any country can be isolationist, lowering $c(N-1)$ impacts both sides of the IC constraint:

efficiency of the punishments. If these ratios aren't equal to each other in the unlinked agreements, there will be welfare gains from using a linked agreement to substitute toward the more efficient punishment.

2.5 Calibration

This section examines whether the gains from issue linkage are likely to be quantitatively large.

2.5.1 Parameter Estimates

In the baseline case, we set $N = 195$, the number of countries that are recognized by the United Nations in 2024.

I take the social cost of carbon estimate from Greenstone (2020): $SCC = \$190$/ton CO2. I take the global abatement cost parameter a_g from Nordhaus (2015). Country-level abatement costs must add up to global abatement costs according to the following formula:

$$a_g(Nc)^2 = Nac^2$$

So, the country level abatement cost parameter is given by:

$$a = a_g N$$

For trade parameters, I take the elasticity estimates from Soderbury (2015) and use linear demand and supply approximations to find b, ϕ and α. The elasticity estimates are for the

it reduces utility if a country deviates and all other $N-1$ countries signed up, but it increases utility if that country signs up and one other country deviates. Since it impacts both the left and the right side of the IC constraint, the net impact on incentives is a bit more complicated, which is why the condition in theorem 2 looks slightly different than in theorem 1. This is the same reason why the numerators on the right sides of the two theorems are slightly different.

United States, so I'm assuming that tariffs placed on other countries have the same welfare impacts as tariffs placed on the US would. Obviously, this is not an accurate assumption. However, since most countries are smaller than the US, using more accurate parameters should make trade sanctions even more efficient punishments (because more of the cost of trade sanctions is internalized by small countries), so using more accurate parameters would likely strengthen my conclusions.

Specifically, Soderbury (2015) estimates that the elasticity of US export supply ρ is about .4 and the elasticity global demand for US goods σ is about 5. Assuming that these demand and supply curves are linear, the trade parameters are given by:

$$\phi = \frac{1/\rho}{(1/\rho) + (1/\sigma)}$$

$$\alpha = Q_0$$

$$b = \frac{Q_0}{(1/\rho) + (1/\sigma)}$$

Where Q_0 is the target country's amount of exports in the absence of tariffs. I set Q_0 to be \$19trillion/N, where \$19 trillion is the total value of global exports in 2019[9].

These parameters give very similar results to Nordhaus (2015) on the welfare costs of the rest of the world raising tariffs on the US (both on the costs to the US and the costs to the rest of the world). Norhaus' calibration, in turn, was done to match results from Ossa (2012)'s full general equilibrium model.

2.5.2 Numerical Strategy

The first step is to transform the problem of choosing the best linked agreement into a convex optimization problem. Recall that the problem of choosing the best linked agreement is to

9. https://www.statista.com/statistics/264682/worldwide-export-volume-in-the-trade-since-1950/

choose $c(n)$ and $\tau(n)$ for all n to maximize:

$$\max_{\{c(n)\},\{\tau(n)\}} E\left[SCCnc(n) - nac(n)^2 - (1/2)(N-n)b\tau(n)^2\right]$$

subject to the IC constraint:

$$E\Big[(SCC/N)(nc(n) + (N-n)c_{nash}) - ac(n)^2 + (N-n)W_h(\tau(n)) + (N-n)W_f(t_{nash})\Big]$$
$$\geq E\Big[(SCC/N)((n-1)c(n-1) + (N-n+1)c_{nash}) - ac_{nash}^2$$
$$+(n-1)W_f(\tau(n-1)) + (N-1)W_h(\tau_{nash}) + (N-n)W_f(\tau_{nash})\Big]$$

This problem is not convex. However, it can be turned into a convex problem by adding choice variables $x(n)$ and $y(n)$ which are meant to capture the square of $c(n)$ and $\tau(n)$, respectively. This also requires adding the constraints $c(n)^2 \leq x(n)$ and $\tau(n)^2 \leq y(n)$. This gives us the new, convex problem:

$$\max_{\{c(n)\},\{\tau(n)\},\{x(n)\},\{y(n)\}} E\left[SCCnc(n) - nax(n) - (1/2)(N-n)by(n)\right]$$

subject to:

$$E\Big[(SCC/N)(nc(n) + (N-n)c_{nash}) - ax(n)^2$$
$$+(N-n)\tilde{W}_h(\tau(n), y(n)) + (N-n)W_f(t_{nash})\Big] \geq$$
$$E\Big[(SCC/N)((n-1)c(n-1) + (N-n+1)c_{nash}) - ac_{nash}^2$$
$$+(n-1)\tilde{W}_f(\tau(n-1), y(n)) + (N-1)W_h(\tau_{nash}) + (N-n)W_f(\tau_{nash})\Big]$$

and, for all n:

$$c(n)^2 \leq x(n)$$

$$\tau(n)^2 \leq y(n)$$

where the new reduced form tariff welfare functions are:

$$\tilde{W}_h(\tau, y) = \alpha(1-\phi)\tau - b(1-(1/2)\phi)y^2$$

$$\tilde{W}_f(\tau, y) = -\alpha(1-\phi)\tau + (1/2)b(1-\phi)y^2$$

The following proposition establishes that the transformed problem gives the same solutions as the linked optimization problem.

Proposition 5: *A vector ($\{c(n)\}, \{\tau(n)\}, \{x(n)\}, \{y(n)\}$) solves the transformed problem if and only if the subvector ($\{c(n)\}, \{\tau(n)\}$) solves the linked optimization problem.*

Proof. First we'll show that the constraints $c(n)^2 \leq x(n)$ and $\tau(n)^2 \leq y(n)$ have to bind with equality. Assume for contradiction that a solution to the transformed problem has $c(n')^2 < x(n')$ for some n'. Then, by lowering $x(n')$ and holding all other variables constant, the objective function could be increased while still satisfying all constraints. So this can't be a solution, which means $c(n)^2 = x(n)$ for all n in any solution.

Similarly, assume for contradiction that a solution to the transformed problem has $\tau(n')^2 < y(n')$. Then, by lowering $y(n')$ while holding all other variables constant, the objective function can be increased while satisfying all constraints. So we have a contradiction, and can conclude that for all n, $\tau(n)^2 = y(n)$ at the optimum.

Now, using the fact that $c(n)^2 = x(n)$ and $\tau(n)^2 = y(n)$ in any solution, we can eliminate the choice variables $x(n)$ and $y(n)$ from the transformed problem, which makes the transformed problem identical to the linked optimization problem. □

Since the transformed problem is convex, it can be easily solved using standard programming packages. I use the python package CVXPY. The same steps can be followed to solve for the unlinked climate and trade agreements.

2.5.3 Calibration Results

Table 2.1 shows calibration results for the baseline case. We see that even if the probability of country being isolationist is relatively small (.01), there are still significant welfare losses from using a linked agreement rather than unlinked agreements.

p	Gains from Issue Linkage
.001	0.20
.01	1.19
.1	2.45

Table 2.1: Calibration results with 195 identical countries.

Notes: Welfare gains are in units of trillions \$ per year.

To get a sense of where the welfare losses are coming from, figure 2.1 shows the level of emissions cuts for participants as a function of the realized number of low types when $p = .01$. We see that emissions cuts are much larger in the linked agreement than the unlinked one as long as there is at least one low type. Now, the linked agreement also has higher tariffs than the unlinked case[10]. However, these higher tariffs cause welfare losses that are an order of magnitude smaller than the gains from having higher emissions cuts, as is shown in figure 2.2.

2.6 Conclusion

This paper makes a novel contribution to the debate regarding issue linkage in international climate agreements. Specifically, it shows that using trade sanctions as a way to incentivize participation in a climate agreement can be welfare improving in contexts where there's a chance that some countries won't join the agreement.

10. Tariffs of around 12% are placed on each nonparticipant in the linked case. In the unlinked case, tariffs on nonparticipants are about 10.5%

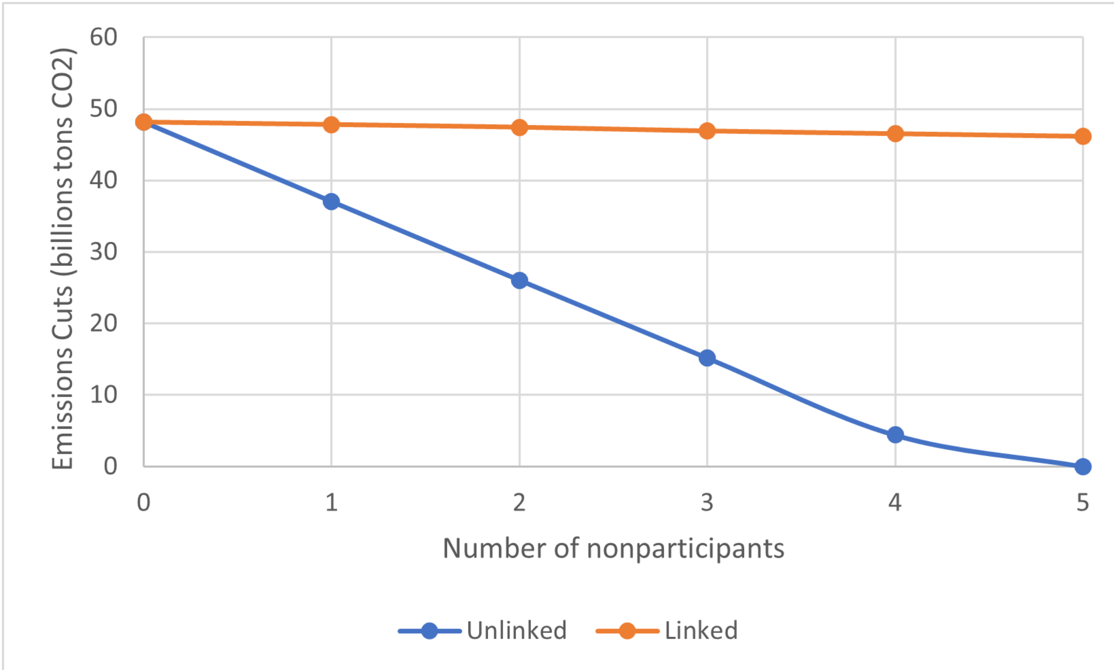

Figure 2.1: Aggregate emissions cuts for participants as a function of the realized number of isolationist types.

The key idea is that without issue linkage, the only way to incentivize participation in a climate agreement is to have participants increase emissions in the event that other countries don't sign up. Since any individual country only internalizes a small percentage of global climate damages, this is a very inefficient punishment from a global perspective. With issue linkage, trade sanctions can be used instead of emissions increases as a way to incentivize participation. Trade sanctions are much more efficient punishments, in the sense that a large fraction of the global welfare losses from trade sanctions will be experienced by the nonparticipant countries that are the target of the sanctions. In this way, global welfare losses will be much lower from using trade sanctions to incentivize participation instead of emissions increases.

This argument is made using a model and calibration that is very similar to Nordhaus (2015)'s well-cited paper on issue linkage. While the model has a number of strong assumptions, it provides at least suggestive evidence that the novel mechanism identified in

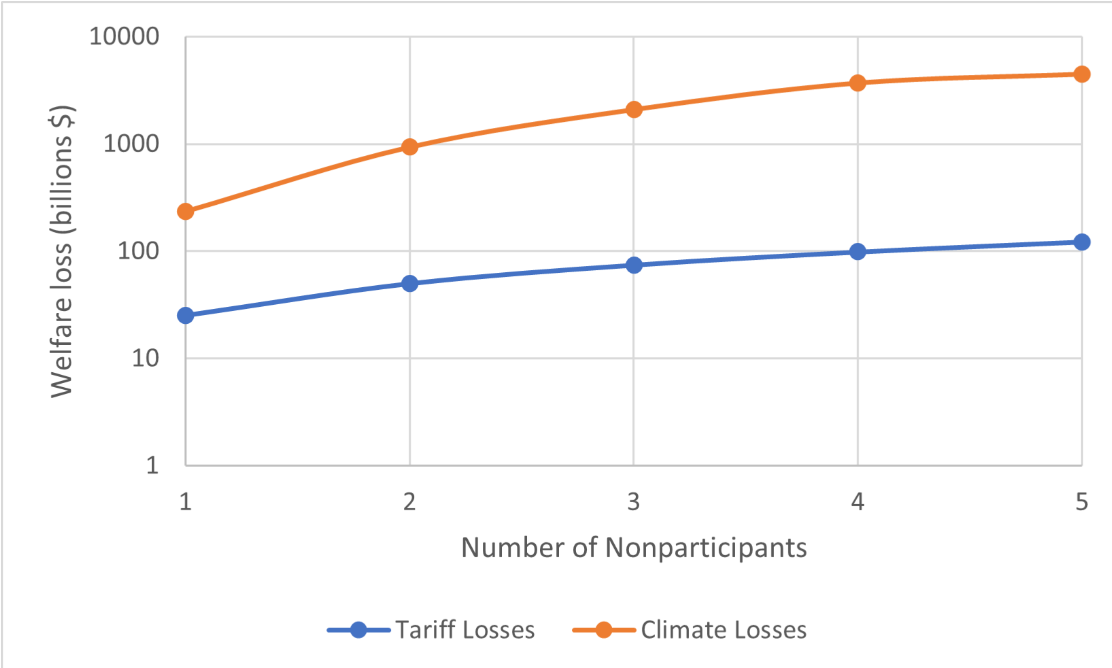

Figure 2.2: Welfare Losses from Increased Carbon Emissions vs Increased Tariffs

Notes: The welfare losses are conditional on the realization of the number of isolationist types. Note that the y-axis is on a log scale.

this paper can lead to large welfare gains from issue linkage. A promising opportunity for future work is to relax many of these assumptions to test the robustness of this conclusion. For example, a more standard general equilibrium climate model which includes country heterogeneity could be used.

REFERENCES

Jan Abrell, Mirjam Kosch, and Sebastian Rausch. The economic cost of carbon abatement with renewable energy policies. *SSRN Electronic Journal*, 2018. doi:10.2139/ssrn.2987006.

Daron Acemoglu and James A Robinson. Economics versus politics: Pitfalls of policy advice. *Journal of Economic perspectives*, 27(2):173–192, 2013.

Claustre Bajona and Josh Ederington. Domestic policies, hidden protection and the gatt/wto. *Hidden Protection and the GATT/WTO (September 3, 2012)*, 2012.

Scott Barrett. The strategy of trade sanctions in international environmental agreements. *Resource and Energy Economics*, 19, 1997. ISSN 09287655. doi:10.1016/s0928-7655(97)00016-x.

Severin Borenstein and Ryan Kellogg. Carbon pricing, clean electricity standards, and clean electricity subsidies on the path to zero emissions. *Environmental and Energy Policy and the Economy*, 4, 2023. ISSN 2689-7857. doi:10.1086/722675.

Richard Chisik et al. *Limited Incremental Linking and Unlinked Trade Agreements*. Citeseer, 2010.

Tatyana Deryugina, Alexander MacKay, and Julian Reif. The long-run dynamics of electricity demand: Evidence from municipal aggregation. *American Economic Journal: Applied Economics*, 12, 2020. ISSN 19457790. doi:10.1257/app.20180256.

Josh Ederington. Policy linkage and uncertainty in international agreements. *Economic Inquiry*, 41(2):305–317, 2003.

Emmanuel Farhi and Ivan Werning. The political economy of nonlinear capital taxation. Technical report, mimeo, 2008.

James D. Fearon. Domestic political audiences and the escalation of international disputes. *American Political Science Review*, 88, 1994. ISSN 0003-0554. doi:10.2307/2944796.

Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54, 1986. ISSN 00129682. doi:10.2307/1911307.

Kenneth Gillingham and James H. Stock. The cost of reducing greenhouse gas emissions. volume 32, 2018. doi:10.1257/jep.32.4.53.

M Greenstone and I Nath. Put a price on it: The how and why of pricing carbon. *U. S Energy & Climate Road Map: Evidence-Based Policies for Effective Action*, 2021.

Gal Hochman and David Zilberman. Optimal environmental taxation in response to an environmentally-unfriendly political challenger. *Journal of Environmental Economics and Management*, 106, 2021. ISSN 10960449. doi:10.1016/j.jeem.2020.102407.

Dale W Jorgenson, Richard J Goettle, Mun S Ho, and Peter J Wilcoxen. Double dividend: Environmental taxes and fiscal reform in the united states. *National Tax Journal*, 68, 2013.

G. Maggi. Issue linkage. 2016. doi:10.1016/bs.hescop.2016.04.017.

Stephen Nalley and Angelina LaRose. Annual energy outlook 2022 (aeo2022). *Energy Information Agency*, 2022, 2022.

William Nordhaus. Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105, 2015. ISSN 00028282. doi:10.1257/aer.15000001.

Torsten Persson and Lars E.O. Svensson. Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *Quarterly Journal of Economics*, 104, 1989. ISSN 15314650. doi:10.2307/2937850.

Torsten Persson and Guido Tabellini. *Political Economics: Explaining Economic Policy*. MIT Press, 2002.

Joshua D. Rhodes, Carey King, Gürcan Gulen, Sheila M. Olmstead, James S. Dyer, Robert E. Hebner, Fred C. Beach, Thomas F. Edgar, and Michael E. Webber. A geographically resolved method to estimate levelized power plant costs with environmental externalities. *Energy Policy*, 102, 2017. ISSN 03014215. doi:10.1016/j.enpol.2016.12.025.

Dani Rodrik. The limits of trade policy reform in developing countries. *Journal of Economic Perspectives*, 6, 1992. ISSN 0895-3309. doi:10.1257/jep.6.1.87.

Dani Rodrik. Green industrial policy. *Oxford Review of Economic Policy*, 30, 2014. ISSN 14602121. doi:10.1093/oxrep/gru025.

Thomas Schelling. *The Strategy of Conflict*. Harvard University Press, 1960.

Alex Schmitt. *Beyond Pigou: climate change mitigation, policy making and distortions*. Department of Economics, Stockholm University, 2014.

Alistair Ulph and David Ulph. Optimal climate change policies when governments cannot commit. *Environmental and Resource Economics*, 56, 2013. ISSN 09246460. doi:10.1007/s10640-013-9682-7.

Asa Watten. *Risk, Uncertainty, and Heterogeneity: Three and a Half Essays in Energy and Environmental Economics*. Michigan State University, 2021.

APPENDIX A

APPENDIX FOR CHAPTER 1

A.1 Further Details

A.1.1 General Equilibrium Microfoundation

Consider a general equilibrium economy with two consumption goods: energy E_t and general consumption y_t. Consumer preferences are quasilinear in y_t:

$$U(\{y(h_t)\}, \{E(h_t)\}) = \sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(y(h_t) + v(E(h_t)))$$

The technology for energy production is the same as in the body of the text. In each period t, there's a total endowment of labor equal to L, which can be used for consumption, brown energy production, or investment according to the following constraint: $y_t + mcE_{bt} + x_t = L$.

The consumer problem is to choose an allocation $\{y(h_t)\}, \{E(h_t)\}$ to maximize utility:

$$\sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(y(h_t) + v(E(h_t)))$$

subject to the budget constraint:

$$\sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)p_y(h_t)(y(h_t) + p(h_t)E(h_t) - T(h_t)) \leq A$$

where $T(h_t)$ are transfers from the government and A is the expected present value of firm profits.

Due to the quasilinear structure of preferences, all prices $p_y(h_t)$ must be constant in any competitive equilibrium[1,2], so we can normalize them to 1 and simplify the constraint to:

$$\sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(y(h_t) + p(h_t)E(h_t) - T(h_t)) \leq A$$

The constraint must hold with equality since utility is strictly increasing in y. We can now rewrite the constraint as:

$$\sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)y(h_t) = \sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(p(h_t)E(h_t) - T(h_t))$$

By plugging this constraint into the objective function, the consumer problem simplifies to:

$$max_{\{E(h_t)\}} \sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(v(E(h_t) - p(h_t)E(h_t))$$

which is identical to the consumer problem in the partial equilibrium model.

The firm problem is to choose an allocation ($\{y(h_t)\}$, $\{E_b(h_t)\}$, $\{K(h_t)\}$, $\{x(h_t)\}$) to maximize expected profits:

$$\sum_t \sum_{h_t} \beta^{t-1}\Pi(h_t)(y(h_t) + p(h_t)(F(K_t) + E_b(h_t)) - \gamma\tau(h_t)E_b(h_t) + s(h_t)x(h_t))$$

---

1. A competitive equilibrium in this GE economy is a set of quantities ($\{y(h_t)\}$, $\{E_b(h_t)\}$, $\{K(h_t)\}$, $\{x(h_t)\}$, $\{E(h_t)\}$), prices $\{p(h_t)\}$, and tax policies ($\{\tau(h_t)\}$, $\{s(h_t)\}$, $\{T(h_t)\}$) which satisfies:

   1. The firm problem
   2. The consumer problem
   3. The market clearing constraint for all h_t: $E_b(h_t) + F(K(h_t)) = E(h_t)$
   4. The government budget constraint for all h_t: $T(h_t) = \tau(h_t)E_b(h_t) - s(h_t)x(h_t)$

2. Proof: The FOC for $y(h_t)$ is: $1 - \lambda p_y(h_t) = 0$, where λ is the lagrange multiplier on the budget constraint. Since we have no nonnegativity constraint here, this FOC must hold with equality for all h_t.

Subject to the resource constraint, law of motion, and irreversibility constraints:

$$y(h_t) + mcE_b(h_t) + x(h_t) = L$$

$$K(h_t) = K(h_{t-1}) + x(h_t)$$

$$x_t \geq 0$$

Using the first constraint to eliminate y_t, and the firm problem simplifies to choosing an allocation $(\{E_b(h_t)\}, \{K(h_t)\}, \{x(h_t)\})$ to maximize:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(p(h_t)(F(K_t) + E_b(h_t)) - (mc + \gamma\tau(h_t))E_b(h_t) - (1 - s(h_t))x(h_t))$$

subject to the law of motion and irreversibility constraints. This is identicial to the firm problem in the partial equilibrium case.

Finally, the social welfare for party j is:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(y(h_t) + v(E(h_t)) - \gamma d_j E_b(h_t))$$

Using the resource constraint to eliminate $y(h_t)$, this becomes:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(v(E_t) - mcE_b(h_t) - x(h_t) - \gamma d_j E_b(h_t))$$

which, again, is equivalent to party j's social welfare function in the partial equilibrium version of the model.

A.1.2 Microfoundation with Brown Capital

In this section we show that the assumption that brown energy production only requires non-durable inputs can be microfounded in a model where the initial stock of brown energy capacity is sufficiently large.

Consider a model where brown energy requires both nondurable inputs L_f and brown capital K_f[3]. The production function is leontief: $E_{bt} = \min\{K_{ft}/k, L_{ft}/mc\}$, where k and mc are the capital and non-durable input requirements to produce a unit of brown energy, respectively. Aside from that change, the setup of the model is the same as the baseline case.

Party j's planner problem in the no-turnover case is now to choose an allocation ($\{x_t\}$, $\{K_t\}$, $\{E_{bt}\}$, $\{x_{ft}\}$, $\{K_{ft}\}$ to maximize social welfare:

$$\sum_t \beta^{t-1}(v(F(K_t) + E_{bt}) - (mc + \gamma d_j)E_{bt} - x_t - x_{ft})$$

subject to the laws of motion, irreversibility constraints, and the capacity constraint for brown energy:

$$K_t = K_{t-1} + x_t$$

$$K_{f,t} = K_{f,t-1} + x_{f,t}$$

$$x_t, x_{ft} \geq 0$$

$$E_{bt} \leq K_{bt}/k$$

Intuitively, if the initial stock of brown capital is sufficiently large, then the capacity constraint on brown capital won't bind. In that case, there will be no reason to invest any more in brown capital, and the solution will coincide with the case where brown energy only requires nondurable inputs (that is, it will coincide with the baseline model). Below we show

3. f is the subscript here since b is already used to denote the brown party

this formally.

Let $E_b^* \equiv D(mc+\gamma d_b) - F(K_b^*)$ be the brown party's no-turnover level of brown production in the baseline model, as is defined in theorem 1. The following proposition establishes that if the initial stock of brown capital is greater than K_b^*, then each party's no-turnover solution in the model with brown capital is identical to that in the baseline model:

Proposition A.3: *If $K_{f0}/k > E_b^*$, then party j's no turnover solution is to set $x_{ft} = 0$ and $K_{ft} = K_{f0}$ for all t, while setting all other quantities to the same levels as in theorem 1.*

Proof. The Bellman equation for this problem is:

$$V(K, K_f) = max_{K', K_f', E_b} v(F(K') + E_b) - (mc + \gamma d_j)E_b - (K' - K) - (K_f' - K_f) + \beta V(K', K_f')$$

Subject to the constraints $K' \geq K$, $K_f' \geq K_f$, $E_b \geq 0$, and $E_b \leq K_f'/k$. Guess that the policy function is to set $K_f'(K, K_f) = K_{f0}$, $K'(K, K_f) = max\{K_j^*, K\}$, and $E_b(K, K_f) = max\{0, D(mc+\gamma d_j) - F(K')\}$. This is consistent with the behavior in the proposition. With this guess the value function doesn't depend on K_f and is identical to the value function in the proof of theorem 1. Plugging this value function into the Bellman equation, we see that all FOCs are uniquely satisfied at the policy function that we guessed. □

Following similar steps, we can show that the equilibrium allocations in the single-election model and the multi-election model will also be the same as in the baseline model.

So, as long as there is enough initial capital to meet the brown party's demand for brown energy (without any additional investment), then the behavior in the model with brown capital is identical to the behavior in the baseline model.

Empirically, in 2022 there was enough installed fossil fuel capacity in the US to meet at least 115% of US electricity demand, assuming that the plants ran at full capacity (Nalley and LaRose (2022)). In practice, fossil fuel only provided 60% of US generation because much of the capacity went unused. This suggests that it is reasonable to assume that there is initially sufficient brown capital available in the economy to fully meet brown energy demand without any new investment.

Notice, however, that this result also hinges on our assumption that capital does not depreciate over time. While we believe this is a reasonable approximation to make as a first step (as was noted above, the lifespan of fossil fuel plants is very long, between 30 and 50 years), we think that relaxing this assumption presents a useful opportunity for future work.

A.1.3 Single-election model with upward sloping marginal costs

In this section we relax the assumption that brown energy has constant marginal costs and show that this leads to the same sufficient statistic for the optimal subsidy.

The cost function for producing brown energy is $c(E_b)$, which is increasing and convex. The consumer problem is unchanged. The firm problem is now to choose an allocation $(\{x(h_t)\},\{K(h_t)\}, \{E_b(h_t)\})$ to maximize expected profits:

$$\sum_{t=1}^{\infty} \sum_{h_t \in H_t} \Pi(h_t)\beta^{t-1}(\underbrace{(F(K(h_t)) + E_b(h_t))p(h_t)}_{\text{Revenue}} - \underbrace{(1 - s(h_t))x(h_t) - c(E_b(h_t)) - \tau(h_t)\gamma E_b(h_t)}_{\text{Cost}})$$

Subject to the law of motion and irreversible investment constraints for all h_t:

$$K(h_t) = K(h_{t-1}) + x(h_t)$$

$$x(h_t) \geq 0$$

where the exogenous initial capital stock K_0 is zero.

The definition of competitive equilibrium is unchanged.

Party j's expected utility is now:

$$\sum_{t=1}^{\infty} \sum_{h_t \in H_t} \Pi(h_t)\beta^{t-1}(\underbrace{v(F(K(h_t)) + E_b(h_t))}_{\text{Consumption Value}} - \underbrace{x(h_t) - c(E_b(h_t))}_{\text{Cost}} - \underbrace{d_j \gamma E_b(h_t)}_{\text{Externality}})$$

Just as we did in the baseline model, here we make assumptions on parameter values which guarantee that positive amounts of both types of energy are used in each party's no-turnover solution. First, define p_{0b} as the equilibrium price if the brown party had permanent control and green energy wasn't available and p_{0b} as the equilibrium price if the green party had full control and brown energy wasn't available[4]. The interior assumptions are then:

1. $F'(0)p_{b0} > 1 - \beta$

2. $p_{g0} > c'(0) + \gamma d_g$

As in the baseline model, let the function $C_b(K_1)$ give C_b^{SS} as a function of the first period capital stock, where C_b^{SS} is the equilibrium level carbon emissions for $t \geq 2$ in the event that the brown party wins the second period election.

The following result establishes that we get the same sufficient statistic formula as in the baseline model:

Theorem A.4 Take any subgame perfect equilibrium allocation $\mathcal{A}$. Any tax policy which implements the allocation $\mathcal{A}$ has the following feature: in the first period, the green party uses a subsidy equal to $\theta \frac{\beta}{1-\beta}(d_g - d_b)\frac{dC_b(K_1)}{dx_1}$.

4. Formally, first let $\tilde{E}_b$ be the unique solution to $v'(\tilde{E}_b) = c'(\tilde{E}_b) + \gamma d_b$. $p_{0b} \equiv v'(\tilde{E}_b)$. Similarly, let $\tilde{K}$ be the unique solution to $v'(F(\tilde{K}))F'(\tilde{K}) = 1 - \beta$. $p_{g0} \equiv v'(F(\tilde{K}_g))$.

A.1.4 Sufficient Statistic Estimation Details

Details on Abrell et al (2018)

To estimate how much increased green energy production reduces carbon emissions ($\frac{dC}{dE_g}$), we use estimates from Abrell et al (2018). This paper uses hourly variation in weather as a source of exogenous variation in green energy production. The data is from Germany and Spain from 2014-2015. There are a few caveats with applying these estimates to our setting:

- There are external validity concerns from using European results from eight years ago to the present US context.
- Because this paper uses hourly variation, it is a short run analysis. Long run crowd out effects could be different since long run demand and supply elasticities could be different.

I would like to note, however, that estimates from have been used in a number of well-cited policy papers as a measure of the impact of clean energy subsidies on emissions (see Gillingham (2018) and Greenstone and Nath (2021)). So, it seems that this is the current best available quasi-experimental evidence on the crowd out effect. As is mentioned in the main text, further empirical research could improve on these estimates, which are important not just for the question of the optimal subsidy under government turnover, but also the more basic question of the impact of clean energy subsidies on carbon emissions.

The lower panel in table 1A shows the relevant estimates from Abrell et al (2017). They give 12 different estimates of the crowd out effect, which vary in:

- The type of clean energy (wind vs solar)
- The country analyzed (Germany or Spain)
- Assumptions about how much reducing imports reduces emissions.

This third bullet point is the largest driver of heterogeneity in their estimates. Because they are analyzing small countries (relative to the size of the US), increased green energy production significantly reduces electricity imports. Abrell et al (2017) don't directly have data on how reducing these imports reduces emissions, so they consider three possible scenarios:

1. Reducing imports causes no reduction in emissions
2. Each MWh of reduced imports decreases global coal production by 1 MWh, which reduces emissions by $\approx$ 1 ton of CO2.
3. Each MWh of reduced imports decreases global natural gas production by 1 MWh, which reduces emissions by $\approx$ .4 ton of CO2.

As can be seen from figure 3, all of their estimates are within the range of .18 to .58 tons of CO2 per MWh. The average of their 12 estimates is .3 tons of CO2 per MWh.

Details on Borenstein and Kellogg (2023)

To estimate the marginal levelized cost of green energy, we use estimates from Borenstein and Kellogg (2023). This paper has detailed cost data on fossil fuel plants. They estimate the marginal cost of fossil fuel plants in 2019 to be $64 per MWh, and infer that the equilibrium marginal cost of green energy must also be equal to $64 per MWh.

Another source, Nalley and LaRose (2022), finds the average cost of new wind and solar energy to be significantly lower, around $35/MWh. This difference could just be due to the difference between average and marginal costs, or due to the fact that Borenstein and Kellogg (2023) try to account for the full marginal levelized cost, including grid integration costs. If we use the Nalley and Larose (2022) estimate instead, then our resulting estimate of the optimal subsidy would be larger.

TABLE 5. Average marginal CO_2 emissions offset from RE e by technology by market (kg CO_2/MWh of intermittent RE e)

	Market r and RE type e			
	Germany		Spain	
	Wind	Solar	Wind	Solar
Offset by technology ($= \varphi_{ir} \Delta X^e_{ir}$)				
Coal	-109.2	-147.6	-150.2	-85.4
	(6.0)	(9.0)	(9.1)	(9.1)
Gas	-21.3	-29.2	-93.9	-91.0
	(1.9)	(2.8)	(3.6)	(7.3)
Lignite	-46.7	-48.8	–	–
	(4.0)	(6.1)	–	–
Net imports				
"Domestic offsets only"	0	0	0	0
"Exports replace coal"	-258.0	-360.1	-69.5	-168.1
	(13.6)	(23.0)	(10.4)	(19.5)
"Exports replace natural gas"	-108.6	-134.9	-27.6	-66.7
	(5.7)	(8.6)	(4.1)	(7.8)
Total annual carbon offset ($= \Delta E^e_r$)				
"Domestic offsets only"	-177.2	-225.6	-244.1	-176.5
	(7.5)	(11.2)	(12.8)	(11.7)
"Exports replace coal"	-388.5	-585.6	-313.6	-344.6
	(15.5)	(25.6)	(14.3)	(22.8)
"Exports replace natural gas"	-239.1	-360.5	-271.7	-243.2
	(9.4)	(14.2)	(10.7)	(14.0)

Notes: Numbers in parentheses refer to robust standard errors.

Figure A.1: Crowd out estimates from Abrell et al (2017)

Notes: The lower panel (starting with the heading "Total Annual Carbon Offset") gives the relevant 12 estimates of the crowd out effect. The units are kg of CO2 per MWh. Divide by 1000 to get this in units of metric tons of CO2 per MWh.

Sensitivity Analyses

Table A.1 shows the optimal subsidy results for discount factors of 3, 5, and 7 percent. We see that the optimal subsidy is robust to this parameter.

As was mentioned earlier, in light of new evidence on the damages from climate change and on changes in capital markets, the Biden administration is considering increasing their social cost of carbon by nearly a factor of 4, to $d_g = \$190$ per ton of CO2. Table A.2 shows how the results change if we set d_g to that updated level.

	$\beta = .93^4$	$\beta = .95^4$	$\beta = .97^4$
Optimal Subsidy	8.8%	9.5%	10.4%

Table A.1: Sufficient statistic estimates for green party's optimal subsidy with different levels of β

	$d_g = \$51/\text{ton CO2}$	$d_g = \$190/\text{ton CO2}$
Optimal Subsidy	9.5%	35.6%

Table A.2: Sufficient statistic estimates for green party's optimal subsidy with different values for the green party's social cost of carbon.

Table A.3 shows sensitivity to the crowding out estimate. In the baseline case we used the average estimate from Abrell et al (2018). Table A.3 shows that the subsidy estimate varies considerably if we use the lowest and highest crowding out estimate from Abrell et al (2018).

	$\frac{dC}{dE_g} = .18$	$\frac{dC}{dE_g} = .3$	$\frac{dC}{dE_g} = .59$
Optimal Subsidy	5.6%	9.5%	18.8%

Table A.3: Sufficient statistic estimates for the green party's optimal subsidy with different values of the crowd out effect $\frac{dC}{dE_g}$.

A.1.5 Calibration Details

Details of Demand Function Calibration

We start by assuming a constant elasticity of demand function:

$$D(p) = Ap^{-\epsilon}$$

This implies an inverse demand function of:

$$v'(E) = (E/A)^{-1/\epsilon}$$

There's a small issue here. Since the inverse demand diverges at $E = 0$, if we try to directly define $v(E) = \int_0^E v'(x)dx$, then $v(E)$ doesn't exist. To get around this, we use a piece-wise function for $v'(E)$:

$$v'(E) = \begin{cases} (E/A)^{-1/\epsilon} & E \geq \alpha \\ (\alpha/A)^{-1/\epsilon} & E < \alpha \end{cases}$$

where α is an arbitrarily small constant. These keeps $v'(E)$ from diverging, so $v(E)$ is well defined. As long as α is smaller than K_b^*, the exact value that we pick won't impact our results.

To calibrate the demand elasticity ϵ, we use micro estimates from Deryugina et al (2020). They use changes in local utility suppliers as a source of long run exogenous variation in electricity prices. Their point estimate for the long run demand elasticity is -.27, with standard errors of .04.

The parameter A in the demand function is set so that with laissez faire policies, the equilibrium level of E equals total electricity consumption in 2022 times four (to account for our period length of four years).

Details of Cost Function Calibration

Following Borenstein and Kellogg (2023), we assume a linear specification for the marginal levelized cost of green energy:

$$MLCOE(E_g) = ME_g + b$$

In words, this is the marginal cost, per period, of producing an additional MWh of green energy each period.

We can find the total present value cost of providing E_g units of green energy by integrating the marginal levelized cost and multiplying by $\frac{1}{1-\beta}$:

$$C(E_g) = (\frac{1}{1-\beta})\int_0^E {}_gMLCOE(x)dx = (.5 * ME_g + bE_g)/(1-\beta)$$
$$= (.5 * M * F(K) + b * F(K))/(1-\beta)$$

where in the last step we plugged in $E_g = F(K)$. We also know that the total present value cost of producing $F(K)$ units now and forever is just the investment cost K. Setting these two equal to each other gives a quadratic equation, which we can solve to obtain the following expression for $F(K)$:

$$F(K) = (1/M)(b^2 + 2(1-\beta)MK)^{1/2} - (b/M)$$

Now, to calibrate M and b we use numbers from Borenstein and Kellogg (2023). Based on engineering cost data, they estimate that the marginal cost of zero emission energy starts at \$64/MWh at 2019 levels of production and rises to \$91/MWh if zero emission sources provided 90% of total electricity generation in 2019[5]. In 2019, wind and solar provided 368 million MWh. Assuming that all new zero emission electricity would come from solar and wind (which is in line with projections from Nalley and LaRose (2022)), then for zero emission sources to provide 90% of 2019 levels, wind and solar generation would have to rise to 2528 million MWh/year. These values, combined with the baseline value for β, imply that $M = \$1.26 * 10^{-8}/MWh^2$ and $b = \$59.40/MWh$.

5. They also consider robustness checks where marginal cost at 90% of 2019 production is \$70/MWh and \$110/MWh. We use these same numbers for the sensitivity analysis in table A7.

Finding the Optimal Subsidy

First, we check whether numerically the calibrated demand and cost functions satisfy the interior assumptions from the baseline model setup. In the baseline case, they do, so the green party's optimal subsidy is just given by the expression in theorem 9:

$$s_g^* = (1/2)(\frac{\beta}{1-\beta})(d_g - d_b)\gamma F'(K_g^*)$$

And K_g^* is the unique solution to the following FOC:

$$(mc + \gamma d_g)F'(K_g^*) = 1 - \beta$$

Solving this for $F'(K_g^*)$ and plugging it into the subsidy expression gives:

$$s_g^* = (1/2)\beta\gamma(d_g - d_b)/(mc + \gamma d_g)$$

This only depends on exogenous parameter values, so it can be used directly to find s_g^*

For some specifications in the sensitivity analyses, the second interior assumption, listed below, doesn't hold:

$$F'(F^{-1}(D(mc + d_g)))(mc + d_g) < 1 - \beta$$

In this case, the theorems 9 and 10 don't apply. Instead, we introduce a new assumption, which can be shown to hold in all of the specifications that we consider. First, define K^{**} as the solution to following equation: $v'(F(K^*{*})) = mc + \gamma d_b$. For any capital stock at or above K^{**}, the brown party finds it optimal to use no brown energy. Our new assumption is:

Assumption A.1:

$$v'(F(K^{**}))F'(K^{**}) - 1 + (1/2)(\frac{\beta}{1-\beta})(v'(F(K^{**}))F'(K^{**}) + (mc + d_g)F'(K^{**})) < 0$$

The following proposition pins down the value of the optimal subsidy in this case:

Proposition A1: *When the first interior assumption and Assumption A.1 hold, then there is a unique equilibrium allocation in the multi-election game which has the following features:*

- *The green party's target capital level (the steady state level) K_g satisfies the following condition:*

$$v'(F(K_g))F'(K_g) - 1 + (1/2)(\frac{\beta}{1-\beta})(v'(F(K_g))F'(K_g) + (mc + d_g)F'(K_g)) = 0$$

- *Any tax policy which implements the allocation involves the green party uses a subsidy s_g in the first period that they gain control which is equal to* $(1/2)(\frac{\beta}{1-\beta})(d_g - d_b)\gamma F'(K_g)$

So, to solve for the optimal subsidy, we first numerically solve for the K_g which satisfies the first condition in Proposition A1. Then plug that into the expression for s_g from proposition A1.

Sensitivity Analysis

As with the sufficient statistic estimate, we first show sensitivity checks for the value of the discount factor and for the two parties' social costs of carbon. In table A.4, we see that higher discount factors (more patience) lead to higher subsidies, but the impact is not large. Table A.5 shows that increasing the green party's social cost of carbon to the Biden Administration's suggested new value leads to significantly higher subsidies.

	$\beta = .95^4$	$\beta = .97^4$	$\beta = .93^4$
Optimal Subsidy	7.7%	8.2%	7.0%

Table A.4: Calibration results for green party's optimal subsidy with different levels of β.

	Baseline	Updated
Optimal Subsidy	7.7%	21.5%

Table A.5: Calibration estimates for green party's optimal subsidy with different values for the green party's social cost of carbon.

Table A.6 shows the sensitivity analysis for the three cost functions for green energy considered in Borenstein and Kellogg (2023). Calibration results for green party's optimal subsidy for different specifications of the green energy cost function. In their baseline specification, the marginal cost of green energy is equal to \$91/MWh at 90% 2019 production levels. In the "High elasticity" case, this number is \$70/MWh. In the "Low elasticity" case, it's \$110/MWh. In the case where the supply curve is more elastic (less steep), we see slightly higher subsidies, as the marginal cost of clean energy doesn't rise by as much when subsidies are introduced. But overall, the subsidy levels are nearly identical to the baseline levels.

	Baseline	High Elasticity	Low Elasticity
Optimal Subsidy	7.7%	7.7%	8.2%

Table A.6: Calibration results for green party's optimal subsidy for different specifications of the green energy cost function.

Solving for the competitive equilibrium in the naive case

First, we define the following function, which equilibrium level of brown production when carbon taxes are equal to d_g and the capital stock is K:

$$E_b^g(K) = D(mc + \gamma d_g) - F(K)$$

The following theorem characterizes the competitive equilibrium in the case where both parties act naively.

Theorem A.2: *If each party j always uses a carbon tax equal to d_j when in power and assumption A.1 holds, there is a unique competitive equilibrium with the following features:*

- *If the brown party is in power in period t, $K_t = max\{0, K_{t-1} - K_b^*\}$*
- *If the green party is in power in period t, $K_t = K_n > K_b$, where K_n is the unique solution to:*

$$-1 + (1 + \frac{\beta}{2(1-\beta)})v'(F(K_n) + E_b^g(K_n))F'(K_n) + \frac{\beta}{2(1-\beta)}(mc + \gamma d_b)F'(K_n) = 0$$

So, to solve for the steady state level K_n, we just need to solve for the condition in the second part of theorem A.2.

A.1.6 Implausible Equilibrium Example

There can be many subgame perfect equilibria in the multiple-election version of the game if parties are sufficiently patient, including one where the no party ever uses green energy. An example of possible strategies which sustain this equilibrium are the following:

1. Phase 1: Whichever party is in power sets $E_b = E_1$ (E_1 is defined below), $x = 0$ and $E_g = 0$. If party i deviates, move to phase 2^i.

2. Phase 2^i: For 1 stage, if party j is in power, they set $E_b = \bar{E}$, $x = 0$, and $E_g = 0$. If party i is in power, they set $x = max\{0, \tilde{K}_i - K_{t-1}\}$, $E_g = F(K_t)$, and $E_b = max\{0, D(mc + \gamma d_i) - F(K_t)\}$. If party j deviates, move to phase 2^j; otherwise, move to phase 1^j.

3. Phase 1^i: Whichever party is in power sets $E_b = E_1 + \epsilon(1\{i = b\} - 1\{i = g\})$, $x = 0$, $E_g = 0$. If any party k deviates, move to phase 2^k.

Where brown production along the equilibrium path E_1 is set to a level that is too high from the green party's perspective and too low from the brown party's perspective:

$$v'(E_1) - (mc + \gamma d_b) > 0 > v'(E_1) - (mc + \gamma d_g)$$

Let *epsilon* be small enough so that the above inequalities are satisfied when we replace E_1 with $E_1 \pm \epsilon$. Define party i's target capital level in phase 2, $\tilde{K}_i$ as the unique solution to $(mc + \gamma d_i)F'(\tilde{K}_i) - 1 = 0$.

Proposition A6: *With sufficiently high β and $\bar{E}$, the above strategy profile is a subgame perfect equilibrium*

Proof. If party i deviates in phase 1, the increase in their flow payoff in that period must be less than $\bar{V} - (v(E_1) - (mc + \gamma d_i))$, where $\bar{V} \equiv lim_{E \to \infty} v(E)$. If they retain power in the next period (the punishment phase), then the increase in their flow payoff is also bounded above by $\bar{V} - (v(E_1) - (mc + \gamma d_i))$. If they lose power in the next period, their flow payoff is $v(\bar{E}) - (mc + \gamma d_i)\bar{E} - (v(E_1) - (mc + \gamma d_i))$. In periods more than two steps ahead, their flow payoff would be weakly lower due to the deviation. So, a sufficient condition for this deviation to not be profitable is:

$$(1 + (\beta/2))\bar{V} + (1/2)\beta(v(\bar{E}) - (mc + \gamma d_i)\bar{E}) < (1 + \beta)(v(E_1) - (mc + \gamma d_i))$$

With sufficiently high $\bar{E}$, this condition will be satisfied. Following the same steps, we can show that with sufficiently high $\bar{E}$, there is no profitable deviation, for either party, from phase 1^i or phase 1^j.

Finally, we need to show that there's no profitable deviation from phase 2^i for either party. For party j, deviating would lead to a one-period benefit of at most $\bar{V} - (v(\bar{E}) - (mc + \gamma d_j)\bar{E})$, but a loss in all future periods due to moving from phase 1^i to phase 1^j. With β close enough to 1, this long term loss is guaranteed to dominate the short term benefit, so the deviation can't be profitable. For party i, their behavior has no impact on future payoffs, so we only need to check that the strategies chosen maximize their one-period payoff. The FOCs for party i are:

$$v'(F(K) + E_b)F'(K) - 1 \leq 0$$

$$v'(F(K) + E_b) - (mc + \gamma d_i) \leq 0$$

Which must hold with equality if the solution is interior. It can easily be verified that these FOCs hold for party i's phase 2 strategy, and that they are strictly decreasing, so they are a sufficient condition for optimality. □

Intuitively, the idea behind this equilibrium is that if any party i deviates from phase 1, they will be punished with a huge amount of brown production in phase 2. To make it rational for party j to carry out the phase 2 punishment, they must be rewarded by moving to phase 1^j afterwards, which is slightly better for them than staying along the equilibrium path.

A.2 Proofs for Chapter 1

Proof of Theorem 1: The Bellman equation for this planning problem is:

$$V(K) = max_{K' \geq K, E_b \geq 0}\ v(E_b + F(K')) - (mc + \gamma d_j)E_b - (K' - K) + \beta V(K')$$

subject to $K' \geq K$, $E_b \geq 0$

The FOCs are:

$$v'(E_b + F(K')) - (mc + \gamma d_j) \leq 0$$

$$v'(E_b + F(K'))F'(K') - 1 + \beta V'(K') \leq 0$$

With equality holding for interior solutions.

Define the function $E_b^j(K') = max\{D(mc + \gamma d_j) - F(K'), 0\}$. Notice that for any value K', setting $E_b = E_b^j(K')$ solves the FOC for E_b.

Now, define K_j^* as the unique solution to:

$$v'(E_b^j(K_j^*) + F(K_j^*))F'(K_j^*) - 1 + \beta = 0$$

There's a unique positive solution since the left side is strictly decreasing, continuous, is greater than zero for $K_j^* = 0$ (due to the first interior assumption) and eventually drops below zero for large enough K_j^* (since $lim_{x \to \infty} v'(x) = 0$).

Guess that the optimal strategy is to set $K' = max\{0, K_j^* - K\}$ and $E_b = E_b^j(K')$. With this guess, the value function becomes:

$$V(K) = \begin{cases} -(K_j^* - K) + \frac{1}{1-\beta}(v(F(K_j^*) + E_b^j(K_j^*)) - (mc + \gamma d_j)E_b^j(K_j^*)) & K \leq K_j^* \\ \frac{1}{1-\beta}(v(F(K) + E_b^j(K)) - (mc + \gamma d_j)E_b^j(K)) & K > K_j^* \end{cases}$$

This satisfies $V'(K) = 1$ for $K \leq K_j^*$ and $V'(K) < 1$ for $K > K_j^*$. From this, we see that the FOC for K' is satisfied at our guessed solution for any K. And, as was already noted, the FOC for E_b is also satisfied. Since the maximization problem in the Bellman equation is concave, we know that the FOCs are sufficient conditions for solving the optimization problem. So, our guessed solution is a solution. Finally, the FOCs have unique solutions, so our guessed solution is the unique solution.

Since the initial capital level is zero, the optimal path has $K_t = K_j^*$ and $E_{bt} = E_b^j(K_j^*)$ in every period. This implies that $x_1 = K_j^* > 0$ and $x_t = 0$ for all $t > 1$, which proves the second part of the theorem.

The next step is to show that $E_b^j(K_j^*) > 0$, which happens iff $D(mc + \gamma d_j) > F(K_j^*)$. Assume otherwise for contradiction. Then the definition of K_j^* implies: $v'(F(K_j^*))F'(K_j^*) = 1 - \beta$. Combining this with the condition that $D(mc + \gamma d_j) \leq F(K_j^*)$ gives the following inequality:

$$(mc + \gamma d_j)F'(F^{-1}(D(mc + \gamma d_j)) \geq 1 - \beta$$

which contradicts the second interior assumption. So, we know that $E_b^j(K_j^*) > 0$. This proves the third part of the theorem.

With this result, the definition of K_j^* simplifies to:

$$(mc + \gamma d_b)F'(K_j^*) = 1 - \beta$$

Which proves the first part of the theorem.

Proof of Theorem 2: Set $p_t = v'(F(K_j^*) + E_b^j(K_j^*)) = mc + \gamma d_j$. At this set of prices and the planner allocation, the consumer's FOCs are satisfied, which implies that the consumer problem is satisfied since it's a convex optimization problem.

The firm's FOC for E_{bt} is:

$$p_t = mc + \gamma \tau_t$$

Plugging in our values for p_t and τ_t this becomes:

$$d_b + \gamma d_j = mc + \gamma d_j$$

So this is satisfied

The firm's FOC for K_1 is:

$$p_t F'(K_j^*) = 1 - \beta$$

And the FOC for K_t for $t \geq 2$ is:

$$p_t F'(K_j^*) \leq 1 - \beta$$

By the planner's FOC, we know that the left side of both of these is equal to $1 - \beta$, so both conditions are satisfied.

Since all FOCs are satisfied, and the firm problem is convex, this is a solution to the firm problem. So, it is a competitive equilibrium.

Proof of Theorem 3: To satisfy the consumer problem, any competitive equilibrium in which quantities are equal to the planner allocation must have $p_t = mc + \gamma d_j$. The firm's net marginal revenue from a marginal investment in period t must be weakly negative for the allocation to solve the firm problem (and must be equal to zero in period 1):

$$\sum_{t'=1}^{\infty} \beta^{t'-1}(mc + \gamma d_j) F'(K_j^*) \leq 1 - s_t$$

From the planner's FOC, we know $(mc + \gamma d_j) F'(K_j^*) = 1 - \beta$, so this condition reduces to:

$$1 \leq 1 - s_t$$

The only way for this to hold is for $s_t \leq 0$.

Proof of Theorem 4: The subgame starting in period 2 after the winner of the election has been realized as party j is just a one-player game, where party j chooses allocations in each period to maximize their social welfare. This is identical to the planner problem considered in theorem 1, except now the initial capital stock is K_1 instead of zero. We already showed in the proof of theorem 1 that the optimal solution in this case is to set K equal to $\max\{K_j^* - K_1, 0\}$ and E_b equal to $max/D(mc + \gamma d_j), 0/$.

Proof of Corollary 1 This follows immediately from the results in theorem 4.

Proof of Theorem 5 According to theorem 4, E_{b1} has no impact on future allocations, so it will just be set to maximize the first period utility, which means $E_{b1} = E_b^g(K_1) = max\{D(mc + \gamma d_g) - F(K_1), 0\}$

So, the green party's optimization problem simplifies to choosing K_1 to maximize:

$$u_g(K_1) - K_1 + \frac{\beta}{1-\beta}((1-\theta)u_g(K_2^g(K_1)) + \theta u_b(K_2^b(K_1))) \\ + \beta((1-\theta)(K_2^g(K_1) - K_1) + \theta(K_2^b(K_1) - K_1))$$

where $u_j(K) \equiv v(F(K) + E_b^j(K)) - (mc + \gamma d_g)E_b^j(K)$ and $K_2^j(K_1) \equiv max\{K_j^* - K_1, 0\}$

Notice that the objective function is continuous. Taking the right derivative of the objective w.r.t. K_1 gives $MB(K-1)$, the net marginal benefit of increasing K_1:

$$\begin{cases} (mc + \gamma d_g)F'(K_1) - (1-\beta) & K_1 < K_b^* \\ (mc + \gamma d_g)F'(K_1) - 1 + \beta(1-\theta) + \frac{\beta}{1-\beta}\theta(mc + \gamma d_g)F'(K_1) & K_b^* \leq K_1 < K_g^* \\ (mc + \gamma d_g)F'(K_1) - 1 + \frac{\beta}{1-\beta}(mc + \gamma d_g)F'(K_1) & K_g^* \leq K_1 < \bar{K}_g \\ v'(F(K_1))F'(K_1) - 1 + \frac{\beta}{1-\beta}((1-\theta)v'(F(K_1)) + \theta(mc + \gamma d_g))F'(K_1) & \bar{K}_g \leq K_1 < \bar{K}_b \\ (1 + \frac{\beta}{1-\beta})v'(F(K_1))F'(K_1) - 1 & K_1 \geq \bar{K}_b \end{cases}$$

where $\bar{K}_j$ is defined as the unique solution to: $v'(F(\bar{K}_j)) = mc + \gamma d_j$.

$MB(K_1)$ is positive for all $K < K_g^*$, negative for all $K > K_g^*$, and equal to zero for $K = K_g^*$. So, $K = K_g^*$ is the unique solution to the optimization problem.

Proof of Theorem 6:

By evaluating the consumer problem FOCs at the equilibrium allocation, we get $p_1 = mc + \gamma d_g$ and $p_t = mc + \gamma d_j$ for all $t \geq 2$ when party j wins the election. With prices at that level, all consumer FOCs are satisfied, so the consumer problem is satisfied.

Note that with the tax policies given in the theorem, the firm's profits don't depend on E_{bt} since price equals the constant marginal cost in each period. So, any choice of E_{bt} can solve the firm problem. All that's left to show is that the firm problem is solved at the subgame perfect allocation of capital.

With this price sequence and the tax policies given in the theorem, the firm problem can be written recursively with the Bellman equations:

$$V_1 = max_{K_1}(mc + \gamma d_g)F(K_1) + (1 - s_1)K_1 + \beta((1 - \theta)V_g(K_1) + \theta V_b(K_1)$$

$$V_j(K) = max_{K'}(mc + \gamma d_j)F(K') - (K' - K) + \beta V_j(K')$$

subject to $K' \geq K$. V_1 corresponds to the first period (note, it doesn't depend on K because the initial capital stock is fixed at zero). $V_j(K)$ is the value function for all periods greater than 1 when party j wins the election.

First we'll solve $V_j(K)$. Guess that the solution is $K' = max\{K, K_j^*\}$. If this guess were correct, the value function would be::

$$V_j(K) = \begin{cases} -(K_j^* - K) + \frac{1}{1-\beta}(mc + \gamma d_j)F(K_j^*) & K \leq K_j^* \\ \frac{1}{1-\beta}\frac{1}{1-\beta}(mc + \gamma d_j)F(K) & K > K_j^* \end{cases}$$

This value function can be easily shown to solve the Bellman. Furthermore, our guess is consistent with the subgame perfect equilibrium allocation. The last step is to show that solves the maximization problem in the first period value function. Since $V_j(K)$ is concave for each j, the FOC in the first period value function is a sufficient conditions for optimality:

$$(mc+\gamma d_g)F'(K_g^*)-(1-s_1)+\beta(\theta V_g'(K_g^*)+(1-\theta)V_b'(K_g^*))=0$$

With our expression for $V_j(K)$, we can show that $V_g'(K_g^*)=1$ and $V_b'(K_b^*)=\frac{1}{1-\beta}(mc+\gamma d_b)F'(K_g^*)$. Plugging these into the FOC for K_1 and solving for s_1:

$$s_1=\theta(\frac{1}{1-\beta})(d_g-d_b)\gamma F'(K_g^*)$$

Which is subsidy in the theorem, so this FOC holds.

Proof of Theorem 7 In the proof of theorem 5, we showed the green party's FOC for increasing K_1 in the first period is:

$$\frac{1}{1-\beta}(mc+\gamma d_g)F'(K_g^*)-1+\beta=0$$

First off, in any competitive equilibrium which has the subgame perfect allocation, prices are pinned down by the consumer problem FOCs: $p_1=mc+\gamma d_g$ and $p_t=mc+\gamma d_j$ for all $t\geq 2$ when party j wins the election. Next, a necessary condition for the firm problem to be solved is that a marginal change in first period investment can't increase expected profits:

$$(mc+\gamma d_g)F'(K_g^*)-(1-s_1)+\frac{\beta}{1-\beta}(\theta(mc+\gamma d_g)+(1-\theta)(mc+\gamma d_g))F'(K_g^*)=0$$

Subtracting this from the green party's FOC gives:

$$s_1=\theta(\frac{1}{1-\beta})(d_g-d_b)\gamma F'(K_g^*)$$

Proof of Theorem 8 In the subgame starting in period 2, the optimal strategy for the party in power is to choose E_{bt}, K_t to maximize their social welfare function:

$$\left[\sum_{t=2}^{\infty}\beta^{t-1}(\underbrace{v(F(K_t)+E_{bt})}_{\text{Consumption Value}}-\underbrace{x_t-mcE_{bt}}_{\text{Cost}}-\underbrace{d_j\gamma E_{bt}}_{\text{Externality}})\right]$$

Subject to the law of motion for all t:

$$K_t = K_{t-1} + x_t$$

conditional on the inherited level of capital K_1.

Notice that there is now no irreversible investment constraint. The Bellman equation associated with this problem is:

$$V_j(K) = max_{K',E_b} v(F(K') + E_b) - (K' - K) - (mc + \gamma d_b)E_b + \beta V_j(K')$$

Guess that the solution is

$$V_j(K) = (K_j^* - K) + \frac{\beta}{1-\beta}(v(F(K_j^*) + E_b^j(K_j^*)) - (mc + \gamma d_j)E_b^j(K_j^*))$$

where $E_b^j(K_j^*) = D(mc + \gamma d_j) - F(K_j^*)$.

This value function solves the bellman equation, and maximization problem within the bellman has a unique solution: $K' = K_j^*$ and $E_b = D(mc + \gamma d_j) - F(K_j^*)$. So, in any subgame perfect equilibrium, $x_2 = K_j^* - K_1$, $x_t = 0$ for all $t \geq 2$, $K_t = K_J^*$ for all $t \geq 2$, and $E_{bt} = D(mc + \gamma d_j) - F(K_j^*)$ for all $t \geq 2$, where j denotes the party who wins the second period election.

The green party's problem in the first period then simplifies to maximizing:

$$v(F(K_1) + E_{b1}) - (mc + \gamma d_g)E_{b1} - K_1 + \beta K_1$$

The FOCs for this are:

$$v'(F(K_1) + E_{b1}) - (mc + \gamma d_g) \leq 0$$

$$v'(F(K_1) + E_{b1})F'(K_1) - (1 - \beta) \leq 0$$

We already showed in the proof of theorem 1 that these have a unique solution of $K_1 = K_g^*$ and $E_{b1} = D(mc + \gamma d_g) - F(K_g^*)$. So, in every period, the party in power sets K_t and E_{bt} to their no-turnover level.

The FOCs from the consumer problem say that if this allocation is part of a competitive equilibrium, $p_t = mc + \gamma d_j$, where j is the party in power in period t. This, combined with the firm's FOCs for E_{bt}, says that $\tau_t = d_j$ for all t, where j is the party in power at time t.

The firm can't benefit from marginally changing investment in the first period and holding it constant in all future periods, which gives the condition:

$$(mc + \gamma d_g)F'(K_g^*) - (1 - s_1) + \frac{\beta}{1 - \beta}(\theta(mc + \gamma d_g)F'(K_g^*) + (1 - \theta)(mc + \gamma d_b)F'(K_b^*)) = 0$$

Since the FOCs from the planner problems tell us that $(mc + \gamma d_j)F'(K_j^*) = 1 - \beta$ for each j, this simplifies to: $s_1 = 0$.

Similarly, the firm can't benefit from marginally changing investment in any period $t \geq 2$:

$$\frac{1}{1 - \beta}(mc + \gamma d_j)F'(K_j^*) - (1 - s_t)$$

which simplifies to $s_t = 0$

Proof of Theorem 9: First, in any markov perfect equilibrium, allocations in each period can only depend on the capital inherited in that period. This means that allocations cannot depend on the history of previous level of brown production. So, brown production in each period is just set to maximize the party in power's period utility, which implies E_{bt} satisfies the FOC $v'(F(K_t) + E_{bt}) = (mc + \gamma d_j)$, which further implies that $E_{bt} = E_b^j(K_t) =$

$max\{D(mc+\gamma d_j) - F(K_t), 0\}$. This is equal to $D(mc+\gamma d_j)$, the value stated in the theorem, as long as $D(mc+\gamma d_j) - F(K_t) \geq 0$, which follows from the later result that along the equilibrium path, $K_t \leq K_g^*$.

Now, to prove the first part of the theorem. Let K_g and K_b be the equilibrium capital targets for the green and brown party, respectively. First we'll show that $K_g > K_b$. Assume otherwise for contradiction. Consider the green party's net marginal benefit from a permanent deviation where they set their capital target to be marginally lower than K_g:

$$1 - [\sum_{t=0}^{\infty} \beta^t (1/2)^t v'(F(K_g) + E_b^g(K_g))F'(K_g) + \sum_{t=1}^{\infty} \beta^t (1/2)^t] = 0$$

This condition simplifies to:

$$v'(F(K_g) + E_b^g(K_g)) - (1-\beta) = 0$$

This is the FOC for investment in the green party's full control solution, so $K_g = K_g^*$.

Now write the brown party's FOC for a marginal increase in K_b:

$$-1 + v'(F(K_b) + E_b^b(K_b))F'(K_b) + (1/2)\sum_{t=0}^{\infty} \beta^t (v'(F(K_b) + E_b^b(K_b))F'(K_b) + u_g^{b\prime}(K_b))$$

where $u_g^b(K_b) \equiv v(f(K_b) + E_b^g(K_b)) - (mc+\gamma d_b)E_b^{g\prime}(K_b)$ is the period utility that the brown party gets when the green party is in power and $K_t = K_b$, which implies $u_g^{b\prime}(K_b) = v'(F(K_b) + E_b^g(K_b))(F'(K_b) + E_b^{g\prime}(K_b)) - (mc+\gamma d_b)E_b^{g\prime}(K_b)$.

There are two possible cases. The first case is $E_b^g(K_b) > 0$. This implies that $E_b^{g\prime}(K_b) = F'(K_b)$, which implies $u_g^{b\prime}(K_b) = v'(F(K_b) + E_b^b(K_b))F'(K_b)$. The brown party's FOC for K_b then simplifies to:

$$-(1-\beta) + v'(F(K_b) + E_b^b(K_b)) = 0$$

We know from the proofs in the no turnover case that this is only satisfied for $K_b = K_b^*$, but we assumed that $K_b > K_b^*$, so this is case is impossible. The second case is $E_b^g(K_b) = 0$. This implies $E_b^b(K_b) = 0$ and $u_g^{b\prime}(K_b) = v'(F(K_b) + E_b^g(K_b))F'(K_b) = v'(F(K_b))F'(K_b)$. So the FOC for K_b simplifies to:

$$-(1-\beta) + v'(F(K_b))F'(K_b) = 0$$

We know from the proofs in the full control case that $v'(F(K_b)+E_b^b(K_b))F'(K_b)$ is strictly decreasing and equals $1-\beta$ when $K_b = K_b^*$. We assumed that $K_b > K_b^*$ here, so $v'(F(K_b))F'(K_b) = v'(F(K_b) + E_b^b(K_b))F'(K_b) < (1-\beta)$, which implies:

$$-(1-\beta) + v'(F(K_b))F'(K_b) < 0$$

But this contradicts our earlier expression for the FOC. In either case we have a contradiction, so it's impossible for $K_b > K_g$.

Now, we can pin down the value of K_b with the brown party's FOC for a permanent deviation which marginally decreases their target level:

$$1 - [\sum_{t=0}^{\infty} \beta^t (1/2)^t v'(F(K_b) + E_b^b(K_b))F'(K_b)] + \sum_{t=1}^{\infty} \beta^t (1/2)^t = 0$$

This condition simplifies to:

$$v'(F(K_b) + E_b^b(K_b)) - (1-\beta) = 0$$

Which we know from the no-turnover case is only satisfied at $K_b = K_b^*$

Now, we can pin down the value of K_g by looking at the green party's net marginal benefit of a permanent marginal increase in K_g, $MB(K_g)$. First, look at the case where

$K_g < K_b^*$. In this case:

$$MB(K_g) = -1 + \sum_{t=0}^{\infty} \beta^t (1/2)^t v'(F(K_g) + E_b^g(K_g))F'(K_g) + \sum_{t=1}^{\infty} \beta^t (1/2)^t$$

which simplifies to:

$$(1 - (1/2)\beta)[-(1 - \beta) + (mc + \gamma d_g)F'(K_g)] > 0$$

where the inequality follows from the fact that $K_g < K_g^*$.

Next the case where $K_g \in [K_b, \bar{K}_g)$, where $\bar{K}_g$ is the unique solution to $v'(F(\bar{K}_g)) = mc + \gamma d_g$. In this case:

$$MB(K_g) = (mc + \gamma d_g)F'(K_g) - 1 + \frac{\beta}{1 - \beta}(mc + \gamma d_g)F'(K_g)$$

which simplifies to:

$$\frac{1}{1 - \beta}[(mc + \gamma d_g)F'(K_g) - (1 - \beta)]$$

We know from the proofs in the no turnover case that this is strictly decreasing and is equal to zero when $K_g = K_g^*$.

Next is the case where $K_g \in [\bar{K}_g, \bar{K}_b)$. In this case:

$$MB(K_g) = v'(F(K_g))F'(K_g) - 1 + \frac{\beta}{1 - \beta}(1/2)(v'(F(K_g)) + (mc + \gamma d_g))F'(K_g)$$

Since $v'(F(K_g)) \leq mc + \gamma d_g$, we find:

$$MB(K_g) \leq \frac{1}{1 - \beta}[(mc + \gamma d_g)F'(K_g) - (1 - \beta)] < 0$$

where the last inequality follows because $K_g > K_g^*$.

The final case is where $K_g \geq \bar{K}_b$. Then we get:

$$MB(K_g) = v'(F(K_g))F'(K_g) - 1 + \frac{\beta}{1-\beta}v'(F(K_g))F'(K_g) < 0$$

Where the inequality follows from $K_g > K_g^*$ and $v'(F(K_g)) \leq mc + \gamma d_g$

So, we've shown that $MB(K)$ is equal to zero at K_g^*, positive for all feasible $K < K_g^*$, and negative for all $K > K_g^*$. This means that the green party's target level must be K_g^*.

Last is to show that this is in fact a subgame perfect equilibrium by checking that there are no profitable one-shot deviations. Start first with the green party. The benefit from a one-shot deviation that increases K_t to $\hat{K} > K_g^*$ is the same as the benefit of permanently increasing the target level. We already showed that this can't be profitable. Now, consider a one shot deviation to $\hat{K} < K_g^*$. If this were profitable, then a permanent deviation of this sort would also be profitable (that is, a deviation where the green party sets their target level to $\hat{K}$ instead of K_g^*). But again, we already showed that this sort of permanent deviation isn't profitable. So the green party has no profitable one shot deviation.

Similarly, for the brown party, we know that if a one-shot deviation to $\hat{K}$ is profitable, then a permanent deviation to $\hat{K}$ must also be profitable. We've already shown that a permanent deviation to $\hat{K} \geq K_g^*$ isn't profitable. Now consider a deviation to $\hat{K} < K_g^*$. The marginal benefit of increasing $\hat{K}$ is given by:

$$-1 + [\sum_{t=0}^{\infty} \beta^t (1/2)^t v'(F(\hat{K}) + E_b^b(\hat{K}))F'(\hat{K})] + \sum_{t=1}^{\infty} \beta^t (1/2)^t$$

which simplifies to:

$$v'(F(\hat{K}) + E_b^b(\hat{K})) - (1-\beta)$$

But, as we already noted previously, this is equal to zero when $\hat{K} = K_b^*$, positive when $\hat{K} < K_b^*$, and negative when $\hat{K} > K_b^*$. So no permanent deviation away from K_b^* can be profitable, which means that no one-shot deviation can be profitable.

Finally, with these equilibrium strategies, and an initial capital level of zero, the capital level never gets above K_g^* along the equilibrium path. So $E_{bt} = D(mc + \gamma d_j) - F(K_t) > 0$ for all t.

Proof of Theorem 10 First, if the competitive equilibrium allocation is equal to the equilibrium allocation from theorem 9, then we know from the consumer problem FOCs that $p_t = mc + \gamma d_j$ for any t where party j is in power.

The FOCs for E_{bt} from the firm problem give:

$$p_t = mc + \gamma \tau_t$$

Plugging in our expression for p_t, this simplifies to $\tau_t = d_j$.

A necessary condition for the firm problem to be satisfied is that the expected impact on profits from a one time marginal increase in investments is zero. In the first period when the green party is in power, that expected increase in profits is:

$$-(1 - s_g^*) + (mc + \gamma d_g)F'(K_g^*) + (1/2)\frac{\beta}{1-\beta}((mc + \gamma d_g)F'(K_g^*) + (c_g + d_b)F'(K_g^*)) = 0$$

The green party's FOC for K_g^* is:

$$-1 + (mc + \gamma d_g)F'(K_g^*) + \frac{\beta}{1-\beta}(mc + \gamma d_g)F'(K_g^*) = 0$$

Subtracting these gives:

$$s_g^* = (d_g - d_b)\frac{(1/2)\beta}{1-\beta}F'(K_g^*)$$

The expected increase in profits from a marginal increase in investment when the brown party has control in the first period is:

$$-(1 - s_b^*) + \frac{1}{1-(1/2)\beta}(mc + \gamma d_b)F'(K_b^*) + \frac{(1/2)\beta}{1-(1/2)\beta}(1 - s_g^*)$$

The brown party's FOC for K_b^* is:

$$-1 + \frac{1}{1-(1/2)\beta}(mc + \gamma d_b)F'(K_b^*) + \frac{(1/2)\beta}{1-(1/2)\beta}$$

Subtracting these gives:

$$s_b^* = \frac{(1/2)\beta}{1-(1/2)\beta} s_g^*$$

Proof of Proposition A.1: First off, in cases where the interior assumptions hold, theorem 9 applies, and so proposition A1 follows directly from theorem 9. Now we need to look at the case where the second interior assumption doesn't hold. Then the green party's full control steady state $K_g^* \geq \bar{K}$, so no brown energy gets used in the green party's full control solution.

Follow identical steps to the proof of theorem 9 until solving for the green party's target level. Here we get the same expressions for the green party's marginal benefit of increasing K_g:

$$\begin{cases} (1-(1/2)\beta)[-(1-\beta) + (mc + \gamma d_g)F'(K)] & K < K_b^* \\ \frac{1}{1-\beta}[(mc+\gamma d_g)F'(K) - (1-\beta)] & K_b^* \leq K < \bar{K}_g \\ v'(F(K))F'(K) - 1 + \frac{\beta}{1-\beta}(1/2)(v'(F(K)) + (mc+\gamma d_g))F'(K) & \bar{K}_g \leq K < \bar{K}_b \\ v'(F(K))F'(K) - 1 + \frac{\beta}{1-\beta}v'(F(K))F'(K) & \bar{K}_g \leq K < \bar{K}_b \end{cases}$$

Following the same steps from Theorem 9, we can show that $MB(K) > 0$ when $K < K_b^*$ and that $MB(K) < 0$ for $K > \bar{K}_b$. Unlike in theorem 9, we can now show that $MB(K) < 0$ for $K \in [K_b, \bar{K}_g)$ since K_g^* is now weakly greater than $\bar{K}_g$.

For $K = \bar{K}_g$, we know that $v'(F(K))F'(K) > 1-\beta 0$ (since $K < K_g^*$ and the green party's full control FOC says $v'(F(K_g^*))F'(K_g^*) = 1-\beta$). This then implies that $MB(\bar{K}_g) > 0$. Assumption A.1 says that $MB(\bar{K}_b) < 0$. Since $MB(K)$ is continuous and strictly decreasing for all $K \in [\bar{K}_g, \bar{K}_b$, there must be a point in that interval where it equals zero:

$$v'(F(K_g))F'(K_g) - 1 + \frac{\beta}{1-\beta}(1/2)(v'(F(K_g)) + (mc + \gamma d_g))F'(K_g)$$

That point must be the green party's steady state K_g, since $MB(K) < 0$ for all $K < K_g$ and $MB(K) > 0$ for all $K > K_g$.

Following the same steps as in theorem 9, we can show that there are no profitable one-shot deviations, so this is an equilibrium. So that proves the first part of the proposition.

Following the proof of theorem 10, we can write the firm's marginal benefit from a one-time increase in investment in the first period the green party has control:

$$-(1 - s_g^*) + v'(F(K_g)F'(K_g) + (1/2)\frac{\beta}{1-\beta}(v(F(K_g))F'(K_g) + (c_g + d_b)F'(K_g^*)) = 0$$

The green party's FOC for K_g^* is:

$$-1 + v'(F(K_g)F'(K_g) + (1/2)\frac{\beta}{1-\beta}(v(F(K_g))F'(K_g) + (c_g + d_g)F'(K_g^*)) = 0$$

Subtracting these gives:

$$s_g^* = (d_g - d_b)\frac{(1/2)\beta}{1-\beta}F'(K_g)$$

The brown party's FOC for K_b^* is:

$$-1 + \frac{1}{1-(1/2)\beta}(mc + \gamma d_b)F'(K_b^*) + \frac{(1/2)\beta}{1-(1/2)\beta}$$

Subtracting these gives:

$$s_b^* = \frac{(1/2)\beta}{1-(1/2)\beta}s_g^*$$

Proof of Proposition 1: Set $p_t(h^t) = mc + \gamma d_j$ in any history where party j has power. The consumer problem FOCs are: $p_t(h^t) = v'(E_t(h^t)) = mc + \gamma d_j$, so these are satisfied, which means the consumer problem is satisfied.

The firm FOCs for E_{bt} are $p_t(h^t) = mc + \gamma\tau_t(h^t)$. With $\tau_t(h^t) = d_j$ and $p_t(h^t) = mc + \gamma d_j$, these are satisfied. E_{bt} drops out of the firm problem since prices exactly equal marginal costs. So the firm problem is now to choose an allocation for K_t and x_t to maximize expected profits:

$$E[\sum_{t=1}^{\infty} \beta^{t-1}(p_t F(K_t) - (1 - s_t)x_t]$$

subject to the law of motion and irreversibility constraints.

This can be written recursively with the following two Bellman equations:

$$V_j(K) = max_{K' \geq K}(mc + \gamma d_j)F(K') - (1 - s_j)(K' - K) + (1/2)\beta(V_b(K') + V_g(K'))$$

Guess that the the policy function when party j is in control is $K'(K) = max\{K, K_j^*\}$. If this guess is correct, then the equilibrium allocation solves the firm problem. With this guess, $V_g(K)$ becomes:

$$\begin{cases} -(K_g^* - K)(1 - s_g) + (mc + \gamma d_g)F(K_g^*) + \frac{(1/2)\beta}{1-\beta}(2mc + \gamma d_g + \gamma d_b)F(K_g^*) & K \leq K_g^* \\ (mc + \gamma d_g)F(K) + \frac{(1/2)\beta}{1-\beta}(2mc + \gamma d_g + \gamma d_b)F(K) & K > K_g^* \end{cases}$$

and $V_b(K)$ is:

$$\begin{cases} -(K_b^* - K)(1 - s_b) + \frac{1}{1-(1/2)\beta}(mc + \gamma d_b)F(K_b^*) & \\ +\frac{(1/2)\beta}{1-(1/2)\beta}(V_g(K_g^*) - (1 - s_g)(K_g^* - K_b^*) & K \leq K_b^* \\ \frac{1}{1-(1/2)\beta}(mc + \gamma d_b)F(K) + \frac{(1/2)\beta}{1-(1/2)\beta}(V_g(K_g^*) - (1 - s_g)(K_g^* - K) & K_b^* < K \leq K_g^* \\ (mc + \gamma d_g)F(K) + \frac{(1/2)\beta}{1-\beta}(2mc + \gamma d_g + \gamma d_b)F(K) & K > K_g^* \end{cases}$$

Taking derivatives:

$$V_g'(K) = \begin{cases} -(1-s_g) & K \leq K_g^* \\ (mc+\gamma d_g)F'(K) + \frac{(1/2)\beta}{1-\beta}(2mc+\gamma d_g+\gamma d_b)F'(K) & K > K_g^* \end{cases}$$

$$V_b'(K) = \begin{cases} -(1-s_b) & K \leq K_b^* \\ \frac{1}{1-(1/2)\beta}(mc+\gamma d_b)F'(K) + \frac{(1/2)\beta}{1-(1/2)\beta}(1-s_g) & K_b^* < K \leq K_g^* \\ (mc+\gamma d_g)F'(K) + \frac{(1/2)\beta}{1-\beta}(2mc+\gamma d_g+\gamma d_b)F'(K) & K > K_g^* \end{cases}$$

These derivatives are continuous and weakly decreasing, which means that the maximization problems within the Bellman equations are convex, so the FOC is a sufficient condition for optimality. Evaluating the FOC of the g bellman at $K' = K_g^*$ and simplifying gives:

$$-(1-s_g) + (mc+\gamma d_g)F'(K_g^*) + \frac{(1/2)\beta}{1-\beta}(2mc+\gamma d_g+\gamma d_b)F'(K_g^*)$$

In the proof of theorem 10, we showed that this is equal to zero. And, since the FOC is strictly decreasing, this means that for $K > K_g^*$ the left side will be ≤ 0, so our guessed solution satisfies the FOC for all values of K.

Evaluating the FOC of the b Bellman at $K' = K_b^*$ and simplifying gives:

$$-(1-s_b) + \frac{1}{1-(1/2)\beta}(mc+\gamma d_b)F'(K_b^*) + \frac{(1/2)\beta}{1-(1/2)\beta}(1-s_g) = 0$$

In the last step of the proof of theorem 10, we showed that this equals zero. And, since the FOC is strictly decreasing, this means that for $K > K_b^*$ the left side will be ≤ 0, so our guessed solution satisfies the FOC for all values of K.

Proof of Theorem A.2: Now, consider a modified version of the G.E. model from appendix A.1. where tax and subsidy revenue does not get transferred to consumers (i.e. taxes represent real costs in this economy). The consumer problem is to choose an allocation $\{y(h_t)\}, \{E(h_t)\}$ to maximize utility:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(y_t + v(E_t))$$

subject to the budget constraint:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(y(h_t) + p(h_t)E(h_t)) \leq A$$

This is the same as in the original GE economy but now transfers equal zero.

The firm problem is to choose an allocation $(\{y(h_t)\}, \{E_b(h_t)\}, \{K(h_t)\}, \{x(h_t)\})$ to maximize expected profits:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(y(h_t) + p(h_t)(F(K_t) + E_b(h_t)))$$

Subject to the resource constraint, law of motion, and irreversibility constraints:

$$y(h_t) + (mc + \tau(h_t))E_b(h_t) + (1 - s(h_t))x(h_t) = L$$

$$K(h_t) = K(h_{t-1}) + x(h_t)$$

$$x_t \geq 0$$

The difference is that taxes and subsidies now appear in the resource constraint.

A competitive equilibrium is now defined as a set of quantities, prices, taxes and subsidies which:

1. Solve the consumer problem

2. Solve the firm problem

3. Satisfy market clearing

The only difference here is that since there are no transfers, there's no government budget constraint.

Note that the consumer and firm problems both reduce to the partial equilibrium problems if we use the constraints to eliminate $y(h_t)$. So, any $(\{E_b(h_t)\}, \{K(h_t)\}, \{x(h_t)\}, \{E(h_t)\})$ which are part of a competitive equilibrium allocation of this economy are also a competitive equilibrium allocation of the PE economy, since they solve the same consumer problem, firm problem, and market clearing constraints.

So, to solve for the set of possible C.E.A. in the PE economy, we just need to solve for the set of possible C.E.A. in the modified GE economy. The first welfare theorem says that any C.E.A. is pareto efficient. To solve for all pareto efficient allocations, find allocations which maximize expected utility:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(y_t + v(E_t))$$

subject to the resource constraint, irreversibility constraint, and law of motion. Using the resource constraint to eliminate $y(h_t)$, the problem simplifies to choosing an allocation to maximize:

$$\sum_t \sum_{h_t} \beta^{t-1} \Pi(h_t)(v(E_t) - (mc + \tau(h_t))E_b(h_t) - (1 - s(h_t))x(h_t))$$

subject to the law of motion and irreversibility constraints.

We are specifically looking at the case where $\tau(h_t) = d_j$ and $s(h_t) = 0$ for all t. In this case, the planner problem can be written recursively as the following two Bellman equations:

$$V_j(K) = max_{K' \geq K, E_b \geq 0} v(F(K')+E_b) - (mc+\gamma d_j)E_b - (K'-K) + (1/2)\beta(V_b(K')+V_g(K'))$$

Guess that the solution is to

1. set $K' = \max\{K, K_n\}$ and $E_b = D(mc + \gamma d_g) - F(K')$ when party g is in power

2. set $K' = \max\{K, K_b^*\}$ and $E_b = D(mc + \gamma d_g) - F(K')$ when party b

The Bellman equations that correspond to this guess are:

$$V_g(K) = \begin{cases} -(K_n - K) + u_g(K_n) + \frac{(1/2)\beta}{1-\beta}(u_g(K_n) + u_b(K_n)) & K \leq K_n \\ u_g(K) + \frac{(1/2)\beta}{1-\beta}(u_b(K) + u_g(K)) & K > K_n \end{cases}$$

$$V_b(K) = \begin{cases} -(K_b^* - K) + \frac{1}{1-(1/2)\beta} u_g(K_b^*) + \frac{(1/2)\beta}{1-(1/2)\beta}(V_g(K_b^*)) & K \leq K_b^* \\ \frac{1}{1-(1/2)\beta} u_g(K) + \frac{(1/2)\beta}{1-(1/2)\beta}(V_g(K)) & K_b^* < K \leq K_n \\ u_b(K) + \frac{(1/2)\beta}{1-\beta}(u_b(K) + u_g(K)) & K > K_g^* \end{cases}$$

where $u_j(K) \equiv v(F(K) + E_b^j(K)) - (mc + \gamma d_j)E_b^j(K)$ and $E_b^j(K) \equiv max\{0, D(mc + \gamma d_j) - F(K)\}$.

Taking derivatives:

$$V_g'(K) = \begin{cases} 1 & K \leq K_n \\ u_g'(K) + \frac{(1/2)\beta}{1-\beta}(u_b'(K) + u_g'(K)) & K > K_n \end{cases}$$

$$V_b'(K) = \begin{cases} 1 & K \leq K_b^* \\ \frac{1}{1-(1/2)\beta}u_g'(K) + \frac{(1/2)\beta}{1-(1/2)\beta} & K_b^* < K \leq K_n \\ u_b'(K) + \frac{(1/2)\beta}{1-\beta}(u_b'(K) + u_g'(K)) & K > K_g^* \end{cases}$$

Both $V_b'(K)$ and $V_g'(K)$ are continuous and weakly decreasing. This means that the maximization problems inside of the Bellmans are strictly convex, so the FOCs are sufficient conditions for optimality. By plugging $V_b'(K)$ and $V_g'(K)$ into the FOCs, we see that they are satisfied at our guessed solution for all values of K. Finally, with these value functions, the FOCs have a unique solution every K, so our guessed solution is the unique pareto efficient allocation in this economy. The welfare theorems tell us, then, that this is the unique competitive equilibrium allocation of this economy.

Proof of Theorem A.4: From period 2 on, the party in power has full control, so it's a one player game. Their planning problem is recursive and has the following Bellman:

$$V(K) = max_{K' \geq K, E_b \geq 0} v(F(K') + E_b) - c(E_b) - \gamma d_j E_b - (K' - K) + \beta V(K')$$

The FOC for E_b is:

$$v'(F(K') + E_b) - c'(E_b) - \gamma d_j \leq 0$$

Where equality holds if $E_b > 0$. For fixed K', this is strictly decreasing and for large enough E_b is negative, so it has a unique solution. Define the function $E_b^j(K')$ to be that unique solution.

As in the proof of theorem 1, guess that the policy functions are $K'(K) = max\{K, K_j^*\}$ and $E_b(K) = E_b^j(K'(K))$, where K_j^* is the unique solution to $v'(F(K_j^*) + E_b^j(K_j^*))F'(K_j^*) = 1 - \beta$. This can be easily shown to solve the bellman equation.

Let the function $K_2^j(K_1) \equiv max\{K, K_j^*\}$ give the capital stock for periods $t \geq 2$ in the case where party j wins the election.

The green party's problem in the first period is then to choose K_1 to maximize:

$$u_g(K_1) - K_1 + \frac{\beta}{1-\beta}[\theta u_b(K_1) + (1-\theta)u^g(K_1)] - \theta(K_2^b(K_1) - K_1) - (1-\theta)(K_2^g(K_1) - K_1)$$

where:

$$u_j(K_1) \equiv v(F(K_1) + E_b^j(K_1)) - c(E_b^j(K_1)) - \gamma d_g E_b^j(K_1)$$

This objective function is continuous. For $K_1 \geq \bar{K}_b$ (where $\bar{K}_b$ is the point where the brown party uses no brown energy in the second period, which is the unique solution to $v'(F(k))F'(K) = c'(0) + \gamma d_b$), the objective function's right derivative is continuous and is equal to $\frac{1}{1-\beta}v'(F(K_1))F'(K_1) - 1 < 0$ since in this range $K_1 > K_g^*$. For $K_1 < K_b^*$, the right derivative is $\frac{1}{1-\beta}v'(F(K_1) + E_b^g(K_1))F'(K_1) - 1 > 0$. So, there must exist at least one finite point $K_1 \in [K_b^*, \bar{K}_b]$ which maximizes the objective.

Now we'll show that the solution must be in the interior of the interval $[K_b^*, \bar{K}_b]$.

First, the right derivative at K_b^* is equal to:

$$v'(E_b^g(K_b^*) + F(K_b^*))F'(K_b^*) - 1 + \beta(1-\theta) + \frac{\beta}{1-\beta}\theta u_b'(K_b^*) > 0$$

Where the inequality follows from the fact that:

$$\begin{aligned} u_b'(K_b^*) &= v'(E_b^b(K_b^*) + F(K_b^*))(F'(K_b^*) + E_b^{b\prime}(K_b^*)) - (c'(E_b^b(K_b^*)) + \gamma d_g)E_b^{b\prime}(K_b^*) \\ &< v'(E_b^b(K_b^*) + F(K_b^*))(F'(K_b^*) + E_b^{b\prime}(K_b^*)) - (c'(E_b^b(K_b^*)) + \gamma d_b)E_b^{b\prime}(K_b^*) = 1-\beta \end{aligned}$$

where the last equality follows from the brown party's FOC in the second period. So the solution can't be at K_b^*.

Next, the left derivative at $\bar{K}_b$ is:

$$v'(F(\bar{K}_b))F'(\bar{K}_b) - 1 + \frac{\beta}{1-\beta}(1-\theta)v'(F(\bar{K}_b))F'(\bar{K}_b) + \frac{\beta\theta}{1-\beta}u_b'(\bar{K}_b)$$
$$< v'(F(\bar{K}_g))F'(\bar{K}_g) - 1 + \frac{\beta}{1-\beta}(1-\theta)v'(F(\bar{K}_g))F'(\bar{K}_g) + \frac{\beta\theta}{1-\beta}u_b'(\bar{K}_b)$$

where $\bar{K}_g$ is the point which satisfies $v'(F(K)) = c'(0) + \gamma d_g$.

We can then put an upper bound on $u_b'(\bar{K}_b)$:

$$u_b'(\bar{K}_b) = v'(F(\bar{K}_b))(F'(\bar{K}_b) + E_b^{b\prime}(\bar{K}_b)) - (c'(E_b^b(\bar{K}_b)) + \gamma d_g)E_b^{b\prime}(\bar{K}_b)$$
$$\leq (c'(0) + \gamma d_g)F'(\bar{K}_b) = v'(F(\bar{K}_b))F'(\bar{K}_b)$$

where the inequality used the fact that crowd out can be at most one-for-one, so $E_b^{b\prime}(\bar{K}_b) \geq -F'(\bar{K}_b)$. This result, combined with our earlier upper bound, says that the left derivative of the objective function at $\bar{K}_b$ is negative.

So, any maximum is in the range $(K_b^*, \bar{K}_b)$. Within that range, the objective function is continuously differentiable.

There are two cases. The first is where $K_1^* < K_g^*$. In this case, the green party makes positive investment when they gain control in the second period. Following the same steps as in the proof of theorem 3, we can easily show that in any implementation of the optimum the green party must use zero subsidies in the second period and a carbon tax equal to d_g in all periods they have control. Similarly, the brown party must use a carbon tax equal to d_b whenever they have control. The green party's FOC for K_1 is:

$$v'(E_b^g(K_1^*) + F(K_b^*))F'(K_1^*) - 1 + \beta(1-\theta) + \frac{\beta\theta}{1-\beta}u_b'(K_1^*) = 0$$

and the firm's FOC is:

$$v'(E_b^g(K_1^*) + F(K_b^*))F'(K_1^*) - 1 + s_1 + \beta(1-\theta) + \frac{\beta}{1-\beta}\theta u'_{bf}(K_1^*) = 0$$

where:

$$u_{bf}(K_1) \equiv v(F(K_1) + E_b^b(K_1)) - c(E_b^b(K_1)) - \gamma d_b E_b^b(K_1)$$

and we used the fact that prices in each period are equal to the marginal utility of consumption. Subtracting the firm and green party FOCs gives:

$$s_1 = \frac{\beta\theta}{1-\beta}(d_g - d_b)E_b^{b\prime}(K_1) = \frac{\beta\theta}{1-\beta}(d_g - d_b)\frac{dC_b(K_1)}{dx_1}$$

The second case is where $K_1 \geq K_g^*$. Here the green party's FOC for K_1 is:

$$v'(E_b^g(K_1^*) + F(K_b^*))F'(K_1^*) - 1 + u'_g(K_1^* + \frac{\beta\theta}{1-\beta}u'_b(K_1^*) = 0$$

And the firm's FOC is given by:

$$v'(E_b^g(K_1^*) + F(K_b^*))F'(K_1^*) - 1 + s_1 + \beta(1-\theta) + \frac{\beta}{1-\beta}\theta u'_{bf}(K_1^*) = 0$$

Subtracting these gives the same subsidy expression:

$$s_1 = \frac{\beta\theta}{1-\beta}(d_g - d_b)E_b^{b\prime}(K_1) = \frac{\beta\theta}{1-\beta}(d_g - d_b)\frac{dC_b(K_1)}{dx_1}$$

APPENDIX B

APPENDIX FOR CHAPTER 2

B.1 Proofs for Chapter 2

Proof of Proposition 3.2:

First, notice that as long as IC constraint 2.2 holds, IC constraint 2.3 can always be satisfied by setting $c(N-2) = c_{nash}$. Since $c(N-2)$ doesn't enter the objective function, there is no cost to setting it equal to c_{nash}. So, the constraint 2.3 can always be trivially satisfied and will never have a positive lagrange multiplier. So, we can restrict our attention to the maximization problem that only includes constraint 2.2.

This maximization problem is convex, so first order conditions are necessary and sufficient.

The FOC for $c(N)$ is:

$$(1-p)NSCC - 2Nac(N) + \lambda(SCC - 2ac(N)) = 0$$

This is only satisfied for $c(N) = c_{fb}$.

The FOC for $c(N-1)$ is:

$$p(SCC - 2ac(N-1)) - \lambda SCC/N \leq 0$$

This must hold with equality if $c(N-1) > 0$.

Both of these FOCs cannot be satisfied for $c(N)$ or $c(N-1)$ which are greater than c_{fb}. So, we can add the constraints that $c(N) \leq c_{fb}$ and $c(N-1) \leq c_{fb}$ without changing the solution. With these added constraints, the optimization problem has a compact feasible set and a continuous objective, so we know that a solution exists.

Assume for contradiction that $c(N-1) = 0$, then constraint 2.2 will not bind, so $\lambda = 0$.

But then the FOC can't hold, since $pSCC > 0$. So, we must have an interior solution where $c(N-1) > 0$. This means that the FOC must hold with equality:

$$p(SCC - 2ac(N-1)) - \lambda SCC/N \leq 0$$

Now, assume for contradiction that the constraint 2.2 didn't bind. Then the FOC says that $c(N-1) = c_{fb}$. But this contradicts proposition 3.1. So, the optimal $c(N-1)$ must be the unique solution to:

$$SCCc(N) - ac(N)^2 = (1/N)SCC((N-1)c(N-1) + c_{nash}) - ac_{nash}^2$$

Proof of Proposition 3.3:

As long as constraint 2.4 holds, then constraint 2.5 can be satisfied at no cost by setting $\tau(N-1) = \tau_n ash$. So, we can focus on the optimization problem that only includes constraint 2.4. This problem is convex, has a continous objective, and has a closed and bounded constraint set. So, there exists a solution and first order conditions are necessary and sufficient for optimality.

The FOC for $\tau(N-1)$ is:

$$-(N-1)pb\tau(N-1) - \lambda(N-1)W_f'(\tau(N-1)) \geq 0$$

This must hold with equality if $\tau(N-1) < \bar{\tau}$. Assume for contradiction that $\tau(N-1) = \bar{\tau}$. Then, since $W_h(\tau_{nash}) + W_f(\bar{\tau}) < 0$, we know that the constraint 2.4 won't bind, so $\lambda = 0$. In that case the IC constraint says that $\tau(N-1) = 0$, which contradicts the earlier assumption. So, the solution must be interior ($\tau(N-1) < \bar{\tau}$) and the FOC must hold with equality.

Now assume for contradiction that the IC constraint 2.4 didn't bind. In that case $\lambda = 0$, so from the FOC we see that $\tau(N-1) = 0$. But this contradicts proposition 3.1. So, the IC constraint must bind. So, the optimal $c(N-1)$ must be the unique solution to:

$$0 = (N-1)(W_h(\tau_{nash}) + W_f(\tau(N-1)))$$

Proof of Proposition 4:

As long as constraint 2.6 holds, constraint 2.7 can be satisfied at no cost by setting $c(N-2) = c_n ash$ and $\tau(N-2) = \tau_{nash}$. So, the optimal $c(N), c(N-1), and \tau(N-1)$ can be found by solving the optimization problem that only has constraint 2.6. That problem is convex, so first order conditions are necessary and sufficient for optimality.

Just as in the proof of proposition 3.2, the FOCs for $c(N)$ and $c(N-1)$ cannot be satisfied for $c(N) > c_{fb}$ or $c(N-1) > c_{fb}$. So, we can add the constraints that $c(N) \leq c_{fb}$ or $c(N-1) \leq c_{fb}$ without changing the solution. With these added constraints, the problem has a continuous objective and a compact constraint set, so we know that a solution must exist.

Since we're assuming the solution is interior, the first order conditions for $c(N-1)$ and $\tau(N-1)$ are:

$$p(SCC - 2ac(N-1)) - \lambda SCC/N = 0$$

$$pb\tau(N-1) - \lambda W_f'(\tau(N-1)) = 0$$

Combining these equations and rearranging gives the equation in the proposition.

Proof of Theorem 1:

The unlinked agreements $c^*_{ul}(N)$ and $\tau^*_{ul}(N)$ are feasible choices in the linked setting. For there to be no gains from issue linkage, the unlinked agreements must be solutions to the linked agreement optimization problem. Since it's a convex problem, these unlinked agreements are a solution to the linked problem if and only if they solve the first order conditions for that problem:

$$p(SCC - 2ac_{ul}^*(N-1)) - \lambda SCC/N = 0$$

$$pb\tau_{ul}^*(N-1) - \lambda W_f'(\tau_{ul}^*(N-1)) = 0$$

where the FOCs must hold with equality since we showed in the proofs of proposition 3.2 and 3.3 that the best unlinked agreements are interior solutions. Combining these FOCs gives the equation in the theorem.

If these FOCs aren't satisfied, then the unlinked agreement doesn't solve the linked optimization problem. Furthermore, since we showed in the proof of proposition 4 that a solution exists to the linked problem, we know in that case that there exists a linked agreement which gives higher expected welfare than the unlinked agreements.

Proof of Theorem 2:

First, we'll show that the solution cannot involve $c(n) > c_{fb}$ for any n. Assume for contradiction that $c(\hat{n}) > c_{fb}$. Then the following FOC would have to hold for $n = \hat{n}$:

$$P(n)(nSCC - 2nac(n)) + \lambda \left[P(n)((SCCn/N) - 2ac(n)) - P(n+1)(SCC/N)n\right] = 0$$

But with $c(n) \geq c_{fb}$, the left side is always negative, so the FOC can't hold, contradicting our earlier assumption. So, we can add the constraints that $c(n) \leq c_{fb}$ for all n without changing the solution. With that addition, the problem has a compact constraint set and a continuous objective, so we know that a solution exists. We know that there is a best linked agreement. So, unless the best unlinked agreement is a solution to the linked optimization problem, then we know that there exists a linked agreement which gives higher expected welfare.

A necessary condition for the best unlinked agreement to solve the linked optimization problem is that the following FOCs are satisfied:

$$P(n)(nSCC - 2nac^*_{ul}(n)) + \lambda \left[P(n)((SCCn/N) - 2ac^*_{ul}(n)) - P(n+1)(SCC/N)n\right] = 0 \tag{B.1}$$

$$- P(n')(N-n')b\tau^*_{ul}(n') + \lambda \left[P(n')(N-n')W'_h(\tau^*_{ul}(n')) - P(n+1)W'_f(\tau^*_{ul}(n'))n'\right] \leq 0 \tag{B.2}$$

where we know that the FOC for $c(n)$ (condition B.1) must hold with equality the theorem states that $c(n) > 0$ and that the FOC for $\tau(n')$ (condition B.2) doesn't include a term for the constraint $\tau(n) \leq \bar{\tau}$ since the theorem states that $\tau^*_{ul}(n') < t\bar{a}u$.

If $\tau^*_{ul}(n') = 0$, then condition B.2 can never hold, so $\tau^*_{ul}(n')$ must be interior, which means that condition B.2 must hold with equality.

Finally, combining the two FOCs to eliminate the lagrange multiplier gives:

$$\frac{P(n)((SCC/N)n - 2ac^*_{ul}(n)) - P(n+1)(SCC/N)n}{P(n)(nSCC - 2nac^*_{ul}(n))} = \frac{P(n')(N-n')W'_h(\tau^*_{ul}(n')) - P(n'+1)W'_f(\tau^*_{ul}(n'))n'}{P(n')(N-n')b\tau^*_{ul}(n')}$$

If this condition doesn't hold, then the unlinked agreements don't solve the linked optimization problem, so there are gains from issue linkage.