Demand- versus Supply-Side Climate Policies with a Carbon Dioxide Ceiling

Thomas Eichner, Gilbert Kollenbach, Mark Schopf*

Abstract

In a dynamic Hotelling model with a climate coalition and free-riding countries, we compare demand-side and supply-side climate policies aimed at keeping the CO_2 concentration below a ceiling which corresponds to a global warming of 2 degrees Celsius. With the demand-side policy the coalition caps its fuel consumption to adhere the ceiling. The associated allocation is intra-temporally distorted and inefficient. With the supply-side policy the coalition purchases fuel deposits to postpone their extraction. When the coalition's initial budget is limited, both the fuel price and the fuel extraction paths can be discontinuous, the supply-side policy causes an inter-temporal distortion and is inefficient. When we add coalition formation, in an empirically calibrated economy the stable coalitions are medium-sized with the demand-side policy, whereas the grand coalition is stable with the supply-side policy. If the coalition acts strategically, the stable grand coalition implements first best.

Keywords: demand-side policy, supply-side policy, climate change, deposit, fossil fuel
JEL Classification: F55, H23, Q54, Q58

*University of Hagen, Department of Economics, Universitätsstr. 41, 58097 Hagen, Germany, emails: thomas.eichner@fernuni-hagen.de, gilbert.kollenbach@fernuni-hagen.de, mark.schopf@fernuni-hagen.de. Financial support from the German Science foundation (DFG grant number EI 847/1-2 and PE 212/9-2) is gratefully acknowledged. We would to thank three anonymous referees for helpful comments. Remaining errors are the authors' sole responsibility.

1. Introduction

In recent years, climate change and its economic consequences have received considerable attention. There is a broad political consensus that the global temperature should not rise by more than two degrees Celsius (UN, 2015). However, even if the parties that ratified the Paris Agreement were to fully implement their nationally determined contributions, the temperature would rise by about three degrees (UN, 2020). Thus, one can doubt whether voluntary contributions to a global climate agreement could guarantee meeting international climate goals. Efforts to mitigate climate change vary substantially across countries. While the European Union committed to reduce greenhouse gas emissions by at least 40% by 2030, compared to 1990 levels, other countries' submitted targets are less ambitious. It is indeed disturbing that worldwide carbon emissions are still increasing. If voluntary contributions to climate agreements cannot stabilize the temperature at safe levels, it is worth considering appropriate unilateral policies to combat global warming at manageable cost.

Unilateral climate policies can be targeted at the demand for fossil fuels (*demand-side policy*) or at the supply of fossil fuels (*supply-side policy*) and aim at ensuring that the carbon dioxide concentration remains below a critical level – a ceiling on the carbon dioxide concentration. According to the IPCC (2013, chapter 8.5 and 10.3), it is very likely that more than half of the global temperature increase between 1951 and 2010 is due to the increase in greenhouse gas concentrations, and it is very likely that carbon dioxide accounted for more than half of the radiative forcing of greenhouse gases between 1750 and 2011 (and between 1980 and 2011). Thus, a ceiling on the carbon dioxide concentration is consistent with both the two-degree target and the UN's (1992) objective to stabilize greenhouse gas concentrations "at a level that would prevent dangerous anthropogenic interference with the climate system".

The present paper is the first to investigate demand- versus supply-side policies in a *dynamic multi-country Hotelling model* with a ceiling on the carbon dioxide concentration. In that model fossil fuels are homogeneous and extracted at marginal costs which decrease over time due to technological progress to focus on the development of scarcity rent and its change due to the various unilateral climate policies.[1] Renewable energy is available as perfect substitute to fossil fuels in energy consumption. There are two groups of countries, a climate coalition which implements climate policies to combat global warming and the fringe which is not concerned about climate change.

There is a large literature that analyzes optimal demand-side policies, particularly cap-and-trade schemes and carbon taxes for adhering to the ceiling in dynamic one-country models. Chakravorty et al. (2006) investigate the implications of increasing or decreasing energy demand over time on optimal abatement and renewable energy utilization. Chakravorty et al. (2008) address the optimal extraction composition of two polluting nonrenewable resources and find that this composition can change several times until the cleaner resource is exhausted. Chakravorty et al. (2012) find that optimal energy prices can decline over time at the ceiling and in the long run, if there is learning-by-doing in the renewable energy sector. Henriet (2012) analyzes the optimal date of backstop invention. Finally, there are several papers which consider carbon taxation in a dynamic two-country model without a ceiling on the carbon dioxide concentration, e.g. Hoel (2011), Hassler and Krusell (2012), Bretschger and Suphaphiphat (2014), Hémous (2016), Ryszka and Withagen (2016), Fischer and Salant (2017), van der Meijden et al. (2018) and Kollenbach (2019).

The first papers that took up unilateral supply-side policies were Bohm (1993) and Hoel (1994). Hoel (1994) analytically characterized the unilaterally optimal mix of supply- and demand-side caps (or taxes) in a static model. Hoel's (1994) model has been further refined in numerical (static) analyses by e.g. Golombek et al. (1995) and Fæhn et al. (2017). Hagem and Storrøsten (2019) derive the unilaterally optimal supply- and demand-side taxes in a dynamic model with free-riders. They find that demand-side policies lead to intertemporal and within-period leakage, which creates a green paradox, whereas supply-side policies lead to negative leakage.

A growing literature has argued that supply-side policies are preferable to demand-side policies (Collier and Venables 2014, Faehn et al., Asheim et al. 2019). Harstad (2012) even

[1] This is a common assumption in the ceiling literature. See, e.g., Amigues et al. (2011), Amigues et al. (2014), Chakravorty et al. (2006), Chakravorty et al. (2008), Henriet (2012), Lafforgue et al. (2008; 2009) and Smulders and Van der Werf (2008).

has shown that a supply-side policy can implement the first best. His supply-side policy consists of purchasing fossil fuel deposits and implements first best by assuming Coasian bargaining in the deposit market, which removes trade and, thus, strategic incentives in the fuel market. To eliminate strategic incentives, not only deposits for preservation, but also those for extraction are traded. Recent literature has revisited this theory under alternative representations of fossil fuel markets (e.g., Eichner and Pethig 2017a, 2017b). Efficiency is violated if the Coasian bargaining is replaced by deposit trade at a uniform price (Eichner and Pethig, 2017b), and efficiency may be violated if deposits are only purchased for preservation purposes, but not for extraction (Eichner and Pethig, 2017a).

We add to this important policy discussion by showing that, on the one hand, (i) in a setting with an emissions ceiling and with homogeneous[2] fossil fuels following Hotelling dynamics, unilateral supply-side policy is unable to decentralize the first best when the coalition's initial budget is not enough to purchase all deposits at the outset, but, on the other hand, that (ii) once we add coalition formation, supply-side policy is actually predicted to result in a stable grand coalition, whereas demand-side policy is predicted to result in medium-sized stable coalitions. If the coalition acts strategically, the grand coalition with supply-side policy achieves the first-best. In any case, allowing for transfers between countries the transition from the stable equilibrium with demand-side policy to the stable equilibrium with supply-side policy is a Pareto improvement. In sum, the results add further support for supply-side policy under both a different policy objective (emissions ceiling) and another fossil fuel market representation (homogeneous Hotelling dynamics) than the prior literature has considered.

[2]In contrast to Harstad (2012) and Eichner and Pethig (2017a; 2017b), whose fuel deposits are heterogeneous and economically exhausted, we consider homogenous fuel deposits and physical exhaustion. The assumption of homogeneous fossil fuels is in line with Chakravorty et al. (2006), Chakravorty et al. (2008) and Hoel (2011). However, whether fossil fuels are homogeneous or heterogeneous is an empirical question. According to IEA (2013, pp. 228), the extraction costs of fuel indeed depend on the characteristic of the deposit. However, considerable amounts can be extracted for almost constant unit costs: 1120 billion barrel oil in North Africa and the Middle East can be extracted for 25$ per barrel or less. 220 trillion cubic meters of conventional gas have extraction costs of 9$ per MBtu or less. Approximately 600 billion tonnes of coal can be extracted for 3$ per MBtu or less. Using the conversion factors of IPCC (2006, chapter 2, table 2.2) shows that the corresponding CO_2 emissions amount to more than 2621Gt. According to IPCC (2018, chapter 2, Tab. 2.2), these emissions imply a violation of the 2°C climate target.

When analyzing the unilaterally optimal demand- and supply-side policies,[3] we characterize the extraction and consumption paths of fossil fuels and the evolution of backstop and total energy, and identify the different distortions caused by these policies. The climate coalition may behave as a price-taker or may act strategically in the fuel market and deposit market. The associated policy is denoted as *competitive* climate policy in the former and as *strategic* climate policy in the latter case. In the analysis of strategic climate policy, we follow Lewis and Schmalensee (1980), Benchekroun et al. (2009), and Benchekroun et al. (2010) and restrict our analysis to the open-loop solution.[4]

If the climate coalition applies a demand-side policy, the social climate costs are not internalized in the fringe. If the coalition is a price taker in the fuel market, the climate coalition limits its fuel consumption below the fringe's fuel consumption to adhere the ceiling. The fringe's fuel consumption is inefficiently high until the ceiling is no longer binding, and the coalition's fuel consumption is inefficiently low when the ceiling is binding. The different fuel consumptions cause an intra-temporal distortion. With the exception of linear energy demands, the demand-side policy also causes an inter-temporal distortion and the associated fuel extraction path does not coincide with the efficient extraction path. The coalition bears the burden of complying with the ceiling by dispensing with fuel consumption. In case of strategic demand-side policy, the strategic effects weaken [can strengthen] the intra-temporal distortion, if the coalition exports [imports] fuel.

If the climate coalition applies a supply-side policy and is a price taker in the fuel market and deposit market, a coalition in general[5] buys deposits successively until it owns the entire fuel stock before the ceiling becomes binding. As long as the coalition buys deposits

[3] A prominent example for the supply-side policy, in especially for the policy of purchasing deposits to prevent their extraction, is the Yasuni-ITT initiative, proposed in 2007 by the Ecuadorian President Correa which was built on the idea that Ecuador leaves oil underground in the Ecuadorian Yasuni National Park, a UNESCO biosphere reserve, in exchange for financial contributions from the international community.

[4] According to Benchekroun et al. (2009), perfect future markets for resources can justify commitment. A comprehensive review of dynamic games of exhaustible resources is given by Van Long (2011). Wirl (1994), Wirl and Dockner (1995), Tahvonen (1996), and Rubio and Escriche (2001) consider polluting non-renewable resources. However, to the best of our knowledge, this is the first paper that investigates a game with one player imposing a ceiling on the stock of emission and natural decay of emissions. As pointed out by Lewis and Schmalensee (1980), feedback equilibria are often intractable, so that this task is left for future research.

[5] An exception is the unrealistic case that the coalition's initial budget is unlimited and it can buy all deposits at the outset.

only firms supply fossil fuel. Since firms do not account for the climate costs of emissions, extraction is inefficiently high and there is an inter-temporal distortion. At the time when the coalition has purchased the last deposit, both the fuel price and fuel extraction path exhibit a jump. With competitive supply-side policy, fuel extraction is inefficiently high in the beginning, and inefficiently low at the end. With strategic supply-side policy, the strategic effects may weaken or strengthen the inter-temporal distortion. If strategic effects are positive, the equilibrium qualitatively has the same properties as the equilibrium with the competitive supply side policy. If strategic effects are negative and sufficiently strong, the fuel extraction path is smooth but still inefficient.

The previously mentioned results hold for two exogenously given groups of countries. Finally, in order to compare the demand-side policy with the supply-side policy we endogenize coalition formation, i.e. we analyze the size of the stable climate coalition with demand-side policy and supply-side policy, respectively, when countries decide to join or to stay outside the climate coalition and conclude a long-term contract.[6] A coalition is stable if no coalition country has an incentive to leave the coalition (internal stability) and no fringe country has an incentive to join the coalition (external stability). In case of competitive demand-side policy, the climate coalition is just so large that the ceiling is adhered to. In an empirically calibrated economy, with demand-side policy the stable coalition is medium-sized, whereas with supply-side policy the grand coalition is stable. If the grand coalition strategically uses its supply-side policy, it purchases the complete fossil stock at the outset and implements the efficient allocation. Comparing both climate policies, the global welfare is higher in the stable equilibrium with competitive [strategic] supply-side policy than in the stable equilibrium with competitive [strategic] demand-side policy.

The remainder of the paper is organized as follows: Section 2 outlines the model. Section 3 characterizes the (constrained) social optimum as a benchmark. Section 4 introduces

[6]Our paper also contributes to the large literature on climate treaties. Economists have analyzed international environmental agreements for a while both in static and dynamic models. In this literature countries decide in a pre-commitment game whether to join the coalition or to stay outside and act as fringe country, and the stability of the climate coalition is analyzed. Reviews of the literature are given by Finus (2001) and Barrett (2003) and some recent outstanding contributions applying dynamic games are Battaglini and Harstad (2016) and Kováč and Schmidt (2021).

the competitive economy in the absence of any regulation (laissez-faire economy). Section 5 analyzes the competitive demand-side policy and Section 6 the competitive supply-side policy. Section 7 turns to demand-side and supply-side policies when the coalition acts strategically in the fuel market and deposit market. Section 8 briefly considers the grand coalition. Section 9 investigates coalition formation in an empirically calibrated economy. Section 10 concludes.

2. The model

This section presents the assumptions of the model that will be maintained throughout the paper. Consider an economy with two (groups of) countries, A and B. Country A is the climate coalition and country B is a (representative) free rider. We refer to the latter also as fringe. Let n_i denote both the size and the population of country $i = A, B$. W.l.o.g. total population is normalized to one, such that $n_A + n_B = 1$. The instantaneous utility of country $i = A, B$ is given by

$$U_i(x_i(t) + q_i(t)) + g_i(t) = n_i U\left(\frac{x_i(t) + q_i(t)}{n_i}\right) + g_i(t) \tag{1}$$

with $U' > 0$ and $U'' < 0$. Each country consumes energy and a consumer good. At each point in time, country i's good consumption is $g_i(t)$. In (1), $\frac{x_i(t)+q_i(t)}{n_i}$ is energy consumption per capita in country i. Energy is generated from fossil fuels, fuel for short, and a renewable (backstop) such as solar energy, wind energy or hydro power. At each point in time, the consumption of fuel and backstop in country i is denoted by $x_i(t)$ and $q_i(t)$, respectively. Both kinds of energy are perfect substitutes.

Each country $i = A, B$ is endowed with fuel and with the consumer good. Country i's good endowment[7] is $\bar{K}_i$ and its fuel endowment is $\upsilon_i S(0)$, where $\upsilon_i \in (0, 1)$ is the share of country i, $\upsilon_A + \upsilon_B = 1$, and $S(0)$ denotes the global fuel endowment. The evolution of the

[7]Our model can be microfounded by production functions with a composite production factor, say land or labor, as input. In that case $\bar{K}_i$ can also be interpreted as country i's factor endowment. For more details we refer to Appendix A.1.

global fuel stock over time is given by[8]

$$\dot{S} = -s, \tag{2}$$

where $s(t)$ is fuel extraction at time t. Fuel extraction at time t causes the cost $C(t)s(t)$. The marginal extraction cost $C(t)$ is constant at time t. Exogenous technological progress decreases the marginal extraction costs over time with the rate χ, i.e.

$$C(t) = C_0 e^{-\chi t}, \tag{3}$$

where C_0 is the initial marginal extraction cost.

Burning fuel unleashes CO_2 emissions, which accumulate in the atmosphere according to[9]

$$\dot{Z} = s - \gamma Z. \tag{4}$$

In (4), Z denotes the emission stock, $\gamma > 0$ a natural regeneration rate and $Z(0) \geq 0$ the emission stock endowment. CO_2 accumulation gives rise to global warming. In line with the ongoing climate protection discussion, in particular the Paris Agreement, we assume that climate damages are controllable, if global temperature does not increase by more than $2°C$ above its pre-industrial level. This climate target translates into a ceiling $\bar{Z}$ on the emission stock, i.e.

$$\bar{Z} - Z(t) \geq 0 \tag{5}$$

must hold at every point in time. Following Chakravorty et al. (2008) and Lafforgue et al. (2009), we assume that the ceiling represents a discrete damage function with negligible damages below and prohibitively high damages above the ceiling.[10]

[8]We use the notation $\dot{z}$ to indicate the derivation of an arbitrary variable z with respect to time t, i.e. $\dot{z} = \frac{\partial z}{\partial t}$. The growth rate $\frac{1}{z}\frac{\partial z}{\partial t}$ is denoted by $\hat{z}$. For the sake of simplicity, the time index t is omitted whenever there is no risk of confusion.

[9]The equation of motion (4) is widely used in the literature, e.g. by Chakravorty et al. (2006), Tsur and Zemel (2009) and Kollenbach (2015a).

[10]Climate scientists argue that above the critical temperature increase of 2° the climate system reaches tipping-points, which initiate non-linear and irreversible processes leading to unbearable consequences for mankind, cf. Graßl et al. (2003). We neglect damages below the ceiling to sharpen our focus on the effects of the ceiling. Amigues et al. (2011) and Dullieux et al. (2011) assume a damage function that reflects manageable damages from emission stocks below the ceiling. In the following, we assume in line with Chakravorty et al. (2006), Chakravorty et al. (2008), Lafforgue et al. (2009), Chakravorty et al. (2012), Kollenbach (2015a) and Kollenbach (2015b) that the ceiling is exogenously given.

Next to the extraction of fuel, each country $i = A, B$ generates renewable energy. The production locations of renewable energy can be ranked according to their costs and the cheapest locations are used first. Thus, the renewable energy cost of country i

$$M_i(q_i) = n_i M\left(\frac{q_i}{n_i}\right), \tag{6}$$

is increasing and convex[11] in the per-capita renewable energy generation $\frac{q_i}{n_i}$. We assume that M satisfies $M_i'(0) = 0$.

The description of the model is completed by the consumer good constraint

$$g_A(t) + g_B(t) = \sum_i \left[\bar{K}_i - M_i(q_i(t)) - \upsilon_i C(t)s(t)\right] \tag{7}$$

and the fuel constraint

$$s = x_A + x_B. \tag{8}$$

In (7), $\bar{K}_i - M_i(q_i(t)) - \upsilon_i C(t)s(t)$ is country i's possibility frontier of transforming energy into the consumer good. Both the consumer good and fuel are internationally mobile as expressed by (7) and (8).

The dynamics of fuel depletion and CO_2 accumulation divide the time line in four different phases, as illustrated in Fig. 1.

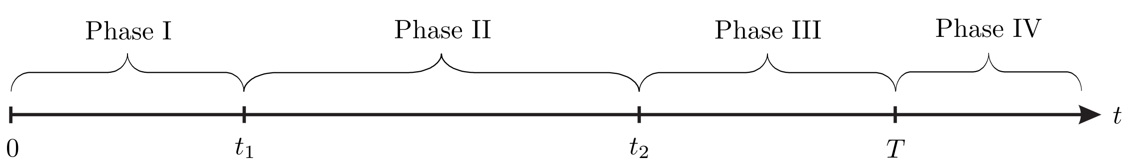

Figure 1: Timeline and sequence of phases

For $t \in [0, t_1)$ the economy is in the pre-ceiling Phase I. During that phase the ceiling is non-binding but CO_2 accumulates in the atmosphere. The ceiling is reached at t_1 and the

[11] The convex cost function can be justified by the rising marginal opportunity cost of access to the land in the deployment process of renewable production locations. Convex costs of renewable energy production are also assumed by Grafton et al. (2012), Fell et al. (2017), and Hoel (2020). Analogously to the marginal extraction costs, backstop costs may decrease because of technological progress. Quantitatively, decreasing backstop costs favor an increasing backstop production and a decreasing fossil fuel extraction. However, this does not change our results qualitatively, so that we abstract from technological progress in the backstop sector for the sake of simplicity.

economy switches into the ceiling Phase II. During Phase II the ceiling binds and limits fuel extraction. At t_2 fuel becomes too scarce so that the ceiling becomes non-binding and the economy switches into the post-ceiling Phase III. This phase ends with the exhaustion of fuel at time T. For all $t \geq T$, the economy is in Phase IV and only renewable energy is used. Def. 1 summarizes the time phases.

Definition 1.
(i) *Phase I — $t \in [0, t_1)$: Ceiling is non-binding but will bind in the future.*
(ii) *Phase II — $t \in [t_1, t_2)$: Ceiling binds.*
(iii) *Phase III — $t \in [t_2, T)$: Ceiling is non-binding and will not bind in the future.*
(iv) *Phase IV — $t \in [T, \infty)$: Only the backstop is used.*

3. The social optimum

In this section we characterize as a benchmark the (constrained) social optimum.[12] The social planner maximizes the intertemporal sum of utility $\int_0^\infty e^{-\rho t}\{\sum_i [U_i(x_i(t)+q_i(t)) + \bar{K}_i - M_i(q_i(t))] - C(t)s(t)\}\,\mathrm{d}t$ subject to the limited fuel stock, the CO_2 ceiling and $s = x_A + x_B$. $\rho > 0$ is the time preference rate. From the first-order conditions we obtain[13]

$$U'\left(\frac{x_i + q_i}{n_i}\right) = C + \tau + \theta - \zeta_{x_i} = M'\left(\frac{q_i}{n_i}\right) - \zeta_{q_i}, \tag{9}$$

$$\tau(t) = \tau_0 e^{\rho t}, \tag{10}$$

$$\dot{\theta} = [\rho + \gamma]\theta - \mu, \tag{11}$$

where the costate variable τ is the scarcity rent of fuel, τ_0 is the initial scarcity rent, and the costate variable θ represents the *social climate costs of emissions*. μ is the multiplier associated with the ceiling, and ζ_{x_i} and ζ_{q_i} are the multipliers of the non-negativity conditions $x_i \geq 0$ and $q_i \geq 0$.

Equation (9) represents the rule for the efficient allocation of energy. It requires the marginal benefit of energy consumption in country $i = A, B$ ($U_i' = U'$) and the *marginal social cost* of energy production to be equal. In case of fuel, the marginal social cost consist of the marginal extraction cost C, the scarcity rent τ, and the social climate costs of emissions

[12] The social optimum is constrained, because the social planner takes the ceiling as exogenously given, see also footnote 10.
[13] The current-value Lagrangian is solved in Appendix A.2.

θ. In case of backstop, the marginal cost is purely private and given by $M'_i = M'$. For an arbitrary point of time $t \in [0, T)$, efficient energy consumption in country $i = A, B$ is illustrated in Fig. 2.

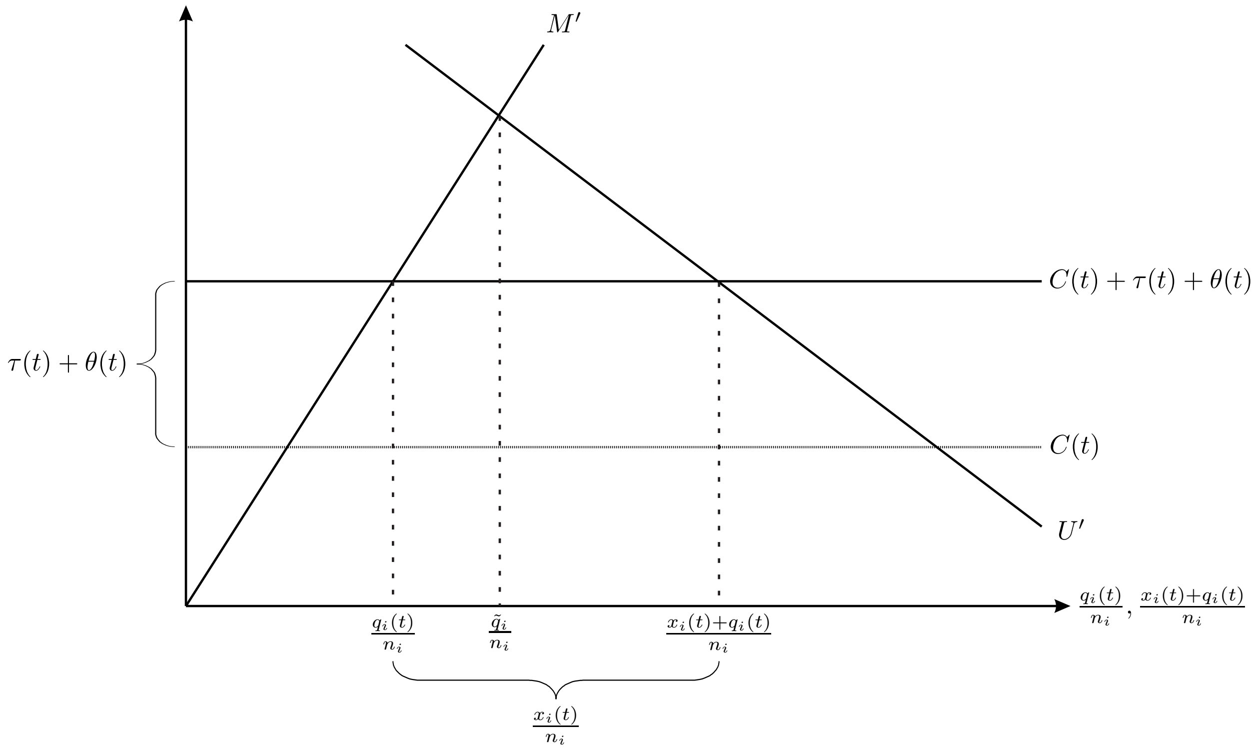

Figure 2: Efficient energy consumption in country $i = A, B$

Total energy consumption $x_i(t) + q_i(t)$ is determined by the intersection of the marginal utility curve U' and the horizontal line $C(t) + \tau(t) + \theta(t)$ which represents the marginal social cost of fuel. Whereas it is efficient to consume backstop energy at any point in time, it is only efficient to consume fuel if the marginal social cost of fuel falls short of the marginal backstop cost $M'\left(\frac{\tilde{q}_i}{n_i}\right)$. In that case, the intersection of the marginal backstop cost curve M' with the horizontal line $C(t) + \tau(t) + \theta(t)$ determines backstop utilization $q_i(t)$ and fuel consumption is given by the difference between total energy consumption and backstop use.[14]

While the marginal utility function and the marginal backstop cost function do not

[14]The energy transition is not the focus of the paper. The assumption $M'(0) = 0$ is a short-cut to get an energy transition from the beginning, and not to run into a out of focus discussion of the characteristics of the efficient allocation and the climate policies throughout the different phases, depending on whether only fossil energy or both energies are consumed.

depend on time, the marginal social cost of fuel develops in accordance with (3), (10) and (11). The marginal extraction costs continuously decrease over time with the rate χ, whereas the scarcity rent grows at the constant rate $\hat{\tau} = \rho$. The growth rate of the social climate costs θ depends on the time phase. During Phase I the ceiling is not binding ($\mu = 0$) and θ evolves in time according to $\theta_I(t) = \theta_0 e^{(\rho+\gamma)t}$. Two opposing effects determine the evolution of fossil fuel and backstop utilization. On the one hand, the increasing scarcity rent and the increasing social climate costs ceteris paribus reduce [raise] fossil fuel extraction [backstop use]. On the other hand, technological progress reduces marginal extraction costs and ceteris paribus leads to a higher [less] fuel extraction [backstop use]. If the latter effect, which is called *technology effect*, dominates, fuel extraction increases. However, the technology effect is the weaker the more advanced the extraction technology, and it can only be dominant until the switching time t^s.[15] In Appendix A.2, we show that a binding ceiling at time t_1 implies a decreasing fossil fuel extraction path at the end of Phase I. Therefore, fossil fuel extraction either decreases for $t \in [0, t_1)$ ($\chi = 0$ is sufficient), or it increases for $t \in [0, t^s)$ and decreases for $t \in [t^s, t_1)$.

In Phase II, the ceiling binds and fuel extraction is constant at rate $\bar{s} := \gamma \bar{Z}$. Fuel consumption is divided over both countries according to $\bar{x}_A = n_A \bar{s}$ and $\bar{x}_B = n_B \bar{s}$ and is time-invariant. The marginal social costs of fuel $C + \tau + \theta$ are also constant during Phase II. The social climate costs $\theta(t)$ are positive during both Phase I and Phase II but they decrease to zero at the end of Phase II, because the ceiling is never reached again for $t \geq t_2$.[16] Solving (11) and making use of $\theta(t_2) = 0$ yields

$$\theta_{II}(t) = \int_t^{t_2} \mu(j) e^{-(\rho+\gamma)(j-t)} \, \mathrm{d}j, \tag{12}$$

where $\mu(j)e^{-(\rho+\gamma)(j-t)}$ represents the present value opportunity costs at time $j > t$ of an additional fuel unit used at time t. If the ceiling binds, the social climate costs at time t are given by the discounted sum of the opportunity costs of the ceiling.

[15] Strictly speaking, t^s denotes a point in time, where the dynamics of $C(t) + \tau(t) + \theta(t)$ switch from $\frac{\mathrm{d}[C(t)+\tau(t)+\theta(t)]}{\mathrm{d}t} > 0$ to $\frac{\mathrm{d}[C(t)+\tau(t)+\theta(t)]}{\mathrm{d}t} < 0$ or vice versa. Because a switch of the dynamics implies a switch in the evolution of fossil fuel extraction, t^s denotes the (local) extrema of the fossil fuel extraction path.

[16] Because the ceiling is non-binding for all $t \geq t_2$ and we abstain from climate damages below the ceiling, the social climate costs are zero.

In Phase III and IV, the ceiling never binds so that both μ and θ equal zero. The marginal social cost of fuel continuously increases, because $\rho\tau(t) > \chi C(t)$ for $t > t_2 > t^s$. At time T, it is equal to $M'\left(\frac{q_i(T)}{n_i}\right) = M'\left(\frac{\tilde{q}_i}{n_i}\right)$ with the consequence that the social planer stops fuel extraction. The path of fuel extraction is summarized in

Lemma 1. *Fossil fuel extraction either decreases or peaks during Phase I, is constant during Phase II, decreases during Phase III, and expires at the end of Phase III.*

The socially optimal evolution of marginal utility and the corresponding fuel extraction path are illustrated in Fig. 3.[17] Phase I lasts from $t = 0$ to t_1^*, Phase II from t_1^* to t_2^*, Phase III from t_2^* to T^* and Phase IV begins at T^*, where the asterisk ($*$) is used to mark the socially optimal values.[18] Consider Phase I. As the ceiling is not binding, U_i' equals $C + \tau + \theta$, where τ and θ grow monotonically over time, while C monotonically decreases. At early points in time, the technology effect dominates, so that fossil fuel extraction peaks before t_1^*. After the peak, fuel extraction decreases and the emission stock increases until the end of Phase I. At $t = t_1^*$ the ceiling is reached and binds until $t = t_2^*$. In Phase II fuel extraction is fixed at $\bar{s}$ and the marginal utility is constant for both countries. As from $t = t_2^*$ the ceiling is non-binding and θ equals zero, so that U_i' equals the sum of marginal extraction cost C and the scarcity rent τ. At $t = T^*$ this sum equals the marginal backstop cost of $q_i(T)$, fuel extraction expires and energy generation relies on the backstop only.

4. The laissez-faire economy

In this section we turn to the laissez-faire economy. In that economy there are perfectly competitive international markets for fuel with price p, deposits with price p_y, the consumer good[19] with price 1, and perfectly competitive national markets for renewable energy with price p_{q_i} in country $i = A, B$ in the absence of any government intervention or regulation. The laissez-faire economy can only be studied for Phase I, i.e. before the ceiling binds. We assume that consumers directly purchase fuel from the fuel firm at the world market price

[17] The numbers in the Figures 3 - 7 stem from our calibration presented in Appendix B. Lemma A.2 proves that the evolution paths are continuous.

[18] The costates τ and θ are part of the social planner's solution of the optimization problem and, therefore, socially optimal by definition.

[19] The consumer good is chosen as numéraire.

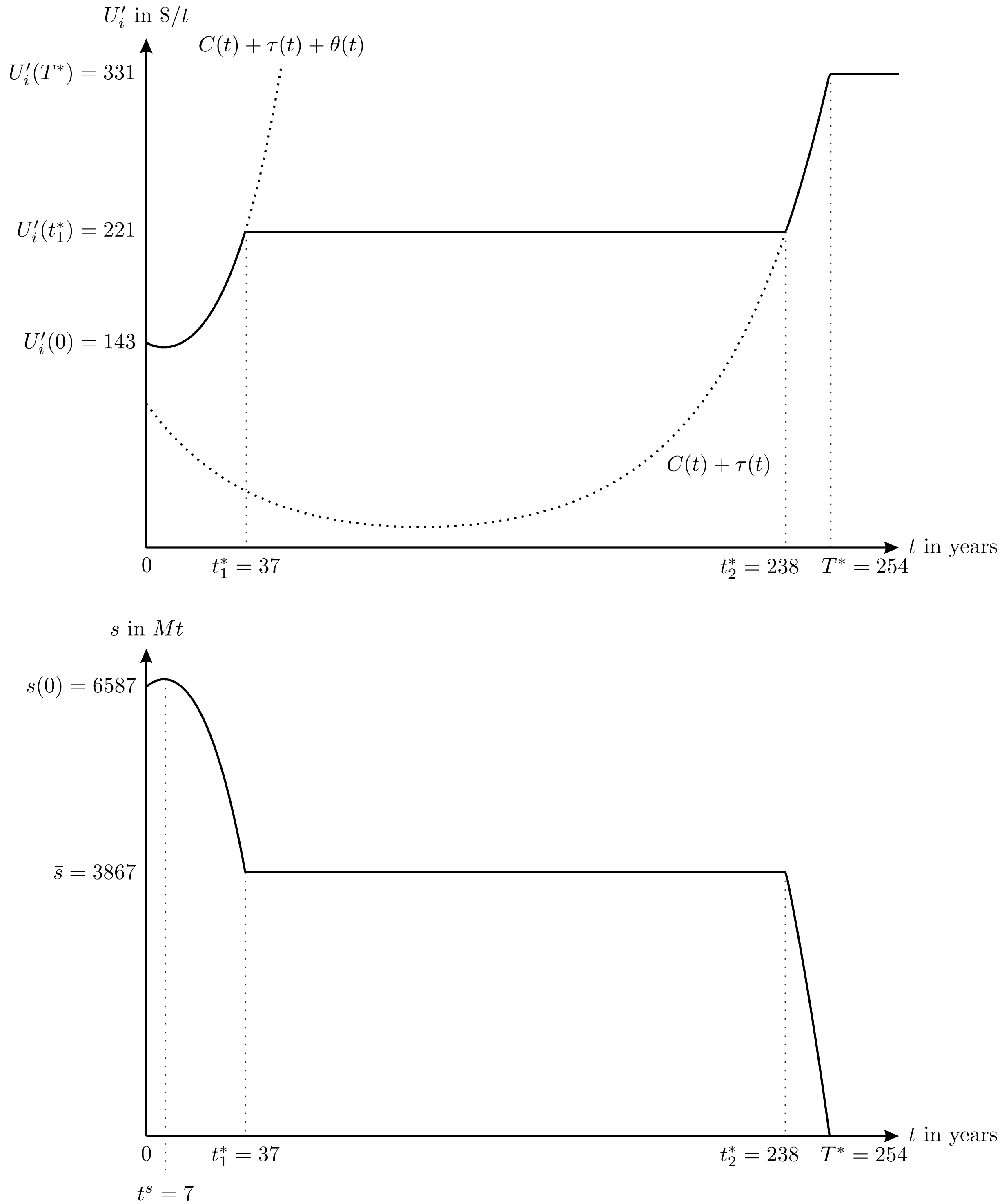

Figure 3: Evolution of U_i' and fuel extraction over time at the social optimum

p. In case of renewable energy, inappropriate cross-border infrastructure, conversion losses of long-range energy transport as well as high transportation costs support the assumption of domestic markets for renewables with prices p_{q_i}. In each country, consumers are identical and backstop firms are identical. W.l.o.g. we consider a representative consumer and a representative backstop firm in country $i = A, B$. In addition, there is a single representative fuel firm.

The backstop firm of country $i = A, B$ maximizes its profit $\Pi_i(t) = p_{q_i}(t)q_i(t) - M_i(q_i(t))$ with respect to $q_i(t)$. The first-order condition

$$M'\left(\frac{q_i(t)}{n_i}\right) = p_{q_i}(t) \tag{13}$$

equates the marginal backstop cost to the renewable energy price p_{q_i} of country i. The fuel firm maximizes its intertemporal profit[20] $\int_0^\infty e^{-\rho t}\Pi_F(t)\,dt$ with respect to fuel and deposit supply, subject to a limited fuel stock. When selling deposits the fuel firm sells the property right of extracting the fuel stored in these deposits, i.e. it sells property rights on mines, oil or gas fields. At each point of time, the profit $\Pi_F(t) = p(t)s_F(t) + p_y(t)y_F(t) - C(t)s_F(t)$ consists of the revenues $p(t)s_F(t)$ from selling fuel and the revenues [costs] $p_y(t)y_F(t) > [<]0$ from selling [of purchasing] deposits diminished by the extraction costs $C(t)s_F(t)$. The corresponding first-order conditions yield the fuel supply correspondence[21]

$$s_F(t) \in [0, S(t)], \text{ if } p(t) = C(t) + \tau_F(t), \tag{14}$$

the deposit supply correspondence

$$y_F(t) \in [0, S(t)], \text{ if } p_y(t) = \tau_F(t) \tag{15}$$

and the Hotelling-rule

$$\tau_F(t) = \tau_{F0}e^{\rho t}, \tag{16}$$

with τ_{F0} as the initial private scarcity rent.[22] (14) shows that the firm is willing to sell any desired amount of fuel if the fuel price $p(t)$ equals the sum of extraction costs $C(t)$ and

[20]In the market economy there exists a capital market (not modeled here) with an interest rate that equilibrates the capital market. We assume that the interest rate is equal to the social discount rate.

[21](14) - (16) are derived in Appendix A.3

[22]To distinguish between the efficient costates of Section 3 and the costates in the competitive economy, we add subscripts to the latter. Thus, τ_F refers to the scarcity rent of the representative fuel firm, while τ denotes the efficient scarcity rent.

scarcity rent $\tau_F(t)$. If the fuel price falls short of this sum, the firm does not sell any fuel at time t. In contrast, the firm sells all remaining reserves if the fuel price exceeds $C(t)+\tau_F(t)$. In a similar manner, all remaining deposits are sold if the deposit price exceeds the scarcity rent. If the deposit price equals the scarcity rent, the firm is willing to sell any desired amount of deposits, and it would like to acquire deposits if the deposit price falls below the scarcity rent. (16) is the well-known Hotelling-rule, i.e. the scarcity rent grows over time at the time preference rate ρ.

Consider the energy demand in the laissez-faire economy. The representative consumer of country $i = A, B$ maximizes her utility (1) subject to the budget constraint $p(t)x_i(t) + p_{q_i}(t)q_i(t) + g_i(t) = \omega$. Her income $\omega = \bar{K}_i + \Pi_i(t) + v_i\Pi_F(t)$ consists of the exogenous income $\bar{K}_i$, the profit $\Pi_i(t)$ of the backstop firm in country i and the share v_i of the fuel firm's profit $\Pi_F(t)$. The first-order condition of utility maximization

$$U'\left(\frac{x_i(t)+q_i(t)}{n_i}\right) = p(t) = p_{q_i}(t) \tag{17}$$

determines country i's demand for fuel and backstop energy in case of an interior solution. Eq. (17) requires at the margin that the benefit of consuming energy in country i equals the international fuel price and the national backstop energy price of country i.

Next, consider the equilibrium of the fuel and backstop energy markets. Assuming an interior solution with respect to fuel supply, i.e. $p(t) = C(t) + \tau_F(t)$, the equilibrium of the fuel market and the national backstop energy market in country $i = A, B$ at some point in time t can be illustrated as in Fig. 2. The intersection of the marginal utility curve U' with the horizontal $p(t) = C(t) + \tau_F(t)$ determines total energy consumption in equilibrium, while backstop consumption is in equilibrium if the fuel price $p(t)$ equals the marginal backstop cost. Formally, fuel and backstop consumption of country $i = A, B$ are given by

$$D_i(p(t)) = \begin{cases} n_i U'^{-1}(p(t)) - n_i M'^{-1}(p(t)), & \text{if } p(t) < M'\left(\frac{\tilde{q}_i}{n_i}\right) \\ 0, & \text{otherwise} \end{cases}, \tag{18}$$

$$Q_i(p(t)) = n_i M'^{-1}(p(t)), \tag{19}$$

with $p(t) = C(t) + \tau_F(t)$ and

$$(a)\ \frac{\mathrm{d}D_i}{\mathrm{d}p} = \frac{n_i}{U''} - \frac{n_i}{M''} < 0, \qquad (b)\ \frac{\mathrm{d}Q_i}{\mathrm{d}p} = \frac{n_i}{M''} > 0. \tag{20}$$

Fuel consumption decreases, whereas backstop consumption increases in the fuel price p. Since the increase of backstop consumption cannot completely compensate the corresponding decrease of fuel consumption, total energy demand decreases in the fuel price p.

5. Competitive demand-side policy

In case of demand-side climate policy, the climate coalition unilaterally caps its fuel consumption. In the economy with unilateral demand-side policy, backstop and fuel supply are determined by (13), (14) and (16). The fringe's[23] fuel and backstop energy demand is given by (18) and (19), and the consumer's demand for backstop energy in the coalition is characterized by $U'_A = p_{q_A}$. The coalition's optimal fuel cap follows from maximizing the welfare $\int_0^\infty e^{-\rho t}\ \{U_A(x_A(t) + q_A(t)) - p(t)x_A(t) - p_{q_A}(t)q_A(t) + \bar{K}_A + \upsilon_A \Pi_F(t) + \Pi_A(t)\}\, \mathrm{d}t$ with respect to x_A, given the CO_2 ceiling. When doing so, the climate coalition neglects its influence on the instant fuel price p and the scarcity rent τ_F. In other words, the coalition is a price taker in the fuel market. The first-order condition[24]

$$U'_A = p(t) + \theta_A(t) \tag{21}$$

characterizes the coalition's optimal fuel cap which is set such that the coalition's marginal utility equals its consumer price. The consumer price is composed of the fuel price p and the coalition's climate costs of emissions θ_A.[25]

In Phase I the ceiling is not binding but the emission stock becomes larger. The ceiling binds between t_1 and t_2 implying $\theta_A(t) > 0$ for $t \in [0, t_2)$. In contrast, $\theta_A = 0$ during Phase III and IV, because the ceiling never binds for $t \geq t_2$. It holds $U'_A > p$ in Phase I and II and $U'_A = p$ in Phase III. Since the sum of fuel extraction costs and scarcity rent $C(t) + \tau_F(t)$ exceeds the marginal backstop cost $M'\left(\frac{\tilde{q}_i}{n_i}\right)$ for all $t \geq T$, the economy's energy use rests on renewable energy in Phase IV. In Appendix A.4 we prove

Proposition 1. *Suppose the coalition applies a demand-side climate policy and is a price taker in the fuel market.*

[23] Throughout the paper the fringe's government is inactive. In particular, we abstract from energy security within the fringe which is beyond the scope of the paper and left for future research.

[24] We add the subscript A to the co-states to differentiate them from the socially optimal co-states from Section 3. See Appendix A.4 for the complete solution of the coalition's optimization problem.

[25] The costate variable θ_A develops over time as in (11).

(i) In Phase I and II fuel and total energy [backstop] consumption per capita are larger [smaller] in the fringe than in the coalition. In Phase III and IV fuel and backstop consumption per capita in both countries coincide.

(ii) Fuel and total energy [backstop] consumption in the fringe increases [declines] until $t^s_B \in [0, t_2)$ and declines [increases] afterwards.

(iii) The coalition's fuel and total energy [backstop] consumption increases [declines] until $t^s_A \in [0, \min\{t^s_B, t_1\})$ and declines [increases] for $t \in [t^s_A, t_1)$. It increases [declines] during Phase II if $t^s_B < t_1$, and it declines [increases] for $t \in [t_1, t^s_B)$ and increases [declines] for $t \in [t^s_B, t_2)$ if $t^s_B > t_1$. Finally, it declines [increases] during Phase III.

(iv) Fuel and backstop consumption in both countries are continuous for all points in time.

The evolution of both marginal utility and fuel consumption per capita for the climate coalition and the fringe are depicted in Fig. 4. Phase I lasts from period 0 to t_1. In Phase I the marginal utility within the coalition exceeds the fringe's marginal utility $\left(U'\left(\frac{x_A+q_A}{n_A}\right) = p + \theta_A > p = U'\left(\frac{x_B+q_B}{n_B}\right)\right)$ such that the fringe consumes per capita more fuel than the coalition.[26] The demand-side policy drives a wedge between consumer prices and a wedge between per-capita fuel consumption in both countries. The coalition limits its per-capita fuel consumption below the fringe's per-capita fuel consumption in order to adhere the ceiling. The evolution of the consumer prices is determined by the technological progress, the growing scarcity rent and the evolution of the coalition's climate costs. In case of the coalition, the technology effect is dominated from the outset, so that per-capita fuel [backstop] consumption in the coalition decreases [increases] during Phase I. For the fringe, the technology effect dominates during the complete Phase I and per-capita fuel [backstop] consumption continuously increases [decreases]. The increasing fuel consumption of the fringe outweighs at early periods, so that total fuel extraction peaks at t^s.

At time t_1, the CO_2 ceiling is reached and the economy switches into Phase II, which lasts until t_2. At the ceiling, aggregated fuel consumption $\bar{s}$ is constant. Because the technology effect dominates the scarcity rent effect until t^s_B, the fringe's per-capita fuel [backstop] consumption increases [decreases] for $t \in [t_1, t^s_B]$, so that the coalition's fuel cap has to decrease to adhere the ceiling. To put it differently, the wedge between the consumer prices still increases during this time. From t^s_B on, the fringe's consumer price increases due to the

[26] Fuel consumption in the coalition is given by $D_A(p(t) + \theta_A(t))$, while fuel consumption in the fringe is given by $D_B(p(t))$.

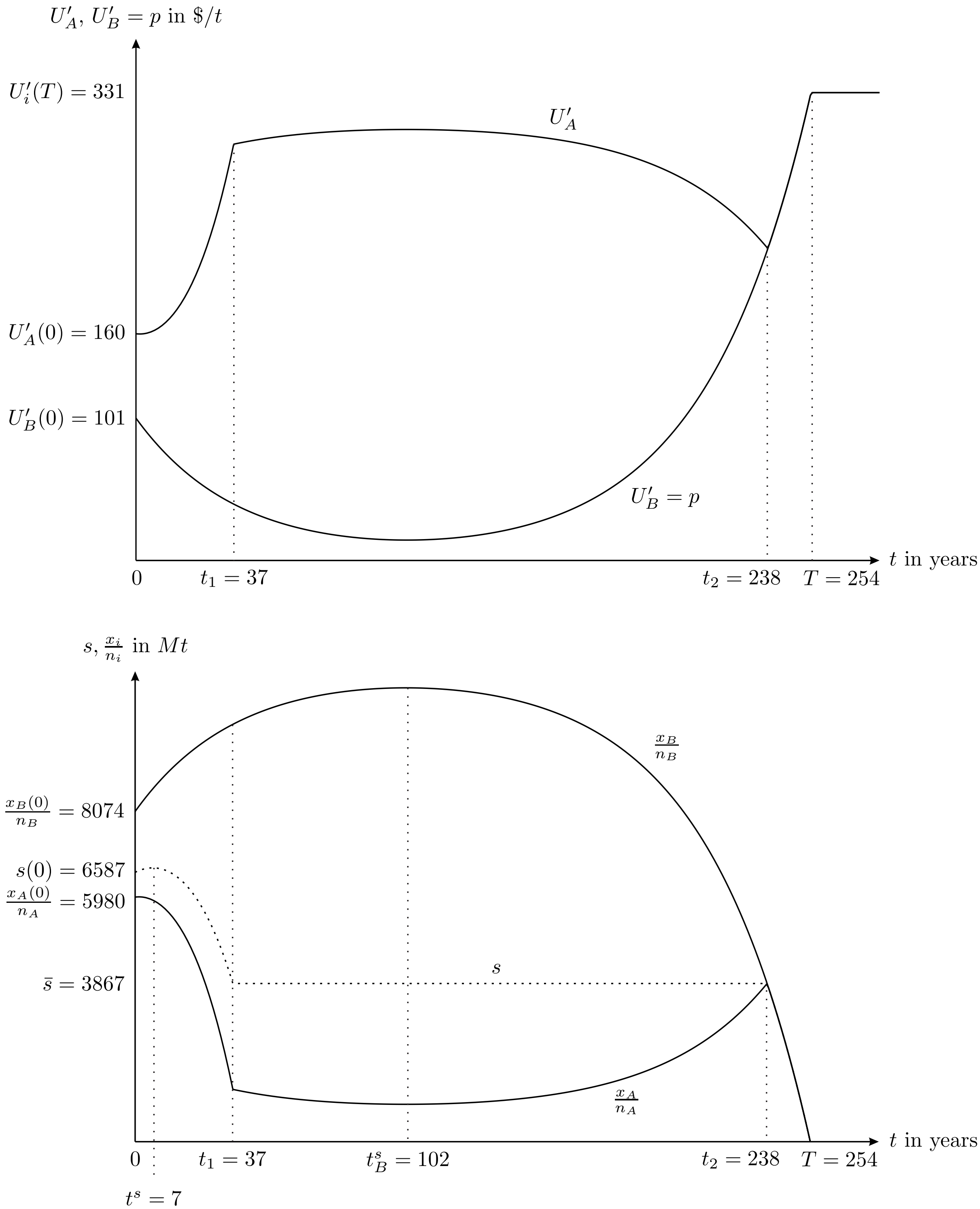

Figure 4: Evolution of U'_A, $U'_B = p$, fuel consumption and extraction over time with competitive demand-side policy

increasing scarcity rent such that the fringe's fuel consumption decreases and its backstop consumption increases which allows the coalition to increase its fuel cap. For $t \in [t_B^s, t_2)$ the consumer price wedge decreases and at time t_2 it vanishes completely.

Actually, the scarcity of fuel renders the CO_2 ceiling irrelevant for all $t \geq t_2$, because the high fuel price induces consumers to reduce their fuel consumption to levels that no longer endanger the ceiling. The ceiling becomes non-binding at time t_2 and the economy switches into Phase III, which lasts until time T. The climate costs θ_A of fuel use are nil for all $t \geq t_2$, so that the coalition and the fringe face the same consumer price and their fuel consumption per capita coincide. Within Phase III, fuel [renewables] consumption decreases [increases] in both countries until the fuel consumption is abandoned at time T. For all $t \geq T$, the economy is in Phase IV and relies completely on renewable energy.

Next, we compare the allocation of the demand-side policy with the socially optimal allocation. In Appendix A.4 we establish

Proposition 2. *Suppose the coalition applies a demand-side climate policy and is a price taker in the fuel market.*

(i) The demand-side policy is inefficient.

(ii) In Phase I and II the fuel price path does not fully internalize the climate costs of emissions.

(iii) The fringe's fuel consumption is higher than at the social optimum for $t \in [0, \max\{t_1, t_1^*\})$. *The coalition's [fringe's] fuel consumption is lower [higher] than at the social optimum when the ceiling is binding.*

(iv) If $\tau_{F0} > [<] \, \tau_0$, *then cumulative fuel extraction* $\int s_F(t)\,\mathrm{d}t$ *is higher [lower] than at the social optimum for* $t \in [0, \max\{t_1, t_1^*\})$ *and lower [higher] than at the social optimum for* $t \in [\min\{t_2, t_2^*\}, \max\{T, T^*\})$. *The switches to Phase III and IV occur earlier [later] than at the social optimum.*

Comparing the fuel consumption path with the socially optimal one reveals an *inter-temporal* and an *intra-temporal* distortion caused by the demand-side climate policy. In Phase I and II, the coalition accounts for the climate costs θ_A whereas the fringe does not account for any climate costs of emissions. Moreover, the coalition is not able to control the fringe's fuel consumption. As a consequence, the fringe's fuel consumption is inefficiently high. In Phase II the ceiling is binding, country A sacrifices fuel consumption in order to adhere the ceiling and its fuel consumption is inefficiently low. For $t \in [0, t_2)$ the demand-side climate policy causes an intra-temporal distortion by violating $\frac{x_A(t)}{n_A} = \frac{x_B(t)}{n_B}$.

The relationship between the efficient scarcity rent τ_0 and the scarcity rent τ_{F0} in the economy with the demand-side policy provides the intuition for the inter-temporal distortion. In Phase II fuel extraction is $\bar{s}$, both at the social optimum and with the demand-side policy. Next, consider Phase I. Ceteris paribus, fuel demand is higher in the economy with the demand-side policy than in the social optimum, which has a positive effect on the scarcity rent τ_{F0}. To compensate the fringe's excess consumption, the coalition sharpens its climate policy. This, in turn, depresses demand and, therefore, has a negative effect on τ_{F0}. The total effect is indeterminate in sign. If $\tau_{F0} > [<] \tau_0$, the energy price is inefficiently high [low], cumulative fuel extraction is inefficiently high [low] in Phase I, and the transition from Phase III to Phase IV occurs earlier [later] with the demand-side policy than at the social optimum. The demand-side policy causes an inter-temporal distortion by antedating [delaying] fuel extraction.

To shed more light on the opposing effects with respect to the inter-temporal distortion, we turn to specific functions.[27] In the special case of linear demand functions, the inter-temporal distortion vanishes, i.e. $\tau_{F0} = \tau_0$, $t_1 = t_1^*$, $t_2 = t_2^*$, $T = T^*$ and cumulative fuel extraction is efficient in all phases. For linear demand functions the non-internalized climate costs in the fringe are exactly offset by an over-internalization in the coalition. It remains the intra-temporal distortion that leads to inefficiently low fuel consumption of the coalition, and inefficiently high fuel consumption of the fringe in Phase I and II.

When utility functions exhibit hyperbolic absolute risk aversion and demand functions become convex, the over-internalization in the coalition becomes more pronounced to ensure that the ceiling is not violated. Furthermore, the fuel consumption paths become flatter, which postpones the switch to Phase II, i.e. $t_1 > t_1^*$, and reduces fuel extraction in Phase I. This depresses the scarcity rent, i.e. $\tau_{F0} < \tau_0$, which postpones the switches to Phase III and IV, i.e. $t_2 > t_2^*$ and $T > T^*$. Intuitively, the coalition's climate policy must overcompensate the fringe's free-riding behavior to adhere the ceiling, which reduces global fuel demand and thereby lowers the value of the exhaustible resource. The effects are reversed, when the demand functions become concave.

[27]The associated analysis can be found in Appendix A.4.

6. Competitive supply-side policy

In case of the supply-side policy, the climate coalition accumulates a state-owned fuel stock by purchasing fuel deposits, and extracts later on fuel from these deposits to sell it to the consumers.[28] The state-owned fuel stock S_A evolves in time according to

$$\dot{S}_A = -s_A + y_A. \tag{22}$$

Purchased deposits are denoted by y_A, while s_A refers to the coalition's fuel supply. To distinguish between s_A and s_F we refer to the latter as *private* fuel supply. The coalition's supply-side policy consists of purchasing deposits, $y_A(t)$, and supplying fuel, $s_A(t)$. At every period t, the coalition's fiscal budget is given by

$$G(t) = p_y(t)y_A(t) - [p(t) - C(t)]s_A(t), \tag{23}$$

where $p_y(t)y_A(t)$ are the expenditures for purchasing deposits and $[p(t) - C(t)]s_A(t)$ are the coalition's profits from selling fuel. The fiscal budget is financed by the lump sum tax $G(t)$ imposed on the coalition's consumers.[29] Due to the quasi-linear utility function (1), an (exogenous) increase in the lump sum tax reduces the consumers' equilibrium good consumption but leaves the consumers' equilibrium energy consumption unchanged. Assuming the coalition's government considers a minimum level of the consumer good $\bar{g}_A(t)$ as necessary to secure a subsistence level for the consumers, the lump sum tax $G(t)$ is constrained by

$$\bar{G}(t) := \bar{K}_A + \Pi_A(t) + \upsilon_A \Pi_F(t) - p(t)x_A(t) - p_{q_A}(t)q_A(t) - \bar{g}_A(t), \tag{24}$$

where $\bar{G}(t)$ is the maximal feasible lump sum tax at time t. Then $\bar{y}_A(t) = \frac{\bar{G}(t)+[p(t)-C(t)]s_A(t)}{p_y(t)}$ are the maximal purchasable deposits at time t.

In the economy with unilateral supply-side policy, backstop energy supply is given by (13). The firm's fuel supply is determined by (14) and (16), and its deposit supply by (15). Both countries' energy demand is given by (18) and (19). The coalition purchases deposits to

[28] Alternatively, the coalition could sell the deposits back to the fuel firm bit by bit. Independently whether the climate coalition sells its fuel to consumers or the fuel firm, by purchasing deposits it indirectly controls fuel supply and fuel price. Selling the fuel to the fuel firm would not change the results.

[29] If $G > 0$ [$G < 0$], the individuals are taxed [receive a transfer].

build up a fuel stock. The supply-side policy influences both the fuel market and the deposit market equilibrium. The unilaterally optimal supply-side policy of the climate coalition follows from maximizing the welfare $\int_0^\infty e^{-\rho t}\{U_A(x_A(t) + q_A(t)) + \bar{K}_A - G(t) - p(t)x_A(t) - p_{q_A}(t)q_A(t) + \upsilon_A \Pi_F(t) + \Pi_A(t)\}\,\mathrm{d}t$ subject to the CO_2 ceiling, the limited fuel stock $S_A(t)$, the fiscal budget (23) and the fiscal budget constraint $G(t) \leq \bar{G}(t)$. When doing so, the climate coalition chooses its supply-side policy $(s_A(t), y_A(t))$ for $t < T_F$ and $s_A(t)$ for $t \geq T_F$, where T_F denotes the point in time the fuel firm's fuel stock becomes exhausted. In addition, the coalition is a price taker in the fuel market and deposit market, i.e. it takes the instant fuel price p and the instant deposit price p_y as given.

As shown in Appendix A.5, the coalition demands deposits according to the correspondence

$$y_A(t) \begin{cases} \in [0, \min\{\bar{y}_A(t), S(t)\}], & \text{if } p_y(t) = \tau_A(t), \\ = \min\{\bar{y}_A(t), S(t)\}, & \text{if } p_y(t) < \tau_A(t). \end{cases} \tag{25}$$

and supplies fuel according to the correspondence

$$s_A(t) \in [0, S_A(t)], \text{ if } p(t) = C(t) + \tau_A(t) + \theta_A(t). \tag{26}$$

The evolution of the coalition's scarcity rent is governed by the Hotelling-rule

$$\tau_A(t) = \tau_{A0} e^{\rho t}, \tag{27}$$

with τ_{A0} as the coalition's initial scarcity rent. The scarcity rent of the fuel firm, $\tau_F(t)$, and the scarcity rent of the climate coalition, $\tau_A(t)$, grow with the time preference rate. In (26), θ_A denotes the coalition's climate costs of emissions.[30] $\theta_A > 0$ during both Phase I and Phase II, and $\theta_A = 0$ in Phase III and Phase IV.

To adhere the ceiling, the climate coalition must buy some deposits. According to (15) and (25), either $\tau_A(t) > \tau_F(t) = p_y(t)$ or $\tau_A(t) = \tau_F(t) = p_y(t)$ holds.[31] In both cases, the fuel firm is willing to sell any desired amount of deposits. If $\tau_A(t) > p_y(t)$, the coalition applies a maximal deposit acquisition policy. It buys the complete fuel stock at time 0 and implements

[30]The costate variable θ_A develops over time as differential equation such as (11).
[31]See also Lemma A.5 of Appendix A.5.

the social optimum presupposed the initial budget $\bar{G}(0)$ is unlimited ($\bar{G}(0) > p_y(0)S(0)$). This result is recorded in

Proposition 3. *Suppose the coalition applies a supply-side climate policy and is a price taker in the fuel market and deposit market. If the coalition's initial budget is unlimited and deposit acquisitions are maximal, the supply-side policy is efficient.*

The assumption of an unlimited initial budget is unrealistic, since the value of the world's proven coal reserves exceeds the world's GDP. According to EIA (2020b), the world's proven coal reserves were 1031 billion tonnes in 2015. Given a price of \$100 per tonne, the reserves have a value of \$103.1 trillion, whereas the world's GDP was \$75.1 trillion in 2015 (The World Bank, 2022).

Next, suppose the initial fiscal budget is limited ($\bar{G}(0) < p_y(0)S(0)$). If $\tau_A(t) > p_y(t)$, the coalition's budget constraint binds[32] and its deposit acquisitions are as before maximal, but it now uses its complete budget $\bar{G}(t)$ to successively buy deposits until the fuel firm's stock will be exhausted at time $T_F \leq t_1$. If $\tau_A(t) = p_y(t)$, the coalition applies a singular deposit acquisition policy. It is indifferent with respect to the amount of purchased deposits. The coalition's fiscal budget constraint does not bind and it purchases in every period $t < T_F$ the deposits $y_A(t) \in [0, \bar{y}_A(t)]$. It chooses T_F so that the marginal value of postponing T_F equals the marginal costs of postponing T_F. Proposition 4 which is proven in Appendix A.5 provides further information about the competitive supply-side policy.

Proposition 4. *Suppose the coalition applies a supply-side climate policy and is a price taker in the fuel market and deposit market. If the coalition's budget constraint binds or deposit acquisitions are singular,*

(i) the private fuel stock becomes exhausted during Phase I, i.e. $T_F \leq t_1$. For $t \in [0, T_F)$, the equilibrium is characterized by $\bar{y}_A(t) \geq y_A(t) \geq 0$, $s_F(t) > 0$ and $s_A(t) = 0$. For $t \in [T_F, T)$, the equilibrium is characterized by $y_A(t) = s_F(t) = 0$ and $s_A(t) > 0$.

(ii) the coalition either uses its complete budget for deposit acquisitions until the firm's stock is exhausted [$y_A(t) = \bar{y}_A(t)\ \forall t < T_F$] or the budget constraint never binds [$y_A(t) \in [0, \bar{y}_A(t)]\ \forall t < T_F$].

(iii) fuel and backstop consumption per capita in both countries coincide for all points in time.

(iv) fuel and total energy [backstop] consumption in both countries increase [decline] until $t^s \in [0, t_1)$ and decline [increase] for $t \in (t^s, t_1)$, they are constant during Phase II, and they decline [increase] during Phase III.

[32] In that case the budget constraint binds in every period $t \in [0, T_F]$.

(v) *fuel and total energy [backstop] consumption in both countries jump downwards [upwards] at* $t = T_F$*, and are continuous for all other points in time.*

To better understand Proposition 4, we illustrate the fuel price path and the fuel extraction path in Fig. 3. Consider the fuel market. For the sake of clarity we distinguish between the consumer price $p(t) = U'$, the coalition's producer fuel price $p_A(t) = C(t) + \tau_A(t) + \theta_A(t)$ and the firm's producer fuel price $p_F(t) = C(t) + \tau_F(t)$. Because the coalition's scarcity rent coincides with or exceeds the firm's scarcity rent ($\tau_A(t) \geq \tau_F(t)$), the firm's producer price falls short of the coalition's producer price during Phase I, so that the fuel firm sells fuel ($s_F(t) > 0$), whereas the coalition's supply is nil ($s_A(t) = 0$) as long as the firm's fuel stock is not exhausted ($t < T_F$). To ensure the ceiling, the coalition purchases deposits ($y_A(t) > 0$) during this time. In the calibrated economy underlying Fig. 3, the effect of technological progress outweighs the increase of the scarcity rent for all $t < T_F$. Thus, the price path decreases, the fuel consumption and total energy consumption paths in both countries increase, while the backstop consumption paths in both countries decline.[33]

At time T_F, the firm's fuel stock becomes exhausted and the coalition takes over fossil fuel supply. Because the fuel firm does not take the ceiling into account, its producer price is lower than the coalition's producer price, the fuel consumer price jumps from $p = C(T_F) + \tau_F(T_F)$ upwards to $p = C(T_F) + \tau_A(T_F) + \theta_A(T_F)$ and fuel [backstop] consumption in both countries jumps downwards [upwards]. The price jump rests on the exhaustion of the firm's fuel stock. Suppose that the (representative) fuel firm tries to exploit the price jump. It may do so in two ways. First, it may withhold some reserves to sell them at time $j > T_F$. Then, perfect competition on the fuel market would yield the equilibrium fuel price $p(j) = p_F(j) = C(j) + \tau_F(j)$. In other words, by trying to exploit the price jump, the firm would eliminate it and would be, therefore, indifferent between selling fuel (and deposits) at time j or some $t < T_F$. Withholding reserves is no option to exploit the price jump. Second, the firm may try to buy some deposits at time T_F to sell them later with profit to the coalition. Perfect competition on the deposit market would increase the equilibrium

[33] If the scarcity rent effect becomes dominant before T_F, fossil fuel extraction peaks at $t^s \in [0, T_F)$. In this case, fuel consumption and total energy consumption increase [decline] and backstop consumption declines [increases] for $t \in [0, t^s)$ $[t \in [t^s, T_F)]$. If fossil fuel extraction has not peaked before T_F, it may peak at $t^s \in [T_F, t_1)$.

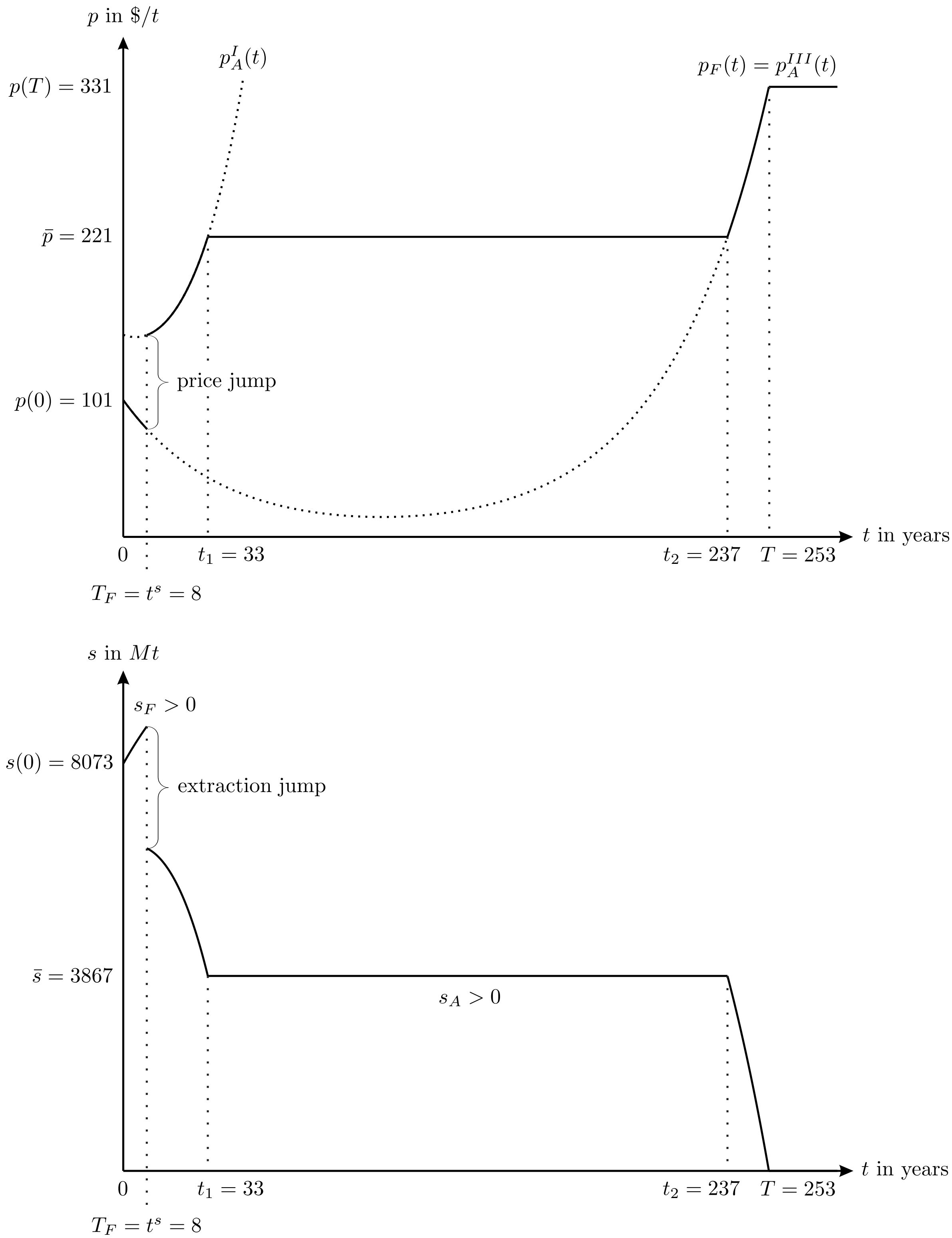

Figure 5: Evolution of p and fuel extraction over time with competitive supply-side policy

deposit price from $p_y(T_F) = \tau_F(T_F)$ up to $p_y(T_F) = \tau_F(T_F) + \frac{\theta_A(T_F)}{1+\zeta_G}$, where $\zeta_G \geq 0$ is the Lagrange multiplier associated to $\bar{y}_A(t) - y_A(t) \geq 0$.[34] Moreover, the firm would not be willing to sell fuel at the price $C(t) + \tau_F(t)$ but only at the price $C(t) + \tau_F(t) + \frac{\theta_A(T_F)}{1+\zeta_G} e^{\rho(t-T_F)}$. If at $t = t_1$ the private fuel supply were above $\bar{s}$, then the ceiling would be violated, so that the climate coalition has to buy deposits implying a price jump.[35] The fuel firm can again try to exploit this price jump by buying deposits to sell them to the coalition. As explained above, this would lead to a further increase in the fuel and deposit prices, and to a price jump until the private extraction path does not violate the ceiling anymore. In this case, the coalition would not buy any deposits, and fuel consumption would end before the firm's stock is exhausted. Because leaving some resources in situ does not maximize the firm's profits, buying deposits is also not an option to exploit the price jump.[36] We conclude that the fuel firm has no possibility to exploit the price jump, which rules out private fuel supply after T_F.

When the ceiling becomes binding at time t_1, the economy switches into Phase II. The constant fuel consumption level at the ceiling is $\bar{s} = D_A(\bar{p}) + D_B(\bar{p})$. The associated constant fuel price is $\bar{p}$. Because the fuel firm does not take the ceiling into account, its producer price is lower than the coalition's producer price. If the fuel firm sells fuel, the ceiling would be violated. To avoid this, the coalition buys the last remaining deposits of the firm before the ceiling becomes binding, so that private fuel supply expires before t_1 and the coalition's fuel supply begins at $T_F \leq t_1$.

At time t_2, the coalition's remaining fuel stock becomes too low to allow for an extraction rate of $\bar{s}$. The economy switches into Phase III and the ceiling becomes non-binding. In Phase III the consumer price still equals the coalition's producer price. Since the latter continuously increases in time, fuel consumption in both countries decreases and backstop consumption increases until fuel consumption is abandoned at time T and the economy switches into Phase IV.

A comparison with the efficient allocation (see Appendix A.5) reveals

[34]Proposition 4(ii) implies that ζ_G is a constant. See the proof of Lemma A.5 in Appendix A.5 for details.
[35]Observe that the coalition's fuel supply price exceeds the private fuel supply price.
[36]See the proof of Lemma A.9 in Appendix A.5 for details.

Proposition 5. *Suppose the coalition applies a supply-side climate policy and is a price taker in the fuel market and deposit market. If the coalition's budget constraint binds or deposit acquisitions are singular,*

(i) the supply-side policy is inefficient.

(ii) in Phase I the fuel price path does not fully internalize the climate costs of emissions until the private fuel stock is exhausted.

(iii) private [governmental] fuel extraction is higher [lower] than at the social optimum directly before [after] the private fuel stock is exhausted.

(iv) cumulative fuel extraction $\int [s_F(t)+s_A(t)]\,\mathrm{d}t$ *is higher than at the social optimum for* $t \in [0, \max\{t_1, t_1^*\})$ *and lower than at the social optimum for* $t \in [\min\{t_2, t_2^*\}, \max\{T, T^*\})$. *The switches to Phase III and IV occur earlier than at the social optimum.*

Observe that during Phase I and Phase II the supply-side policy is characterized by $U'\left(\frac{x_i(t)+q_i(t)}{n_i}\right) = C(t) + \tau_F(t)$ until $t = T_F$ and by $U'\left(\frac{x_i(t)+q_i(t)}{n_i}\right) = C(t) + \tau_A(t) + \theta_A(t)$ for $t \geq T_F$. In Phase III, the policy is determined by $U'\left(\frac{x_i(t)+q_i(t)}{n_i}\right) = C(t) + \tau_A(t)$. Both fuel consumption per capita and backstop consumption per capita are identical in the coalition and the fringe for all points in time, formally $\frac{x_A(t)}{n_A} = \frac{x_B(t)}{n_B}$ and $\frac{q_A(t)}{n_A} = \frac{q_B(t)}{n_B}$. In contrast to the demand-side policy, there is no intra-temporal distortion. However, the climate costs of emissions are not internalized until $t = T_F$, which gives rise to an inter-temporal distortion. Until the exhaustion of the private fuel stock, the growth rate of the fuel price path $p(t) = C(t) + \tau_F(t)$ is smaller than the growth rate of the optimal price path $C(t) + \tau(t) + \theta(t)$. As a consequence, private fuel extraction is inefficiently high when the coalition takes over fossil fuel supply. Afterwards, governmental fuel extraction is inefficiently low to adhere the ceiling. Nevertheless, cumulative fuel extraction is inefficiently high in Phase I and inefficiently low in Phase III. Finally, in Appendix A.5, we prove that the coalition's scarcity rent exceeds the socially optimal scarcity rent ($\tau_A(t) > \tau(t)$), such that the switches to Phase III and IV occur earlier than at the social optimum.

7. Strategic climate policy

In this section we assume that the coalition acts strategically, i.e. uses its climate policy both to comply with the ceiling and to influence prices and scarcity rents in its favor. We consider a form of a dynamic Stackelberg game where the coalition is the leader and takes the impact of its climate policy on the reaction of all consumers and producers (in its own jurisdiction and in the fringe) as well as on the market equilibria into account.[37]

7.1. Strategic demand-side policy

As before, we first consider the demand-side policy. Abandoning the technical details to Appendix A.6 we prove that the unilaterally optimal fuel cap path $x_A(t)$ is now determined by

$$U'_A = p(t) + \theta_A(t) + \mathrm{SE}e^{\rho t} \tag{28}$$

for $t \in [0, T_A)$ and by $x_A(t) = 0$ for $t \in [T_A, \infty)$, where

$$\mathrm{SE} := \frac{\int_0^{T_A} x_A(t)\,\mathrm{d}t - \upsilon_A S(0)}{\left|\int_0^T e^{\rho t} D'_B(t)\,\mathrm{d}t\right|} + \frac{\int_0^{t_2} \theta_A(t) D'_B(t)\,\mathrm{d}t}{\left|\int_0^T e^{\rho t} D'_B(t)\,\mathrm{d}t\right|}, \tag{29}$$

represents the coalition's *strategic effects*:

- The *terms-of-trade effect* $\frac{\int_0^{T_A} x_A(t)\,\mathrm{d}t - \upsilon_A S(0)}{\left|\int_0^T e^{\rho t} D'_B(t)\,\mathrm{d}t\right|} \gtreqless 0$ induces the coalition to reduce [raise] its fuel caps if the coalition is a fuel importer [exporter], to reduce [increase] its import bill [export revenues] by reducing [increasing] the fuel price.

- The *emission effect* $\frac{\int_0^{t_2} \theta_A(t) D'_B(t)\,\mathrm{d}t}{\left|\int_0^T e^{\rho t} D'_B(t)\,\mathrm{d}t\right|} < 0$ induces the coalition to raise its fuel caps to reduce the fringe's fuel demand and carbon leakage, which mitigates the climate costs of emissions in Phase I and II, by increasing the fuel price.

At first, suppose there is no climate problem, so that only the terms-of-trade effect exists. Then, the coalition uses its fuel caps to reduce [increase] its import costs [export revenues] by depressing [increasing] the fuel price if it is fuel importer [exporter]. The coalition reduces [increases] its fuel demand below [above] the fringe's demand for all $t \leq T_B$ [$t \leq T_A$] and it abandons fuel consumption earlier [later] than the fringe. In the presence of the climate problem the emission effect emerges and the coalition switches to the backstop later than the fringe ($T_A > T_B$), if the coalition exports fuel ($n_A \leq \upsilon_A$) or if the emission effect outweighs

[37]By following Lewis and Schmalensee (1980), Benchekroun et al. (2009), Benchekroun et al. (2010), and Battaglini and Harstad (2016), we assume throughout this section that the coalition can commit itself to its strategy, i.e. we assume open-loop strategies. The time consistency problem of our approach is discussed in Appendix A.6. See also Bergstrom (1982), and, e.g., Karp (1984), Karp and Newbery (1992) and Maskin and Newbery (1990).

the terms-of-trade effect. This case is illustrated in Fig. 6. Ceteris paribus, the strategic effects alleviate the inefficiency of the demand side policy (see Proposition 2) by weakening the intra-temporal distortion. Relaxing the coalition's fuel cap counters the inefficiently high fuel consumption in the fringe. However, the fuel price path in the fringe is $p(t) = c + \tau_F(t)$, so that the intra-temporal distortion still exists and efficiency is not achieved.

If the coalition imports fuel ($n_A > \upsilon_A$) and the terms-of-trade effect outweighs the emission effect, fuel utilization terminates earlier in the coalition than in the fringe ($T_A < T_B$) and fuel consumption per capita in the coalition falls short of per-capita consumption in the fringe during the first three phases. Ceteris paribus, the strategic effects amplify the inefficiency of the demand-side policy, because the lower fuel consumption in the coalition reduces the fuel price, and therefore increases the already inefficiently high fuel consumption in the fringe during Phase I and II, i.e. the strategic effects strengthen the intra-temporal distortion. Irrespective of whether the coalition imports or exports fuel, it stops consuming fuel at another point in time than the fringe ($T_A \neq T_B$) indicating that the inter-temporal distortion still is present. For the knife-edge case that the strategic effects cancel out, both countries abandon fuel use at the same time. We summarize our results in[38]

Proposition 6. *Suppose the coalition applies a demand-side climate policy, acts strategically in the fuel market and is committed to its strategy.*

(i) The demand-side policy inefficient.

(ii) In Phase I-III [III] fuel and total energy consumption per capita are larger [smaller] in the fringe than in the coalition if the positive [negative] strategic effects dominate, which implies $T_B >[<] T_A$.

(iii) If the positive strategic effects dominate, Propositions (iii) and (iv) continue to hold. Fuel and total energy [backstop] consumption in the fringe increases [declines] until $t_B^s \in [0, T_B)$ and declines [increases] afterwards.

(iv) If the negative strategic effects dominate, Propositions 1(ii) and (iv) continue to hold. The coalition's fuel and total energy [backstop] consumption declines [increases] for $t \in [t_1, \max\{t_1, t_B^s\})$, increases [declines] for $t \in [\max\{t_1, t_B^s\}, \min\{t_2, T_B\})$ and is constant for $t \in [\min\{t_2, T_B\}, t_2)$.

(v) The negative strategic effects dominate if the coalition exports fuel ($n_A \leq \upsilon_A$).

In the special case of linear demand functions, the switch to Phase II [III] occurs later

[38]For $\chi = 0$, the technology effect vanishes. If the coalition then imports fuel ($n_A > \upsilon_A$), the emission effect is outweighed by the coalition's climate cost of emissions at the end of Phase I, so that fuel consumption per capita is larger in the fringe than in the coalition at $t = t_1$.

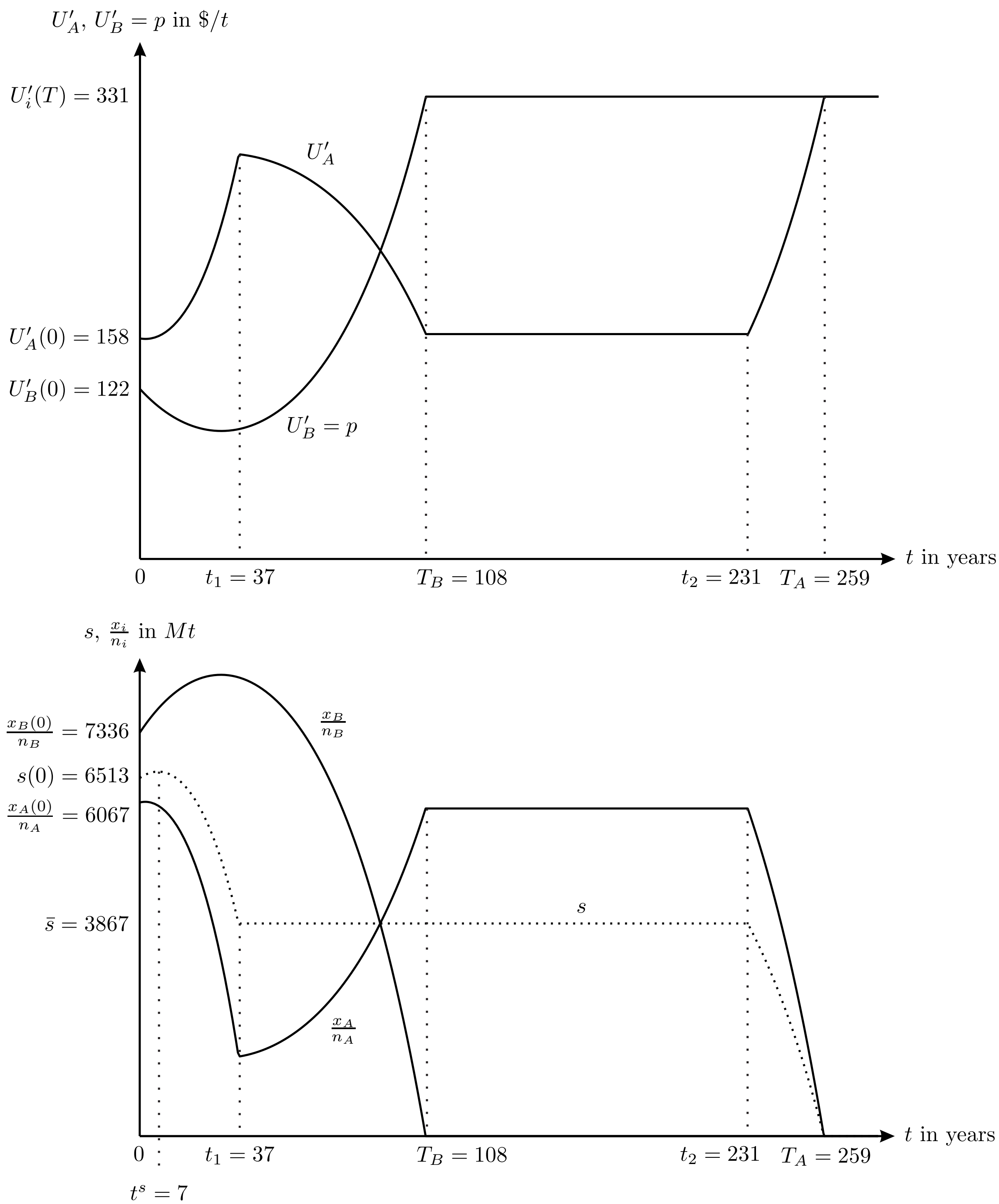

Figure 6: Evolution of U'_A, $U'_B = p$, fuel consumption and extraction over time with strategic demand-side policy

[earlier] than at the social optimum (or without strategic action), and cumulative fuel extraction is lower [higher] during Phase I [III]. If the positive [negative] strategic effects dominate,

the coalition reduces [raises] fuel demand, which reduces [raises] the scarcity rent. In both cases, the initial extraction declines, i.e. the smaller initial fuel consumption in country A [B] outweighs the larger initial fuel consumption in country B [A] in case of SE$>$ [$<$]0.[39]

7.2. Strategic supply-side policy

With the supply-side policy the coalition acts strategically both in the fuel market and deposit market. For the sake of more specific results, we assume[40] that the (absolute) price elasticity of demand, $\epsilon(p) := -\frac{D'(p)p}{D(p)} > 0$, is weakly increasing in the price. The technical details of deriving the unilaterally optimal supply-side policy are delegated to Appendix A.7.

If the coalition's initial budget is unlimited and it applies a maximal deposit acquisition policy, it buys all deposits at $t = 0$. In this case, the coalition internalizes the climate costs of emissions from the outset and exploits its monopoly position on the fuel market subsequently. The corresponding fuel price path is given by

$$p_A(t) = C(t) + \tau_A(t) + \theta_A(t) + \mathrm{ME}(t). \tag{30}$$

In (30), $\mathrm{ME}(t) := -\frac{D_B(t)}{D'(t)} > 0$ reflects the *monopoly effect*, which induces the coalition to reduce its fuel supply in order to increase the fuel price and the coalition's export revenues. This result is recorded in[41]

Proposition 7. *Suppose the coalition applies a supply-side climate policy, acts strategically in the fuel market and deposit market and is committed to its strategy. If the coalition's initial budget is unlimited and deposit acquisitions are maximal,*

(i) the supply-side policy is inefficient.
(ii) the equilibrium is characterized by $s_F(t) = 0$ and $s_A(t) > 0$ for $t \in [0, T)$.
(iii) Propositions 4(iii) and (iv) continue to hold.
(iv) fuel and backstop consumption in both countries are continuous for all points in time.

Next, suppose the coalition's initial budget is limited. Then the coalition does not buy

[39] See Lemma A.11 in Appendix A.6.

[40] The sign of $\epsilon'(p)$ also plays an important role in the literature on monopolistic competition. Bertoletti and Etro (2017) use $\epsilon'(p) > 0$ as standard assumption, and Mrázová and Neary (2017) find empirical evidence supporting $\epsilon'(p) > 0$.

[41] Applying the proof of Lemma A.18 in Appendix A.7 shows that $s^*(j) \leq s_A(j)$ for some j during Phase III implies $s^*(t) < s_A(t)$ for all $t > j$, such that $s^*(j) = s_A(j)$ holds for at most one j during Phase III and the supply-side policy cannot be efficient.

the initial fuel stock at once, and its deposit demand correspondence is given by

$$y_A(t) \begin{cases} \in [0, \bar{y}_A(t)], & \text{if } p_y(t) = \tau_A(t) - \mathrm{SE}e^{\rho t} \\ = \bar{y}_A(t), & \text{if } p_y(t) < \tau_A(t) - \mathrm{SE}e^{\rho t}. \end{cases} \tag{31}$$

During Phase I, the firm's fuel supply is positive in the time interval $[t_a, t_b)$, while the coalition's fuel supply correspondence is given by

$$s_A(t) \in [0, S_A(t)], \text{ if } p(t) = C(t) + \tau_A(t) - \mathrm{SE}e^{\rho t}. \tag{32}$$

In (31) and (32),

$$\mathrm{SE} := \mathrm{BE} + \mathrm{ToT} + \mathrm{EE} \tag{33}$$

represents the coalition's *strategic effects*:

- The *budget effect* $\mathrm{BE} := \zeta_G \left\{ \upsilon_A \tau_{F0} - \frac{\int_{t_a}^{t_b} [C(t)+\tau_F(t)] \left| D'_A(t) + Q'_A(t) \right| \mathrm{d}t}{\left| \int_{t_a}^{t_b} e^{\rho t} D'(t) \, \mathrm{d}t \right|} \right\} \gtreqless 0$ emerges if the coalition's budget constraint binds at some point in time. The budget effect is ambiguous in sign. It induces the coalition to use its deposits acquisitions to weaken the fiscal budget constraint.

- The *terms-of-trade effect* $\mathrm{ToT} := [1 + \zeta_G] \frac{\upsilon_B S(0) - \int_{t_a}^{t_b} D_B(t) \, \mathrm{d}t}{\left| \int_{t_a}^{t_b} e^{\rho t} D'(t) \, \mathrm{d}t \right|} \gtreqless 0$ induces the coalition to reduce its deposit acquisitions if the coalition is a fuel and deposit importer, such that the deposit price and the firm's fuel price and the import costs decline. If the coalition is a fuel exporter, the effect increases deposit acquisitions to increase the fuel price and, therefore, the export revenues.

- The *emission effect* $\mathrm{EE} := \frac{\int_{t_a}^{t_b} \theta_A(t) D'(t) \, \mathrm{d}t}{\left| \int_{t_a}^{t_b} e^{\rho t} D'(t) \, \mathrm{d}t \right|} < 0$ induces the coalition to raise its deposit acquisitions, such that the fuel and deposit prices increase. This reduces the fringe's fuel demand and carbon leakage, which mitigates the climate costs of emissions during Phase I for $s_F > 0$.

Suppose there is no climate problem and the coalition's budget constraint does not bind. Then, only the monopoly effect and the terms-of-trade effect exist. If the coalition is a fuel exporter, both effects work in the same direction, i.e. the terms-of-trade effect leads to higher deposit acquisitions, so that the depletion of the firm's resource stock is antedated and the coalition can longer enjoy its monopoly on the fuel market. If the terms-of-trade effect is sufficiently strong, the coalition's monopoly price can undercut the firm's producer price before T_F, so that there is a smooth transition. In contrast, the effects work in opposite directions if the coalition is a fuel importer. On the one hand, the coalition has an incentive to drive the fuel firm out of the market to enjoy its monopoly position. On the other hand, the terms-of-trade effect induces the coalition to reduce its import costs by reducing its deposit acquisitions, which spares the firm's resource stock.

Suppose next there is a climate problem. If the coalition applies a singular deposit acquisition strategy, its fuel price is given by (30). If the coalition's budget constraint binds,[42] it applies a maximal deposit acquisition strategy and its fuel price reads

$$p_A(t) = C(t) + \mathrm{ME}(t) + \tau_A(t) + \theta_A(t) + \widetilde{\mathrm{BE}}(t), \tag{34}$$

where $\widetilde{\mathrm{BE}}(t) := \frac{\zeta_G}{D'+\zeta_G[D'_B - Q'_A]} \Big\{ [D'_A + Q'_A] \big[C(t) + \mathrm{ME}(t)\big] - [D'_B - Q'_A] \big[\tau_A(t) + \theta_A(t)\big] \Big\}$ is also a budget effect of ambiguous sign. The budget effect composes of two partial effects. The first product in curly brackets induces the coalition to increase its fossil fuel revenues by increasing its fuel supply price. The second product reflects that the coalition can reduce its energy costs by reducing its supply price.

In Appendix A.7 we prove

Proposition 8. *Suppose the coalition applies a supply-side climate policy, acts strategically in the fuel market and deposit market and is committed to its strategy.*

(i) If the coalition's budget constraint binds, the supply-side policy is inefficient and the equilibrium is characterized by $s_F(t) = y_A(t) = 0$ and $s_A(t) > 0$ for $t \in [T_F, T)$. Propositions 4(ii) and (iii) continue to hold, and fuel and backstop consumption are constant during Phase II.

- *If $\theta_{A0} + SE \geq 0$, the equilibrium is characterized by $s_F(t) > 0$ and $s_A(t) = 0$ for $t \leq T_F \leq t_1$, and Propositions 4(iv) and (v) continue to hold.*
- *If $\theta_{A0} + SE < 0$, the equilibrium is characterized by either $s_F(t) > 0$ and $s_A(t) = 0$ or by $s_F(t) = 0$ and $s_A(t) > 0$ for $t \leq T_F$, and fuel and backstop consumption are continuous for $t \neq T_F$.*

(ii) If the coalition's deposit acquisitions are singular, the supply-side policy is inefficient and the equilibrium is characterized by $s_F(t), s_A(t) \geq 0$ for $t \leq t_1$, and by $s_F(t) = 0$, $y_A(t) \geq 0$ and $s_A(t) > 0$ for $t \in [t_1, T)$. Propositions 4(ii) and (iii) continue to hold.

[42] If the budget constraint binds at $t = 0$, it binds in every period until the coalition has purchased the last deposit of the fuel firm.

- *If $\theta_{A0} + SE \geq 0$, the equilibrium is characterized by $s_F(t) > 0$ for $t \leq T_F \leq t_1$, and Propositions 4(iv) and (v) continue to hold.*
- *If $\theta_{A0} + SE < 0$, fuel and total energy [backstop] consumption in both countries are constant during Phase II, and they decline [increase] during Phase III. Fuel and total energy [backstop] consumption in both countries jump downwards [upwards] at $t = T_F$ if the coalition imports fuel and deposits ($n_A \geq \upsilon_A$), and they are continuous for all other points in time.*

Both for the binding budget constraint and the singular deposit acquisition strategy, the consumer price and, thus, the fuel and backstop consumption per capita in both countries coincide for all points in time. Furthermore, there is no private supply at the ceiling to ensure a constant fuel consumption level. Because both the private and governmental price paths are continuous and $p(t) = \min\{p_F(t), p_A(t)\}$ holds for $t < T_F$, the consumer price path can jump at $t = T_F$ but is continuous for all other points in time.

Consider the singular deposit acquisition strategy. If the positive strategic effects dominate (SE > 0), the firm's producer price (14) is always below the coalition's producer price (30), as with the competitive supply-side policy ($p_F(t) < p_A(t)$ for all $t \in [0, T_F)$). Consequently, there is private fuel supply at early points in time and fuel consumption in both countries jumps downwards, while backstop consumption jumps upwards, at exactly the time $T_F \leq t_1$ when the private fuel stock becomes exhausted. Qualitatively, this equilibrium coincides with that of the competitive supply-side policy illustrated in Fig. 5. Quantitatively, the positive monopoly effect depresses the private scarcity rent below its level with the competitive supply-side policy.[43]

Even if the negative strategic effects dominate (SE < 0), the internalization of the climate costs can raise the coalition's producer price above the firm's producer price, at least when the private fuel stock becomes exhausted. This is definitely the case when the coalition imports fuel and deposits ($n_A \geq \upsilon_A$ is sufficient). Due to $p_F(T_F) < p_A(T_F)$, again fuel [backstop] consumption in both countries jumps downwards [upwards] at $T_F \leq t_1$.

If the negative strategic effects are sufficiently strong ($n_A < \upsilon_A$ is necessary), there are time periods in which fuel is only supplied by the firm[44] and there are time periods in

[43]This is the case if the budget constraint does not bind without strategic action, and the private fuel stock is exhausted at the same time or later with strategic action than without.

[44]Otherwise, the negative parts of the strategic effects $-\int_{t_a}^{t_b} D_B(t)\,\mathrm{d}t$ and $\int_{t_a}^{t_b} \theta_A(t) D'(t)\,\mathrm{d}t$ vanish.

which fuel is only supplied by the coalition before the private fuel stock is exhausted. In the former case the consumer price is temporarily determined by the firm's producer price and in the latter case the coalition has the monopoly on fuel and sets the consumer price. For strong negative effects the coalition's fuel supply price may undercut the firm's price ($p_F(T_F) > p_A(T_F)$), only the coalition supplies fuel when the firm's stock gets exhausted, and the fuel price and the extraction paths are smooth, as illustrated in Fig. 7.

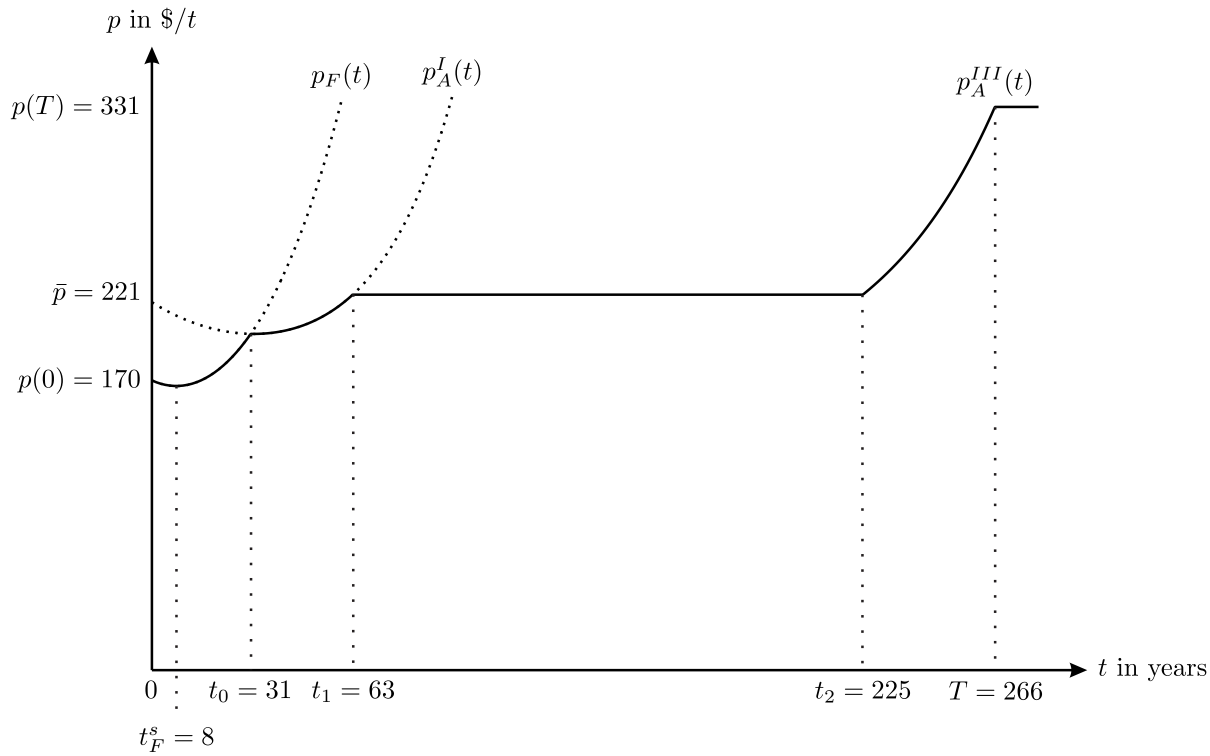

Figure 7: Evolution of p over time with strategic supply-side policy and strong negative strategic effects

Finally, suppose the budget constraint binds. In that case, the budget effects alter the coalition's producer price (34). If $\widetilde{\text{BE}}$ is sufficiently small, the coalition undercuts the firm's producer price temporary during Phase I,[45] and both the fuel price and the fuel extraction paths are smooth. If $\widetilde{\text{BE}}$ is sufficiently large, the coalition does not sell fuel at the firm's producer price, because the coalition values fuel more than the firm. The fuel and extraction paths have a jump and the equilibrium qualitatively coincides with that of the competitive

[45]Lemma A.18 rules out that the coalition undercuts the firm for all t.

supply-side policy.

Similar to the competitive supply-side policy the strategic supply-side policy causes no intra-temporal distortions which prevail in the economy with the demand-side policy. However, both the strategic effects and private fuel supply during Phase I lead to inter-temporal distortions and render the strategic supply-side policy inefficient.

Neither the demand- nor the supply-side climate policy in general implements the social optimum. There are two reasons for this negative result. First, the climate costs of emissions are not fully internalized, either internationally (demand-side regime) or intertemporally (supply-side regime). Second, the coalition may have strategic incentives to manipulate the fuel and deposit prices in its favor.

8. Grand coalition

So far we have assumed that there are two groups of countries, a coalition and a group of free riders. In this subsection, we briefly report on the performance of the grand coalition (a coalition encompassing all countries) with the different policies. In case of both competitive and strategic demand-side policy the grand coalition implements the efficient allocation by choosing the efficient fuel caps in all countries. Since there are no free-riders, there is no possibility to benefit from strategic behavior and the competitive and the strategic demand-side policy coincide.[46] In case of the supply-side policy, the grand coalition also implements the efficient allocation when it acts strategically. The coalition's initial budget constraint does not bind, it purchases all deposits at $t = 0$ with both maximal and singular deposit acquisitions, and the extraction path is smooth and efficient.[47] In contrast, with the competitive supply-side policy the coalition's budget constraint does also not bind, but the coalition only implements the efficient allocation with maximal deposit acquisitions. If deposit acquisitions are singular, the extraction path jumps at $t = T_F$ and the allocation is inefficient.[48]

---

[46] In the competitive [strategic] demand-side regime, the grand coalition sets $\theta_A(t) = \theta(t)$ [and SE $= 0$], which yields $p(t) = C(t) + \tau(t)$ and $U'_A = C(t) + \tau(t) + \theta(t)$ in equilibrium.

[47] In the strategic supply-side regime, the grand coalition sets $T_F = 0$ and $p(t) = C(t) + \tau(t) + \theta(t)$ with singular deposit acquisitions.

[48] In the competitive supply-side policy with singular deposit acquisitions the coalition sets $T_F > 0$. The firm's profit increases when T_F decreases, and the coalition delays T_F until the instantaneous profit of the

9. Coalition formation

In this section, we endogenize coalition formation, i.e. at $t = 0$ countries decide whether to join the climate coalition or stay outside and act as fringe country. Coalition countries conclude a long-term contract. Applying the stability concept of d'Aspremont et al. (1983), a coalition is stable if no coalition member has an incentive to leave the coalition (internal stability) and no fringe country has an incentive to join the coalition (external stability). In Appendix A.8 we prove

Proposition 9. *In the competitive demand-side regime with linear demand functions and quadratic cost functions, any stable coalition is just so large that the ceiling is adhered to.*

In the competitive demand-side regime with quadratic cost functions and linear demand functions, per-capita welfare of any country increases [decreases] if it leaves [joins] the coalition as long as the ceiling is adhered to. As a consequence, any country has an incentive to leave a coalition if the remaining coalition countries can ensure the ceiling. By contrast, no country has an incentive to leave a coalition if the remaining countries cannot ensure the ceiling, because this would lead to prohibitively high climate damages and definite welfare losses. Thus, any stable coalition is just large enough to ensure the ceiling. The grand coalition implements the efficient allocation, but is not stable.

For the competitive supply-side policy and the strategic climate policies, even for linear demand functions and quadratic cost functions we are not able to characterize the size of the stable coalition. Therefore, we now turn to an empirically calibrated economy. The model is calibrated to the world coal market in the year 2015.[49] Following Hassler and Krusell (2012) and Hassler et al. (2021), we divide the world into nine asymmetric regions, consider the five countries with the greatest coal reserves and divide the rest of the world into four regions with comparable coal reserves. The world regions are listed in Table 1.

We begin with the competitive demand-side policy. In line with Proposition 9, any stable coalition is just large enough to ensure the ceiling in the competitive demand-side regime. In the calibrated economy the coalition countries' share of global consumption

fuel firm is so small that it is perceived as worth taking over fossil fuel supply.

[49]The calibration is described in more detail in Appendix B.

Table 1: World regions in 2015 (υ_i, n_i in %; $\bar{K}_i$ in \$trillion).

	υ_i	n_i	$\bar{K}_i$
USA	22.4	9.4	18.2
Russia (Rus)	15.6	2.7	1.4
Australia (Aus)	14.0	1.5	1.4
China (Chi)	13.0	51.3	11.1
India (Ind)	9.2	10.3	2.1
EU	7.2	8.4	13.6
Rest of Asia (ROA)	6.6	9.6	12.2
Rest of Europe (ROE)	5.7	3.0	5.2
Rest of the World (ROW)	6.3	3.8	9.3

Note: υ_i is region i's share of global coal reserves, n_i is region i's share of global coal consumption, and $\bar{K}_i$ is region i's exogenous income, i.e. its GDP.

to adhere the ceiling is $n_A \geq 0.651$. Closer inspection of Table 1 reveals that there are multiple stable coalitions and any stable coalition comprises three or four world regions including China. The smallest stable coalition consists of China, Rest of Asia, Russia and Australia ($n_A = 0.651$), and the largest stable coalition consists of China, India and USA ($n_A = 0.710$).[50] In the smallest [largest] stable coalition, global energy welfare[51] amounts to 70.5\$trillion [72.3\$trillion], each coalition country's per-capita energy welfare net of its fuel firm's profit share amounts to 58.2\$trillion [63.5\$trillion], and each fringe country's per-capita energy welfare net of its fuel firm's profit share amounts to 91.7\$trillion [91.7\$trillion].[52]

Table 2 lists the grand coalition (line 2) and[53] coalitions consisting of eight regions for the

[50]The 24 stable coalitions are {Chi + USA + Ind}, {Chi + USA + EU}, {Chi + USA + ROA}, {Chi + Ind + EU}, {Chi + Ind + ROA}, {Chi + Ind + ROW}, {Chi + EU + ROA}, {Chi + USA + Rus + ROE}, {Chi + USA + Rus + ROW}, {Chi + USA + Aus + ROE}, {Chi + USA + Aus + ROW}, {Chi + USA + ROE + ROW}, {Chi + Ind + Rus + Aus}, {Chi + Ind + Rus + ROE}, {Chi + Ind + Aus + ROE}, {Chi + EU + Rus + ROE}, {Chi + EU + Rus + ROW}, {Chi + EU + ROE + ROW}, {Chi + ROA + Rus + Aus}, {Chi + ROA + Rus + ROE}, {Chi + ROA + Rus + ROW}, {Chi + ROA + Aus + ROE}, {Chi + ROA + Aus + ROW}, {Chi + ROA + ROE + ROW}.

[51]Here, we report the energy welfare of country $i \in N$, which is defined as welfare net of exogenous income $W_i|_{\bar{K}_i=0} = W_i - \int_0^\infty e^{-\rho t}\bar{K}_i \,\mathrm{d}t$, instead of welfare, because the welfare related to energy consumption and production only amounts to 3% of welfare and the difference between countries becomes more visible when considering energy welfare. Global energy welfare is then defined by $W|_{\bar{K}=0} = \sum_{i\in N} W_i|_{\bar{K}_i=0}$.

[52]An increase in the coalition size raises global welfare and per-capita welfare of the coalition countries, and it does not affect per-capita welfare of the fringe countries. See Lemma A.21 in Appendix A.8.

[53]In competitive supply-side regime underlying Table 2 we have assumed that the equilibrium with singular deposit acquisition sets in. However, the results do not change qualitatively when the equilibrium with maximal deposit acquisition sets in. For more details we refer to Table B.6 in Appendix B.1.

Table 2: Large coalitions with the competitive supply-side policy (υ_A, n_A in %; τ_{F0}, θ_{A0} in \$/t; $W|_{\bar{K}=0}, \Delta\frac{W_i}{n_i}$ in \$trillion).

	υ_A	n_A	t_1	t_2	T	τ_{F0}	θ_{A0}	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$
WORLD	100	100	33	237	253	0.59	47.9	77.7	–
WORLD-Aus	86.0	98.5	34	237	253	0.59	46.8	77.8	16.2
WORLD-Rus	84.4	97.3	34	237	253	0.59	46.5	77.8	18.4
WORLD-ROE	94.3	97.0	34	237	253	0.59	46.8	77.8	20.7
WORLD-ROW	93.7	96.2	34	237	253	0.59	46.6	77.8	20.9
WORLD-EU	92.8	91.6	35	237	253	0.59	45.8	77.8	21.6
WORLD-USA	77.6	90.6	35	237	254	0.58	45.1	77.9	21.0
WORLD-ROA	93.4	90.4	35	237	253	0.59	45.6	77.8	21.8
WORLD-Ind	90.8	89.7	35	237	253	0.58	45.4	77.9	21.7
WORLD-Chi	87.0	48.7	36	238	254	0.58	43.4	78.0	23.0

Note: υ_A is the coalition's share of global coal reserves, n_A is the coalition's share of global coal consumption, t_1, t_2 and T is the end of Phase I, II and III, respectively, τ_{F0} is the initial private scarcity rent, θ_{A0} is the coalition's initial cost of emissions, $W|_{\bar{K}=0}$ is global energy welfare, and $\Delta\frac{W_i}{n_i}$ is the fringe country's increase in per-capita welfare by joining the coalition.

competitive supply-side regime (line 3 - 11) and shows that the grand coalition is stable.[54] In Table 2, $\Delta\frac{W_i}{n_i}$ is the increase in per-capita welfare of the respective fringe country by joining the coalition. The grand coalition does not purchase all deposits at $t = 0$ but successively up to the time $T_F = 9$, and thus does not implement the first-best solution. For coalitions of eight countries, the private fuel stock becomes exhausted earlier at $T_F \in [3, 8]$, which shortens the period of private supply and alleviates the climate costs of emissions.[55] This decreases the magnitude of the extraction jump and increases global welfare. More importantly, it is always beneficial to join the grand coalition since $\Delta\frac{W_i}{n_i} > 0$ holds. The main reason why a country looses welfare when it leaves the grand coalition is that a free rider receives no revenues from the coalition's fuel sales. Consequently, the grand coalition is stable and coalitions comprising eight regions are unstable. In the grand coalition, global energy welfare amounts to 77.7\$trillion, and each coalition country's per-capita energy welfare net of its fuel firm's

[54]For coalitions of eight world regions, the budget constraint does not bind in the competitive supply-side regime. In particular, any of these coalitions could purchase the entire private fuel stock within a year, whereas the optimal exhaustion date of the private fuel stock is $T_F \geq 3$.

[55]The smaller the coalition, the earlier the private fuel stock becomes exhausted, because the difference in consumption welfare between private and governmental supply becomes less relevant ($n_A \downarrow$) and the profit of the fuel firm becomes less relevant ($\upsilon_A \downarrow$).

Table 3: Stable coalitions with the strategic demand-side policy (υ_A, n_A in %; τ_{F0}, θ_{A0}, SE, ToT in \$/t; $W|_{\bar{K}=0}$ in \$trillion).

	υ_A	n_A	t_1	t_2	T_B	T_A	τ_{F0}	θ_{A0}	SE	ToT	$W\|_{\bar{K}=0}$
Chi+USA+Rus+Aus	65.0	64.9	37	231	108	259	21.6	57.3	−21.1	148	75.7
Chi+ROW+Rus+Aus	48.9	59.3	37	228	119	261	16.6	63.7	−16.1	219	74.3
Chi+ROE+Rus+Aus	48.3	58.5	37	227	118	261	16.9	64.3	−16.4	218	74.2
Chi+ROW+ROE+Rus	40.6	60.8	37	229	132	261	12.2	64.5	−11.7	282	73.6
Chi+ROW+ROE+Aus	39.0	59.6	37	228	131	261	12.4	65.5	−11.9	283	73.4
Chi+EU+Aus	34.2	61.2	37	229	140	260	9.88	65.1	−9.39	327	73.1
Chi+EU+ROW	26.5	63.5	37	230	156	260	6.71	64.3	−6.21	400	72.7
Chi+EU+ROE	25.9	62.7	37	230	154	260	7.00	64.9	−6.50	396	72.6
Chi+ROA	24.6	60.9	37	229	156	260	6.66	66.9	−6.17	419	72.0
Chi+Ind	22.2	61.6	37	229	155	260	6.81	66.1	−6.31	409	72.3

Note: υ_A is the coalition's share of global coal reserves, n_A is the coalition's share of global coal consumption, t_1, t_2 and T_i is the end of Phase I, II and III for group $i = A, B$, respectively, τ_{F0} is the initial private scarcity rent, θ_{A0} is the coalition's initial cost of emissions, SE is the strategic effect, ToT is the terms-of-trade effect, and $W|_{\bar{K}=0}$ is global energy welfare.

profit share amounts to 77.1\$trillion.[56] Comparing the welfare levels of coalition countries shows that the countries' welfare levels increase when moving from the stable equilibrium of the demand-side regime to the stable equilibrium of the supply-side regime. We summarize our results in

Proposition 10. *Suppose the climate coalition is a price taker. In the calibrated economy*

(i) the grand coalition is stable in the supply-side regime, while the stable coalition comprises only three or four world regions including China with 65.1% *to* 71% *of the global energy demand in the demand-side regime,*

(ii) the per-capita welfare of each coalition country is higher in the stable equilibrium of the supply-side regime than in the stable equilibrium of the demand-side regime,

(iii) the global welfare is higher in the stable equilibrium of the supply-side regime than in the stable equilibrium of the demand-side regime.

Table 3 provides information about the stable coalitions in the strategic demand-side regime. As in the competitive demand-side regime in the strategic demand-side regime there are multiple stable coalitions and stable coalitions are small. In view of Table 3 the stable coalition comprises two, three or four world regions including China. The smallest stable

[56] Note that the scarcity rent and, thus, the per-capita profit of the fuel firm is smaller in the demand-side regime than in the supply-side regime. In particular, $\tau_{F0}S(0) = 592$\$billion in the demand-side regime and $\tau_{F0}S(0) = 608$\$billion in the supply-side regime.

Table 4: Large coalitions with the strategic supply-side policy (υ_A, n_A in %; $\tau_{F0}, \theta_{A0}, \mathrm{SE}, \mathrm{ME}(0)$ in \$/t; $W|_{\bar{K}=0}, \Delta\frac{W_i}{n_i}$ in \$trillion).

	υ_A	n_A	t_1	t_2	T	τ_{F0}	θ_{A0}	SE	ME(0)	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$
WORLD	100	100	37	238	254	0.57	42.4	–	–	78.11	–
WORLD-Aus	86.0	98.5	37	237	254	0.40	41.4	0.18	2.79	78.10	25.8
WORLD-Rus	84.4	97.3	38	238	255	0.38	40.5	0.19	4.99	78.10	25.3
WORLD-ROE	94.3	97.0	38	238	255	0.48	40.3	0.09	5.53	78.10	24.3
WORLD-ROW	93.7	96.2	38	238	255	0.47	39.8	0.10	6.97	78.10	24.3
WORLD-EU	92.8	91.6	39	237	255	0.45	36.8	0.10	15.0	78.08	24.6
WORLD-USA	77.6	90.6	40	237	255	0.32	36.2	0.23	16.7	78.08	25.1
WORLD-ROA	93.4	90.4	40	237	255	0.46	36.1	0.10	17.0	78.08	24.6
WORLD-Ind	90.8	89.7	40	237	256	0.43	35.6	0.13	18.2	78.07	24.7
WORLD-Chi	87.0	48.7	55	234	262	0.34	16.0	0.14	72.6	77.37	26.9

Note: υ_A is the coalition's share of global coal reserves, n_A denotes the coalition's share of global coal consumption, t_1, t_2 and T is the end of Phase I, II and III, respectively, τ_{F0} is the initial private scarcity rent, θ_{A0} is the coalition's initial cost of emissions, SE is the strategic effect, ME(0) is the initial monopoly effect, $W|_{\bar{K}=0}$ is global energy welfare, and $\Delta\frac{W_i}{n_i}$ is the fringe country's increase in per-capita welfare by joining the coalition.

coalition consists of China, ROE, Russia and Australia ($n_A = 0.585$), and the largest stable coalition consists of China, USA, Russia and Australia ($n_A = 0.649$). These stable coalitions are slightly smaller than the smallest stable coalition in the competitive demand-side regime. Strategic action enables coalition countries to increase their welfare. More importantly, it enables the coalition to increase the fuel firm's scarcity rent, which reduces the carbon leakage and makes it easier to adhere to the ceiling.[57]

Table 4 provides information about the grand coalition and coalitions of eight regions in the strategic supply-side regime. Common to the competitive supply-side regime and the strategic supply-side regime is that the grand coalition is stable, but in contrast to the competitive supply-side regime in the strategic supply-side regime the grand coalition buys all deposits at $t = 0$ and implements the first-best solution. For coalitions of eight countries the fiscal budget constraint is binding and the private fuel stock becomes exhausted later. The strategic effects are positive, and in particular the monopoly effect leads to large profits from fuel sales in coalition countries at the fringe country's expense. More specifically, joining

[57]The coalition consisting of China and USA could adhere the ceiling, but Russia and Australia are better off joining this coalition to raise the scarcity rent and, thus, their export revenues.

Table 5: Per-capita energy welfare in \$trillion in the stable coalitions with the strategic climate policies.

Region \ v_A	100	65.0	48.9	48.3	40.6	39.0	34.2	26.5	25.9	24.6	22.2
USA	**78.9**	**98.4**	113	114	106	106	103	98.5	98.9	98.5	98.7
Russia	**80.9**	**174**	**146**	**147**	**124**	150	137	122	123	122	122
Australia	**83.0**	**253**	**206**	**208**	193	**170**	**148**	147	149	146	147
China	**77.7**	**51.0**	**51.1**	**50.4**	**54.5**	**53.5**	**55.6**	**58.1**	**57.5**	**56.0**	**56.6**
India	**78.0**	88.4	87.6	87.6	87.4	87.4	87.6	88.2	88.2	88.3	**61.1**
EU	**78.0**	87.6	86.9	87.0	86.9	86.9	**61.2**	**62.3**	**61.8**	88.0	88.0
Rest of Asia	**77.9**	83.8	84.0	84.0	84.8	84.8	85.5	86.8	86.7	**59.0**	86.8
Rest of Europe	**78.6**	111	105	**79.1**	**75.1**	**74.5**	97.8	95.2	**69.4**	95.2	95.3
Rest of the World	**78.5**	105	**75.1**	101	**72.1**	**71.4**	95.4	**67.9**	93.7	93.5	93.6

Note: Bold numbers refer to coalition countries.

the coalition of eight countries raises the per-capita welfare of the respective fringe country by about one third of the global per-capita welfare which proves the stability of the grand coalition.

Finally, Table 5 compares the welfare of coalition countries in the stable coalition of the strategic supply-side regime with the coalition countries' welfare in stable coalitions of the strategic demand-side regime. For that purpose, Table 5 lists country's per-capita energy welfare. The column with $v_A = 100$ belongs to the (stable) grand coalition in the strategic supply-side regime. All other columns belong to stable coalitions in the strategic demand-side regimes. The welfare levels of the countries which are in the stable coalition are highlighted in bold letters. Closer inspection of Table 5 reveals that USA, Russia, Australia are always better off in the demand-side regime, China is always better off in the supply-side regime, and India, EU, Rest of Asia, Rest of World are better off in the supply-side regime if and only if they participate in the demand-side coalition.[58] We summarize our results in

Proposition 11. *Suppose that any climate coalition acts strategically and is committed to its strategy. In the calibrated economy*

(i) the grand coalition is stable in the supply-side regime, while the stable coalition com-

[58] The latter also holds for Rest of Europe with the exception that it is better off in the demand-side regime if it participates in the {Chi+ROE+Rus+Aus} coalition. Combining Table 5 with the population shares n_i of Table 1 it turns out that both the average and the median per-capita welfare of coalition countries increase when moving from the stable equilibrium of the demand-side regime to the stable equilibrium of the supply-side regime.

prises two, three or four world regions including China with 58.5% *to* 64.9% *of the global energy demand in the demand-side regime,*

(ii) the average and median per-capita welfare of the coalition countries is higher in the stable equilibrium of the supply-side regime than in the stable equilibrium of the demand-side regime,

(iii) the global welfare is higher in the stable equilibrium of the supply-side regime than in the stable equilibrium of the demand-side regime.

Propositions 10 and 11 have a striking implication. Suppose that any stable coalition has formed in the competitive or strategic demand-side regime, and that the respective coalition countries can propose (consumer good) transfers to change the policy and to establish the stable grand coalition of the competitive or strategic supply-side regime. Since global [each fringe country's] welfare is smaller [greater] in any stable equilibrium with demand-side policy than in the stable equilibrium with supply-side policy,[59] there is a transfer scheme for any stable equilibrium with demand-side policy which ensures that <u>all</u> countries benefit from the transition to the stable grand coalition of the supply-side regime.[60] This grand coalition implements the efficient allocation if the climate coalition acts strategically in the fuel market and deposit market.

10. Conclusion

This paper compares the fuel price paths, fuel consumption paths and welfare levels in a dynamic Hotelling model with unilateral demand- and supply-side climate policies. In our model, the climate coalition ensures that a ceiling on the carbon dioxide concentration is not violated either by limiting domestic fuel consumption or by buying fuel deposits to postpone their extraction. In the case of unilateral demand-side policy, the fringe does not fully internalize the climate costs of emissions. The climate coalition must reduce domestic fuel consumption to the benefit of the fringe whose consumption is inefficiently high until the ceiling is no longer binding. The unilateral demand-side policy distorts the intra-temporal allocation.

The unilateral supply-side policy is also inefficient under the realistic assumption that

[59]See Propositions 10(iii) and 11(iii), page 38, and Table 5.

[60]That is, $\sum_{i\in A}\left(W_i^{supply} - W_i^{demand}\right) > \sum_{i\in B}\left(\underbrace{W_i^{demand} - W_i^{supply}}_{>0}\right) \Longleftrightarrow W^{supply} > W^{demand}$.

the coalition's budget is limited and the coalition cannot purchase all deposits at the outset, since the fuel firm does not internalize the climate costs of emissions. Before the ceiling becomes binding, the coalition has purchased the entire fuel stock and the coalition is from this time on the sole supplier of fuel. When the ceiling becomes binding, the fuel firm does not own any deposits and the coalition restricts its fuel extraction so as to guarantee the ceiling. If the coalition is a price taker, both the fuel price and the fuel consumption paths are discontinuous when the private fuel stock is exhausted. If the coalition acts strategically, depending on the strength of the strategic effects the fuel price and the fuel consumption paths can be continuous or discontinuous. The unilateral supply-side policy distorts the inter-temporal allocation and is inefficient.

Endogenizing coalition formation and comparing the demand-side policy with the supply-side policy in an empirically calibrated economy, we find that the stable coalition is medium-sized and just as large to adhere the ceiling with the demand-side policy. In contrast, with the supply-side policy the grand coalition forms and is stable. In case of strategic supply-side policy, the coalition purchases all deposits from the outset, the fuel price and fuel extraction paths are continuous and the (stable) grand coalition implements the efficient allocation.

The policy conclusion of our paper is a recommendation in favor of the supply-side policy. The global welfare is higher in the stable equilibrium with supply-side policy than in the stable equilibrium with demand-side policy. In particular, the climate ambitious European Union should switch from the demand-side policy to the supply-side policy, or augment the demand-side policy with the supply-side policy, and begin with purchasing fuel deposits. In this way, it sets incentives for other countries to participate in an international climate agreement and reduces the inefficiency of the climate treaty.

In the present paper, we have assumed that the fringe does not pursue any policy. Introducing a fuel cap in the fringe or regulating the deposit-demand of fuel producers in the fringe would not change the results as long as the fringe behaves as price-taker. If the fringe acts strategically, we expect that it improves and the coalition worsens in terms of welfare. With the demand-side policy the fringe can increase its welfare lead over the coalition. With supply-side policy, the fringe can restrain selling deposits or can jack up the deposit price and reduce the welfare gap to the coalition or overtake the coalition. However, we also assumed that the fringe has no climate goals and climate policy. Introducing such goals would redistribute the climate mitigation burden from the coalition to the fringe with both policies.

Our analysis can be extended in various directions. First, one could replace the CO_2 ceiling by a climate-damage function to determine the robustness of the results. Second, it may be important in the future to check the robustness of our results when deposits are heterogeneous. Third, one could use subgame-perfect strategies to analyze the strategic behavior of the coalition.

References

Amigues, J.P., Lafforgue, G., Moreaux, M., 2014. Optimal timing of CCS policies with heterogeneous energy consumption sectors. Environmental and Resource Economics 57, 345–366.

Amigues, J.P., Moreaux, M., Schubert, K., 2011. Optimal use of a polluting non-renewable resource generating both manageable and catastrophic damages. Annals of Economics and Statistics/Annales d'Économie et de Statistique 103/104, 107–141.

Barrett, S., 2003. Environment and statecraft: The strategy of environmental treaty-making. Oxford University Press.

Battaglini, M., Harstad, B., 2016. Participation and duration of environmental agreements. Journal of Political Economy 124, 160–204.

Benchekroun, H., Halsema, A., Withagen, C., 2009. On nonrenewable resource oligopolies: The asymmetric case. Journal of Economic Dynamics and Control 33, 1867–1879.

Benchekroun, H., Halsema, A., Withagen, C., 2010. When additional resource stocks reduce welfare. Journal of Environmental Economics and Management 59, 109–114.

Bergstrom, T.C., 1982. On capturing oil rents with a national excise tax. The American Economic Review 72, 194–201.

Bertoletti, P., Etro, F., 2017. Monopolistic competition when income matters. The Economic Journal 127, 1217–1243.

Bloch, H., Rafiq, S., Salim, R., 2015. Economic growth with coal, oil and renewable energy consumption in China: Prospects for fuel substitution. Economic Modelling 44, 104–115.

Bohm, P., 1993. Incomplete international cooperation to reduce CO_2 emissions: Alternative policies. Journal of Environmental Economics and Management 24, 258–271.

BP, 2021. BP statistical review of world energy 2021. Online at: https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

Bretschger, L., Suphaphiphat, N., 2014. Effective climate policies in a dynamic North–South model. European Economic Review 69, 59–77.

Burke, P.J., Liao, H., 2015. Is the price elasticity of demand for coal in China increasing? China Economic Review 36, 309–322.

Chakravorty, U., Leach, A., Moreaux, M., 2012. Cycles in nonrenewable resource prices with pollution and learning-by-doing. Journal of Economic Dynamics and Control 36, 1448–1461.

Chakravorty, U., Magné, B., Moreaux, M., 2006. A Hotelling model with a ceiling on the stock of pollution. Journal of Economic Dynamics and Control 30, 2875–2904.

Chakravorty, U., Moreaux, M., Tidball, M., 2008. Ordering the extraction of polluting nonrenewable resources. The American Economic Review 98, 1128–1144.

Chiang, A., 1999. Elements of dynamic optimization. Waveland Press.

d'Aspremont, C., Jacquemin, A., Gabszewicz, J.J., Weymark, J.A., 1983. On the stability of collusive price leadership. Canadian Journal of economics , 17–25.

Department of Commerce, 2021. Trends in atmospheric carbon dioxide. Online at: https://www.esrl.noaa.gov/gmd/ccgg/trends/gl_data.html.

Dullieux, R., Ragot, L., Schubert, K., 2011. Carbon tax and OPEC's rents under a ceiling constraint. The Scandinavian Journal of Economics 113, 798–824.

EIA, 2020a. International energy statistics: Coal and coke consumption. Online at: https://www.eia.gov/international/data/world/coal-and-coke/coal-and-coke-consumption.

EIA, 2020b. More coal and coke data. Online at: https://www.eia.gov/international/data/world/coal-and-coke/more-coal-and-coke-data.

Eichner, T., Pethig, R., 2017a. Buy coal and act strategically on the fuel market. European Economic Review 99, 77–92.

Eichner, T., Pethig, R., 2017b. Trade in fossil fuel deposits for preservation and strategic action. Journal of Public Economics 147, 50–61.

Fæhn, T., Hagem, C., Lindholt, L., Mæland, S., Rosendahl, K.E., 2017. Climate policies in a fossil fuel producing country–demand versus supply side policies. The Energy Journal 38, 77–102.

Feichtinger, G., Hartl, R., 1986. Optimale Kontrolle ökonomischer Prozesse: Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften. Walter de Gruyter.

Fell, H., Kaffine, D., Steinberg, D., 2017. Energy efficiency and emissions intensity standards. Journal of the Association of Environmental and Resource Economists 4, 201–226.

Finus, M., 2001. Game theory and international environmental cooperation. Edward Elgar Publishing.

Fischer, C., Salant, S.W., 2017. Balancing the carbon budget for oil: The distributive effects of alternative policies. European Economic Review 99, 191–215.

Golombek, R., Hagem, C., Hoel, M., 1995. Efficient incomplete international climate agreements. Resource and Energy Economics 17, 25–46.

Grafton, Q., Kompas, T., Van Long, N., et al., 2012. Substitution between biofuels and fossil fuels: Is there a Green Paradox? Journal of Environmental Economics and Management 64, 328–341.

Graßl, H., Kokott, J., Kulessa, M., Luther, J., Nuscheler, F., Sauerborn, R., Schellnhuber, H.J., Schubert, R., Schulze, E.D., 2003. Climate protection strategies for the 21st century: Kyoto and beyond. WBGU Berlin.

Hagem, C., Storrøsten, H.B., 2019. Supply-versus demand-side policies in the presence of carbon leakage and the green paradox. The Scandinavian Journal of Economics 121, 379–406.

Hang, L., Tu, M., 2007. The impacts of energy prices on energy intensity: Evidence from China. Energy Policy 35, 2978–2988.

Harstad, B., 2012. Buy coal! A case for supply-side environmental policy. Journal of Political Economy 120, 77–115.

Hassler, J., Krusell, P., 2012. Economics and climate change: Integrated assessment in a multi-region world. Journal of the European Economic Association 10, 974–1000.

Hassler, J., Krusell, P., Olovsson, C., 2021. Presidential Address 2020 Suboptimal Climate Policy. Journal of the European Economic Association 19, 2895–2928.

Hémous, D., 2016. The dynamic impact of unilateral environmental policies. Journal of International Economics 103, 80–95.

Henriet, F., 2012. Optimal extraction of a polluting nonrenewable resource with R&D toward a clean backstop technology. Journal of Public Economic Theory 14, 311–347.

Hoel, M., 1994. Efficient climate policy in the presence of free riders. Journal of Environmental Economics and Management 27, 259–274.

Hoel, M., 2011. The supply side of CO_2 with country heterogeneity. The Scandinavian Journal of Economics 113, 846–865.

Hoel, M., 2020. The rise and fall of bioenergy. Journal of Environmental Economics and Management 101, 102314.

IEA, 2013. Resources to reserves. Online at: `https://iea.blob.core.windows.net/assets/afc6bec5-22e8-4105-a1b3-ceb89eaec1e9/Resources2013.pdf`.

IPCC, 2006. 2006 IPCC guidelines for national greenhouse gas inventories. Online at: `https://www.ipcc-nggip.iges.or.jp/public/2006gl/`.

IPCC, 2013. Climate change 2013: The physical science basis. Online at: `https://www.ipcc.ch/report/ar5/wg1`.

IPCC, 2018. Global warming of 1.5°C. Online at: `https://www.ipcc.ch/site/assets/uploads/sites/2/2019/02/SR15_Chapter2_Low_Res.pdf`.

Joos, F., Roth, R., Fuglestvedt, J.S., Peters, G.P., Enting, I.G., Bloh, W.v., Brovkin, V., Burke, E.J., Eby, M., Edwards, N.R., Friedrich, T., Frölicher, T.L., Halloran, P.R., Holden, P.B., Jones, C., Kleinen, T., Mackenzie, F.T., Mastsumoto, K., Meinshausen, M., Plattner, G., Reisinger, A., Segschneider, J., Shaffer, G., Steinacher, M., Strassmann, K., Tanaka, K., Timmermann, A., Weaver, A.J., 2013. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: A multi-model analysis. Atmospheric Chemistry and Physics 13, 2793–2825.

Kamien, M., Schwartz, N., 2012. Dynamic optimization: The calculus of variations and optimal control in economics and management. 2 ed., Elsevier Science.

Karp, L., 1984. Optimality and consistency in a differential game with non-renewable resources. Journal of Economic Dynamics and Control 8, 73–97.

Karp, L., Newbery, D.M., 1992. Dynamically consistent oil import tariffs. Canadian Journal of Economics 25, 1–21.

Kollenbach, G., 2015a. Abatement, R&D and growth with a pollution ceiling. Journal of Economic Dynamics and Control 54, 1–16.

Kollenbach, G., 2015b. Endogenous growth with a ceiling on the stock of pollution. Environmental and Resource Economics 62, 615–635.

Kollenbach, G., 2019. Unilateral climate policy and the green paradox: Extraction costs matter. Canadian Journal of Economics 52, 1036–1083.

Kováč, E., Schmidt, R.C., 2021. A simple dynamic climate cooperation model. Journal of Public Economics 194, 104329.

Lafforgue, G., Magné, B., Moreaux, M., 2008. Energy substitutions, climate change and carbon sinks. Ecological Economics 67, 589–597.

Lafforgue, G., Magné, B., Moreaux, M., 2009. Optimal sequestration policy with a ceiling on the stock of carbon in the atmosphere, in: Guesnerie, R., Tulkens, H. (Eds.), The design of climate policy. MIT Press, pp. 273–306.

Lewis, T.R., Schmalensee, R., 1980. On oligopolistic markets for nonrenewable natural resources. The Quarterly Journal of Economics 95, 475–491.

Ma, H., Oxley, L., Gibson, J., Kim, B., 2008. China's energy economy: Technical change, factor demand and interfactor/interfuel substitution. Energy Economics 30, 2167–2183.

Maskin, E., Newbery, D., 1990. Disadvantageous oil tariffs and dynamic consistency. The American Economic Review 80, 143–156.

McKinsey & Company, 2020. Has global mining productivity reversed course? Online at: `https://www.mckinsey.com/industries/metals-and-mining/our-insights/has-global-mining-productivity-reversed-course`.

van der Meijden, G., Ryszka, K., Withagen, C., 2018. Double limit pricing. Journal of Environmental Economics and Management 89, 153–167.

Mrázová, M., Neary, J.P., 2017. Not so demanding: Demand structure and firm behavior. The American Economic Review 107, 3835–3874.

Nordhaus, W.D., 2017. DICE 2016R2. Online at: `https://williamnordhaus.com/dicerice-models`.

Rogelj, J., Hare, W., Lowe, J., Van Vuuren, D.P., Riahi, K., Matthews, B., Hanaoka, T., Jiang, K., Meinshausen, M., 2011. Emission pathways consistent with a 2° c global temperature limit. Nature Climate Change 1, 413–418.

Rubio, S.J., Escriche, L., 2001. Strategic pigouvian taxation, stock externalities and polluting non-renewable resources. Journal of Public Economics 79, 297–313.

Ryszka, K., Withagen, C., 2016. Unilateral Climate Policies: Incentives and Effects. Environmental and Resource Economics 63, 471–504.

Seierstad, A., Sydsaeter, K., 1987. Optimal control theory with economic applications. North-Holland Amsterdam.

Serletis, A., Timilsina, G., Vasetsky, O., 2011. International evidence on aggregate short-run and long-run interfuel substitution. Energy Economics 33, 209–216.

Serletis, A., Timilsina, G.R., Vasetsky, O., 2010. Interfuel substitution in the United States. Energy Economics 32, 737–745.

Smulders, S., Van der Werf, E., 2008. Climate policy and the optimal extraction of high-and low-carbon fossil fuels. Canadian Journal of Economics 41, 1421–1444.

Tahvonen, O., 1996. Trade with polluting nonrenewable resources. Journal of Environmental Economics and Management 30, 1–17.

The World Bank, 2022. Gross domestic product. Online at: `https://data.worldbank.org/indicator/NY.GDP.MKTP.CD`.

Tsur, Y., Zemel, A., 2009. Endogenous discounting and climate policy. Environmental and Resource Economics 44, 507–520.

UN, 1992. United nations framework convention on climate change. Online at: `https://unfccc.int/resource/docs/convkp/conveng.pdf`.

UN, 2015. Paris agreement. Online at: `http://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf`.

UN, 2020. The emissions gap report. Online at: `https://wedocs.unep.org/xmlui/bitstream/handle/20.500.11822/34426/EGR20.pdf`.

Van Long, N., 2011. Dynamic games in the economics of natural resources: A survey. Dynamic Games and Applications 1, 115–148.

Wirl, F., 1994. Pigouvian taxation of energy for flow and stock externalities and strategic, noncompetitive energy pricing. Journal of Environmental Economics and Management 26, 1–18.

Wirl, F., Dockner, E., 1995. Leviathan governments and carbon taxes: Costs and potential benefits. European Economic Review 39, 1215–1236.

A. Appendix

A.1. Microfoundation

The model and social optimum: Let the utility of country $i = A, B$ at time t be given by the quasi-linear function

$$U_i\left(x_i(t) + q_i(t)\right) + g_i(t). \tag{A.1}$$

Denoting by ℓ a composite production factor, say land or labor, in the economy there are the (inverse) production functions

$$\ell_{gi}(t) = g_i^s(t), \tag{A.2}$$

$$\ell_{qi}(t) = M_i(q_i(t)), \tag{A.3}$$

$$\ell_{si}(t) = v_i C(t) s(t), \tag{A.4}$$

where $\ell_{gi}(t)$ is the input in the consumer good production of country i, $\ell_{qi}(t)$ is the input in the backstop generation of country i and $\ell_{si}(t)$ is the input in the fuel extraction in country i. The resource constraints are

$$\ell_{gi}(t) + \ell_{qi}(t) + \ell_{si}(t) = \bar{K}_i, \tag{A.5}$$

$$g_A(t) + g_B(t) = g_A^s(t) + g_B^s(t), \tag{A.6}$$

where $\bar{K}_i$ is country i's endowment of land. Land is immobile, whereas the consumption good is mobile. Inserting (A.2)-(A.6) into (A.1) yields the total welfare

(A.7)
$$\sum_i \left[U_i\left(x_i(t) + q_i(t)\right) + g_i(t)\right] = \sum_i \left[U_i\left(x_i(t) + q_i(t)\right) + \bar{K}_i - M_i(q_i(t)) - v_i C(t) s(t)\right].$$

The laissez-faire economy: Labor is numeraire. The price of the consumer good is also unity because of the linear production function (A.2). The firms' profits are

$$\Pi_{gi}(t) = g_i^s(t) - \ell_{gi}(t) = 0, \tag{A.8}$$

$$\Pi_i(t) = p_{q_i}(t) q_i(t) - \ell_{qi}(t) = p_{q_i}(t) q_i(t) - M_i(q_i(t)), \tag{A.9}$$

$$\begin{aligned} v_i \Pi_F(t) &= v_i[p(t) s_F(t) + p_y(t) y_F(t)] - \ell_{si}(t) \\ &= v_i\left[p(t) s_F(t) + p_y(t) y_F(t) - C(t) s_F(t)\right]. \end{aligned} \tag{A.10}$$

The consumer of country i maximizes (A.1) subject to the budget constraint

$$p(t)x_i(t) + p_{q_i}(t)q_i(t) + g_i(t) = \bar{K}_i + \underbrace{\Pi_{gi}(t)}_{=0} + \Pi_i(t) + v_i\Pi_F(t). \tag{A.11}$$

Demand-side policy: Derivation of country A's welfare: Solving for $g_A(t)$ and inserting the budget constraint (A.11) for $i = A$ into (A.1) yields the welfare

$$U_A\left(x_A(t) + q_A(t)\right) - p(t)x_A(t) - p_{q_A}(t)q_A(t) + \bar{K}_A + \underbrace{\Pi_{gA}(t)}_{=0} + \Pi_A(t) + v_A\Pi_F(t). \tag{A.12}$$

Supply-side policy: To model governmental extraction we consider the (inverse) production function[61]

$$\ell_A(t) = C(t)s_A(t). \tag{A.13}$$

The coalition's expenditures for purchasing deposits less their profits from selling fuel are

$$G(t) = p_y(t)y_A(t) - [p(t)s_A(t) - \ell_A(t)] = p_y(t)y_A(t) - [p(t) - C(t)]s_A(t). \tag{A.14}$$

Consumer A's budget constraint is then given by

$$px_A + p_{q_A}q_A + g_A = \bar{K}_A - G(t) + \underbrace{\Pi_{gA}(t)}_{=0} + \Pi_A(t) + v_A\Pi_F(t). \tag{A.15}$$

Derivation of country A's welfare: Solving for g_A and inserting the budget constraint (A.15) into (A.1) yields the welfare

$$U_A\left[x_A(t) + q_A(t)\right] - p(t)x_A(t) - p_{q_A}(t)q_A(t) + \bar{K}_A - G(t) + \Pi_A(t) + v_A\Pi_F(t). \tag{A.16}$$

A.2. The social optimum

The current-value Lagrangian reads[62]

$$\begin{aligned} L = & \sum_i [U_i(x_i + q_i) + \bar{K}_i - M_i(q_i)] - [C + \tau][x_A + x_B] - \theta[x_A + x_B - \gamma Z] - \iota\chi C \\ & + \mu[\bar{Z} - Z] + \sum_i \zeta_{x_i}x_i + \sum_i \zeta_{q_i}q_i, \end{aligned} \tag{A.17}$$

[61] Note that the coalition's resource constraint (A.5) turns into $\ell_{gA}(t) + \ell_{qA}(t) + \ell_{sA}(t) + \ell_A(t) = \bar{K}_A$.

[62] We apply the direct approach to solve dynamic optimization problems with state space constraints. Cf. Chiang (1999, chap. 10), Feichtinger and Hartl (1986, chap. 6), Kamien and Schwartz (2012, part II, chap. 17), and Seierstad and Sydsaeter (1987, chap. 5,6).

where ι is the extraction technology's costate. From the first-order conditions we get (9)-(11) and

$$\dot{\iota} = [\rho + \chi]\iota + s. \tag{A.18}$$

The complementary slackness conditions are

$$\mu \geq 0, \quad \mu[\bar{Z} - Z] = 0, \tag{A.19}$$

$$(a): \ \zeta_{x_i} \geq 0, \ \zeta_{x_i} x_i = 0, \qquad (b): \ \zeta_{q_i} \geq 0, \ \zeta_{q_i} q_i = 0. \tag{A.20}$$

(A.19) implies that μ is zero during Phase I, III and IV but positive during Phase II, so that μ is discontinuous at t_1 and t_2. Finally, the transversality conditions read[63]

$$(a): \ \lim_{t\to\infty} e^{-\rho t}\tau(t)[S(t) - S^{opt}(t)] \geq 0, \quad (b): \ \lim_{t\to\infty} e^{-\rho t}\theta(t)[-Z(t) + Z^{opt}(t)] \geq 0, \tag{A.21}$$

$$\lim_{t\to\infty} e^{-\rho t}\iota(t)[C(t) - C^{opt}(t)] \geq 0. \tag{A.22}$$

The variables marked with the superscript ^{opt} denote optimal values, while the unmarked variables of (A.21) refer to any feasible path.

Lemma A.1. *There is at most one switching time in Phase I, i.e. $t^s \in [0, t_1)$.*

Proof of Lemma A.1. We have $\dot{U}'(t) = \rho\tau_0 e^{\rho t} + (\rho+\gamma)\theta_0 e^{(\rho+\gamma)t} - \chi C_0 e^{-\chi t}$ and $\ddot{U}'(t) = \rho^2\tau_0 e^{\rho t} + (\rho+\gamma)^2\theta_0 e^{(\rho+\gamma)t} + \chi^2 C_0 e^{-\chi t}$, such that $\dot{U}'(t)$ is either positive for $t \in [0, t_1)$, or it is positive for $t \in [0, \tilde{t})$ and negative for $t \in (\tilde{t}, t_1)$, or it is negative for $t \in [0, t_1)$. However, a smooth transition in Phase I and $\dot{U}'(t)$ being negative for $t \in [0, t_1)$ implies $s(t) < \gamma\bar{Z}$ for $t \in [0, t_1)$ and, thus,

$$Z(t_1) = \left[Z_0 + \int_0^{t_1} s(j)e^{\gamma j}\,\mathrm{d}j\right]e^{-\gamma t_1} < \left[Z_0 + \bar{Z}\left(e^{\gamma t_1} - 1\right)\right]e^{-\gamma t_1} \leq \bar{Z} \iff Z_0 \leq \bar{Z}, \tag{A.23}$$

such that the ceiling would bind either from the beginning or never. $\square$

Lemma A.2. *Suppose that the economy is in Phase II, i.e. $t \in [t_1, t_2)$ and $\bar{Z} = Z(t)$. $\tau(t)$, $\theta(t)$ and $\iota(t)$ are continuous for all $t \in [t_1, t_2)$.*

[63] The transversality conditions belong to the sufficient conditions. We write the transversality conditions in the form used by Feichtinger and Hartl (1986, chapter 7.2). Note that we defined θ as non-negative.

Proof of Lemma A.2. For $t \in [t_1, t_2)$, the costate-variables may jump according to the condition[64]

$$\Lambda_c(t^-) = \Lambda_c(t^+) + \eta_\Lambda(t) \frac{\partial[\bar{Z} - Z(t)]}{\partial \Lambda} \quad \text{with} \quad \eta_\Lambda(t) \geq 0,\ \eta_\Lambda(t)[\bar{Z} - Z(t)] = 0, \tag{A.24}$$

where Λ is an arbitrary state-variable and Λ_c the corresponding costate-variable. t^- and t^+ refer to the value just before and just after time t, respectively. Applying (A.24) to S and C shows that τ and ι are continuous. In case of Z, we get

$$\theta(t^-) = \theta(t^+) - \eta_Z(t) \quad \text{with} \quad \eta_Z(t) \geq 0,\ \eta_Z(t)[\bar{Z} - Z(t)] = 0. \tag{A.25}$$

Suppose that θ jumps at time t, i.e. suppose that $\eta_Z(t) > 0$. Then, $\theta(t^+) > \theta(t^-)$ implying an increase of the climate costs of emissions and, therefore, a reduction of fuel extraction. Since $s(t) = \bar{s}$ during Phase II, the reduction implies $\bar{Z} > Z(t)$, which in turns requires $\eta_Z(t) = 0$. The contradiction proves Lemma A.2. □

Lemma A.3. *There is no switching time t^s in Phase III or Phase IV, i.e. $t^s \in [0, t_2)$.*

Proof of Lemma A.3. Consider Phase II, such that $s(t) = \bar{s}$ is fixed for all $t \in [t_1, t_2)$. If $\dot{\tau}(t) < -\dot{C}(t)$, $\theta(t)$ increases to ensure $s(t) = \bar{s}$. At the end of Phase II, $\theta(t_2) = 0$ holds. According to Lemma A.2, $\theta(t)$ is continuous for all $t \in [t_1, t_2)$, so that $\theta(t_2) = 0$ requires $\dot{\theta}(t) < 0$ for some $t \in [t_1, t_2)$. To ensure $s(t) = \bar{s}$, $\dot{\tau}(t) > -\dot{C}(t)$ holds at these points in time. Because $\hat{\tau}(t) = \rho$ and $\hat{C}(t) = -\chi$, $\dot{\tau}(j) > -\dot{C}(j)$ implies that $\dot{\tau}(t) > -\dot{C}(t)$ holds for all $t \geq j$. $\theta(t) = 0$ for $t \geq t_2$, so that $t^s < t_2$. □

A.3. The laissez-faire economy

The current-value Lagrangian of the representative fuel firm reads

$$L = p_F s_F + p_y y_F - C s_F - \tau_F[s_F + y_F] - \iota_F \chi C + \zeta_{s_F} s_F + \zeta_S[S - s_F - y_F], \tag{A.26}$$

where ζ_{s_F} and ζ_S are the multipliers of the non-negativity conditions $s_F \geq 0$ and $S - s_F - y_F \geq 0$. Because (A.26) is linear in both s_F and y_F, the optimal solution maximizes

[64] Cf. Chiang (1999, pp. 298) and Feichtinger and Hartl (1986, chap. 6.2).

$H = p_F s_F + p_y y_F - C s_F - \tau_F[s_F + y_F] - \iota_F \chi C$. The corresponding complete fuel supply correspondence and complete deposit supply correspondence read

$$s_F(t) \begin{cases} = 0 & \text{if } p(t) < C + \tau_F(t), \\ \in [0, S(t)] & \text{if } p(t) = C + \tau_F(t), \\ = S(t) & \text{if } p(t) > C + \tau_F(t), \end{cases} \tag{A.27}$$

and

$$y_F(t) \begin{cases} \leq 0 & \text{if } p_y(t) < \tau_F(t), \\ \in [0, S(t)] & \text{if } p_y(t) = \tau_F(t), \\ = S(t) & \text{if } p_y(t) > \tau_F(t). \end{cases} \tag{A.28}$$

The complete Hotelling-rule is

$$\tau_F(t) \begin{cases} = \tau_{F0} e^{\rho t} & \text{if } p(j) \leq C(j) + \tau_F(j) \wedge p_y(j) \leq \tau_F(j) \; \forall \; j \leq t, \\ \leq \tau_{F0} e^{\rho t} & \text{otherwise.} \end{cases} \tag{A.29}$$

The first order condition with respect to C gives

$$i_F = [\rho + \chi]\iota_F + s_F. \tag{A.30}$$

The transversality conditions read

$$(a): \lim_{t\to\infty} e^{-\rho t} \tau_F(t)[S(t) - S^{opt}(t)] \geq 0, \quad (b): \lim_{t\to\infty} e^{-\rho t} \iota(t)[C(t) - C^{opt}(t)] \geq 0. \tag{A.31}$$

The first ensures that no fuel is left in situ. Solving (A.30) and making use of the second transversality condition yield $\iota_F(t) = -\int_t^T e^{-[\rho+\chi](j-t)} s_F(j)\,\mathrm{d}j$, so that the firm is not willing to pay for any exogenous improvement of extraction technology after its resource stock is exhausted.

A.4. Competitive demand-side policy

Derivation of (21). The current-value Lagrangian of country A reads

$$\begin{aligned} L = {} & U_A(x_A + q_A) + \bar{K}_A - p x_A - p_{q_A} q_A + \upsilon_A \Pi_F + \Pi_A - \theta_A[x_A + D_B - \gamma Z] \\ & + \mu_A[\bar{Z} - Z] + \zeta_x x_A. \end{aligned} \tag{A.32}$$

From the first-order conditions we get (21) and

$$\dot{\theta}_A = [\rho + \gamma]\theta_A - \mu_A. \tag{A.33}$$

The complementary slackness conditions are

$$\mu_A \geq 0, \quad \mu_A[\bar{Z} - Z] = 0, \tag{A.34}$$

$$\zeta_x \geq 0, \quad \zeta_x x = 0. \tag{A.35}$$

Finally, the transversality condition reads

$$\lim_{t\to\infty} e^{-\rho t}\theta_A(t)[-Z(t) + Z^{opt}(t)] \geq 0. \tag{A.36}$$

Proof of Proposition 1. The fuel cap of country A is given by (A.38), while the fuel demand of the fringe is given by (18) with $p(t) = C(t) + \tau_F(t)$. For $t \in [0, t_2)$ the climate costs of emissions $\theta_A(t)$ are positive, so that $\frac{x_A(t)}{n_A} < \frac{x_B(t)}{n_B}$. For $t \geq t_2$ the emission costs term vanishes, i.e. $\theta_A(t) = 0$, implying $\frac{x_A(t)}{n_A} = \frac{x_B(t)}{n_B}$.

Fuel demand of the fringe is given by (18) with $p(t) = C(t) + \tau_F(t)$. Because of $\hat{\tau}_F(t) = \rho > 0$, $\hat{C} = -\chi < 0$ and (20), fuel utilization of the fringe increases for $t \in [0, t_B^s)$ and decreases for $t \in [t_B^s, T)$. Using $U'\left(\frac{x_i+q_i}{n_i}\right) = M'\left(\frac{q_i}{n_i}\right)$ yields

$$\frac{\mathrm{d}q_i}{\mathrm{d}x_i} = \frac{U''}{M'' - U''} < 0. \tag{A.37}$$

Consequently, backstop consumption in the fringe evolves according to $\dot{q}_B(t) < 0$ for $t \in [0, t_B^s)$ and $\dot{q}_B(t) > 0$ for $t \in [t_B^s, T)$.

For $t \in [0, T)$, an equilibrium on the energy market of country A requires $U'\left(\frac{x_A(t)+q_A(t)}{n_A}\right) = C(t) + \tau_F(t) + \theta_A(t) = p_{q_A}(t) = M'\left(\frac{q_A(t)}{n_A}\right)$, so that the fuel cap is given by

$$x_A(t) = n_A[U'^{-1}(C(t) + \tau_F(t) + \theta_A(t)) - M'^{-1}(C(t) + \tau_F(t) + \theta_A(t))]. \tag{A.38}$$

Consider Phase I, such that the growth rates $\hat{\tau}_F = \rho$, $\hat{C} = -\chi$ and $\hat{\theta}_A = \rho + \gamma$ are constants. If the dynamics of $C(t) + \tau_F(t) + \theta_A(t)$ switch during Phase I, the switching time $t_{A_I}^s \in [0, t_1)$ is unique within Phase I due to the constant growth rates and $\dot{\tau}_F(t) + \dot{\theta}_A(t) > \dot{\tau}_F(t)$ implies $t_{A_I}^s < t_B^s$. In this case, $\dot{x}_A(t) > 0$ for $t \in [0, t_{A_I}^s)$ and $\dot{x}_A(t) < 0$ for $t \in [t_{A_I}^s, t_1)$. Using (A.37) implies $\dot{q}_A(t) < 0$ for $t \in [0, t^s_{A_I})$ and $\dot{q}_A(t) > 0$ for $t \in [t^s_{A_I}, t_1)$. If the dynamics do not switch, $\dot{x}_A(t) > 0$ and $\dot{q}_A(t) < 0$ for all $t \in [0, t_1)$.

Consider Phase II, such that $\bar{s} = x_A + D_B$. If $\dot{x}_A(t) > 0$, $\dot{D}_B(t) < 0$ implying $\rho\tau_F(t) > \chi C(t)$, and $\rho\tau_F(t) < \chi C(t) - \dot{\theta}_A(t)$, so that $\dot{\theta}_A(t) < 0$. Analogously, if $\dot{x}_A(t) < 0$, $\dot{D}_B(t) > 0$ implying $\rho\tau_F(t) < \chi C(t)$, and $\rho\tau_F(t) > \chi C(t) - \dot{\theta}_A(t)$, so that $\dot{\theta}_A(t) > 0$. Because of $\theta_A(t) > 0$ for $t \in [0, t_2)$, $\theta_A(t) = 0$ for all $t \geq t_2$ and Lemma A.2, $\dot{\theta}_A(t) < 0$ has to hold for some $t \in [t_1, t_2)$ ruling out an increasing fuel consumption in the fringe at these points in time. Due to the constant growth rates $\hat{\tau}_F = \rho$ and $\hat{C} = -\chi$, fuel consumption of the fringe is constant for only one point in time, so that $\dot{\theta}_A(t) < 0$ implies $\rho\tau_F(t) > \chi C(t)$ and, therefore, $t^s_B < t_2$. If $t^s_B \in [t_1, t_2)$, the constant fuel extraction $\bar{s}$ implies a switch in the dynamics of $C(t) + \tau_F(t) + \theta_A(t)$ from $\frac{\mathrm{d}[C(t)+\tau_F(t)+\theta_A(t)]}{\mathrm{d}t} > 0$ to $\frac{\mathrm{d}[C(t)+\tau_F(t)+\theta_A(t)]}{\mathrm{d}t} < 0$ at time $t^s_{A_{II}} = t^s_B$. In this case, $\dot{x}_A(t) < 0$ for $t \in [t_1, t^s_{A_{II}})$ and $\dot{x}_A(t) > 0$ for $t \in [t^s_{A_{II}}, t_2)$, while (A.37) implies $\dot{q}_A(t) > 0$ for $t \in [t_1, t^s_{A_{II}})$ and $\dot{q}_A(t) < 0$ for $t \in [t^s_{A_{II}}, t_2)$. If $t^s_B < t_1$, $\dot{x}_A(t) > 0$ and $\dot{q}_A(t) < 0$ for all $t \in [t_1, t_2)$.

Consider Phase III, so that $\theta_A(t) = 0 \; \forall \, t \geq t_2$. Because $t^s_B < t_2$, $C(t) + \tau_F(t)$ monotonically increases in time for all $t \in [t_2, T)$. Consequently, $\dot{x}_i(t) < 0$ and $\dot{q}_i(t) > 0$ for $i = A, B$ and $t \in [t_2, T)$.

The fuel cap path $x_A(t)$ is continuous if $C(t)$, $\tau_F(t)$ and $\theta_A(t)$ evolve continuously in time. In case of $C(t)$, the continuous evolution is ensured by assumption. According to (16), $\tau_F(t)$ has no discontinuities. For $\theta_A(t)$ we find a continuous evolution during Phase I, Phase III and Phase IV, because (11) holds in similar manner for θ_A, and $\theta_A(t) = 0$ for all $t \geq t_2$. For $t \in [t_1, t_2)$ the proof of Lemma A.2 can be applied in a similar manner. □

Proof of Proposition 2. The fuel price reads $p(t) = c + \tau_F(t)$ for all t. In Phase I and II it does not internalize the climate costs of emissions. In general, $U'(x_A(t) + q_A(t)) = p(t) + \theta_A(t) > p(t) = U'(x_B(t) + q_B(t))$ and $x_A(t) + x_B(t) = \bar{s}$ for $t \in [t_1, t_2)$ imply $x_B(t) > x^*_B(t)$ and $x_A(t) < x^*_A(t)$ at the ceiling.

Now suppose $\tau_{F0} < \tau_0$, such that $p(t) < C(t) + \tau(t) + \theta(t)$ and, therefore, $x_B(t) > x^*_B(t)$ for all t. $U'\left(\frac{x_i(t_2)+q_i(t_2)}{n_i}\right) = U'\left(\frac{x^*_i(t^*_2)+q^*_i(t^*_2)}{n_i}\right) = \bar{p} = C_0 e^{-\chi t_2} + \tau_{F0} e^{\rho t_2} = C_0 e^{-\chi t^*_2} + \tau_0 e^{\rho t^*_2}$ and $M'\left(\frac{q_i(T)}{n_i}\right) = M'\left(\frac{q^*_i(T^*)}{n_i}\right) = C_0 e^{-\chi T} + \tau_{F0} e^{\rho T} = C_0 e^{-\chi T^*} + \tau_0 e^{\rho T^*}$ hold. The first [second] equality and $\tau_{F0} < \tau_0$ imply $p(t_2) > p(t^*_2)$ [$p(T) > p(T^*)$]. Because $T > t_2 > t^s$, $T^* > t^*_2 > t^{s^*}$

and $\dot{p}(t) > 0$ for $t > t^s, t^{s^*}$, we get $t_2 > t_2^*$ and $T > T^*$. Then, $x_B(t) > x_B^*(t)$ for all $t < T$, $x_A(t) > x_A^*(t)$ for $t \in [t_2^*, T]$ and $s(t) = s^*(t) = \bar{s}$ for $t \in [\max\{t_1, t_1^*\}, t_2^*)$ imply $\int_0^{\max\{t_1,t_1^*\}} x_A(t)\,\mathrm{d}t < \int_0^{\max\{t_1,t_1^*\}} x_A^*(t)\,\mathrm{d}t$ and $\int_0^{\max\{t_1,t_1^*\}} s(t)\,\mathrm{d}t < \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$. Finally, $\theta_{A0} \leq \theta_0$ implies that the extraction paths do not cut during Phase I, such that the ceiling is violated. Therefore, $\theta_{A0} > \theta_0$ holds.

Next suppose $\tau_{F0} > \tau_0$, such that $t_2 < t_2^*$ and $T < T^*$ by $U'\left(\frac{x_i(t_2)+q_i(t_2)}{n_i}\right) = U'\left(\frac{x_i^*(t_2^*)+q_i^*(t_2^*)}{n_i}\right) = \bar{p} = C_0 e^{-\chi t_2} + \tau_{F0} e^{\rho t_2} = C_0 e^{-\chi t_2^*} + \tau_0 e^{\rho t_2^*}$ and $M'\left(\frac{q_i(T)}{n_i}\right) = M'\left(\frac{q_i^*(T^*)}{n_i}\right) = C_0 e^{-\chi T} + \tau_{F0} e^{\rho T} = C_0 e^{-\chi T^*} + \tau_0 e^{\rho T^*}$, respectively. Then, $x_i(t) < x_i^*(t)$ for $t \in [t_2, T^*)$ and $s(t) = s^*(t) = \bar{s}$ for $t \in [\max\{t_1, t_1^*\}, t_2)$ imply $\int_0^{\max\{t_1,t_1^*\}} s(t)\,\mathrm{d}t > \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$. Furthermore, $\frac{x_A(t)}{n_A} < \frac{x_B(t)}{n_B}$ and $\frac{x_A^*(t)}{n_A} = \frac{x_B^*(t)}{n_B}$ during Phase I imply $\int_0^{\max\{t_1,t_1^*\}} x_B(t)\,\mathrm{d}t > \int_0^{\max\{t_1,t_1^*\}} x_B^*(t)\,\mathrm{d}t$. Finally, $x_A(0) > x_A^*(0) \Leftrightarrow \tau_{F0} + \theta_{A0} \leq \tau_0 + \theta_0$ implies that the extraction paths do not cut during Phase I, such that the ceiling is violated. Therefore, $x_A(0) < x_A^*(0) \Leftrightarrow \tau_{F0} + \theta_{A0} > \tau_0 + \theta_0$ holds. $\square$

Now consider the quadratic cost functions

$$M_i(t) = \frac{1}{2}\frac{n_i}{m}\left[\frac{q_i(t)}{n_i}\right]^2, \tag{A.39}$$

and the HARA utility functions

$$U_i(t) = n_i \beta^{1-\phi} \frac{1-\phi}{\phi}\left\{\left[\frac{\frac{x_i(t)+q_i(t)}{n_i}}{1-\phi} + \alpha\right]^{\phi} - \alpha^{\phi}\right\}, \tag{A.40}$$

which yields the demands

$$D_i(t) = n_i(\phi - 1)\left[\alpha - \beta U_i'(t)^{\frac{1}{\phi-1}}\right] - n_i m U_i'(t) \qquad \text{for } \phi \neq 1. \tag{A.41}$$

Lemma A.4. *Suppose the coalition applies a demand-side climate policy and is a price taker in the fuel market. Consider the quadratic cost functions* (A.39) *and the HARA utility functions* (A.40)*, which yields the demands* (A.41)*. If $\phi = 2$ (linear demands), then $\tau_{F0} = \tau_0$, $t_1 = t_1^*$, $t_2 = t_2^*$, $T = T^*$ and $\theta_{A0} = \theta_0/n_A$, and a marginal increase [decrease] in ϕ implies $\tau_{F0} < [>]\tau_0$, $t_1 > [<]t_1^*$, $t_2 > [<]t_2^*$, $T > [<]T^*$ and $\theta_{A0} > [<]\theta_0/n_A$ for $\gamma = \rho$ and $C_0 = 0$. If $\phi = 3/2$ (quadratic demands), then $\tau_{F0} > \tau_0$, $t_2 < t_2^*, T < T^*$ and $\theta_{A0} \in (\theta_0, \theta_0/n_A)$ for $\gamma = \rho$ and $\chi = 0$ or $\chi = \rho$.*

Proof of Lemma A.4. The equilibrium is characterized by

$$\text{(A.42)}\quad \begin{aligned} S(0) &= \int_0^{t_1} D_A\left(C_0e^{-\chi t}+\tau_{F0}e^{\rho t}+\theta_{A0}e^{(\rho+\gamma)t}\right)\mathrm{d}t+\int_0^{t_1} D_B\left(C_0e^{-\chi t}+\tau_{F0}e^{\rho t}\right)\mathrm{d}t \\ &\quad +[t_2-t_1]\gamma\bar{Z}+\int_{t_2}^{T} D\left(C_0e^{-\chi t}+\tau_{F0}e^{\rho t}\right)\mathrm{d}t, \end{aligned}$$

$$\text{(A.43)}\quad \begin{aligned} \bar{Z} &= Z(0)e^{-\gamma t_1}+\int_0^{t_1} D_A\left(C_0e^{-\chi t}+\tau_{F0}e^{\rho t}+\theta_{A0}e^{(\rho+\gamma)t}\right)e^{\gamma t}\,\mathrm{d}te^{-\gamma t_1} \\ &\quad +\int_0^{t_1} D_B\left(C_0e^{-\chi t}+\tau_{F0}e^{\rho t}\right)e^{\gamma t}\,\mathrm{d}te^{-\gamma t_1}, \end{aligned}$$

$$\text{(A.44)}\quad \gamma\bar{Z}=D_A\left(C_0e^{-\chi t_1}+\tau_{F0}e^{\rho t_1}+\theta_{A0}e^{(\rho+\gamma)t_1}\right)+D_B\left(C_0e^{-\chi t_1}+\tau_{F0}e^{\rho t_1}\right),$$

$$\text{(A.45)}\quad \gamma\bar{Z}=D\left(C_0e^{-\chi t_2}+\tau_{F0}e^{\rho t_2}\right),$$

$$\text{(A.46)}\quad 0=D\left(C_0e^{-\chi T}+\tau_{F0}e^{\rho T}\right).$$

Differentiating with respect to n_B and using $D_A(t_1)+D_B(t_1)=\gamma\bar{Z}=D_A(t_2)+D_B(t_2)$ yields

$$\text{(A.47)}\quad \begin{aligned} 0 &= \left\{\int_0^{t_1}[D_A'(t)+D_B'(t)]e^{\rho t}\,\mathrm{d}t+\int_{t_2}^{T} D'(t)e^{\rho t}\,\mathrm{d}t\right\}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B} \\ &\quad +\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}+\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]\mathrm{d}t, \end{aligned}$$

$$\text{(A.48)}\quad \begin{aligned} 0 &= \int_0^{t_1}[D_A'(t)+D_B'(t)]e^{(\rho+\gamma)t}\,\mathrm{d}te^{-\gamma t_1}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B} \\ &\quad +\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}te^{-\gamma t_1}\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}+\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]e^{\gamma t}\,\mathrm{d}te^{-\gamma t_1}, \end{aligned}$$

$$\text{(A.49)}\quad \begin{aligned} 0 &= [D_A'(t_1)+D_B'(t_1)]e^{\rho t_1}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}+D_A'(t_1)e^{(\rho+\gamma)t_1}\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}+\frac{D_B(t_1)}{n_B}-\frac{D_A(t_1)}{n_A} \\ &\quad +\left\{[D_A'(t_1)+D_B'(t_1)][\rho\tau_{F0}e^{\rho t_1}-\chi C_0e^{-\chi t_1}]+D_A'(t_1)(\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t_1}\right\}\tfrac{\mathrm{d}t_1}{\mathrm{d}n_B}, \end{aligned}$$

$$\text{(A.50)}\quad 0=D'(t_2)\left[e^{\rho t_2}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}+[\rho\tau_{F0}e^{\rho t_2}-\chi C_0e^{-\chi t_2}]\tfrac{\mathrm{d}t_2}{\mathrm{d}n_B}\right],$$

$$\text{(A.51)}\quad 0=D'(T)\left[e^{\rho T}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}+[\rho\tau_{F0}e^{\rho T}-\chi C_0e^{-\chi T}]\tfrac{\mathrm{d}T}{\mathrm{d}n_B}\right].$$

Using (A.51) in (A.47) and solving (A.47), (A.48) and (A.49) for $\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}$, $\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}$ and $\frac{\mathrm{d}t_1}{\mathrm{d}n_B}$ yields

$$\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}= \\ \frac{\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]e^{\gamma t}\,\mathrm{d}t-\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]\mathrm{d}t}{X\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}[D_A'(t)+D_B'(t)]e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t},$$

$$\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B} =$$

$$-\frac{X\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]e^{\gamma t}\,\mathrm{d}t-\int_0^{t_1}[D_A'(t)+D_B'(t)]e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}\left[\frac{D_B(t)}{n_B}-\frac{D_A(t)}{n_A}\right]\mathrm{d}t}{X\int_0^{t_1}D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}[D_A'(t)+D_B'(t)]e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t},$$

$$\frac{\mathrm{d}t_1}{\mathrm{d}n_B} = -\frac{[D_A'(t_1)+D_B'(t_1)]e^{\rho t_1}\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}+D_A'(t_1)e^{(\rho+\gamma)t_1}\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}+\frac{D_B(t_1)}{n_B}-\frac{D_A(t_1)}{n_A}}{[D_A'(t_1)+D_B'(t_1)][\rho\tau_{F0}e^{\rho t_1}-\chi C_0e^{-\chi t_1}]+D_A'(t_1)(\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t_1}},$$

where $X \coloneqq \int_0^{t_1}[D_A'(t)+D_B'(t)]e^{\rho t}\,\mathrm{d}t-\frac{\gamma\bar{Z}}{\rho\tau_{F0}}$.

Next, consider the quadratic cost functions (A.39) and the HARA utility functions (A.40), which yields the demands (A.41). For $\phi = 2$, i.e. linear-quadratic utility functions with $D_i'' = 0$, we get

$$\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B} = \frac{\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}\tilde{\beta}\theta_{A0}e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+2\gamma)t}\,\mathrm{d}t\int_0^{t_1}\tilde{\beta}\theta_{A0}e^{(\rho+\gamma)t}\,\mathrm{d}t}{X\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t} = 0,$$

$$\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B} = \frac{X\int_0^{t_1}\tilde{\beta}\theta_{A0}e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}\tilde{\beta}\theta_{A0}e^{(\rho+\gamma)t}\,\mathrm{d}t}{X\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+2\gamma)t}\,\mathrm{d}t-\int_0^{t_1}\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1}n_A\tilde{\beta}e^{(\rho+\gamma)t}\,\mathrm{d}t} = \frac{\theta_{A0}}{n_A},$$

$$\frac{\mathrm{d}t_1}{\mathrm{d}n_B} = \frac{0-n_A\tilde{\beta}e^{(\rho+\gamma)t_1}\frac{\theta_{A0}}{n_A}+\tilde{\beta}\theta_{A0}e^{(\rho+\gamma)t_1}}{\tilde{\beta}[\rho\tau_{F0}e^{\rho t_1}-\chi C_0e^{-\chi t_1}]+n_A\tilde{\beta}(\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t_1}} = 0,$$

and $\frac{\mathrm{d}t_2}{\mathrm{d}n_B}, \frac{\mathrm{d}T}{\mathrm{d}n_B} = 0$ by (A.50) and (A.51), where $\tilde{\beta} \coloneqq \beta + m$.

Substituting (A.41) into $\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}$, differentiating with respect to ϕ, and evaluating at $\phi = 2$, $C_0 = 0$ and $\gamma = \rho$ yields

$$\left.\frac{\partial\left(\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}\right)}{\partial\phi}\right|_{\phi=2,C_0=0,\gamma=\rho} = -\frac{\beta\tau_{F0}^2\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\left[e^{2\rho t_1}+e^{\rho t_1}+4+3\left(e^{\rho t_1}+1\right)\frac{\tau_{F0}}{\theta_{A0}}\right]\mathcal{A}}{\tilde{\beta}\tau_{F0}\left(e^{\rho t_1}-1\right)^3+4\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)\rho\bar{Z}},$$

where

$$\mathcal{A} \coloneqq \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right)+\frac{\left(e^{\rho t_1}-1\right)^3}{\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\left[e^{2\rho t_1}+e^{\rho t_1}+4+3\left(e^{\rho t_1}+1\right)\frac{\tau_{F0}}{\theta_{A0}}\right]}\ln\left(\frac{1+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right)$$
$$-\frac{\left(e^{\rho t_1}-1\right)\left[5e^{2\rho t_1}+2e^{\rho t_1}+5+6\left(e^{\rho t_1}+1\right)\frac{\tau_{F0}}{\theta_{A0}}\right]}{2\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\left[e^{2\rho t_1}+e^{\rho t_1}+4+3\left(e^{\rho t_1}+1\right)\frac{\tau_{F0}}{\theta_{A0}}\right]},$$

$$\frac{\partial\mathcal{A}}{\partial t_1} = \frac{\rho e^{\rho t_1}\left(e^{\rho t_1}-1\right)^2\left(e^{\rho t_1}+2\right)}{\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)^2\left[e^{2\rho t_1}+e^{\rho t_1}+4+3\left(e^{\rho t_1}+1\right)\frac{\tau_{F0}}{\theta_{A0}}\right]^2}\Bigg\{e^{2\rho t_1}+2e^{\rho t_1}-3$$
$$+2\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(2e^{\rho t_1}+1+3\frac{\tau_{F0}}{\theta_{A0}}\right)\left[\ln\left(\frac{1+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right)-1+\frac{\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right]\Bigg\} > 0.$$

Since $\lim_{t_1 \to 0} \mathcal{A} = 0$ holds, $\frac{\partial \mathcal{A}}{\partial t_1} > 0$ implies $\mathcal{A} > 0$, such that $\left.\frac{\partial\left(\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} < 0$. Furthermore, $\left.\frac{\partial\left(\frac{\mathrm{d}t_2}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho}, \left.\frac{\partial\left(\frac{\mathrm{d}T}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} > 0$ by (A.50) and (A.51).

Substituting (A.41) into $\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}$, differentiating with respect to ϕ, and evaluating at $\phi = 2$, $C_0 = 0$ and $\gamma = \rho$ yields

$$\left.\frac{\partial\left(\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} = \frac{6\beta\tau_{F0}\left[\tilde{\beta}\tau_{F0}\left(e^{\rho t_1}-1\right)\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\mathcal{B}_1 + \rho\bar{Z}\mathcal{B}_2\right]}{n_A\tilde{\beta}\left(e^{\rho t_1}-1\right)\left[\tilde{\beta}\tau_{F0}\left(e^{\rho t_1}-1\right)^3 + 4\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)\rho\bar{Z}\right]},$$

where

$$\mathcal{B}_1 \coloneqq \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right) + \frac{\left(e^{\rho t_1}-1\right)\left[\left(e^{\rho t_1}-1\right)^2 - 3\left(e^{\rho t_1}+1+2\frac{\tau_{F0}}{\theta_{A0}}\right)\frac{\tau_{F0}}{\theta_{A0}}\right]}{6\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\frac{\tau_{F0}}{\theta_{A0}}},$$

$$\frac{\partial \mathcal{B}_1}{\partial t_1} = \frac{\rho e^{\rho t_1}\left(e^{\rho t_1}-1\right)^2\left(2e^{\rho t_1}+1\right)}{6\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)^2\frac{\tau_{F0}}{\theta_{A0}}} > 0,$$

$$\begin{aligned}\mathcal{B}_2 \coloneqq\ & \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right)\left(\frac{\tau_{F0}}{\theta_{A0}}\right)^2 - \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right)e^{2\rho t_1} + \ln\left(\frac{1+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right)\\ & + \frac{\left(e^{\rho t_1}-1\right)\left[4\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)+3\left(e^{\rho t_1}+1-2\frac{\tau_{F0}}{\theta_{A0}}\right)\frac{\tau_{F0}}{\theta_{A0}}\right]}{6\frac{\tau_{F0}}{\theta_{A0}}},\end{aligned}$$

$$\frac{\partial \mathcal{B}_2}{\partial t_1} = 2\rho e^{2\rho t_1}\left[\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}} - 1 - \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right)\right] > 0.$$

Since $\lim_{t_1 \to 0} \mathcal{B}_1 = 0$ and $\lim_{t_1 \to 0} \mathcal{B}_2 = 0$ hold, $\frac{\partial \mathcal{B}_1}{\partial t_1} > 0$ and $\frac{\partial \mathcal{B}_2}{\partial t_1} > 0$ imply $\mathcal{B}_1 > 0$ and $\mathcal{B}_2 > 0$, such that $\left.\frac{\partial\left(\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} > 0$.

Finally, substituting (A.41) into $\frac{\mathrm{d}t_1}{\mathrm{d}n_B}$, differentiating with respect to ϕ, and evaluating at $\phi = 2$, $C_0 = 0$ and $\gamma = \rho$ yields

$$\left.\frac{\partial\left(\frac{\mathrm{d}t_1}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} = \frac{\left(e^{\rho t_1}-1\right)\beta\tau_{F0}\left[\tilde{\beta}\tau_{F0}\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(2e^{\rho t_1}+1+3\frac{\tau_{F0}}{\theta_{A0}}\right)\mathcal{C}_1 + 2\left(e^{\rho t_1}+2\right)\rho\bar{Z}\mathcal{C}_2\right]}{\rho\tilde{\beta}\left(\tau_{F0}+2n_A\theta_{A0}e^{\rho t_1}\right)\left[\tilde{\beta}\tau_{F0}\left(e^{\rho t_1}-1\right)^3 + 4\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)\rho\bar{Z}\right]},$$

where

$$\mathcal{C}_1 \coloneqq \frac{\left(e^{\rho t_1}-1\right)\left(e^{\rho t_1}+5+6\frac{\tau_{F0}}{\theta_{A0}}\right)}{2\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(2e^{\rho t_1}+1+3\frac{\tau_{F0}}{\theta_{A0}}\right)} - \ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right),$$

$$\frac{\partial \mathcal{C}_1}{\partial t_1} = \frac{\rho e^{\rho t_1}\left(e^{\rho t_1}-1\right)^3}{\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(2e^{\rho t_1}+1+3\frac{\tau_{F0}}{\theta_{A0}}\right)^2} > 0,$$

$$\mathcal{C}_2 \coloneqq \ln\left(\frac{1+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right) + \frac{e^{3\rho t_1}+2-3e^{\rho t_1}\left(\frac{\tau_{F0}}{\theta_{A0}}\right)^2}{\left(e^{\rho t_1}-1\right)^2\left(e^{\rho t_1}+2\right)}\ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right) - \frac{3e^{\rho t_1}\left(e^{\rho t_1}+1-2\frac{\tau_{F0}}{\theta_{A0}}\right)}{2\left(e^{\rho t_1}-1\right)\left(e^{\rho t_1}+2\right)},$$

$$\frac{\partial \mathcal{C}_2}{\partial t_1} = \frac{6\rho e^{\rho t_1}\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\mathcal{C}_3}{\left(e^{\rho t_1}-1\right)^3\left(e^{\rho t_1}+2\right)^2},$$

$$\mathcal{C}_3 \coloneqq \left(\frac{\tau_{F0}}{\theta_{A0}}-1\right)\ln\left(\frac{e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}}\right) + \frac{\left(e^{\rho t_1}-1\right)\left[e^{2\rho t_1}+e^{\rho t_1}+4-3\left(e^{\rho t_1}-1+2\frac{\tau_{F0}}{\theta_{A0}}\right)\frac{\tau_{F0}}{\theta_{A0}}\right]}{6\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)},$$

$$\frac{\partial \mathcal{C}_3}{\partial t_1} = \frac{\rho e^{\rho t_1}\left(e^{\rho t_1}-1\right)^2\left(e^{\rho t_1}+2\right)}{3\left(1+\frac{\tau_{F0}}{\theta_{A0}}\right)\left(e^{\rho t_1}+\frac{\tau_{F0}}{\theta_{A0}}\right)^2} > 0.$$

Since $\lim_{t_1\to 0}\mathcal{C}_1 = 0$, $\lim_{t_1\to 0}\mathcal{C}_2 = \ln\left(\frac{1+\frac{\tau_{F0}}{\theta_{A0}}}{\frac{\tau_{F0}}{\theta_{A0}}}\right) + 1 - \frac{\frac{\tau_{F0}}{\theta_{A0}}}{1+\frac{\tau_{F0}}{\theta_{A0}}} > 0$ and $\lim_{t_1\to 0}\mathcal{C}_3 = 0$ hold, $\frac{\partial \mathcal{C}_1}{\partial t_1} > 0$, $\frac{\partial \mathcal{C}_2}{\partial t_1} > 0$ and $\frac{\partial \mathcal{C}_3}{\partial t_1} > 0$ imply $\mathcal{C}_1 > 0$, $\mathcal{C}_2 > 0$ and $\mathcal{C}_3 > 0$, such that $\left.\frac{\partial\left(\frac{\mathrm{d}t_1}{\mathrm{d}n_B}\right)}{\partial \phi}\right|_{\phi=2, C_0=0, \gamma=\rho} > 0$.

For $\gamma = \rho$, $\chi = 0$ and $\phi = \frac{3}{2}$, i.e. quadratic demand functions with $D_i'' < 0$, we get $\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B} = \frac{\mathcal{D}_1}{\mathcal{D}_3} > 0$, $\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B} = \frac{\mathcal{D}_2}{\mathcal{D}_3} \in (0, \frac{\theta_{A0}}{n_A})$ and $\frac{\mathrm{d}t_2}{\mathrm{d}n_B}, \frac{\mathrm{d}T}{\mathrm{d}n_B} < 0$ by (A.50) and (A.51), where

$$\begin{aligned}
\mathcal{D}_1 = \frac{n_A\beta\theta_{A0}^2\left(e^{\rho t_1}-1\right)^4}{480\rho^2}\Bigg[&4\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)(m+\beta C_0)\\
&+\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta\tau_{F0}\Bigg] > 0,
\end{aligned}$$

$$\begin{aligned}
\mathcal{D}_2 = \frac{\gamma\bar{Z}}{\rho\tau_{F0}}\frac{\theta_{A0}}{60\rho}&\left\{20\left(e^{3\rho t_1}-1\right)(m+\beta C_0)+15\left(e^{4\rho t_1}-1\right)\beta\tau_{F0}+6\left(e^{5\rho t_1}-1\right)\beta\theta_{A0}\right\}\\
&+\frac{\theta_{A0}\left(e^{\rho t_1}-1\right)^4}{1440\rho^2}\Bigg\{20n_B\left(e^{2\rho t_1}+4e^{\rho t_1}+1\right)(m+\beta C_0)\beta\theta_{A0}\\
&+3n_A\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta^2\theta_{A0}^2+120(m+\beta C_0)^2\\
&+\left[120\left(e^{\rho t_1}+1\right)\tau_{F0}+\left(34e^{2\rho t_1}-8e^{\rho t_1}+34\right)\theta_{A0}\right](m+\beta C_0)\beta\\
&+\left[20\left(e^{2\rho t_1}+4e^{2\rho t_1}+1\right)\tau_{F0}+12\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)\theta_{A0}\right]\beta^2\tau_{F0}\Bigg\} > 0,
\end{aligned}$$

$$\mathcal{D}_3 = \frac{\gamma\bar{Z}}{\rho\tau_{F0}}\frac{n_A}{60\rho}\left\{20\left(e^{3\rho t_1}-1\right)(m+\beta C_0)+15\left(e^{4\rho t_1}-1\right)\beta\tau_{F0}+12\left(e^{5\rho t_1}-1\right)\beta\theta_{A0}\right\}$$

$$+\frac{n_A\left(e^{\rho t_1}-1\right)^4}{1440\rho^2}\Bigg\{20n_B\left(e^{2\rho t_1}+4e^{\rho t_1}+1\right)(m+\beta C_0)\beta\theta_{A0}$$

$$+6n_A\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta^2\theta_{A0}^2+120(m+\beta C_0)^2$$

$$+\Big[120\left(e^{\rho t_1}+1\right)\tau_{F0}+\left(88e^{2\rho t_1}+64e^{\rho t_1}+88\right)\theta_{A0}\Big](m+\beta C_0)\beta$$

$$+\Big[20\left(e^{2\rho t_1}+4e^{2\rho t_1}+1\right)\tau_{F0}+24\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)\theta_{A0}\Big]\beta^2\tau_{F0}\Bigg\}>\frac{n_A}{\theta_{A0}}\mathcal{D}_2.$$

Finally, for $\gamma=\rho$, $\chi=\rho$ and $\phi=\frac{3}{2}$, we get $\frac{\mathrm{d}\tau_{F0}}{\mathrm{d}n_B}=\frac{\mathcal{E}_1}{\mathcal{E}_3}>0$, $\frac{\mathrm{d}\theta_{A0}}{\mathrm{d}n_B}=\frac{\mathcal{E}_2}{\mathcal{E}_3}\in(0,\frac{\theta_{A0}}{n_A})$ and $\frac{\mathrm{d}t_2}{\mathrm{d}n_B},\frac{\mathrm{d}T}{\mathrm{d}n_B}<0$ by (A.50) and (A.51), where

$$\mathcal{E}_1=\frac{n_A\beta\theta_{A0}^2\left(e^{\rho t_1}-1\right)^4}{480\rho^2}\Big[4\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)m$$

$$+6\left(3e^{2\rho t_1}+4e^{\rho t_1}+3\right)\beta C_0+\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta\tau_{F0}\Big]>0,$$

$$\mathcal{E}_2=\frac{\gamma\bar{Z}}{\rho\tau_{F0}}\frac{\theta_{A0}}{60\rho}\Bigg\{20\left(e^{3\rho t_1}-1\right)m+30\left(e^{2\rho t_1}-1\right)\beta C_0+15\left(e^{4\rho t_1}-1\right)\beta\tau_{F0}$$

$$+6\left(e^{5\rho t_1}-1\right)\beta\theta_{A0}\Bigg\}+\frac{\theta_{A0}\left(e^{\rho t_1}-1\right)^4}{1440\rho^2}\Bigg\{20n_B\left(e^{2\rho t_1}+4e^{\rho t_1}+1\right)m\beta\theta_{A0}$$

$$+3n_A\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta^2\theta_{A0}^2+120m^2$$

$$+\Big[120\left(e^{\rho t_1}+1\right)\tau_{F0}+\left(34e^{2\rho t_1}-8e^{\rho t_1}+34\right)\theta_{A0}\Big]\beta m$$

$$+\Big[20\left(e^{2\rho t_1}+4e^{2\rho t_1}+1\right)\tau_{F0}+12\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)\theta_{A0}\Big]\beta^2\tau_{F0}\Bigg\}$$

$$+\frac{\beta C_0\theta_{A0}\left(e^{\rho t_1}-1\right)}{60\rho^2\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\Bigg\{\Bigg[\rho t_1-\frac{5\left(e^{\rho t_1}-1\right)\left(5e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+5\right)}{12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\Bigg]$$

$$\cdot\Big[20\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)m+30\left(e^{\rho t_1}+1\right)\beta C_0+15\left(e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)\beta\tau_{A0}$$

$$+6\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)\beta\theta_{A0}\Big]+\frac{5\left(e^{\rho t_1}-1\right)^3}{12}\Big[12\left(e^{\rho t_1}+1\right)$$

$$\cdot\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)n_B\beta\theta_{A0}+4\left(7e^{3\rho t_1}+7e^{2\rho t_1}+7e^{\rho t_1}+7\right)m$$

$$+6\left(e^{2\rho t_1}+8e^{\rho t_1}+1\right)\beta C_0+3\left(5e^{4\rho t_1}+12e^{3\rho t_1}+6e^{2\rho t_1}+12e^{\rho t_1}+5\right)\beta\tau_{A0}\Big]\Bigg\}>0,$$

$$\mathcal{E}_3=\frac{\gamma\bar{Z}}{\rho\tau_{F0}}\frac{n_A}{60\rho}\Bigg\{20\left(e^{3\rho t_1}-1\right)m+30\left(e^{2\rho t_1}-1\right)\beta C_0+15\left(e^{4\rho t_1}-1\right)\beta\tau_{F0}$$

$$+ 12\left(e^{5\rho t_1}-1\right)\beta\theta_{A0}\Bigg\} + \frac{n_A\left(e^{\rho t_1}-1\right)^4}{1440\rho^2}\Bigg\{20 n_B\left(e^{2\rho t_1}+4e^{\rho t_1}+1\right)m\beta\theta_{A0}$$

$$+ 6n_A\left(e^{4\rho t_1}+4e^{3\rho t_1}+10e^{2\rho t_1}+4e^{\rho t_1}+1\right)\beta^2\theta_{A0}^2 + 120m^2$$

$$+ \left[120\left(e^{\rho t_1}+1\right)\tau_{F0} + \left(88e^{2\rho t_1}+64e^{\rho t_1}+88\right)\theta_{A0}\right]\beta m$$

$$+ \left[20\left(e^{2\rho t_1}+4e^{2\rho t_1}+1\right)\tau_{F0} + 24\left(e^{3\rho t_1}+4e^{2\rho t_1}+4e^{\rho t_1}+1\right)\theta_{A0}\right]\beta^2\tau_{F0}\Bigg\}$$

$$+ \frac{\beta C_0\theta_{A0}\left(e^{\rho t_1}-1\right)}{60\rho^2\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\Bigg\{\left[\rho t_1 - \frac{5\left(e^{\rho t_1}-1\right)\left(5e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+5\right)}{12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\right]$$

$$\cdot\Bigg[20\left(e^{2\rho t_1}+e^{\rho t_1}+1\right)m + 30\left(e^{\rho t_1}+1\right)\beta C_0 + 15\left(e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)\beta\tau_{A0}$$

$$+ 12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)\beta\theta_{A0}\Bigg] + \frac{5\left(e^{\rho t_1}-1\right)^3}{12}\Bigg[12\left(e^{\rho t_1}+1\right)$$

$$\cdot\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)(1+n_B)\beta\theta_{A0} + 4\left(7e^{3\rho t_1}+7e^{2\rho t_1}+7e^{\rho t_1}+7\right)m$$

$$+ 6\left(e^{2\rho t_1}+8e^{\rho t_1}+1\right)\beta C_0 + 3\left(5e^{4\rho t_1}+12e^{3\rho t_1}+6e^{2\rho t_1}+12e^{\rho t_1}+5\right)\beta\tau_{A0}\Bigg]\Bigg\} > \frac{n_A}{\theta_{A0}}\mathcal{E}_2,$$

and where

$$\left[\rho t_1 - \frac{5\left(e^{\rho t_1}-1\right)\left(5e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+5\right)}{12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\right]_{\rho=0} = 0,$$

$$\partial\left[\rho t_1 - \frac{5\left(e^{\rho t_1}-1\right)\left(5e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+5\right)}{12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)}\right]\Bigg/\partial\rho$$
$$= t_1\frac{\left(e^{\rho t_1}-1\right)^4\left(12e^{4\rho t_1}+27e^{3\rho t_1}+22e^{2\rho t_1}+27e^{\rho t_1}+12\right)}{12\left(e^{4\rho t_1}+e^{3\rho t_1}+e^{2\rho t_1}+e^{\rho t_1}+1\right)^2} > 0.$$

□

A.5. Competitive supply-side policy

Derivation of (25)-(27). The current-value Lagrangian of country A reads

$$\begin{aligned} L = {} & U_A\left(x_A+q_A\right) + \bar{K}_A - px_A - p_{q_A}q_A - p_y y_A + [p-C]s_A + \upsilon_A\Pi_F + \Pi_A \\ & + \tau_A[y_A - s_A] - \theta_A[s_A + s_F - \gamma Z] - \iota_A\chi C + \mu_A[\bar{Z} - Z] \\ & + \zeta_{y_A}y_A + \zeta_{s_A}s_A + \zeta_Y[S - y_A] + \zeta_{S_A}[S_A - s_A] \\ & + \zeta_G\left\{\bar{K}_A + \Pi_A + \upsilon_A\Pi_F - px_A - p_{q_A}q_A - \bar{g}_A + [p-C]s_A - p_y y_A\right\}, \end{aligned} \tag{A.52}$$

where ζ_{y_A}, ζ_{s_A}, ζ_Y, ζ_{S_A} and ζ_G are the multipliers of the non-negativity conditions $y_A \geq 0$, $s_A \geq 0$, $S - y_A \geq 0$, $S_A - s_A \geq 0$ and $\bar{y}_A - y_A \geq 0$.

The Lagrangian is linear in both s_A and y_A. The optimal strategies satisfy the first-order conditions

$$\text{(A.53)} \qquad \frac{\partial L}{\partial s_A} = p - C - \tau_A - \theta_A + \zeta_{s_A} - \zeta_{S_A} + \zeta_G[p - C] = 0,$$

$$\text{(A.54)} \qquad \frac{\partial L}{\partial y_A} = \tau_A - p_y + \zeta_{y_A} - \zeta_Y - \zeta_G p_y = 0.$$

and maximize the Hamiltonian $H = U_A(x_A + q_A) + \bar{K}_A - px_A - p_{q_A}q_A - p_y y_A + [p - C]s_A + \upsilon_A \Pi_F + \Pi_A + \tau_A[y_A - s_A] - \theta_A[s_A + s_B - \gamma Z] - \iota_A \chi C$. We get coalition's fossil fuel supply and deposit demand correspondence

$$\text{(A.55)} \qquad s_A(t) \begin{cases} = 0, & \text{if } p(t) < C(t) + \tau_A(t) + \theta_A(t), \\ \in [0, S_A(t)], & \text{if } p(t) = C(t) + \tau_A(t) + \theta_A(t), \\ = S_A(t), & \text{if } p(t) > C(t) + \tau_A(t) + \theta_A(t), \end{cases}$$

$$\text{(A.56)} \qquad y_A(t) \begin{cases} = 0, & \text{if } p_y(t) > \tau_A(t), \\ \in \left[0, \min\{S(t), \bar{y}_A(t)\}\right], & \text{if } p_y(t) = \tau_A(t), \\ = \min\{S(t), \bar{y}_A(t)\}, & \text{if } p_y(t) < \tau_A(t), \end{cases}$$

If $p_y(t) < \tau_A(t)$, the coalition either buys the complete remaining fuel stock $S(t)$ or is constrained by its budget $\bar{y}_A(t)$. In the former case, $\tau_A(t) = p_y(t) + \zeta_Y(t)$ and in the later case $\tau_A(t) = [1 + \zeta_G(t)]p_y(t)$. The evolution of the coalition's scarcity rent is governed by

$$\text{(A.57)} \qquad \dot{\tau}_A = \rho \tau_A - \zeta_{S_A},$$

which yields the Hotelling-rule

$$\text{(A.58)} \qquad \tau_A(t) \begin{cases} = \tau_{A0} e^{\rho t} & \text{if } p(t) \leq C(t) + \tau_A(t) + \theta_A(t), \\ \leq \tau_{A0} e^{\rho t} & \text{if } p(t) > C(t) + \tau_A(t) + \theta_A(t). \end{cases}$$

The first order condition with respect to C and Z yield

$$\text{(A.59)} \qquad i_A = \iota[\rho + \chi] + [1 + \zeta_G]s_A,$$

(A.60) $$\dot{\theta}_A = \theta_A[\rho + \gamma] - \mu_A.$$

The transversality conditions read

(A.61) $$(a): \lim_{t\to\infty} e^{-\rho t}\tau_A(t)[S_A(t) - S_A^{opt}(t)] \geq 0, \quad (b): \lim_{t\to\infty} e^{-\rho t}\iota_A(t)[C(t) - C^{opt}(t)] \geq 0,$$

where the first ensures that no fuel is left in situ. Solving (A.59) and taking account of the second transversality condition gives $\iota_A(t) = -\int_t^T e^{-[\rho+\chi](j-t)}[1 + \zeta_G(j)]s_A(j)\,\mathrm{d}j$.

Proof of Proposition 3. Suppose that $\tau_A > p_y$ and $\bar{G} = \infty$, such that $\zeta_G = 0$ and $y_A(0) = S(0)$. By taking account of $U' = p$, the coalition's first-order conditions (A.53), (A.57), (A.59) and (A.60) are identical to the first-order conditions of the social planner (9)-(11) and (A.18), so that the coalition will implement the efficient solution. □

Proof of Proposition 4.

Lemma A.5. *The equilibrium on the deposit market at time $t \in [0, T_F)$ is given by $\tau_A(t) = p_y(t) = \tau_F(t)$ or $\tau_A(t) > p_y(t) = \tau_F(t)$. In the latter case, $y_A(t) = \bar{y}_A(t)$ for all $t \in [0, T_F)$.*

Proof of Lemma A.5. At first, suppose $p_y(t) > \max\{\tau_A(t), \tau_F(t)\}$ or $p_y(t) < \min\{\tau_A(t), \tau_F(t)\}$. In the former case, the coalition does not buy any deposits, while the firm wants to sell all deposits. In the latter case, the coalition wants to buy all remaining deposits but the firm does not sell any deposits. Both cases cannot be an equilibrium. Thus, $p_y(t) \in [\min\{\tau_A(t), \tau_F(t)\}, \max\{\tau_A(t), \tau_F(t)\}]$.

Suppose that $\tau_A(t) < \tau_F(t)$. Because the firm only sells deposits if $p_y(t) \geq \tau_F(t)$, (A.56) implies $y_A(t) = 0$. If the coalition has not acquired some deposit at time $j < t$, $y_A(t) = 0$ implies that fuel is only supplied by the firm and therefore, $p(t) = C(t) + \tau_F(t)$, $\hat{\tau}_A(t) \leq \rho$ and $\hat{\tau}_F(t) = \rho$. Consequently, $\tau_A < \tau_F$ holds in the next moment in time. This allows us the repeat the argument for all following points in time implying $y_A(t) = 0$ $\forall t$. However, the ceiling is violated without deposit acquisitions.

If the coalition bought some deposits at time $j < t$, (A.28) and (A.56) imply $\tau_A(j) \geq \tau_F(j)$. Thus, at $\tilde{j} \in (j, t)$ the growth rate of τ_A has to be lower than ρ, which requires that the coalition sold its complete fuel stock at once. If $\tilde{j} \in [0, t_1)$ or $\tilde{j} \in [t_2, T)$, this is not possible, because $\tau_A(\tilde{j}) \geq \tau_F(\tilde{j})$, and $\theta_A(\tilde{j}), \hat{\theta}_A(\tilde{j}) \geq 0$ imply that the coalition's supply price equals or exceeds the firm's price, so that $p(\tilde{j}) \leq C(\tilde{j}) + \tau_A(\tilde{j}) + \theta_A(\tilde{j})$ ruling out $s_A(\tilde{j}) = S_A(\tilde{j})$. If $\tilde{j} \in [t_1, t_2)$, $s_A(\tilde{j}) = S_A(\tilde{j})$ implies that the firm supplies fuel in the next moment of time. However, the Hotelling-rule (16) and $\dot{C} < 0$ are not compatible with a constant fuel supply of $\bar{s}$.[65] Consequently, $\tau_A(t) < \tau_F(t)$ cannot be part of an equilibrium. Thus, $p_y(t) \in [\tau_F(t), \tau_A(t)]$.

Suppose $\tau_F(t) < p_y(t) < \tau_A(t)$, so that the firm wants to sell all remaining deposits. If the coalition is not constrained by its funds, the firm can increase its deposit price up to $\tau_A(t)$ without loosing revenues.

Consider now the case with a binding budget constraint. On the fossil fuel market, $s_F(t) > 0$ and $s_A(t) = 0$ holds, because the coalition's supply price $p_A(t) = C(t) + \tau_A(t) + \theta_A(t)$ exceeds the firm's price $p_F(t) = C(t) + \tau_F(t)$. Consequently, $\hat{\tau}_A = \rho$ and $\hat{\tau}_F \leq \rho$, so that $\tau_F < \tau_A$ also holds in the next moment of time implying maximal deposit acquisition by the coalition. Therefore, the firm can increase its price up to $\tau_A(t)$ without loosing revenues. Consequently, $\tau_F(t) < p_y(t) < \tau_A(t)$ cannot be an equilibrium.

The remaining candidates for an equilibrium are

(i) $\tau_F(t) = p_y(t) = \tau_A(t)$,

(ii) $\tau_F(t) = p_y(t) < \tau_A(t)$,

(iii) $\tau_F(t) < p_y(t) = \tau_A(t)$.

Consider case (iii), where the firm wants to sell all remaining deposits, while the coalition is indifferent with respect to the amount of bought deposits at time t. If the coalition's funds are sufficiently high, it can buy all deposits and (iii) is a special case of (i), where both the coalition and the firm are indifferent. If the coalition's funds are not high enough, (iii) is not an equilibrium.

Consider case (ii), where the coalition applies a maximal acquisition regime. Because of $\theta_A \geq 0$, $\tau_F(t) < \tau_A(t)$ implies $s_A(t) = 0$ and $s_F(t) > 0$ and, therefore, $\hat{\tau}_A(t) = \hat{\tau}_F(t) = \rho$. We can rewrite $\tau_A(t) = [1 + \zeta_G(t)]\tau_F(t)$ as $\zeta_G(t) = \frac{\tau_A(0)}{\tau_F(0)} - 1 > 0$ for all $t \in [0, T_F)$, so that acquisition are maximal for all points in time until T_F. □

Lemma A.6. *The equilibrium on the fossil fuel market is characterized by* $s_F(t) > 0$, $s_A(t) =$

[65] See the proof of Lemma A.6.

0 *for* $t \in [0, T_F)$ *and* $s_F(t) = 0$, $s_A(t) > 0$ *for* $t \in [T_F, T)$, *and* $T_F \leq t_1$.

Proof of Lemma A.6. According to Lemma A.5, $\tau_A(t) \geq \tau_F(t)$ for all $t \in [0, T_F)$. Because $\theta_A(t) > 0$ for $t \in [0, t_2)$ and $\theta_A(t) = 0$ for $t \geq t_2$, we find the following relations of the fuel producer prices $p_A(t)$ and $p_F(t)$:

$$p_A(t) = C(t) + \tau_A(t) + \theta_A(t) > C(t) + \tau_F(t) = p_F(t) \qquad \text{for } t \in [0, t_2), \tag{A.62}$$

$$p_A(t) = C(t) + \tau_A(t) \geq C(t) + \tau_F(t) = p_F(t) \qquad \text{for } t \geq t_2. \tag{A.63}$$

$p_A(t) > p_F(t)$ for $t \in [0, t_2)$ implies $s_F(t) > 0$ and $s_A(t) = 0$ for $t \in [0, t_2)$ if $S(t) > 0$. During Phase II, $s(t) = \bar{s}$ implies a constant price $p(t) = \bar{p}$. Since $\dot{p_F} = \rho\tau_F(t) - \chi C(t)$ is zero at only one point in time and $s_F(t) > 0$ for $t \in [0, t_2)$ if $S(t) > 0$, $T_F \leq t_1$ must hold. □

Lemma A.7. *Fuel and backstop consumption per capita in the coalition is equal to fuel and backstop consumption per capita in the fringe for all* t.

Proof of Lemma A.7. The energy price is given by either $p_A(t)$ or $p_F(t)$ and equal in both countries. According to (13) and (17), $U'\left(\frac{x_i(t)+q_i(t)}{n_i}\right) = M'\left(\frac{q_i(t)}{n_i}\right) = p(t)$, $i = A, B$, so that per-capita consumptions are identical. □

Lemma A.8. *Fuel and total energy [backstop] consumption in both countries increase [decline] until* $t^s \in [0, t_1)$ *and decline [increase] for* $t \in (t^s, t_1)$, *they are constant during Phase II, and they decline [increase] during Phase III.*

Proof of Lemma A.8. During Phase I, A.62 implies $\dot{p}_A(t) = -\chi C(t) + \rho\tau_A(t) + (\rho+\gamma)\theta_A(t) > -\chi C(t) + \rho\tau_F(t) = \dot{p}_F(t)$. Consequently, if $\dot{p}_F(t) = 0$ during Phase I at $t = t_F^s$, then $\dot{p}_A(t) = 0$ during Phase I at $t = t_G^s < t_F^s$. Suppose $t_F^s \leq T_F$. Then, $\dot{p}_F(t) \lesseqgtr 0$ for $t \lesseqgtr t_F^s$ with $t_F^s < t_1$, and $\dot{p}_A(t) > 0$ for $t \in [T_F, t_1)$. Next, suppose $t_F^s > T_F$. Then, $\dot{p}_F(t) < 0$ for $t \in [0, T_F)$ and, applying Lemma A.1 for $t \in [T_F, t_1)$, $\dot{p}_A(t) \lesseqgtr 0$ for $t \lesseqgtr t_G^s$ with $t_G^s \in [T_F, t_1)$. Consequently, $\dot{p}(t) \lesseqgtr 0$ for $t \lesseqgtr \max\{T_F, t_G^s\}$ with $t_G^s < t_1$. $\dot{p}(t) = 0$ holds during Phase II. Note that Lemma A.2 and Lemma A.3 hold, so that $\theta_A(t)$ is continuous for $t \in [t_1, t_2)$ and $t^s < t_2$. Consequently, $\dot{p}_A(t) = -\chi C(t) + \rho\tau_A(t) > 0$ holds during Phase III. □

Lemma A.9. *At* $t = T_F$ *the energy price* $p(t)$ *is discontinuous and jumps from* $p_F(T_F)$ *to* $p_A(T_F)$.

Proof of Lemma A.9. Because of (A.62), the price jumps upwards when country A becomes the sole fuel supplier at $T_F \leq t_1$. The fuel firm may try to exploit the jump by withholding some fuel or by buying some deposits. In the first case, the firm tries to sell the fuel at $p_A(T_F)$. However, perfect competition on the fuel market implies then $p(T_F) = p_F(T_F)$.

In the second case, the firm buys deposits at $p_y(t)$ and sells them at time T_F for $\frac{\tau_A(T_F)+\theta_A(T_F)}{1+\zeta_G(T_F)}$. Then, perfect competition implies a deposit price of $p_y(T_F) = \frac{\tau_A(T_F)+\theta_A(T_F)}{1+\zeta_G(T_F)}$, so that the firm's fuel price at time $t \leq t_1$ reads

$$p_F(t) = C(t) + \frac{\tau_{A0}e^{\rho t} + \theta_{A0}e^{\rho t}e^{\gamma T_F}}{1+\zeta_G(T_F)} > C(t) + \tau_{F0}e^{\rho t}. \tag{A.64}$$

If the new price path implies a violation of the ceiling, the coalition buys deposits implying a price jump. Thus, the argument can be repeated until a violation is ruled out. However, the coalition has then no incentive to buy deposits. Then, Lemma A.10 implies that some fuel is left in situ, which is not optimal. □

The Lemmata A.5 to A.8 prove Proposition 4. □

Exhaustion date of the private fuel stock. Let T_F divide the planning horizon into two periods. For $t \geq T_F$, the coalition's value function is

$$\begin{aligned} V^S(S_A(T_F), Z(T_F), C(T_F)) = \max_{s_A} \int_0^\infty e^{-\rho t}\Big\{& U_A(x_A(t) + q_A(t)) + \bar{K}_A - p(t)x_A(t) \\ & - p_{q_A}(t)q_A(t) + [p(t) - C(t)]s_A(t) + \Pi_A(t)\Big\}\,\mathrm{d}t \\ & \text{subject to (3), (4), (5), and (22).} \end{aligned}$$

Assuming singular deposit acquisitions, the transversality conditions at time T_F are

$$(a): \tau_A(T_F) = \frac{\partial V^S}{\partial S_A}, \quad (b): -\theta_A(T_F) = \frac{\partial V^S}{\partial Z} \quad (c): \iota_A(T_F) = \frac{\partial V^S}{\partial C}, \tag{A.65}$$

$$H(T_F) = \rho V^S(S_A(T_F), Z(T_F), C(T_F)). \tag{A.66}$$

Taking account of (A.65), we get

$$\begin{aligned} \rho V^S(S_A(T_F), Z(T_F), C(T_F)) = {}& U_A(x_A(T_F^+) + q_A(T_F^+)) + \bar{K}_A - p(T_F^+)x_A(T_F^+) - p_{q_A}(T_F^+)q_A(T_F^+) \\ & [p(T_F^+) - C(T_F)]s_A(T_F^+) + \Pi_A(T_F^+) - \tau_A(T_F)s_A(T_F^+) \\ & - \theta_A(T_F)[s_A(T_F^+) - \gamma Z] - \iota_A(T_F)\chi C(T_F). \end{aligned}$$

Substituting into (A.66) yields

$$\begin{aligned}
&U_A(x_A(T_F^-)+q_A(T_F^-))-[C(T_F)+\tau_F(T_F)]x_A(T_F^-)-M_A(q_A(T_F^-))\\
&-y_A(T_F^-)[\tau_F(T_F)-\tau_A(T_F)]+\upsilon_A\tau_F(T_F)[s_F(T_F^-)+y_A(T_F^-)]-\theta_A(T_F)s_F(T_F^-)\\
&=U_A(x_A(T_F^+)+q_A(T_F^+))-[C(T_F)+\tau_A(T_F)+\theta_A(T_F)]x_A(T_F^+)-M_A(q_A(T_F^+)),
\end{aligned}\tag{A.67}$$

which determines the optimal T_F. If the left-hand side is greater [smaller] than the right-hand side, we get $T_F = T$ [$T_F = 0$].

For $\zeta_G = \zeta_Y = 0$, (A.67) can be written as

(A.68)

$$\begin{aligned}
&U_A(D_A(T_F^-)+Q_A(T_F^-))-p(T_F^-)D_A(T_F^-)-M_A(Q_A(T_F^-))-\left[p(T_F^+)-p(T_F^-)\right]D(T_F^-)\\
&+\upsilon_A\Pi_F(T_F^-)-\left[U_A(D_A(T_F^+)+Q_A(T_F^+))-p(T_F^+)D_A(T_F^+)-M_A(Q_A(T_F^+))\right]=0.
\end{aligned}$$

Differentiating the left-hand side of (A.68) with respect to $p(T_F^+)$ yields

$$\begin{aligned}
&-D(T_F^-)+D_A(T_F^+)-\underbrace{\left[U_A'(D_A(T_F^+)+Q_A(T_F^+))-p(T_F^+)\right]}_{=0}D_A'(T_F^+)\\
&-\underbrace{\left[U_A'(D_A(T_F^+)+Q_A(T_F^+))-M_A'(Q_A(T_F^+))\right]}_{=0}Q_A'(T_F^+)<0 \iff p(T_F^+)>p(T_F^-).
\end{aligned}$$

Since the left-hand side of (A.68) equals $\upsilon_A\Pi_F(T_F^-)>0$ for $p(T_F^+)=p(T_F^-)$ and the left-hand side of (A.68) decreases with $p(T_F^+)$ for $p(T_F^+)>p(T_F^-)$, $T_F>0$ with $p(T_F^+)>p(T_F^-)$ can be an equilibrium. In particular, $T_F>0$ with $p(T_F^+)>p(T_F^-)$ is an equilibrium if the deposit acquisitions are constant over time, i.e.

$$y_A(t)=\frac{1}{T_F}\left[S(0)-\int_0^{T_F}s_F(t)\,\mathrm{d}t\right],$$

such that $\upsilon_A\Pi_F(T_F^-)$ converges to infinity when T_F converges to zero, i.e.

$$\lim_{T_F\to 0}\upsilon_A\Pi_F(T_F^-)=\lim_{T_F\to 0}\upsilon_A\tau_F(T_F)\left\{D(T_F^-)+\frac{1}{T_F}\left[S(0)-\int_0^{T_F}s_F(t)\,\mathrm{d}t\right]\right\}=\infty.$$

Proof of Proposition 5.

Lemma A.10. *Suppose that $s(0)>s'(0)$, that the extraction paths intersect only once until $t=\max[t_1,t_1']$, and that the extraction paths do not intersect between $t=\min[t_2,t_2']$ and $t=\max[T,T']$. Then, $t_1<t_1'$ and $\int_0^{t_1'}s(t)\,\mathrm{d}t>\int_0^{t_1'}s'(t)\,\mathrm{d}t$.*

Proof of Lemma A.10. Suppose that $s(0) > s'(0)$ and that the extraction paths intersect only once until $t = \max[t_1, t_1']$ at $t = \tilde{t}$. Then, $Z(t) > Z'(t)$ for all $t \in [0, \tilde{t}]$ by $Z(t) = Z(0)e^{-\gamma t} + \int_0^t s(t)e^{\gamma t}\,\mathrm{d}te^{-\gamma t}$. If $t_2 \le t_1'$, then $t_1 < t_2 \le t_1'$. If $t_2 > t_1'$, then $s(t) < s'(t)$ for all $t \in (\tilde{t}, t_1']$ and $Z(t) \le Z'(t)$ for some $t \in (\tilde{t}, t_1')$ would imply $Z(t_1') < Z'(t_1') = \bar{Z}$ and, thus, $s(t) > s'(t)$ for some $t \in (t_1', t_1]$, so that the extraction paths would intersect twice. Consequently, $Z(t) > Z'(t)$ for all $t \in [0, t_1']$ and, thus, $t_1 < t_1'$ and $\int_0^{t_1'}[Z(t) - Z'(t)]\,\mathrm{d}t \ge 0$.

Suppose that the extraction paths do not intersect between $t = \min[t_2, t_2']$ and $t = \max[T, T']$. If $t_2 \le t_1'$, then $\int_{t_1'}^{T'}[s(t) - s'(t)]\,\mathrm{d}t < 0$ and, thus, $\int_0^{t_1'}[s(t) - s'(t)]\,\mathrm{d}t > 0$. If $t_2 > t_1'$, then $\int_0^{t_1'} \dot{Z}(t)\,\mathrm{d}t = \int_0^{t_1'} \dot{Z}'(t)\,\mathrm{d}t = \bar{Z} - Z(0)$, so that (4) becomes

$$\begin{aligned}\int_0^{t_1'}[s(t) - s'(t)]\,\mathrm{d}t &= \gamma \int_0^{t_1'}[Z(t) - Z'(t)]\,\mathrm{d}t + \int_0^{t_1'}[\dot{Z}(t) - \dot{Z}'(t)]\,\mathrm{d}t \\ &= \gamma \int_0^{t_1'}[Z(t) - Z'(t)]\,\mathrm{d}t > 0.\end{aligned} \tag{A.69}$$

□

According to Proposition 4, $s_A(t) = 0$, $s_F(t) > 0$ and $y_A(t) \ge 0$ for $t \in [0, T_F)$. Therefore, the price path is given by $p(t) = p_F(t) = C(t) + \tau_F(t)$, with $\hat{\tau}_F = \rho$. In the social optimum, the price path (marginal utility path) in Phase I reads $p(t) = C(t) + \tau(t) + \theta(t)$, with $\hat{\tau} = \rho$ and $\hat{\theta} = \rho + \gamma$. First, the social climate costs of emissions $\theta(t) > 0$ are missing under supply side policy. Second, the growth rates of the price paths are not identical.

Suppose that $\tau_{A0} \le \tau_0$. Then, $t_2 > t_2^*$ and $T > T^*$, so that $D(p_A(t)) = s_A(t) > s^*(t)$ for $t \in [t_2^*, T)$. At the ceiling, $s_A(t) = s^*(t) = \bar{s}$ for $t \in [\max\{t_1, t_1^*\}, t_2^*)$. Consequently, $\int_0^{\max\{t_1,t_1^*\}}[s_F(t) + s_A(t)]\,\mathrm{d}t < \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$. However, an equilibrium on the deposit market requires $\tau_{F0} \le \tau_{A0}$, so that $C_0 + \tau_{F0} < C_0 + \tau_0 + \theta_0$ and $\dot{C}(t) + \dot{\tau}_F(t) < \dot{C}(t) + \dot{\tau}(t) + \dot{\theta}_0$ hold, which implies $s_F(t) > s^*(t)$ for $t \in [0, T_F)$. Consequently, $s_A(t) < s^*(t)$ must hold for some $t \in [T_F, \max\{t_1, t_1^*\})$ to ensure $\int_0^{\max\{t_1,t_1^*\}}[s_F(t) + s_A(t)]\,\mathrm{d}t < \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$, which in turn implies $\theta_{A0} > \theta_0$. Consequently, $C(t) + \tau_A(t) + \theta_A(t) \ge C(t) + \tau(t) + \theta(t)$ implies $\dot{C}(t) + \dot{\tau}_A(t) + \dot{\theta}_A(t) > \dot{C}(t) + \dot{\tau}(t) + \dot{\theta}(t)$, so that the extraction paths intersect only once until $t = \max\{t_1, t_1^*\}$. According to Lemma A.10, this contradicts $\int_0^{\max\{t_1,t_1^*\}}[s_F(t) + s_A(t)]\,\mathrm{d}t < \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$. Consequently, $\tau_{A0} > \tau_0$, such that $t_2 < t_2^*$ and $T < T^*$ by $\bar{p} = C_0e^{-\chi t_2} + \tau_{A0}e^{\rho t_2} = C_0e^{-\chi t_2^*} + \tau_0e^{\rho t_2^*}$ and $M'\left(\frac{q_i(T)}{n_i}\right) = M'\left(\frac{q_i(T^*)}{n_i}\right) = C_0e^{-\chi T} + \tau_{A0}e^{\rho T} = C_0e^{-\chi T^*} + \tau_0e^{\rho T^*}$, respectively. Furthermore, $\int_{t_2}^{T^*} s(t)\,\mathrm{d}t < \int_{t_2}^{T^*} s^*(t)\,\mathrm{d}t$, which implies $\int_0^{\max\{t_1,t_1^*\}} s(t)\,\mathrm{d}t > \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$.

Suppose that $\tau_A(T_F)+\theta_A(T_F) \leq \tau(T_F)+\theta(T_F)$. Then, $\tau_{A0} > \tau_0$ implies $\theta_{A0} < \theta_0$ and, thus, $C(t)+\tau_A(t)+\theta_A(t) < C(t)+\tau(t)+\theta(t)$ for $t \in [T_F, \max\{t_1,t_1^*\})]$. If $\tau_{F0} \leq \tau_0+\theta_0$, then $s_A(t)+s_F(t) > s^*(t)$ would hold for $t \in [0, \max\{t_1,t_1^*\})$, and the ceiling would be violated. If $\tau_{F0} > \tau_0+\theta_0$, then the extraction paths would intersect only once until $t = \max\{t_1,t_1^*\}$, and $\int_0^{\max\{t_1,t_1^*\}} [s_F(t)+s_A(t)]\,\mathrm{d}t > \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$ would be contradicted by Lemma A.10. Consequently, $\tau_A(T_F)+\theta_A(T_F) > \tau(T_F)+\theta(T_F)$. If $\theta_{A0} \geq \theta_0$, then $\tau_A(t)+\theta_A(t) > \tau(t)+\theta(t)$ for all $t \in [0, \max\{t_1,t_1^*\})$.

Finally, suppose that $\tau_F(T_F) \geq \tau(T_F)+\theta(T_F)$, which implies $\tau_F(t) > \tau(t)+\theta(t)$ for all $t \in [0, T_F)$. Since $\tau_A(T_F)+\theta_A(T_F) > \tau(T_F)+\theta(T_F)$, the extraction paths would either never intersect and the ceiling would never bind, or the extraction paths would intersect only once until $t = \max\{t_1,t_1^*\}$, and $\int_0^{\max\{t_1,t_1^*\}} [s_F(t)+s_A(t)]\,\mathrm{d}t > \int_0^{\max\{t_1,t_1^*\}} s^*(t)\,\mathrm{d}t$ would be contradicted by Lemma A.10. Consequently, $\tau_F(T_F) < \tau(T_F)+\theta(T_F)$.

□

A.6. Strategic demand-side policy

Derivation of (28). The current-value Lagrangian of country A reads

$$\begin{aligned}L = {} & U_A(x_A+q_A(x_A)) + \bar{K}_A - [C+\tau_F]x_A - M_A(q_A(x_A)) \\ & + \upsilon_A\tau_F[x_A+D_B(C+\tau_F)] + \lambda\rho\tau_F - \kappa[x_A+D_B(C+\tau_F)] - \iota_A\chi C \\ & - \theta_A[x_A+D_B(C+\tau_F)-\gamma Z] + \mu_A[\bar{Z}-Z] + \zeta_x x_A.\end{aligned} \tag{A.70}$$

By definition, $x_A(t) = 0$ for all $t \geq T_A$, so that

$$V = \begin{cases} \int_{T_A}^{\infty} e^{-\rho(t-T_A)}[U_A(q_A(t)) + \bar{K}_A - M_A(q_A(t))]\,\mathrm{d}t = n_A\frac{\bar{U}}{\rho} & \text{if } T_A \geq T_B, \\ n_A\frac{\bar{U}}{\rho} + \upsilon_A\int_{T_A}^{T_B} e^{-\rho(t-T_A)}\tau_F(t)D_B(t)\,\mathrm{d}t & \text{if } T_A < T_B. \end{cases} \tag{A.71}$$

The first-order conditions give

$$\frac{\partial L}{\partial x_A} = U_A' - C - \tau_F + \upsilon_A\tau_F - \kappa - \theta_A + \zeta_x = 0, \tag{A.72}$$

$$\frac{\partial L}{\partial \tau_F} = -x_A + \upsilon_A[x_A+D_B] + \upsilon_A\tau_F D_B' + \rho\lambda - \kappa D_B' - \theta_A D_B' = \rho\lambda - \dot{\lambda}, \tag{A.73}$$

$$\frac{\partial L}{\partial S} = 0 = \rho\kappa - \dot{\kappa}, \tag{A.74}$$

$$\frac{\partial L}{\partial Z} = \theta_A \gamma - \mu_A = -\rho\theta_A + \dot{\theta}_A, \tag{A.75}$$

$$\frac{\partial L}{\partial C} = -x_A + \upsilon_A \tau_F D_B' - \kappa D_B' - \iota_A \chi - \theta_A D_B' = \rho\iota_A - \dot{\iota}_A. \tag{A.76}$$

The complementary slackness conditions are

$$\mu_A \geq 0,\ \ \mu_A[\bar{Z} - Z] = 0, \tag{A.77}$$

$$\zeta_x \geq 0, \quad \zeta_x x_A = 0. \tag{A.78}$$

and the transversality conditions read

$$(a)\ \lim_{t\to\infty} e^{-\rho t}\kappa(t)[S(t) - S^{opt}(t)] \geq 0, \qquad (b)\ \lim_{t\to\infty} e^{-\rho t}\theta_A(t)[-Z(t) + Z^{opt}(t)] \geq 0, \tag{A.79}$$

$$(a)\ \lim_{t\to\infty} e^{-\rho t}\lambda(t)[\tau_F(t) - \tau_F^{opt}(t)] \geq 0, \qquad (b)\ \lim_{t\to\infty} e^{-\rho t}\iota_A(t)[C(t) - C^{opt}(t)] \geq 0. \tag{A.80}$$

Solving (A.73), (A.74) and (A.76), and taking account of (A.80)(b) yield

$$\begin{aligned} \lambda(t) = & \int_0^t x_A(j)\,\mathrm{d}j + \int_0^t \kappa(j) D_B'(j)\,\mathrm{d}j + \int_0^t \theta_A(j) D_B'(j)\,\mathrm{d}j \\ & - \upsilon_A \int_0^t (x_A(j) + D_B(j))\,\mathrm{d}j - \upsilon_A \int_0^t \tau_F(j) D_B'(j)\,\mathrm{d}j, \end{aligned} \tag{A.81}$$

$$\kappa(t) = \kappa_0 e^{\rho t}, \tag{A.82}$$

$$\iota_A(t) = -\int_t^T e^{-[\rho+\chi](j-t)} \left[x_A(j) + \kappa(j) D_B'(j) + \theta_A(j) D_B'(j) - \upsilon_A \tau_F(j) D_B'(j)\right] \mathrm{d}j, \tag{A.83}$$

where $T = \max\{T_A, T_B\}$. T divides the planning horizon into two periods, where the coalition's value function for $t \geq T$ reads

$$V = \int_0^\infty e^{-\rho t}\left[U_A(q_A(t)) + \bar{K}_A - M_A(q_A(t))\right] \mathrm{d}t = \frac{n_A \bar{U}}{\rho}$$

By taking account of $\iota(T) = 0$ and $\lim_{t\to T} x_A(t) = \lim_{t\to T} D_B(t) = 0$, the transversality condition for T gives

$$\begin{aligned} H(T) = U_A(q_A(T)) + \bar{K}_A - M_A(q_A(T)) + \lambda(T)\rho\tau_F(T) = \rho V \\ \Leftrightarrow \lambda(T)\rho\tau_F(T) = 0. \end{aligned} \tag{A.84}$$

The substitution of (3), (16) and (A.81) - (A.83) and the consideration of $S(0) = \int_0^T [x_A(t) + D_B(t)]\,\mathrm{d}t$ yield

$$\kappa_0 = \upsilon_A \tau_{F0} + \frac{\int_0^{T_A} x_A(t)\,\mathrm{d}t - \upsilon S(0) + \int_0^{t_2} \theta_A(t) D_B'(t)\,\mathrm{d}t}{\left| \int_0^{T_B} e^{\rho t} D_B'(t)\,\mathrm{d}t \right|}. \tag{A.85}$$

Substituting $p_F(t) = C(t) + \tau_F(t)$, (A.82) and (A.85) into (A.72) yields (28).

Proof of Proposition 6. If $\mathrm{SE}(t) > [<]0$, then $U'_A(t) = C(t) + \tau_F(t) + \theta_A(t) + \mathrm{SE}e^{\rho t} > [<]C(t) + \tau_F(t) = p(t)$ in Phase I-III [Phase III], so that $\frac{x_A}{n_A} < [>]\frac{x_B}{n_B}$ in Phase I-III [Phase III].

Fuel and backstop consumption in both countries are continuous for all points in time if $C(t)$, $\tau_F(t)$, $\theta_A(t)$ and $\mathrm{SE}(t)$ are continuous for all points in time. $C(t)$ is continuous by assumption, $\tau_F(t)$ is continuous from (16), and $\mathrm{SE}(t) = \mathrm{SE}e^{\rho t}$ is continuous. For the continuous evolution of $\theta_A(t)$, the proof of Proposition 1 can be applied in a similar manner. Thus, Proposition 1(iv) continues to hold.

Consider Phase I. Applying Lemma A.1 reveals $\dot{s}(t_1) < 0$. If $\theta_{A0} + \mathrm{SE} \geq 0$, then $\dot{U}'_A(t) - \dot{p}(t) = (\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t} + \rho\mathrm{SE}e^{\rho t} > 0$, so that $t^s_A < t^s_B$ holds. Consequently, $\dot{x}_A \gtreqless 0$ for $t \lesseqgtr t^s_A$ with $t^s_A \in [0, \min\{t^s_B, t_1\})$.

Consider Phase II. If $t^s_B \in [0, t_1)$, then x_B decreases for $t \in [t_1, \min\{t_2, T_B\})$, so that x_A increases for $t \in [t_1, \min\{t_2, T_B\})$ and is constant for $t \in [\min\{t_2, T_B\}, t_2)$. By contrast, if $t^s_B > t_1$, then x_B increases for $t \in [t_1, t^s_B)$ and decreases for $t \in [t^s_B, T_B)$, so that x_A decreases for $t \in [t_1, \min\{t^s_B, t_2\})$. If $t^s_B < t_2$, then x_A increases for $t \in [t^s_B, \min\{t_2, T_B\})$ and is constant for $t \in [\min\{t_2, T_B\}, t_2)$.

Consider Phase III. $\ddot{U}'_A(t) > 0$ and $\ddot{p}(t) > 0$ hold, so that $\dot{U}'_A(t) > 0$ or $\dot{p}(t) > 0$ must hold for $t \geq t_2$ to ensure $\dot{s}(t_2) < 0$. If $\mathrm{SE} > [<]0$, then $\dot{U}'_A(t) - \dot{p}(t) = \rho\mathrm{SE}e^{\rho t} > [<]0$, so that $\dot{U}'_A(t) > 0$ [$\dot{p}(t) > 0$] must hold for $t \geq t_2$ implying $\dot{x}_A < 0$ [$\dot{x}_B < 0$] during Phase III. Consequently, $t^s_B = \frac{\ln\left(\frac{\chi C_0}{\rho\tau_{F0}}\right)}{\rho+\chi} \in [0, T_B)$ [$\in [0, t_2)$]. Thus, Proposition 1(iv) [iii] continues to hold.

Suppose $\mathrm{SE}(t) \geq 0$, such that $\frac{x_A}{n_A} < \frac{x_B}{n_B}$ in Phase I and II and $\frac{x_A}{n_A} \leq \frac{x_B}{n_B}$ in Phase III. If $n_A \leq \upsilon_A$, the terms-of-trade effect would then be negative, which contradicts $\mathrm{SE}(t) \geq 0$. Consequently, $n_A \leq \upsilon_A$ implies $\mathrm{SE}(t) < 0$.

Suppose $\chi = 0$, such that $\dot{x}_B(t) = D'_B(t)\rho\tau_{F0}e^{\rho t} < 0$ implies $\dot{x}_A(t) = D'_A(t)[\rho(\tau_{F0} + \mathrm{SE})e^{\rho t} + \dot{\theta}_A(t)] > 0$ in Phase II and, thus, $\dot{\theta}_A(t) < 0$ in Phase II, such that $\theta_A(t) \leq \theta_{A0}e^{(\rho+\gamma)t_1}$. Then, we have

$$\mathrm{SE}(t) = \mathrm{ToT}(t) - e^{\rho t}\frac{\int_0^{t_2} \theta_A(t)D'_B(t)\,\mathrm{d}t}{\int_0^{T} e^{\rho t}D'_B(t)\,\mathrm{d}t} > \mathrm{ToT}(t) - \theta_{A0}e^{\rho t+\gamma t_1}\frac{\int_0^{t_2} e^{\rho t}D'_B(t)\,\mathrm{d}t}{\int_0^{T} e^{\rho t}D'_B(t)\,\mathrm{d}t},$$

where $\mathrm{ToT}(t) := -\frac{\int_0^{T_A} x_A(t)\,\mathrm{d}t - \upsilon_A S(0)}{\int_0^{T} e^{\rho t}D'_B(t)\,\mathrm{d}t}$. Suppose $\theta_A(t_1) + \mathrm{SE}(t_1) \leq 0$, such that $\frac{x_A}{n_A} > \frac{x_B}{n_B}$ in Phase I-III. If $n_A \geq \upsilon_A$, the terms-of-trade effect would then be positive, which contradicts

$\theta_A(t_1) + \mathrm{SE}(t_1) \leq 0$. Consequently, $\chi = 0$ and $n_A \geq \upsilon_A$ implies $\theta_A(t_1) + \mathrm{SE}(t_1) > 0$. $\square$

Lemma A.11. *Suppose the coalition applies a demand-side climate policy, acts strategically in the fuel market and is committed to its strategy. With quadratic cost functions and linear demand functions, the climate costs are lower, the switch to Phase II occurs later and the switch to Phase III occurs earlier than without strategic action. Furthermore, a positive [negative] strategic effect implies that the scarcity rent is lower [higher], the coalition's switch to Phase IV occurs earlier [later] and the fringe's switch to Phase IV occurs later [earlier] than without strategic action. Finally, the initial fuel extraction is lower and cumulative fuel extraction* $\int s(t)\,\mathrm{d}t$ *is lower [higher] than without strategic action for* $t \in [0, t_1|_{SE\neq 0})$ *[*$t \in [t_2|_{SE\neq 0}, \max\{T_A, T_B\})$*].*

Proof of Lemma A.11. The equilibrium is characterized by

$$\begin{aligned} S(0) = &\int_0^{t_1} D_A\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + \theta_{A0} e^{(\rho+\gamma)t} + \mathrm{SE} e^{\rho t}\right) \mathrm{d}t \\ &+ \int_0^{t_1} D_B\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right) \mathrm{d}t + (t_2 - t_1)\gamma \bar{Z} \\ &+ \int_{t_2}^{T_A} D_A\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + \mathrm{SE} e^{\rho t}\right) \mathrm{d}t + \int_{t_2}^{T_B} D_B\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right) \mathrm{d}t, \end{aligned} \tag{A.86}$$

$$\begin{aligned} \bar{Z} = &Z(0) e^{-\gamma t_1} + \int_0^{t_1} D_A\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + \theta_{A0} e^{(\rho+\gamma)t} + \mathrm{SE} e^{\rho t}\right) e^{\gamma t}\, \mathrm{d}t e^{-\gamma t_1} \\ &+ \int_0^{t_1} D_B\left(C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right) e^{\gamma t}\, \mathrm{d}t e^{-\gamma t_1}, \end{aligned} \tag{A.87}$$

$$\gamma \bar{Z} = D_A\left(C_0 e^{-\chi t_1} + \tau_{F0} e^{\rho t_1} + \theta_{A0} e^{(\rho+\gamma)t_1} + \mathrm{SE} e^{\rho t_1}\right) + D_B\left(C_0 e^{-\chi t_1} + \tau_{F0} e^{\rho t_1}\right), \tag{A.88}$$

$$\gamma \bar{Z} = D_A\left(C_0 e^{-\chi t_2} + \tau_{F0} e^{\rho t_2} + \mathrm{SE} e^{\rho t_2}\right) + D_B\left(C_0 e^{-\chi t_2} + \tau_{F0} e^{\rho t_2}\right), \tag{A.89}$$

$$0 = D_A\left(C_0 e^{-\chi T_A} + \tau_{F0} e^{\rho T_A} + \mathrm{SE} e^{\rho T_A}\right), \tag{A.90}$$

$$0 = D_B\left(C_0 e^{-\chi T_B} + \tau_{F0} e^{\rho T_B}\right). \tag{A.91}$$

Differentiating with respect to SE and using $D_A(t_1) + D_B(t_1) = D_A(t_2) + D_B(t_2) = \gamma \bar{Z}$ and $D_A(T_A) = D_B(T_B) = 0$ yields

$$\begin{aligned} 0 = &\left\{ \int_0^{t_1} \left[D'_A(t) + D'_B(t)\right] e^{\rho t}\, \mathrm{d}t + \int_{t_2}^{T_A} D'_A(t) e^{\rho t}\, \mathrm{d}t + \int_{t_2}^{T_B} D'_B(t) e^{\rho t}\, \mathrm{d}t \right\} \tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} \\ &+ \int_0^{t_1} D'_A(t) e^{(\rho+\gamma)t}\, \mathrm{d}t \tfrac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}} + \int_0^{t_1} D'_A(t) e^{\rho t}\, \mathrm{d}t + \int_{t_2}^{T_A} D'_A(t) e^{\rho t}\, \mathrm{d}t, \end{aligned} \tag{A.92}$$

$$
\begin{aligned}
0 = &\int_0^{t_1} \left[D_A'(t)+D_B'(t)\right] e^{(\rho+\gamma)t}\,\mathrm{d}t e^{-\gamma t_1}\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + \int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t e^{-\gamma t_1}\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}} \\
&+ \int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t e^{-\gamma t_1},
\end{aligned} \tag{A.93}
$$

$$
\begin{aligned}
0 = &D_A'(t_1)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + 1 + \rho(\tau_{F0}+\mathrm{SE})\tfrac{\mathrm{d}t_1}{\mathrm{dSE}}\right]e^{\rho t_1} + D_B'(t_1)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + \rho\tau_{F0}\tfrac{\mathrm{d}t_1}{\mathrm{dSE}}\right]e^{\rho t_1} \\
&+ D_A'(t_1)\left[\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}} + (\rho+\gamma)\theta_{A0}\tfrac{\mathrm{d}t_1}{\mathrm{dSE}}\right]e^{(\rho+\gamma)t_1} - \left[D_A'(t_1)+D_B'(t_1)\right]\chi C_0\tfrac{\mathrm{d}t_1}{\mathrm{dSE}}e^{-\chi t_1},
\end{aligned} \tag{A.94}
$$

$$
\begin{aligned}
0 = &D_A'(t_2)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + 1 + \rho(\tau_{F0}+\mathrm{SE})\tfrac{\mathrm{d}t_2}{\mathrm{dSE}}\right]e^{\rho t_2} + D_B'(t_2)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + \rho\tau_{F0}\tfrac{\mathrm{d}t_2}{\mathrm{dSE}}\right]e^{\rho t_2} \\
&- \left[D_A'(t_2)+D_B'(t_2)\right]\chi C_0\tfrac{\mathrm{d}t_2}{\mathrm{dSE}}e^{-\chi t_2},
\end{aligned} \tag{A.95}
$$

$$
0 = D_A'(T_A)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + 1 + \rho(\tau_{F0}+\mathrm{SE})\tfrac{\mathrm{d}T_A}{\mathrm{dSE}}\right]e^{\rho T_A} - D_A'(T_A)\chi C_0\tfrac{\mathrm{d}T_A}{\mathrm{dSE}}e^{-\chi T_A}, \tag{A.96}
$$

$$
0 = D_B'(T_B)\left[\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + \rho\tau_{F0}\tfrac{\mathrm{d}T_B}{\mathrm{dSE}}\right]e^{\rho T_B} - D_B'(T_B)\chi C_0\tfrac{\mathrm{d}T_B}{\mathrm{dSE}}e^{-\chi T_B}. \tag{A.97}
$$

Solving (A.92)-(A.97) for $\frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}$, $\frac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}}$, $\frac{\mathrm{d}t_1}{\mathrm{dSE}}$, $\frac{\mathrm{d}t_2}{\mathrm{dSE}}$, $\frac{\mathrm{d}T_A}{\mathrm{dSE}}$ and $\frac{\mathrm{d}T_B}{\mathrm{dSE}}$ and using $D_A'(t_1)[\rho(\tau_{F0}+\mathrm{SE})e^{\rho t_1} + (\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t_1} - \chi C_0 e^{-\chi t_1}] + D_B'(t_1)[\rho\tau_{F0}e^{\rho t_1} - \chi C_0 e^{-\chi t_1}] = \dot{s}(t_1^-)$, $D_A'(t_2)[\rho(\tau_{F0}+\mathrm{SE})e^{\rho t_2} - \chi C_0 e^{-\chi t_2}] + D_B'(t_2)[\rho\tau_{F0}e^{\rho t_2} - \chi C_0 e^{-\chi t_2}] = \dot{s}(t_2^+)$, $\rho(\tau_{F0}+\mathrm{SE})e^{\rho T_A} - \chi C_0 e^{-\chi T_A} = U_A'(T_A^-)$ and $\rho\tau_{F0}e^{\rho T_B} - \chi C_0 e^{-\chi T_B} = U_B'(T_B^-)$ yields

$$
\begin{aligned}
\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} &= -\frac{\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t\left[\int_0^{t_1} D_A'(t)e^{\rho t}\,\mathrm{d}t + \int_{t_2}^{T_A} D_A'(t)e^{\rho t}\,\mathrm{d}t\right] - \left[\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2}{X\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t - \int_0^{t_1}\left[D_A'(t)+D_B'(t)\right]e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t}, \\
\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}} &= -\frac{X\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t - \int_0^{t_1}\left[D_A'(t)+D_B'(t)\right]e^{(\rho+\gamma)t}\,\mathrm{d}t\left[\int_0^{t_1} D_A'(t)e^{\rho t}\,\mathrm{d}t + \int_{t_2}^{T_A} D_A'(t)e^{\rho t}\,\mathrm{d}t\right]}{X\int_0^{t_1} D_A'(t)e^{(\rho+2\gamma)t}\,\mathrm{d}t - \int_0^{t_1}\left[D_A'(t)+D_B'(t)\right]e^{(\rho+\gamma)t}\,\mathrm{d}t\int_0^{t_1} D_A'(t)e^{(\rho+\gamma)t}\,\mathrm{d}t}, \\
\tfrac{\mathrm{d}t_1}{\mathrm{dSE}} &= -\frac{D_A'(t_1)[1 + e^{\gamma t_1}\tfrac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}}] + [D_A'(t_1)+D_B'(t_1)]\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{s}(t_1^-)e^{-\rho t_1}}, \\
\tfrac{\mathrm{d}t_2}{\mathrm{dSE}} &= -\frac{D_A'(t_2) + [D_A'(t_2)+D_B'(t_2)]\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{s}(t_2^+)e^{-\rho t_2}}, \\
\tfrac{\mathrm{d}T_A}{\mathrm{dSE}} &= -\frac{1 + \tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{U}_A'(T_A^-)e^{-\rho T_A}}, \\
\tfrac{\mathrm{d}T_B}{\mathrm{dSE}} &= -\frac{\tfrac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{U}_B'(T_B^-)e^{-\rho T_B}},
\end{aligned}
$$

where $X := \int_0^{t_1} \left[D_A'(t) + D_B'(t)\right] e^{\rho t}\,\mathrm{d}t + \int_{t_2}^{T_A} D_A'(t)e^{\rho t}\,\mathrm{d}t + \int_{t_2}^{T_B} D_B'(t)e^{\rho t}\,\mathrm{d}t$. For quadratic cost functions and linear demand functions, we have

(A.98)

$$\frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} = -\frac{n_A\left\{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + \int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2\right\}}{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + n_A\int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t + n_B\int_{t_2}^{T_B} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2} \in (-1,0),$$

(A.99)

$$\frac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}} = -\frac{n_B\frac{e^{\rho T_B}-e^{\rho T_A}}{\rho}\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t}{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + n_A\int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t + n_B\int_{t_2}^{T_B} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2},$$

(A.100)

$$\frac{\mathrm{d}t_1}{\mathrm{dSE}} = -\frac{\frac{n_A n_B\beta}{\dot{s}(t_1^-)e^{-\rho t_1}}\frac{e^{\rho T_B}-e^{\rho T_A}}{\rho}\int_0^{t_1} e^{(\rho+\gamma)t}[e^{\gamma t_1} - e^{\gamma t}]\,\mathrm{d}t}{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + n_A\int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t + n_B\int_{t_2}^{T_B} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2},$$

(A.101)

$$\frac{\mathrm{d}t_2}{\mathrm{dSE}} = \frac{\frac{n_A n_B\beta}{\dot{s}(t_2^+)e^{-\rho t_2}}\frac{e^{\rho T_B}-e^{\rho T_A}}{\rho}\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t}{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + n_A\int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t + n_B\int_{t_2}^{T_B} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2},$$

(A.102)

$$\frac{\mathrm{d}T_A}{\mathrm{dSE}} = -\frac{1 + \frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{U}_A'(T_A^-)e^{-\rho T_A}} < 0,$$

(A.103)

$$\frac{\mathrm{d}T_B}{\mathrm{dSE}} = -\frac{\frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}}{\dot{U}_B'(T_B^-)e^{-\rho T_B}} > 0.$$

The numerator and the denominator of $\frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}}$ are positive by the Cauchy-Schwarz inequality. The continuous evolution of $s(t)$ implies $\dot{s}(t_2^+) < 0$ and $\dot{U}_A'(T_A^-), \dot{U}_B'(T_B^-) > 0$. Consequently, $\frac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}}, \frac{\mathrm{d}t_2}{\mathrm{dSE}} \lesseqqgtr 0$ if and only if $T_B \gtreqqless T_A \Leftrightarrow \mathrm{SE} \gtreqqless 0$. Finally,

$$\begin{aligned}\frac{\mathrm{d}s(0)}{\mathrm{dSE}} &= -\beta\left[n_A + \frac{\mathrm{d}\tau_{F0}}{\mathrm{dSE}} + n_A\frac{\mathrm{d}\theta_{A0}}{\mathrm{dSE}}\right]\\ &= -\frac{n_A n_B\beta\frac{e^{\rho T_B}-e^{\rho T_A}}{\rho}\left[\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]}{\left[\int_0^{t_1} e^{\rho t}\,\mathrm{d}t + n_A\int_{t_2}^{T_A} e^{\rho t}\,\mathrm{d}t + n_B\int_{t_2}^{T_B} e^{\rho t}\,\mathrm{d}t\right]\int_0^{t_1} e^{(\rho+2\gamma)t}\,\mathrm{d}t - \left[\int_0^{t_1} e^{(\rho+\gamma)t}\,\mathrm{d}t\right]^2}. \qquad \text{(A.104)}\end{aligned}$$

Consequently, $\frac{\mathrm{d}s(0)}{\mathrm{dSE}} \lesseqqgtr 0$ if and only if $T_B \gtreqqless T_A \Leftrightarrow \mathrm{SE} \gtreqqless 0$. Since the extraction paths for $\mathrm{SE} = 0$ and $\mathrm{SE} \neq 0$ intersect only once during Phase I with quadratic cost functions and linear demand functions, $s(0)|_{\mathrm{SE}=0} > s(0)|_{\mathrm{SE}\neq 0}$ implies $t_1|_{\mathrm{SE}=0} < t_1|_{\mathrm{SE}\neq 0}$ and $\int_0^{t_1|_{\mathrm{SE}\neq 0}} s(t)|_{\mathrm{SE}=0}\,\mathrm{d}t > \int_0^{t_1|_{\mathrm{SE}\neq 0}} s(t)|_{\mathrm{SE}\neq 0}$ by Lemma A.10 and, thus, $\int_{t_2|_{\mathrm{SE}\neq 0}}^{\max\{T_A,T_B\}} s(t)|_{\mathrm{SE}=0}\,\mathrm{d}t < \int_{t_2|_{\mathrm{SE}\neq 0}}^{\max\{T_A,T_B\}} s(t)|_{\mathrm{SE}\neq 0}$. □

Time consistency. Suppose that the strategic effects in (28) are negative. Then, for all $t \in [T_B, T_A)$ country A uses fuel although the fuel price $p(t) = C(t) + \tau_F(t)$ exceeds the marginal backstop costs of $\tilde{q}_A$ given by $M'\left(\frac{\tilde{q}_A}{n_A}\right)$. If country A cannot commit to its strategy, it has an incentive to reevaluate its policy and abandon fuel use for $t \geq T_B$. If the strategic effects are positive, country A abandons fuel use at $T_A < T_B$. However, for $t \in [T_A, T_B)$ fuel consumption would reduce energy costs below $M'\left(\frac{\tilde{q}_A}{n_A}\right)$, so that country A has an incentive to extend its fuel utilization period.

A.7. Strategic supply-side policy

Derivation of (30)-(34). The current-value Lagrangian of country A reads

$$\begin{aligned}
L = {} & U_A(D_A(p) + Q_A(p)) + \bar{K}_A - pD_A(p) - M_A(Q_A(p)) + [p - C]s_A \\
& + \upsilon_A[p - C][D(p) - s_A] - [1 - \upsilon_A]p_y y_A + \lambda\rho\tau_F + \tau_A[y_A - s_A] - \iota_A\chi C \\
& - \kappa[D(p) + y_A - s_A] - \theta_A[D(p) - \gamma Z] + \mu_A[\bar{Z} - Z] \\
& + \zeta_{y_A} y_A + \zeta_{s_A} s_A + \zeta_Y[S - y_A] + \zeta_{S_A}[S_A - s_A] + \zeta_G\Big[\bar{K}_A - M_A(Q_A(p)) \\
& - pD_A(p) - \bar{g}_A + [p - C]s_A + \upsilon_A[p - C][D(p) - s_A] - [1 - \upsilon_A]p_y y_A\Big].
\end{aligned} \tag{A.105}$$

By definition, $p(T) = M'\left(\frac{q_i(T)}{n_i}\right)$ and $x_A(t) = x_B(t) = 0$ for all $t \geq T$, so that the coalition's value function for $t \geq T$ reads

$$V = \int_0^\infty e^{-\rho t}\left[U_A(q_A(t)) + \bar{K}_A - M_A(q_A(t))\right]\mathrm{d}t = n_A\frac{\bar{U}}{\rho}. \tag{A.106}$$

The Lagrangian (A.105) is linear in y_A, so that the bang-bang solution satisfies

$$\frac{\partial L}{\partial y_A} = -[1 - \upsilon_A]p_y + \tau_A - \kappa + \zeta_{y_A} - \zeta_Y - \zeta_G[1 - \upsilon_A]p_y = 0, \tag{A.107}$$

and maximizes $H = U_A(D_A(p) + Q_A(p)) + \bar{K}_A + pD_A(p) - M_A(Q_A(p)) + [p - C]s_A + \upsilon_A[p - C][D(p) - s_A] - [1 - \upsilon_A]p_y y_A + \lambda\rho\tau_F + \tau_A[y_A - s_A] - \kappa[D(p) + y_A - s_A] - \theta_A[D(p) - \gamma Z] - \iota_A\chi C$, which gives the deposit demand correspondence

$$
\text{(A.108)} \qquad y_A(t) \begin{cases} = 0 & \text{if } p_y > \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}, \\ \in [0, \min\{S(t), \bar{y}_A(t)\}] & \text{if } p_y = \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}, \\ = \min\{S(t), \bar{y}_A(t)\} & \text{if } p_y < \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}. \end{cases}
$$

The first-order conditions with respect to S, S_A and Z yield

$$
\text{(A.109)} \qquad \frac{\partial L}{\partial S} = \zeta_Y = \rho\kappa - \dot{\kappa},
$$

$$
\text{(A.110)} \qquad \frac{\partial L}{\partial S_A} = \zeta_{S_A} = \rho\tau_A - \dot{\tau}_A,
$$

$$
\text{(A.111)} \qquad \frac{\partial L}{\partial Z} = \theta_A\gamma - \mu_A = -\rho\theta_A + \dot{\theta}_A,
$$

while the corresponding transversality conditions read

$$
\text{(A.112)} \qquad (a)\ \lim_{t\to\infty} e^{-\rho t}\kappa(t)[S(t) - S^{opt}(t)] \geq 0, \qquad (b)\ \lim_{t\to\infty} e^{-\rho t}\tau_A(t)[S_A(t) - S_A^{opt}(t)] \geq 0,
$$

$$
\text{(A.113)} \qquad \lim_{t\to\infty} e^{-\rho t}\theta_A(t)[-Z(t) + Z^{opt}(t)] \geq 0.
$$

The complementary slackness conditions are

$$
\text{(A.114)} \qquad \zeta_Y \geq 0,\ \zeta_Y[S - y_A] = 0,
$$

$$
\text{(A.115)} \qquad \zeta_{S_A} \geq 0,\ \zeta_{S_A}[S_A - s_A] = 0,
$$

$$
\text{(A.116)} \qquad \zeta_{s_A} \geq 0,\ \zeta_{s_A} s_A = 0,
$$

$$
\text{(A.117)} \qquad \zeta_{y_A} \geq 0,\ \zeta_{y_A} y_A = 0,
$$

$$
\text{(A.118)} \qquad \begin{aligned} &\zeta_G \geq 0, \\ &\zeta_G\Big[\bar{K}_A - M_A(Q_A) - pD_A + [p - C]s_A + \upsilon_A[p - C][D - s_A] - [1 - \upsilon_A]p_y y_A\Big] = 0. \end{aligned}
$$

Consider the case $s_F(t) > 0$, so that $p(t) = p_F(t) = C(t) + \tau_F(t)$. The first-order conditions with respect to s_A , τ_F and C yield

$$
\text{(A.119)} \qquad \frac{\partial L}{\partial s_A} = [1 - \upsilon_A]\tau_F - \tau_A + \kappa + \zeta_{s_A} - \zeta_{S_A} + \zeta_G[1 - \upsilon_A]\tau_F = 0,
$$

$$
\text{(A.120)} \qquad \begin{aligned} \frac{\partial L}{\partial \tau_F} = &- D_A + s_A + \upsilon_A[D - s_A] + \upsilon_A\tau_F D' - [1 - \upsilon_A]y_A + \rho\lambda - \kappa D' - \theta_A D' \\ &+ \zeta_G\Big\{ - p[Q_A' + D_A'] - D_A + s_A + \upsilon_A[D - s_A] + \upsilon_A D'\tau_F - [1 - \upsilon_A]y_A\Big\} \\ &= \rho\lambda - \dot{\lambda} \end{aligned}
$$

$$\Leftrightarrow \dot{\lambda} = [1+\zeta_G]\left[D_A - s_A - \upsilon_A s_F - \upsilon_A \tau_F D' + [1-\upsilon_A] y_A\right] + [\kappa + \theta_A] D' + \zeta_G [C + \tau_F][Q'_A + D'_A] \tag{A.121}$$

$$\frac{\partial L}{\partial C} = -D_A + \upsilon_A \tau_F D' - \iota_A \chi - \kappa D' - \theta_A D' + \zeta_G \Big\{ -p[Q'_A + D'_A] - D_A + \upsilon_A \tau_F D' \Big\} = \rho \iota_A - i_A \tag{A.122}$$

$$\Leftrightarrow i_A = [\rho + \chi]\iota_A + [1+\zeta_G][D_A - \upsilon_A \tau_F D'] + [\kappa + \theta_A] D' + \zeta_G [C + \tau_F][Q'_A + D'_A] \tag{A.123}$$

Because (A.105) is linear in s_A if $p(t) = p_F(t)$, the optimal $s_A(t)$ is given by the coalition's fuel supply correspondence

$$s_A(t) \begin{cases} = 0 & \text{if } \tau_F(t) < \frac{\tau_A(t) - \kappa(t)}{1-\upsilon_A}, \\ \in [0, S_A(t)] & \text{if } \tau_F(t) = \frac{\tau_A(t) - \kappa(t)}{1-\upsilon_A}, \\ = S_A(t) & \text{if } \tau_F(t) > \frac{\tau_A(t) - \kappa(t)}{1-\upsilon_A}. \end{cases} \tag{A.124}$$

Consider the case $s_F(t) = 0$. The first-order conditions with respect to s_A, τ_F and C yield

$$\frac{\partial L}{\partial s_A} = -D_A \frac{\mathrm{d}p}{\mathrm{d}s_A} + s_A \frac{\mathrm{d}p}{\mathrm{d}s_A} + p - C - \tau_A - \theta_A + \zeta_{s_A} - \zeta_{S_A} + \zeta_G \Big\{ -p[Q'_A + D'_A] \frac{\mathrm{d}p}{\mathrm{d}s_A} - D_A \frac{\mathrm{d}p}{\mathrm{d}s_A} + s_A \frac{\mathrm{d}p}{\mathrm{d}s_A} + p - C \Big\} = 0 \tag{A.125}$$

$$\frac{\partial L}{\partial \tau_F} = -[1-\upsilon_A] y_A + \rho \lambda - \zeta_G [1-\upsilon_A] y_A = \rho \lambda - \dot{\lambda} \tag{A.126}$$

$$\Leftrightarrow \dot{\lambda} = [1+\zeta_G][1-\upsilon_A] y_A, \tag{A.127}$$

$$\frac{\partial L}{\partial C} = -[1+\zeta_G] s_A - \iota_A \chi = \rho \iota_A - i_A \tag{A.128}$$

$$\Leftrightarrow i_A = [\rho + \chi]\iota_A + [1+\zeta_G] s_A, \tag{A.129}$$

where $\frac{\mathrm{d}p}{\mathrm{d}s_A} = \frac{1}{D'}$. From (A.125) we get the coalition's supply price

$$p_A = \frac{[1+\zeta_G]C + \tau_A + \theta_A - [1+\zeta_G]\frac{D_B}{D'}}{1 + \zeta_G - \zeta_G \frac{Q'_A + D'_A}{D'}}. \tag{A.130}$$

Using (20) yields (30) and (34). Suppose $S(t) > 0$. Then, the coalition can only sell fuel at the price $p_A(t)$, if $p_A(t) < C(t) + \tau_F(t) \leq C(t) + \frac{\tau_A(t) - \kappa(t)}{1-\upsilon_A}$, which yields $C(t)\zeta_G(t)\frac{Q'_A(t) + D'_A(t)}{D'(t)} +$

$\tau_A(t) + \theta_A(t) - [1+\zeta_G(t)]\frac{D_B(t)}{D'(t)} < \frac{[\tau_A(t)-\kappa(t)]\left[1+\zeta_G(t)-\zeta_G(t)\frac{Q'_A(t)+D'_A(t)}{D'(t)}\right]}{1-\upsilon_A}$. For $\upsilon_A \to 1$ and $\tau_A(t) - \kappa(t) > 0$, the inequality holds.

For both cases $s_F(t) > 0$ and $s_F(t) = 0$, the remaining transversality conditions read

$$\text{(A.131)} \qquad (a)\ \lim_{t\to\infty} e^{-\rho t}\lambda(t)[\tau_F(t) - \tau_F^{opt}(t)] \geq 0, \qquad (b)\ \lim_{t\to\infty} e^{-\rho t}\iota_A(t)[C(t) - C^{opt}(t)] \geq 0.$$

Solving (A.121) and (A.127) yields

$$\begin{aligned} \lambda(T) = [1-\upsilon_A]\int_0^T [1+\zeta_G(t)]y_A(t)\,\mathrm{d}t + \sum_{i=I,III}\int_{t_a^i}^{t_b^i}\Big\{&\theta_A(t)D'(t) \\ &+ \kappa(t)D'(t) + [1+\zeta_G(t)][D_A(t) - s_A(t)] - [1+\zeta_G(t)]\upsilon_A\tau_F(t)D'(t) \\ &- [1+\zeta_G(t)]\upsilon_A s_F(t) + \zeta_G(t)[C(t)+\tau_F(t)][Q'_A(t) + D'_A(t)]\Big\}\,\mathrm{d}t, \end{aligned} \tag{A.132}$$

Solving (A.123) and (A.129), and taking account of (A.131) yields

$$\iota_A(t) = -\int_t^T e^{-[\rho+\chi](j-t)}[1+\zeta_G(j)]s_A(j)\,\mathrm{d}j \tag{A.133}$$

for $t > t_b^{III}$, so that $\iota_A(T) = 0$. Taking advantage of $T > t_2 \Rightarrow \theta_A(T) = 0$ and $\lim_{t\to T} s_A(T) = \lim_{t\to T} s_F(T) = \lim_{t\to T} y_A(T) = 0$, the optimal T is determined by

$$\begin{aligned} H(T) = U_A(Q_A(T)) + \bar{K}_A - M_A(Q_A(T)) + \lambda(T)\rho\tau_F(T) = \rho V \\ \Leftrightarrow \lambda(T)\rho\tau_F(T) = 0. \end{aligned} \tag{A.134}$$

Substituting (A.132) and taking account of $D_A(t) - s_A(t) = s_F(t) - D_B(t)$, $\zeta_G(t) = \zeta_G[0]$ for $t < [\geq]T_F$, and $\sum_{i=I,III}\int_{t_a^i}^{t_b^i} s_F(t)\,\mathrm{d}t = S(0) - \int_0^T y_A(t)\,\mathrm{d}t$ yield

$$\begin{aligned} \kappa_0 = \upsilon_A\tau_{F0} &+ \frac{[1+\zeta_G]\left[\upsilon_B S(0) - \sum_{i=I,III}\int_{t_a^i}^{t_b^i} D_B(t)\,\mathrm{d}t\right] + \int_{t_a^I}^{t_b^I}\theta_A(t)D'(t)\,\mathrm{d}t}{\left|\sum_{i=I,III}\int_{t_a^i}^{t_b^i} e^{\rho t}D'(t)\,\mathrm{d}t\right|} \\ &+ \zeta_G\left\{\upsilon_A\tau_{F0} + \frac{\sum_{i=I,III}\int_{t_a^i}^{t_b^i}\left[C(t)+\tau_F(t)\right]\frac{n_A}{U''(t)}\,\mathrm{d}t}{\left|\sum_{i=I,III}\int_{t_a^i}^{t_b^i} e^{\rho t}D'(t)\,\mathrm{d}t\right|}\right\}. \end{aligned} \tag{A.135}$$

Substituting into (A.108) and (A.124) and using (20) yield (31) and (32), where we haven taken account of Proposition 8.

Proof of Proposition 7. Suppose that $\frac{\tau_A-\kappa_A}{1-\upsilon_A} > p_y$ and $\bar{G} = \infty$, such that $\zeta_G = 0$ and $y_A(0) = S(0)$. Because of $S(t) = 0$ for all points in time, $s_F(t) = 0$ and $s_A(t) > 0$ for all $t \in [0, T)$. Because of the monopoly effect, the first-order conditions (9) and (A.125) are not identical, so that the coalition does not implement the social optimum.

The coalition's fuel supply price is given by (30), so that $U_i' = U'\left(\frac{x_i+q_i}{n_i}\right) = M'\left(\frac{q_i}{n_i}\right) = p_A$ implies $\frac{x_A(t)}{n_A} = \frac{x_B(t)}{n_B}$ and $\frac{q_A(t)}{n_A} = \frac{q_B(t)}{n_B}$ for all $t \in [0, T)$. For $t \geq T$, $M'\left(\frac{q_i(t)}{n_i}\right) = M'\left(\frac{\tilde{q}_i}{n_i}\right)$ implies $\frac{q_A(t)}{n_A} = \frac{q_B(t)}{n_B}$.

$C(t)$ is continuous by definition. Applying the proof of Lemma A.2 to τ_A and θ_A shows that these costates are continuous. Finally, the monopoly effect is a function of the price path $p_A(t)$. Thus, all elements of (30) which depend on time are continuous implying a continuous evolution of $p_A(t)$ and, therefore, of fuel and backstop consumption.

Consider Phase I. Differentiating (30) yields $\dot{p}_A = \frac{-\chi C(t)+\rho\tau_A+[\rho+\gamma]\theta_A}{1-\frac{n_B}{\epsilon}\left[1-\frac{\partial\epsilon}{\partial p}\frac{p}{\epsilon}\right]}$. The denominator is positive because $\frac{\partial\epsilon}{\partial p} \geq 0$ and $p_A(t) > 0$ implies $1 - \frac{n_B}{\epsilon} > 0$. The nominator is negative [positive] if the technology effects dominates [is dominated]. Applying Lemma A.1 implies $t^s < t_1$. During Phase II, fuel consumption is constant. During Phase III, fuel consumption declines, because of $t^s < t_1$ and the constant growth rates of $\tau_A(t)$ and $C(t)$. □

Proof of Proposition 8.

Lemma A.12. *The equilibrium on the deposit market at time $t \in [0, t_2)$ [$t \in [0, T_F)$] is given by $\tau_F(t) = p_y(t) = \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$ or $\tau_F(t) = p_y(t) < \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$[, if $\zeta_{S_A}(t) = 0\ \forall t$].*

Proof of Lemma A.12. The proof applies arguments used for the proof of Lemma A.5. Therefore, we give only a sketch, where identical arguments are used.

Neither $p_y(t) < \min\left\{\frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}, \tau_F(t)\right\}$ nor $p_y(t) > \max\left\{\frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}, \tau_F(t)\right\}$ can be an equilibrium, because minimal demand meets maximal supply or vice versa. Thus, $p_y(t) \in [\min\{\tau_A(t), \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}\}, \max\{\frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}, \tau_F(t)\}]$.

Suppose $\tau_F(t) > \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$. Because $p_y(t) \geq \tau_F(t)$ for positive deposit supply, (A.108) implies $y_A(t) = 0$. If the coalition has not bought any deposits before t, (A.28), (A.109) and (A.110) imply $\tau_F(j) > \frac{\tau_A(j)-\kappa(j)}{1-\upsilon_A}$ and, therefore, $y_A(j) = 0$ for all $j > t$, so that the ceiling is violated.

If the coalition has bought some deposit before t, (A.108) and (A.110) imply $s_A(\tilde{j}) = S_A(\tilde{j})$

at some $\tilde{j} < t$ and $y_A(j) = 0$ for all $j > \tilde{j}$. However, $s(t) = \bar{s}$ for $t \in [t_1, t_2)$ and $\dot{p}_F(t) \neq 0$ for all t but one moment in time rule out $\tilde{j} < t_2$. If $\zeta_{S_A}(t) = 0 \; \forall t$, $s_A(\tilde{j}) = S_A(\tilde{j})$ is generally ruled out. Thus, $p_y(t) \in [\tau_F(t), \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}]$.

Suppose $\frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A} > p_y(t) > \tau_F(t)$. If $S(t) < \bar{y}_A(t)$, the fuel firm can increase $\tau_F(t)$ up to $\frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$ without loosing revenues. If the budget constraint binds, the firm can increase $\tau_F(t)$ up to $\frac{\tau_A(t)-\kappa(t)}{[1-\upsilon_A][1+\zeta_G(t)]}$. Then, either $\frac{\tau_A(t^+)-\kappa(t^+)}{1-\upsilon_A} > \tau_F(t^+)$ or $S_A(t) = s_A(t)$ holds, so that $y_A(t^+) = \bar{y}_A(t^+)$ or the firm sells fuel at the price $p_F(t^+) = C(t^+) + \tau_F(t^+)$. Consequently, $\tau_F(t) = \frac{\tau_A(t)-\kappa(t)}{[1-\upsilon_A][1+\zeta_G(t)]}$ does not imply a loss of revenues.

$\tau_F(t) = p_y(t) = \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$, $\tau_F(t) = p_y(t) < \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$ and $\tau_F(t) < p_y(t) = \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$ remain as possible equilibria, where $y_A(t) \in [0, \min\{S(t), \bar{y}_A(t)\}]$ and $y_F(t) \in [0, S(t)]$ in the first case, and $y_A(t) = S(t) \in [0, \min\{S(t), \bar{y}_A(t)\}]$ in the third case, so the latter is a special case of the former. $\square$

Lemma A.13. $\zeta_{S_A}(t) = 0 \; \forall t$.

Proof of Lemma A.13. Suppose that $s_F(t) = 0$. Then, the fuel price is given by (A.130) implying an interior solution. Suppose that $s_F(t) > 0$. Then, (A.124) and (A.135) require $\tau_F(t) > \tau_A(t) - \mathrm{SE}(t)$ for $s_A(t) = S_A(t)$. However, $y_A(t) > 0$ for some $t < t_2$, Lemma A.12 and (A.108) imply $\tau_F(t) \leq \tau_A(t) - \mathrm{SE}(t)$. Finally, $\hat{\tau}_F(t) \leq \rho$, $\hat{\mathrm{SE}} = \rho$ and $\hat{\tau}_A = \rho$ as long as $\tau_F(t) \leq \tau_A(t) - \mathrm{SE}(t)$ holds or the fuel price is given by (A.130). $\square$

Lemma A.14. $\zeta_G(t) >$ [=] 0 *for some* $t \in [0, T_F)$ *implies that* $\zeta_G(t)$ *is a constant [zero] for all* $t \in [0, T_F)$.

Proof of Lemma A.14. For $t \in [0, T_F)$, (A.28) and (A.109) imply $\hat{\tau}_F = \hat{\kappa} = \rho$, and (A.110) yields $\hat{\tau}_A = \rho$. If $y_A(t) = \bar{y}_A(t)$, (A.107), (A.108) and Lemma A.12 imply $\tau_F(t) = \frac{\tau_A(t)-\kappa(t)}{[1-\upsilon_A][1+\zeta_G(t)]}$ and, therefore

$$\zeta_G(t) = \frac{\tau_{A0} - \kappa_0}{[1-\upsilon_A]\tau_{F0}} - 1, \tag{A.136}$$

which is a constant for all $t \in [0, T_F)$. If $y_A(t) < \bar{y}_A(t)$, we get $\tau_{F0} = \frac{\tau_{A0}-\kappa_0}{1-\upsilon_A}$. $\square$

Lemma A.15. *The coalition's supply side policy is not efficient.*

Proof of Lemma A.15. From Lemmata A.12 and A.13 we know that $y_A(t) = s_F(t) = 0$ for all $t \geq T_F$, so that the fuel price is governed by (A.125). Due to the monopoly effect, the equation deviates from (9), so that the coalition's supply side policy is not efficient. $\square$

Lemma A.16. *Suppose that $p(t) = p_F(t)$ and $s_A(t) \in [0, S_A(t)]$. Then, $y_A(t) < \bar{y}_A(t)$.*

Proof of Lemma A.16. $s_A(t) \in [0, S_A(t)]$ requires $\tau_F(t) = \frac{\tau_A(t)-\kappa(t)}{1-\upsilon_A}$, while (A.119) yields $\tau_F(t) = \frac{\tau_A(t)-\kappa(t)}{[1-\upsilon_A][1+\zeta_G(t)]}$. The equality of both terms implies $\zeta_G(t) = 0$. $\square$

Lemma A.17. *Fuel and backstop consumption per capita in both countries coincide for all points in time. Fuel and backstop consumption in both countries is constant during Phase II. Suppose $\chi = 0$ and $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$. Then, fuel and total energy [backstop] consumption in both countries declines [increases] over time during Phase I and III. Suppose $\chi > 0$ and $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$. Then, there is at most one switching time t_G^s during Phase I and III.*

Proof of Lemma A.17. The fuel demand per capita $D(p(t))$ and the fuel price $p(t)$ are the same in both countries, which proves $\frac{x_A(t)}{n_A} = \frac{x_B(t)}{n_B}$. During Phase II, $p(t) = \bar{p}$ to adhere the ceiling, which proves $\dot{x}_i(t) = 0$ during Phase II.

For $s_F > 0$ and $\chi = 0$, the price path is given by $p_F(t) = c + \tau_{F0}e^{\rho t}$ and $\dot{p}_F(t) = \rho\tau_{F0}e^{\rho t} > 0$, which proves $\dot{x}_i(t) < 0$ for $s_F > 0$ and $\chi = 0$. During Phase I and III for $s_F = 0$, the price path from (34) and $\epsilon = -\frac{D'(p)p}{D(p)}$ is given by

$$p_A(t) = \frac{\tau_{A0}e^{\rho t} + \theta_{A0}e^{(\rho+\gamma)t} + [1+\zeta_G]\left[C_0 e^{-\chi t} - n_B \frac{D(p_A(t))}{D'(p_A(t))p_A(t)} p_A(t)\right]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}}$$

$$= \frac{\tau_{A0}e^{\rho t} + \theta_{A0}e^{(\rho+\gamma)t} + [1+\zeta_G]C_0 e^{-\chi t}}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} - [1+\zeta_G]\frac{n_B}{\epsilon(t)}}. \tag{A.137}$$

Differentiating with respect to t yields

$$\dot{p}_A(t) = \frac{\rho\tau_{A0}e^{\rho t} + (\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t} - [1+\zeta_G]\left[\chi C_0 e^{-\chi t} + n_B \frac{[D'(t)]^2 - D(t)D''(t)}{[D'(t)]^2}\dot{p}_A\right]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}}$$
$$+ \frac{\zeta_G n_A \frac{Q''(t)D'(t) - Q'(t)D''(t)}{[D'(t)]^2} p_A \dot{p}_A}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}}$$

(A.138)

$$= \frac{\rho\tau_{A0}e^{\rho t} + (\rho+\gamma)\theta_{A0}e^{(\rho+\gamma)t} - [1+\zeta_G]\chi C_0 e^{-\chi t}}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} - [1+\zeta_G]\frac{n_B}{\epsilon(t)}\left[1 - \frac{\partial\epsilon(t)}{\partial p}\frac{p_A}{\epsilon(t)}\right] - \zeta_G n_A \frac{Q''(t)D'(t)-Q'(t)D''(t)}{[D'(t)]^2}p_A},$$

where

$$Q''(t)D'(t) - Q'(t)D''(t) = n_i^2 \frac{\left[M''(t)\right]^2 U'''(t) - \left[U''(t)\right]^2 M'''(t)}{\left[M''(t)\right]^3 \left[U''(t)\right]^3} \leq 0 \iff U''' \geq 0, M''' \leq 0$$

from (20). The denominator of (A.137) must be positive by $p_A(t) > 0$, and the denominator of (A.138) is then positive if $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$ by $\frac{\partial\epsilon(t)}{\partial p} \geq 0$, which proves $\dot{p}_A(t) > 0$ and, thus $\dot{x}_i(t) < 0$ during Phase I and III for $s_F = 0$, $\chi = 0$ and $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$. Furthermore, the numerator of (A.138) increases with t, such that $\dot{p}_A(t)$ is either always positive, first negative and then positive, or always negative during Phase I and III for $s_F = 0$, $\chi > 0$ and $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$. □

Lemma A.18. *Suppose $\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$. Then, the equilibrium is characterized by $s_F(t) = 0$ and $s_A(t) > 0$ for $t \in [t_1, T)$, and by $s_F(t) > 0$ for some $t < t_1$. Fuel and total energy [backstop] consumption in both countries decline [increase] over time during Phase III.*

Proof of Lemma A.18. During Phase II, $p(t) = \bar{p}$ to adhere the ceiling, such that $s_F(t) > 0$ with $p_F(t) = C_0 e^{-\chi t} + \tau_{F0}e^{\rho t}$ and $\dot{p}_F(t) = -\chi C_0 e^{-\chi t} + \rho\tau_{F0}e^{\rho t} \gtrless 0$ cannot hold. Suppose $S(t_2) > 0$. Then, $p_F(t_2) < p_A(t_2)$ would imply a violation of the ceiling. Furthermore, $p_F(t_2) = p_A(t_2)$ and $\dot{p}_F(t_2) > \dot{p}_A(t_2)$ would imply $p_F(t) < \bar{p} = p_A(t)$ for some $t \in [t_1, t_2)$ and, thus, a violation of the ceiling. Consequently, $p_F(t_2) > p_A(t_2)$ or $p_F(t_2) = p_A(t_2)$ and $\dot{p}_F(t_2) < \dot{p}_A(t_2)$ must hold. From (A.137) and (A.138) for $\theta_{A0} = 0$ we get

(A.139)
$$p_F(t) - p_A(t) = \tau_{F0}e^{\rho t} + C_0 e^{-\chi t} - \frac{\tau_{A0}e^{\rho t} + [1+\zeta_G]C_0 e^{-\chi t}}{\mathcal{N}_1},$$

$$\begin{aligned}\dot{p}_F(t) - \dot{p}_A(t) &= \rho\tau_{F0}e^{\rho t} - \chi C_0 e^{-\chi t} - \frac{\rho\tau_{A0}e^{\rho t} - [1+\zeta_G]\chi C_0 e^{-\chi t}}{\mathcal{N}_2} \\ &= \rho[p_F(t) - p_A(t)] + \frac{\left[\rho\tau_{A0}e^{\rho t} - [1+\zeta_G]\chi C_0 e^{-\chi t}\right][\mathcal{N}_2 - \mathcal{N}_1]}{\mathcal{N}_1\mathcal{N}_2} \\ &\quad + \frac{(\rho+\chi)C_0 e^{-\chi t}\left[\zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} + [1+\zeta_G]\frac{n_B}{\epsilon(t)}\right]}{\mathcal{N}_1}\end{aligned}$$

$$
\begin{aligned}
&= \rho[p_F(t) - p_A(t)] + \frac{\dot{p}_A(t)\,[\mathcal{N}_2 - \mathcal{N}_1]}{\mathcal{N}_1} \\
&\quad + \frac{(\rho+\chi)C_0 e^{-\chi t}\left[\zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} + [1+\zeta_G]\frac{n_B}{\epsilon(t)}\right]}{\mathcal{N}_1},
\end{aligned} \tag{A.140}
$$

where

$$
\begin{aligned}
\mathcal{N}_1 &= 1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} - [1+\zeta_G]\frac{n_B}{\epsilon(t)}, \\
\mathcal{N}_2 &= 1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)} - [1+\zeta_G]\frac{n_B}{\epsilon(t)}\left[1 - \frac{\partial\epsilon(t)}{\partial p}\frac{p_A}{\epsilon(t)}\right] \\
&\quad - \zeta_G n_A \frac{Q''(t)D'(t) - Q'(t)D''(t)}{[D'(t)]^2} p_A.
\end{aligned}
$$

$\zeta_G = 0$ or $U''' \geq 0, M''' \leq 0$ implies $\mathcal{N}_2 \geq \mathcal{N}_1 (> 0)$.

$\dot{p}_A(t_2) > 0$ to leave the ceiling implies $\rho\tau_{A0}e^{\rho t} - [1+\zeta_G]\chi C_0 e^{-\chi t} > 0$ for $t \in [t_2, T)$ and, therefore, $\dot{p}_A(t) > 0$ for $t \in [t_2, T)$. Thus, $p_F(t_2) \geq p_A(t_2)$ implies $\dot{p}_F(t) > \dot{p}_A(t)$ for $t \in [t_2, T)$, such that $p_F(t_2) > p_A(t_2)$ must hold to adhere the ceiling. $p_F(t_2) > p_A(t_2)$ and $\dot{p}_F(t) > \dot{p}_A(t)$ for $t \in [t_2, T)$ then imply $p_F(t) > p_A(t)$ for $t \in [t_2, T)$, such that $s_F(t) = 0$ for $t \in [t_2, T)$.

If $s_F(t) = 0$ holds for all $t < t_1$, then $s_F = 0$ during Phase III, $\lambda(T)\rho\tau_F(T) = 0$ and (A.132) would imply $\int_0^T y_A(t)\,\mathrm{d}t = 0$, which contradicts $y_A(t) > 0$ for some t. □

Lemma A.19. *The price path jumps upwards at $t = T_F$ if $s_F(T_F^-) > 0$, is continuous at $t = T_F$ if $s_F(T_F^-) = 0$ and $\zeta_G(T_F^-) = 0$, and is continuous for all other points in time.*

Proof of Lemma A.19. Suppose $s_F(T_F^-) > 0$. Then, $p_F(T_F) < p_A(T_F)$ implying a price jump when the firm's fuel stock becomes exhausted. Suppose $s_F(T_F^-) = 0$ and $\zeta_G(T_F^-) = 0$. Then, $p_F(T_F) \geq p_A(T_F)$ and the coalition's producer price is given by (30). Applying the proof of Lemma A.2 to τ_A, λ, κ, ι and θ_A shows that the costates are continuous, so that all elements of (30) are continuous functions. From Lemma A.14, $\zeta_G(t) > 0$ for some $t \in [0, T_F)$ implies $\zeta_G(t) = \zeta_G > 0$ for all $t \in [0, T_F)$, so that all elements of (34) are continuous functions for $t \in [0, T_F)$. Finally, all elements of the firm's fuel price $p_F(t) = C(t) + \tau_F(t)$ are continuous functions. Consequently, the price path is continuous for $t \in [0, T) \backslash T_F$. □

Lemma A.20. *Suppose $\theta_{A0} + SE \geq 0$. Then, the equilibrium is characterized by $s_F(t) > 0$ and $s_A(t) \geq [=]0$ for $t \in [0, T_F)$ and $\zeta_G = [>]0$, and by $s_F(t) = 0$ and $s_A(t) > 0$ for*

$t \in [T_F, T)$, where $T_F \leq t_1$. Fuel and total energy [backstop] consumption in both countries increase [decline] until $t^s \in [0, t_1)$ and decline [increase] for $t \in (t^s, t_1)$. Suppose $\theta_{A0} e^{\gamma T_F} + SE \geq 0$ ($n_A \geq \upsilon_A$ and $\zeta_G = 0$ are sufficient). Then, the price path jumps upwards at $t = T_F$.

Proof of Lemma A.20. For $s_F(t) = 0$, we get

$$
\begin{aligned}
p_A(t) - p_F(t) &= \frac{\tau_A(t) + \theta_A(t) + [1 + \zeta_G]\big[C(t) + \mathrm{ME}(t)\big]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}} - \big[\tau_F(t) + C(t)\big] \\
&= \frac{\theta_A(t) + \mathrm{SE} e^{\rho t} + [1 + \zeta_G]\big[\tau_F(t) + C(t) + \mathrm{ME}(t)\big]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}} - \big[\tau_F(t) + C(t)\big] \\
&= \frac{\theta_A(t) + \mathrm{SE} e^{\rho t} + [1 + \zeta_G]\mathrm{ME}(t) + \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}\big[\tau_F(t) + C(t)\big]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}} \qquad \text{(A.141)}
\end{aligned}
$$

during Phase I, which is positive if $\theta_{A0} + \mathrm{SE} \geq 0$. Thus, $s_F(t) > 0$ during Phase I, and since private fuel supply is ruled out at the ceiling, $S(t_1) = 0$ implies $T_F \leq t_1$. If $\zeta_G = 0$, (31), (32) and Lemma A.12 imply $s_A(t) \geq 0$ for $t \in [0, T_F)$. If $\zeta_G > 0$, Lemma A.16 implies $s_A(t) = 0$ for $t \in [0, T_F)$. Note that Lemma A.8 holds for $\theta_{A0} + \mathrm{SE} \geq 0$ and $\zeta_G = 0$, which proves the third sentence of Lemma A.20. Finally, $p_A(T_F) > p_F(T_F)$ holds if $\theta_{A0} e^{\gamma T_F} + \mathrm{SE} \geq 0$, and the price path jumps upwards at $t = T_F$. Substituting (33) into (A.141) yields

$$
\begin{aligned}
p_A(t) - p_F(t) = &\frac{\theta_{A0}\left[e^{\gamma t} - \frac{\int_{t_a}^{t_b} e^{(\rho+\gamma)t} D'(t)}{\int_{t_a}^{t_b} e^{\rho t} D'(t)\,\mathrm{d}t}\right] + [1 + \zeta_G]\frac{\upsilon_B S(0) - \int_{t_a}^{t_b} D_B(t)\,\mathrm{d}t}{\left|\int_{t_a}^{t_b} e^{\rho t} D'(t)\,\mathrm{d}t\right|}}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}} e^{\rho t} \\
&+ \frac{\mathrm{BE} e^{\rho t} + [1 + \zeta_G]\mathrm{ME}(t) + \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}\big[\tau_F(t) + C(t)\big]}{1 + \zeta_G - \zeta_G n_A \frac{D'(t)+Q'(t)}{D'(t)}}. \qquad \text{(A.142)}
\end{aligned}
$$

The environmental part is positive if $t \geq t_b$, and the terms-of-trade part is positive if $\upsilon_B S(0) > \int_{t_a}^{t_b} D_B(t)\,\mathrm{d}t \Leftarrow n_A \int_0^T D(t) \geq \upsilon_A S(0) \Leftrightarrow n_A \geq \upsilon_A$. Thus, $p_A(T_F) > p_F(T_F)$ holds if $n_A \geq \upsilon_A$ and $\zeta_G = 0 \Rightarrow \mathrm{BE} = 0$, and the price path jumps upwards at $t = T_F$. □

The Lemmata A.12 to A.20 prove Proposition 8. □

Exhaustion date of the private fuel stock. Time T_F divides the planning horizon into two periods, where the coalition's value function for $t \geq T_F$ is given by

$$
V^S(S_A(T_F), C(T_F), Z(T_F)) = \max_{s_A} \int_0^\infty e^{-\rho t}\Big\{U_A(D_A(t) + Q_A(t)) + \bar{K}_A - p(t) D_A(t)
$$

$$- M_A(Q_A(t)) + [p(t) - C(t)]s_A(t)\Big\} \,\mathrm{d}t$$

subject to (3), (4), (5), and (22).

Assuming singular deposit acquisitions, the transversality conditions at time T_F are

$$(a): \tau_A(T_F) = \frac{\partial V^S}{\partial S_A}, \quad (b): -\theta_A(T_F) = \frac{\partial V^S}{\partial Z}, \quad (c): \iota_A(T_F) = \frac{\partial V^S}{\partial C}, \tag{A.143}$$

$$H(T_F) = \rho V^S(S_A(T_F), C(T_F), Z(T_F)). \tag{A.144}$$

Taking account of (A.143), we get

$$\begin{aligned}\rho V^S(S_A(T_F), C(T_F), Z(T_F)) =& U_A(D_A(T_F^+) + Q_A(T_F^+)) + \bar{K}_A - p(T_F^+)D_A(T_F^+) - M_A(Q_A(T_F^+)) \\ &+ [p(T_F^+) - C(T_F)]s_A(T_F^+) - \tau_A(T_F)s_A(T_F^+) - \iota_A(T_F)\chi C(T_F) \\ &- \theta_A(T_F)[D(T_F^+) - \gamma Z(T_F)].\end{aligned}$$

Substituting into (A.144), and taking $\lambda(T_F) = \lambda(T) = 0$ and $p(T_F^+) = C(T_F) + \tau_A(T_F) + \theta_A(T_F) + \mathrm{ME}(T_F)$ into account yields

$$\begin{aligned}&U_A(D_A(T_F^-) + Q_A(T_F^-)) - p(T_F^-)D_A(T_F^-) - M_A(Q_A(T_F^-)) \\ &+ [p(T_F^-) - C(T_F)]s_A(T_F^-) + \upsilon_A[p(T_F^-) - C(T_F)][D(T_F^-) - s_A(T_F^-)] \\ &- [1-\upsilon_A]p_y(T_F^-)y_A(T_F^-) + \tau_A(T_F)[y_A(T_F^-) - s_A(T_F^-)] \\ &- \kappa(T_F)[D(T_F^-) + y_A(T_F^-) - s_A(T_F^-)] - \theta_A(T_F)D(T_F^-) \\ =& U_A(D_A(T_F^+) + Q_A(T_F^+)) - p(T_F^+)D_A(T_F^+) - M_A(Q_A(T_F^+)) + \mathrm{ME}(T_F^+)s_A(T_F^+),\end{aligned} \tag{A.145}$$

which determines the optimal T_F. If the left-hand side is greater [smaller] than the right-hand side, we get $T_F = T$ $[T_F = 0]$.

For $s_F(T_F^-) > 0 \vee \zeta_G = 0$, (A.145) can be written as

$$\begin{aligned}&U_A(D_A(T_F^-) + Q_A(T_F^-)) - p(T_F^-)D_A(T_F^-) - M_A(Q_A(T_F^-)) \\ &- \left[p(T_F^+) - p(T_F^-) - \mathrm{ME}(T_F^+)\right] D(T_F^-) - \left[U_A(D_A(T_F^+) + Q_A(T_F^+))\right. \\ &\left. - p(T_F^+)D_A(T_F^+) - M_A(Q_A(T_F^+)) + \mathrm{ME}(T_F^+)D(T_F^+)\right] = 0.\end{aligned} \tag{A.146}$$

Differentiating the left-hand side of (A.146) with respect to $p(T_F^+)$ yields

$$- D(T_F^-) + D_A(T_F^+) - \underbrace{\mathrm{ME}(T_F^+)D'(T_F^+)}_{=-D_B(T_F^+)} + \tfrac{\partial \mathrm{ME}(T_F^+)}{\partial p}\left[D(T_F^-) - D(T_F^+)\right]$$

$$
\begin{aligned}
&- \underbrace{\left[U_A'(D_A(T_F^+) + Q_A(T_F^+)) - p(T_F^+)\right]}_{=0} D_A'(T_F^+) \\
&- \underbrace{\left[U_A'(D_A(T_F^+) + Q_A(T_F^+)) - M_A'(Q_A(T_F^+))\right]}_{=0} Q_A'(T_F^+) \\
= &- \left[D(T_F^-) - D(T_F^+)\right] \left[1 - \tfrac{\partial \mathrm{ME}(T_F^+)}{\partial p}\right] \lesseqqgtr 0 \iff \begin{cases} p(T_F^+) \gtreqless p(T_F^-) \text{ if } 1 - \frac{\partial \mathrm{ME}}{\partial p} > 0, \\ p(T_F^+) \lesseqgtr p(T_F^-) \text{ if } 1 - \frac{\partial \mathrm{ME}}{\partial p} < 0. \end{cases}
\end{aligned}
$$

Since the left-hand side of (A.146) equals zero for $p(T_F^+) = p(T_F^-)$, and the left-hand side of (A.146) decreases [increases] with $p(T_F^+)$ for $p(T_F^+) > [<] p(T_F^-)$, $T_F = 0$ $[T_F = T]$ is an equilibrium if $1 - \frac{\partial \mathrm{ME}}{\partial p} > [<] 0$. Note that

$$
1 - \tfrac{\partial \mathrm{ME}}{\partial p} = 1 + n_B \frac{[D']^2 - DD''}{[D']^2} = 1 + \frac{n_B}{\epsilon}\left[1 - \frac{\partial \epsilon}{\partial p}\frac{p}{\epsilon}\right],
$$

such that $n_B = 0$ or $D'' \leq 0 \impliedby U''' \leq 0, M''' \leq 0$ is sufficient for $1 - \frac{\partial \mathrm{ME}}{\partial p} > 0$.

Finally, for $s_F(T_F^-) = 0 \implies \zeta_G = 0$, (A.145) can be written as

$$
\begin{aligned}
&U_A(D_A(T_F^-) + Q_A(T_F^-)) - p(T_F^-)D_A(T_F^-) - M_A(Q_A(T_F^-)) + \mathrm{ME}(T_F^-)D(T_F^-) = \\
&U_A(D_A(T_F^+) + Q_A(T_F^+)) - p(T_F^+)D_A(T_F^+) - M_A(Q_A(T_F^+)) + \mathrm{ME}(T_F^+)D(T_F^+).
\end{aligned} \tag{A.147}
$$

Since $s_F(T_F^-) = 0$ implies $p(T_F^+) = p(T_F^-)$, and the left-hand side of (A.147) equals the right-hand side for $p(T_F^+) = p(T_F^-)$, $T_F \in [0, T]$ is an equilibrium.

A.8. Coalition formation

The coalition's welfare and the fringe's welfare, respectively, are given by

$$
\begin{aligned}
W_A = &\int_0^{T_A} e^{-\rho t}\Bigg\{ n_A U\left(\frac{x_A(t) + q_A(t)}{n_A}\right) - [C(t) + \tau_F(t)]x_A(t) - n_A M\left(\frac{q_A(t)}{n_A}\right) \\
&+ \bar{K}_A \Bigg\} \,\mathrm{d}t + \frac{e^{-\rho T_A}}{\rho} n_A \bar{U} + \tau_{F0}\upsilon_A S(0) \\
&+ \int_0^T e^{-\rho t}[p(t) - C(t) - \tau_F(t)][s_A(t) + \upsilon_A s_F(t) - x_A(t)] \,\mathrm{d}t,
\end{aligned} \tag{A.148}
$$

$$
\begin{aligned}
W_B = \int_0^{T_B} e^{-\rho t} \Bigg\{ & n_B U\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - [C(t)+\tau_F(t)]x_B(t) - n_B M\left(\frac{q_B(t)}{n_B}\right) \\
& + \bar{K}_B \Bigg\} \mathrm{d}t + \frac{e^{-\rho T_B}}{\rho} n_B \bar{U} + \tau_{F0} \upsilon_B S(0) \\
& + \int_0^T e^{-\rho t} [p(t) - C(t) - \tau_F(t)][\upsilon_B s_F(t) - x_B(t)]\, \mathrm{d}t.
\end{aligned} \tag{A.149}
$$

In the demand-side regime, $p(t) - C(t) - \tau_F(t) = 0$ for all t. In the supply-side regime, $p(t) - C(t) - \tau_F(t) = 0$ for $t \in [t_a^i, t_b^i)$ with $i = I, III$. For $t \notin [t_a^i, t_b^i)$, $p(t) = p_A(t)$, $s_F(t) = 0$, and $s_A(t) - x_A(t) = x_B(t)$.

Lemma A.21. *In the competitive demand-side regime with quadratic cost functions and linear demand functions, an increase in the coalition size raises per-capita welfare of the coalition countries and global welfare, and it does not affect per-capita welfare of the fringe countries.*

Proof of Lemma A.21. In the competitive demand-side regime, global welfare from (A.148) and (A.149) is given by

$$
\begin{aligned}
W = \int_0^T e^{-\rho t} \Bigg\{ & n_A U\left(\frac{x_A(t)+q_A(t)}{n_A}\right) + n_B U\left(\frac{x_B(t)+q_B(t)}{n_B}\right) \\
& - C(t)[x_A(t)+x_B(t)] - n_A M\left(\frac{q_A(t)}{n_A}\right) - n_B M\left(\frac{q_B(t)}{n_B}\right) \Bigg\} \mathrm{d}t + \frac{e^{-\rho T}}{\rho} \bar{U}.
\end{aligned} \tag{A.150}
$$

Differentiating with respect to n_A yields

$$
\begin{aligned}
\tfrac{\mathrm{d}W}{\mathrm{d}n_A} = \int_0^T e^{-\rho t} \Bigg[& U\left(\frac{x_A(t)+q_A(t)}{n_A}\right) - U\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - M\left(\frac{q_A(t)}{n_A}\right) + M\left(\frac{q_B(t)}{n_B}\right) \\
& - \frac{x_A(t)+q_A(t)}{n_A} U'\left(\frac{x_A(t)+q_A(t)}{n_A}\right) + \frac{x_B(t)+q_B(t)}{n_B} U'\left(\frac{x_B(t)+q_B(t)}{n_B}\right) \\
& + \frac{q_A(t)}{n_A} M'\left(\frac{q_A(t)}{n_A}\right) - \frac{q_B(t)}{n_B} M'\left(\frac{q_B(t)}{n_B}\right) \Bigg] \mathrm{d}t \\
& + \int_0^T e^{-\rho t} \Bigg\{ \left[U'\left(\frac{x_A(t)+q_A(t)}{n_A}\right) - C(t) \right] \tfrac{\mathrm{d}x_A}{\mathrm{d}n_A} + \left[U'\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - C(t) \right] \tfrac{\mathrm{d}x_B}{\mathrm{d}n_A} \\
& + \left[U'\left(\frac{x_A(t)+q_A(t)}{n_A}\right) - M'\left(\frac{q_A(t)}{n_A}\right) \right] \tfrac{\mathrm{d}q_A}{\mathrm{d}n_A} + \left[U'\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - M'\left(\frac{q_B(t)}{n_B}\right) \right] \tfrac{\mathrm{d}q_B}{\mathrm{d}n_A} \Bigg\} \mathrm{d}t
\end{aligned}
$$

(A.151)

$$= \int_0^{t_2} e^{-\rho t} \Bigg\{ U\left(\frac{x_A(t)+q_A(t)}{n_A}\right) - U\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - M\left(\frac{q_A(t)}{n_A}\right) + M\left(\frac{q_B(t)}{n_B}\right)$$
$$- \left[\frac{D_A(t)}{n_A} - \frac{D_B(t)}{n_B}\right][C(t)+\tau_F(t)+\theta_A(t)]\Bigg\}\,\mathrm{d}t$$
$$+ \int_0^{t_2} e^{-\rho t}\theta_A(t)\left\{\frac{D_A(t)}{n_A} - \frac{D_B(t)}{n_B} + D_A'(t)\tfrac{\mathrm{d}[\tau_F(t)+\theta_A(t)]}{\mathrm{d}n_A}\right\}\mathrm{d}t.$$

For $t \in [t_1, t_2)$, $\frac{\mathrm{d}[\tau_F(t)+\theta_A(t)]}{\mathrm{d}n_A}$ is given by

$$\bar{s} = n_A D(C(t)+\tau_F(t)+\theta_A(t)) + n_B D(C(t)+\tau_F(t))$$
$$\Rightarrow \quad 0 = \frac{D_A(t)}{n_A} - \frac{D_B(t)}{n_B} + D_A'(t)\tfrac{\mathrm{d}[\tau_F(t)+\theta_A(t)]}{\mathrm{d}n_A} + D_B'(t)\tfrac{\mathrm{d}\tau_F(t)}{\mathrm{d}n_A}$$
$$\Leftrightarrow \quad \tfrac{\mathrm{d}[\tau_F(t)+\theta_A(t)]}{\mathrm{d}n_A} = \frac{1}{D_A'(t)}\left[\frac{D_B(t)}{n_B} - \frac{D_A(t)}{n_A} - D_B'(t)\tfrac{\mathrm{d}\tau_F(t)}{\mathrm{d}n_A}\right]. \tag{A.152}$$

Using (A.152) in (A.151) yields

(A.153)

$$\tfrac{\mathrm{d}W}{\mathrm{d}n_A} = \int_0^{t_2} e^{-\rho t} \Bigg\{ U\left(\frac{x_A(t)+q_A(t)}{n_A}\right) - U\left(\frac{x_B(t)+q_B(t)}{n_B}\right) - M\left(\frac{q_A(t)}{n_A}\right) + M\left(\frac{q_B(t)}{n_B}\right)$$
$$- \left[\frac{D_A(t)}{n_A} - \frac{D_B(t)}{n_B}\right][C(t)+\tau_F(t)+\theta_A(t)]\Bigg\}\,\mathrm{d}t$$
$$+ \int_0^{t_1} e^{-\rho t}\theta_A(t)\left\{\frac{D_A(t)}{n_A} - \frac{D_B(t)}{n_B} + D_A'(t)\tfrac{\mathrm{d}[\tau_F(t)+\theta_A(t)]}{\mathrm{d}n_A}\right\}\mathrm{d}t - \int_{t_1}^{t_2} e^{-\rho t}\theta_A(t)D_B'(t)\tfrac{\mathrm{d}\tau_F(t)}{\mathrm{d}n_A}\,\mathrm{d}t.$$

For quadratic cost functions and linear demand functions, we have $\frac{\mathrm{d}\tau_F(t)}{\mathrm{d}n_A} = 0$ and $\frac{\mathrm{d}\theta_A(t)}{\mathrm{d}n_A} = -\frac{\theta_A(t)}{n_A} < 0$ for $t \in [0, t_1)$ from the proof of Lemma A.4. Using this in (A.153) yields

$$\tfrac{\mathrm{d}W}{\mathrm{d}n_A} = \frac{\tilde{\beta}}{2}\int_0^{t_2} e^{-\rho t}\theta_A(t)^2\,\mathrm{d}t > 0. \tag{A.154}$$

□

Proof of Proposition 9. Consider the competitive demand-side regime with quadratic cost functions and linear demand functions. From Lemma A.21, the per-capita welfare of any country increases [decreases] if it leaves [joins] the coalition because the per-capita welfare outside the coalition is independent of the coalition size and greater than the per-capita welfare inside the coalition for a given coalition size. Consequently, all coalitions that cannot ensure the ceiling and coalitions that can just ensure the ceiling are internally stable, and all coalitions that can ensure the ceiling are externally stable. Thus, coalitions that can just ensure the ceiling are the only stable coalitions in the competitive demand-side regime with quadratic cost functions and linear demand functions. □

B. Calibration

In this section, we calibrate the model to the world coal market in the year 2015. From IPCC (2013, p. 491), the natural CO_2 stock is $278ppm$ (parts per million), and from Department of Commerce (2021), the current CO_2 stock is $400ppm$. Following IPCC (2013, p. 471), we set the conversion factor $1ppm = 2,120MtC$ (million tonnes of carbon), and following EIA (2020b), we set the conversion factor $1C =$ two units of coal. The current excess CO_2 stock can then be expressed as $(400 - 278)ppm \cdot 2,120MtC/ppm \cdot 2\text{coal}/C = 517,280Mt\text{coal}$. For the the ceiling, we apply $(468 - 278)ppm \cdot 2,120MtC/ppm \cdot 2\text{coal}/C = 805,600Mt\text{coal}$, which limits the temperature increase below 2°C with a probability of 66% (Rogelj et al., 2011, p. 4). Since coal is responsible for about 1/3 of total CO_2 emissions (IPCC, 2013, p. 487), we assume that each unit of CO_2 from coal is accompanied by two units of CO_2 from oil, gas, cement and land use change.[66]

According to Joos et al. (2013, pp. 2801), the fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t \in [0, 1000]$ is equal to $0.2173+0.2240e^{-0.0025t}+0.2824e^{-0.0274t}+0.2763e^{-0.2323t}$. Since the ceiling is typically reached within 50 years, we set $\gamma = 1.44\%$ such that the fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t = 50$ is the same in Joos et al. (2013) and our calibration.[67]

For the recoverable reserves of coal, we take $S(0) = 1,030,859Mt$ from EIA (2020b).[68] The current extraction costs of coal are on average $100\$/t$ for about 90% of the recov-

[66]In fact, we equivalently assume that the current excess CO_2 stock and the ceiling are 1/3 of the values above.

[67]We thank an anonymous reviewer for this suggestion. A comparison of the fraction of CO_2 that will remain in the atmosphere between Joos et al. (2013) and our calibration can be found in Appendix B.2. There we also carry out a sensitivity analysis on the natural regeneration rate.

[68]The actual coal resources are about twenty times as high (IEA, 2013, pp. 42). However, the extraction

erable reserves (IEA, 2013, pp. 232), such that the initial marginal extraction costs are $C_0 = 100,000,000\$/Mt$. According to McKinsey & Company (2020), the global mining productivity has increased by 2.6% per year from 2013 to 2018, such that we set $\chi = 2.6\%$. The backstop cost functions are specified by the quadratic functions $M_i = \frac{1}{2}\frac{n_i}{m}\left(\frac{q_i}{n_i}\right)^2$, which yields the linear supplies $q_i = n_i m p$ for $i = A, B$.[69] The utility functions are specified by the quadratic functions $U_i = \frac{n_i}{\beta}\left[\alpha\frac{x_i+q_i}{n_i} - \frac{1}{2}\left(\frac{x_i+q_i}{n_i}\right)^2\right]$, which yields the linear demands $x_i = n_i(\alpha - \beta p) - q_i = n_i(\alpha - \tilde{\beta}p)$ for $i = A, B$ with $\tilde{\beta} := \beta + m$, and an increasing price elasticity of coal demand, $\epsilon(p) = \frac{1}{\frac{\alpha}{\tilde{\beta}p} - 1}$.

The model is calibrated to the laissez-faire economy, whereby we choose $\rho = 2.5\%$ and take $\epsilon(p(0)) = 0.5$,[70] $\frac{\sum_i q_i(0)}{\sum_i [x_i(0)+q_i(0)]} = 14\%$[71] and $s(0) = 7734Mt$ (EIA, 2020b) from the literature. Solving the equation system

$$S(0) = \int_0^T \left[\alpha - \tilde{\beta}[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}]\right] \mathrm{d}t, \tag{B.1}$$

$$0 = \alpha - \tilde{\beta}[C_0 e^{-\chi T} + \tau_{F0} e^{\rho T}], \tag{B.2}$$

$$\epsilon(p(0)) = \frac{1}{\frac{\alpha}{\tilde{\beta}[C_0 + \tau_{F0}]} - 1}, \tag{B.3}$$

$$s(0) = \alpha - \tilde{\beta}[C_0 + \tau_{F0}], \tag{B.4}$$

$$\frac{\sum_i q_i(0)}{\sum_i [x_i(0) + q_i(0)]} = \frac{m[C_0 + \tau_{F0}]}{\alpha - \beta[C_0 + \tau_{F0}]}, \tag{B.5}$$

with respect to α, β, m, τ_0 and T yields $\alpha = 11601$, $\beta = 0.000023653, m = 0.000011419, \tau_{F0} = 10.26\$/t$ and $T = 139$. Next, we consider the five countries with the greatest coal reserves and divide the rest of the world into four regions with comparable coal reserves. Using

costs of these resources are greater than $350\$/t$ (IEA, 2013, pp. 232), which is above the prohibitive price in our calibration ($309\$/t$). In other words, it is not reasonable to extract any coal resources, so that we focus on the coal reserves.

[69] The exponent of 2 is close to the exponent of 2.6 in Nordhaus's (2017) abatement cost function.

[70] The price elasticity of coal demand is in the range of 0.5–1.6 in China (Hang and Tu, 2007; Ma et al., 2008; Bloch et al., 2015; Burke and Liao, 2015), about 0.22 in the US (Serletis et al., 2010) and about 0.13 in selected OECD countries (Serletis et al., 2011). Since China's consumption share is greater than 50% (EIA, 2020a) and its price elasticity seems to be an upper bound for the global price elasticity, we set the global price elasticity equal to the lower bound of China's price elasticity.

[71] The global share of fossil fuels in primary energy consumption was 86% in 2015 (BP, 2021).

data on coal reserves and coal consumption from EIA (2020b) and data on GDP from The World Bank (2022), Table 1 summarizes these word regions. Finally, we compute the social optimum, the competitive climate policies and the strategic climate policies in Appendix B.1.

Figure 3 depicts the social optimum, Figure 4 depicts the competitive demand-side regime with the largest stable coalition (China, India, USA), Figure 5 depicts the competitive supply-side regime with the grand coalition being stable, and Figure 6 depicts the strategic demand-side regime with the largest stable coalition (China, USA, Russia, Australia). In the strategic supply-side regime, the grand coalition is stable and the allocation is efficient, so that we choose an exogenous coalition for Figure 7. Note that the extraction and, thus, the switches to Phases II-IV coincide in the social optimum and in the competitive demand side-regime.[72] Furthermore, the extraction is antedated and the switches to Phases III and IV occur earlier in the supply-side regime.[73] Finally, the extraction is postponed and the switches to Phase II and IV [III] occur later [earlier] in the strategic demand-side regime than in the social optimum or in the competitive demand-side regime.[74]

B.1. Equilibria characterization

The social optimum is characterized by

$$\begin{aligned} S(0) = &\int_0^{t_1} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_0 e^{\rho t} + \theta_0 e^{(\rho+\gamma)t} \right] \right\} \mathrm{d}t + [t_2 - t_1]\gamma \bar{Z} \\ &+ \int_{t_2}^{T} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_0 e^{\rho t} \right] \right\} \mathrm{d}t, \end{aligned} \tag{B.6}$$

$$\bar{Z} = Z(0) e^{-\gamma t_1} + \int_0^{t_1} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_0 e^{\rho t} + \theta_0 e^{(\rho+\gamma)t} \right] \right\} e^{-\gamma(t_1 - t)} \,\mathrm{d}t, \tag{B.7}$$

$$\gamma \bar{Z} = \alpha - \tilde{\beta} \left[C_0 e^{-\chi t_1} + \tau_0 e^{\rho t_1} + \theta_0 e^{(\rho+\gamma)t_1} \right], \tag{B.8}$$

$$\gamma \bar{Z} = \alpha - \tilde{\beta} \left[C_0 e^{-\chi t_2} + \tau_0 e^{\rho t_2} \right], \tag{B.9}$$

[72]See the discussion on linear demand functions and quadratic cost functions after Proposition 2. In particular, we find $\tau_{F0} = 0.57\$/t, t_1 = 37, t_2 = 238$ and $T = 254$.

[73]See Proposition 5. In particular, we find $\tau_{F0} = 0.59\$/t, t_1 = 33, t_2 = 237$ and $T = 253$.

[74]See the discussion on linear demand functions and quadratic cost functions after Proposition 6. In particular, we find $\tau_{F0} = 21.61\$/t, t_1 = 37, t_2 = 231$ and $T = 259$.

$$\text{(B.10)} \qquad 0 = \alpha - \tilde{\beta}\left[C_0 e^{-\chi T} + \tau_0 e^{\rho T}\right],$$

which can be solved for $\tau_0, \theta_0, t_1, t_2$ and T.

The competitive demand-side policy is characterized by

$$\text{(B.11)} \qquad \begin{aligned} S(0) = &\int_0^{t_1} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + n_A \theta_{A0} e^{(\rho+\gamma)t}\right]\right\} \mathrm{d}t + [t_2 - t_1]\gamma \bar{Z} \\ &+ \int_{t_2}^{T} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right]\right\} \mathrm{d}t, \end{aligned}$$

$$\text{(B.12)} \qquad \bar{Z} = Z(0)e^{-\gamma t_1} + \int_0^{t_1} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + n_A \theta_{A0} e^{(\rho+\gamma)t}\right]\right\} e^{-\gamma(t_1 - t)} \, \mathrm{d}t,$$

$$\text{(B.13)} \qquad \gamma \bar{Z} = \alpha - \tilde{\beta}\left[C_0 e^{-\chi t_1} + \tau_{F0} e^{\rho t_1} + n_A \theta_{A0} e^{(\rho+\gamma)t_1}\right],$$

$$\text{(B.14)} \qquad \gamma \bar{Z} = \alpha - \tilde{\beta}\left[C_0 e^{-\chi t_2} + \tau_{F0} e^{\rho t_2}\right],$$

$$\text{(B.15)} \qquad 0 = \alpha - \tilde{\beta}\left[C_0 e^{-\chi t_t} + \tau_{F0} e^{\rho T}\right],$$

which can be solved for $\tau_{F0}, \theta_{A0}, t_1, t_2$ and T.

The competitive supply-side policy is characterized by

$$\text{(B.16)} \qquad \begin{aligned} S(0) = &\int_0^{T_F} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right]\right\} \mathrm{d}t \\ &+ \int_{T_F}^{t_1} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + (1+\zeta_G)\tau_{F0} e^{\rho t} + \theta_{A0} e^{(\rho+\gamma)t}\right]\right\} \mathrm{d}t \\ &+ [t_2 - t_1]\gamma \bar{Z} + \int_{t_2}^{T} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + (1+\zeta_G)\tau_{F0} e^{\rho t}\right]\right\} \mathrm{d}t, \end{aligned}$$

$$\text{(B.17)} \qquad \begin{aligned} \bar{Z} = Z(0)e^{-\gamma t_1} &+ \int_0^{T_F} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right]\right\} e^{-\gamma(t_1 - t)} \, \mathrm{d}t \\ &+ \int_{T_F}^{t_1} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + (1+\zeta_G)\tau_{F0} e^{\rho t} + \theta_{A0} e^{(\rho+\gamma)t}\right]\right\} e^{-\gamma(t_1 - t)} \, \mathrm{d}t \end{aligned}$$

$$\text{(B.18)} \qquad \gamma \bar{Z} = \alpha - \tilde{\beta}\left[C_0 e^{-\chi t_1} + (1+\zeta_G)\tau_{F0} e^{\rho t_1} + \theta_{A0} e^{(\rho+\gamma)t_1}\right],$$

$$\text{(B.19)} \qquad \gamma \bar{Z} = \alpha - \tilde{\beta}\left[C_0 e^{-\chi t_2} + (1+\zeta_G)\tau_{F0} e^{\rho t_2}\right],$$

$$\text{(B.20)} \qquad 0 = \alpha - \tilde{\beta}\left[C_0 e^{-\chi T} + (1+\zeta_G)\tau_{F0} e^{\rho T}\right],$$

and, from (A.67), by

$$
\begin{aligned}
& - 0.5 n_A \tilde{\beta} \left[\zeta_G \tau_{F0} e^{\rho T_F} + \theta_{A0} e^{(\rho+\gamma) T_F}\right]^2 \\
& + \left\{ n_A \left[\zeta_G \tau_{F0} e^{\rho T_F} + \theta_{A0} e^{(\rho+\gamma) T_F}\right] - \theta_{A0} e^{(\rho+\gamma) T_F} \right\} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi T_F} + \tau_{F0} e^{\rho T_F} \right] \right\} \\
& + \zeta_G \tau_{F0} e^{\rho T_F} y_A(T_F^-) + \upsilon_A \tau_{F0} e^{\rho T_F} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi T_F} + \tau_{F0} e^{\rho T_F} \right] + y_A(T_F^-) \right\} = 0.
\end{aligned} \tag{B.21}
$$

With a non-binding budget constraint and constant deposit acquisitions over time, we have $\zeta_G = 0$ and

$$
y_A(t) = \frac{1}{T_F} \left\{ S(0) - \int_0^{T_F} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] \right\} \mathrm{d}t \right\}
$$

for $t \in [0, T_F)$, and the equation system can be solved for $\tau_{F0}, \theta_{A0}, T_F, t_1, t_2$ and T. With a binding budget constraint, we have

$$
\int_0^{T_F} \bar{y}_A(t)\, \mathrm{d}t = S(0) - \int_0^{T_F} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] \right\} \mathrm{d}t \tag{B.22}
$$

and

$$
\begin{aligned}
\bar{y}_A(t) &= \frac{\bar{K}_A + \upsilon_A \Pi_F(t) - p(t) x_A(t) + \Pi_A(t) - p_{q_A}(t) q_A(t)}{p_y(t)} \\
&= \frac{1}{(1-\upsilon_A)\tau_{F0} e^{\rho t}} \Bigg\{ \bar{K}_A - \left\{ n_A \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] - \upsilon_A \tau_{F0} e^{\rho t} \right\} \\
&\qquad \cdot \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] \right\} - 0.5 n_A m \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right]^2 \Bigg\}
\end{aligned}
$$

for $t \in [0, T_F)$ from (24), and the equation system can be solved for $\zeta_G, \tau_{F0}, \theta_{A0}, T_F, t_1, t_2$ and T. However, for coalitions of eight world regions, the budget constraint is non-binding since $\zeta_G < 0$. For the grand coalition, the initial budget is unlimited and deposit acquisitions could be maximal. Then, the grand coalition would implement the social optimum and Table 2 would become Table B.6. The second lines and the last columns of these tables differ only quantitatively, such that Proposition 10 would still hold.

The strategic demand-side policy is characterized by

$$
\begin{aligned}
S(0) = & \int_0^{t_1} \left\{ \alpha - \tilde{\beta} \left\{ n_A U_A'(t) + n_B \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] \right\} \right\} \mathrm{d}t + [t_2 - t_1] \gamma \bar{Z} \\
& + n_A \int_{t_2}^{T_A} \left[\alpha - \tilde{\beta} U_A'(t) \right] \mathrm{d}t + n_B \int_{t_2}^{T_B} \left\{ \alpha - \tilde{\beta} \left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} \right] \right\} \mathrm{d}t,
\end{aligned} \tag{B.23}
$$

Table B.6: Large coalitions with the competitive supply-side policy and $T_F = 0$ for $n_A = 100\%$ (υ_A, n_A in %; τ_{F0}, θ_{A0} in \$/t; $W|_{\bar{K}=0}, \Delta\frac{W_i}{n_i}$ in \$trillion).

	υ_A	n_A	t_1	t_2	T	τ_{F0}	θ_{A0}	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$
WORLD	100	100	37	238	254	0.57	42.4	78.1	–
WORLD-Aus	86.0	98.5	34	237	253	0.59	46.8	77.8	16.6
WORLD-Rus	84.4	97.3	34	237	253	0.59	46.5	77.8	18.9
WORLD-ROE	94.3	97.0	34	237	253	0.59	46.8	77.8	21.1
WORLD-ROW	93.7	96.2	34	237	253	0.59	46.6	77.8	21.3
WORLD-EU	92.8	91.6	35	237	253	0.59	45.8	77.8	22.1
WORLD-USA	77.6	90.6	35	237	254	0.58	45.1	77.9	21.4
WORLD-ROA	93.4	90.4	35	237	253	0.59	45.6	77.8	22.2
WORLD-Ind	90.8	89.7	35	237	253	0.58	45.4	77.9	22.2
WORLD-Chi	87.0	48.7	36	238	254	0.58	43.4	78.0	23.3

Note: υ_A is the coalition's share of global coal reserves, n_A is the coalition's share of global coal consumption, t_1, t_2 and T is the end of Phase I, II and III, respectively, τ_{F0} is the initial private scarcity rent, θ_{A0} is the coalition's initial cost of emissions, $W|_{\bar{K}=0}$ is global energy welfare, and $\Delta\frac{W_i}{n_i}$ is the fringe country's increase in per-capita welfare by joining the coalition.

(B.24) $$\bar{Z} = Z(0)e^{-\gamma t_1} + \int_0^{t_1} \left\{\alpha - \tilde{\beta}\left\{n_A U_A'(t) + n_B\left[C_0 e^{-\chi t} + \tau_{F0}e^{\rho t}\right]\right\}\right\} e^{-\gamma(t_1-t)}\,\mathrm{d}t,$$

(B.25) $$\gamma\bar{Z} = \alpha - \tilde{\beta}\left\{n_A U_A'(t_1) + n_B\left[C_0 e^{-\chi t_1} + \tau_{F0}e^{\rho t_1}\right]\right\},$$

(B.26) $$\gamma\bar{Z} = \alpha - \tilde{\beta}\left\{n_A U_A'(t_2) + n_B\left[C_0 e^{-\chi t_2} + \tau_{F0}e^{\rho t_2}\right]\right\},$$

(B.27) $$0 = \alpha - \tilde{\beta}U_A'(T_A),$$

(B.28) $$0 = \alpha - \tilde{\beta}\left[C_0 e^{-\chi T_B} + \tau_{F0}e^{\rho T_B}\right],$$

where

$$U_A'(t) = C_0 e^{-\chi t} + \tau_{F0}e^{\rho t} + \theta_A(t) + \mathrm{SE}e^{\rho t}$$

from (28). Thereby, $\theta_A(t) = \theta_{A0}e^{(\rho+\gamma)t}$ for $t \in [0, t_1)$,

$$\gamma\bar{Z} = \alpha - \tilde{\beta}\left\{C_0 e^{-\chi t} + \tau_{F0}e^{\rho t} + n_A\left[\theta_A(t) + \mathrm{SE}e^{\rho t}\right]\right\}$$
$$\Leftrightarrow \quad \theta_A(t) = -\frac{D_B(t) - n_B\gamma\bar{Z}}{n_A D_B'(t)} - \mathrm{SE}e^{\rho t}$$

for $t \in [t_1, t_2)$, and $\theta_A(t) = 0$ for $t \geq t_2$. Furthermore,

$$\mathrm{SE} = \frac{\upsilon_B S(0) - \int_0^{T_B} D_B(t)\,\mathrm{d}t + \int_0^{t_1}\theta_{A0}e^{(\rho+\gamma)t}D_B'(t)\,\mathrm{d}t + \int_{t_1}^{t_2}\theta_A(t)D_B'(t)\,\mathrm{d}t}{\left|\int_0^{T_B} e^{\rho t}D_B'\,\mathrm{d}t\right|}$$

$$= \frac{\upsilon_B S(0) - \int_0^{T_B} D_B(t)\,\mathrm{d}t + \int_0^{t_1} \theta_{A0} e^{(\rho+\gamma)t} D'_B(t)\,\mathrm{d}t - \frac{1}{n_A}\int_{t_1}^{t_2}\left[D_B(t) - n_B\gamma\bar{Z}\right]\mathrm{d}t}{\left|\int_0^{T_B} e^{\rho t} D'_B\,\mathrm{d}t\right| - \left|\int_{t_1}^{t_2} e^{\rho t} D'_B\,\mathrm{d}t\right|}.$$

The equation system can then be solved for $\tau_{F0}, \theta_{A0}, t_1, t_2, T_A$ and T_B.

Our nine world regions imply 512 possible coalitions. We first analyze which coalitions without China adhere the ceiling ({World - China}, {World - China - ROW}, {World - China - ROE}, {World - China - Russia}, {World - China - Australia}), and find these coalitions to be externally unstable because China has an incentive to join. Consequently, any stable coalition comprises China and we are left with 256 possible coalitions. We then analyze the grand coalition, the eight $n = 8$ coalitions, the twenty-eight $n = 7$ coalitions, the fifty-six $n = 6$ coalitions and the seventy $n = 5$ coalitions, and find these coalitions to be internally unstable because, e.g., India, ROA, USA, EU or ROW have an incentive to leave. Next, we analyze which of the eight $n = 2$ coalitions adhere the ceiling and find that this is the case for {China + India}, {China + ROA} and {China + USA}. {China + India} and {China + ROA} turn out to be internally and externally stable, whereas {China + USA} is externally unstable because Australia has an incentive to join. Concerning the twenty-eight $n = 3$ coalitions, we find that only those with India, ROA, USA or EU adhere the ceiling. {China + USA + Australia} is externally unstable because Russia has an incentive to join, whereas the remaining coalitions with India, ROA or USA are internally unstable. {China + EU + ROW}, {China + EU + ROE} and {China + EU + Australia} are internally and externally stable to adhere the ceiling, whereas {China + EU + Russia} is externally unstable because Australia has an incentive to join. Finally, we analyze the fifty-six $n = 4$ coalitions and find that all coalitions without India, ROA, USA and EU are internally and externally stable to adhere the ceiling ({China + ROW + ROE + Russia}, {China + ROW + ROE + Australia}, {China + ROW + Russia + Australia}, {China + ROE + Russia + Australia}). Furthermore, {China + USA + Russia + Australia} is internally and externally stable, whereas all other coalitions are internally unstable.

The strategic supply-side policy with dominating positive strategic effects, i.e. $p_A(t) \geq$

$p_F(t)$ for $t \in [0, t_1)$, is characterized by

$$S(0) = \int_0^{T_F} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}\right]\right\} \mathrm{d}t + \int_{T_F}^{t_1} \left[\alpha - \tilde{\beta} p_A(t)\right] \mathrm{d}t$$
$$+ [t_2 - t_1]\gamma \bar{Z} + \int_{t_2}^{T} \left[\alpha - \tilde{\beta} p_A(t)\right] \mathrm{d}t, \tag{B.29}$$

$$\bar{Z} = Z(0)e^{-\gamma t_1} + \int_0^{T_F} \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi t} + \tau_{A0} e^{\rho t}\right]\right\} e^{-\gamma(t_1 - t)} \mathrm{d}t$$
$$+ \int_{T_F}^{t_1} \left[\alpha - \tilde{\beta} p_A(t)\right] e^{-\gamma(t_1 - t)} \mathrm{d}t, \tag{B.30}$$

$$\gamma \bar{Z} = \alpha - \tilde{\beta} p_A(t_1), \tag{B.31}$$

$$\gamma \bar{Z} = \alpha - \tilde{\beta} p_A(t_2), \tag{B.32}$$

$$0 = \alpha - \tilde{\beta} p_A(T), \tag{B.33}$$

and, from (A.145), by

(B.34)

$$- (1 - 0.5 n_A)\tilde{\beta}\left[p_A(T_F) - C_0 e^{-\chi T_F} - \tau_{F0} e^{\rho T_F}\right]^2$$
$$+ \left[p_A(T_F) - C_0 e^{-\chi T_F} - \tau_{F0} e^{\rho T_F} - \theta_{A0} e^{(\rho+\gamma)T_F} - \mathrm{SE} e^{\rho T_F} - \mathrm{ME}(T_F) - \zeta_G \upsilon_A \tau_{F0} e^{\rho T_F}\right]$$
$$\cdot \left\{\alpha - \tilde{\beta}\left[C_0 e^{-\chi T_F} + \tau_{F0} e^{\rho T_F}\right]\right\} + \zeta_G \upsilon_B y_A(T_F^-) = 0,$$

where

$$p_A(t) = \frac{(1 + \zeta_G)\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + \mathrm{ME}(t)\right] + \theta_A(t) + \mathrm{SE} e^{\rho t}}{1 + \zeta_G - \zeta_G n_A \frac{\beta}{\tilde{\beta}}}$$
$$= \frac{(1 + \zeta_G)\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + n_B \frac{\alpha}{\tilde{\beta}}\right] + \theta_A(t) + \mathrm{SE} e^{\rho t}}{(1 + \zeta_G)(1 + n_B) - \zeta_G n_A \frac{\beta}{\tilde{\beta}}}$$

from (A.141). Thereby, $\theta_A(t) = \theta_{A0} e^{(\rho+\gamma)t}$ for $t \in [0, t_1)$,

$$\gamma \bar{Z} = \alpha - \tilde{\beta} p_A(t)$$
$$\Leftrightarrow \quad \theta_A(t) = \frac{\alpha - \gamma \bar{Z}}{\tilde{\beta}} - \frac{(1 + \zeta_G)\left[C_0 e^{-\chi t} + \tau_{F0} e^{\rho t} + n_B \frac{\alpha}{\tilde{\beta}}\right] + \mathrm{SE} e^{\rho t}}{(1 + \zeta_G)(1 + n_B) - \zeta_G n_A \frac{\beta}{\tilde{\beta}}}$$

for $t \in [t_1, t_2)$, and $\theta_A(t) = 0$ for $t \geq t_2$. Furthermore,

$$\mathrm{SE} = \frac{(1 + \zeta_G)\left[\upsilon_B S(0) - \int_0^{T_F} D_B(t)\,\mathrm{d}t\right] + \int_0^{T_F} \theta_{A0} e^{(\rho+\gamma)t} D'(t)\,\mathrm{d}t}{\left|\int_0^{T_F} e^{\rho t} D'(t)\,\mathrm{d}t\right|}$$

$$+\frac{\zeta_G \int_0^{T_F} [C_0 e^{-\chi t} + \tau_{F0} e^{\rho t}] \left[D'_A(t) + Q'_A(t)\right] \mathrm{d}t}{\left| \int_0^{T_F} e^{\rho t} D'(t) \, \mathrm{d}t \right|}.$$

With a non-binding budget constraint, we have $\zeta_G = 0$ and $T_F = 0$, since the left-hand side of (B.34) then equals $-(1 - 0.5n_A)\tilde{\beta}\{p_A(T_F) - C_0 e^{-\chi T_F} - \tau_{F0} e^{\rho T_F}\}^2 < 0$, and the equation system can be solved for $\tau_{F0}, \theta_{A0}, t_1, t_2$ and T. Consequently, the budget constraint is non-binding if and only if the coalition's initial budget is unlimited. With a binding budget constraint, we have (B.22), and the equation system can be solved for $\zeta_G, \tau_{F0}, \theta_{A0}, T_F, t_1, t_2$ and T.

B.2. Sensitivity analysis

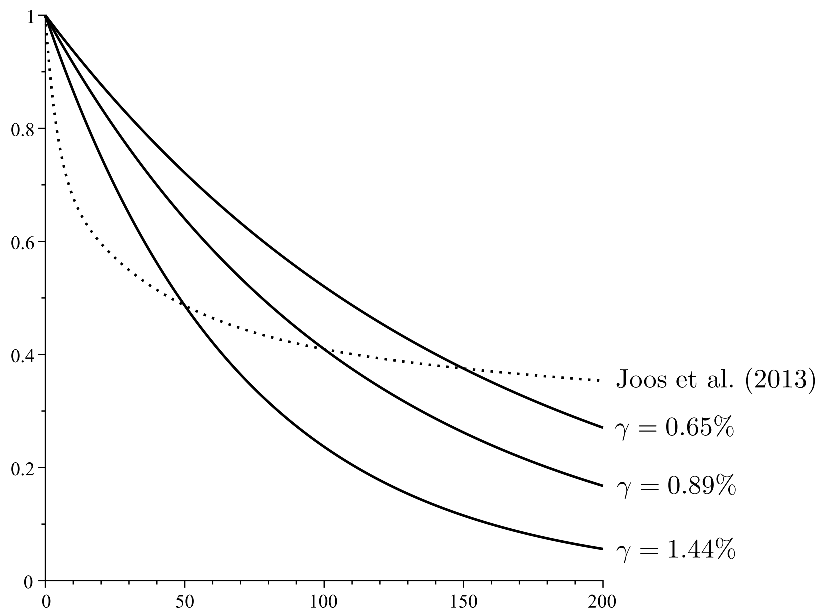

Figure B.8: Fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t \in [0, 200]$ in Joos et al. (2013) (dotted) and our calibration (solid) for different values of γ.

In Section B, we set $\gamma = 1.44\%$ such that the fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t = 50$ is the same in Joos et al. (2013) and our calibration. In this section, we use $\gamma = 0.89\%$ and $\gamma = 0.65\%$ such that the fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t = 100$ or $t = 150$ is the same in Joos et al. (2013) and our calibration. Figure B.8 illustrates the fraction of CO_2 emitted at $t = 0$ that will remain in the atmosphere at $t \in [0, 200]$ in Joos et al. (2013) and our calibration for $\gamma = 1.44\%$, $\gamma = 0.89\%$ and $\gamma = 0.65\%$.

Table B.7: Large coalitions with the competitive supply-side policy (υ_A, n_A in %; $W|_{\bar{K}=0}, \Delta\frac{W_i}{n_i}$ in \$trillion).

			$\gamma = 0.89\%$		$\gamma = 0.65\%$	
	υ_A	n_A	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$
WORLD	100	100	69.691	–	65.357	–
WORLD-Aus	86.0	98.5	69.694	22.417	65.357	20.580
WORLD-Rus	84.4	97.3	69.695	22.475	65.357	20.581
WORLD-ROE	94.3	97.0	69.694	22.531	65.357	20.583
WORLD-ROW	93.7	96.2	69.695	22.537	65.357	20.583
WORLD-EU	92.8	91.6	69.697	22.557	65.357	20.583
WORLD-USA	77.6	90.6	69.714	22.526	65.357	20.583
WORLD-ROA	93.4	90.4	69.716	22.541	65.357	20.584
WORLD-Ind	90.8	89.7	69.715	22.543	65.357	20.584
WORLD-Chi	87.0	48.7	69.710	22.585	65.357	20.585

Note: υ_A is the coalition's share of global coal reserves, n_A is the coalition's share of global coal consumption, $W|_{\bar{K}=0}$ is global energy welfare, and $\Delta\frac{W_i}{n_i}$ is the fringe country's increase in per-capita welfare by joining the coalition.

In the competitive demand-side regime, any stable coalition is just large enough to ensure the ceiling ($n_A \geq 0.793$ for $\gamma = 0.89\%$ and $n_A \geq 0.849$ for $\gamma = 0.65\%$). Closer inspection of Table 1 reveals that any stable coalition comprises four to seven [six or seven] world regions including China, the smallest stable coalition consists of China, India, USA, Rest of the World, Rest of Europe, and Australia ($n_A = 0.793$) [China, India, Rest of Asia, EU, Rest of Europe, Russia ($n_A = 0.849$)], and the largest stable coalition consists of China, Rest of Asia, USA, EU, and Rest of the World ($n_A = 0.825$) [China, India, Rest of Asia, USA, Rest of World, Rest of Europe ($n_A = 0.874$)] for $\gamma = 0.89\%$ [$\gamma = 0.65\%$]. In the smallest and largest stable coalition, global energy welfare amounts to 63.6\$trillion and 64.7\$trillion [60.4\$trillion and 61.3\$trillion], respectively, each coalition country's per-capita energy welfare net of its fuel firm's profit share amounts to 55.8\$trillion and 58.6\$trillion [54.5\$trillion and 56.7\$trillion], respectively, and each fringe country's per-capita energy welfare net of its fuel firm's profit share amounts to 93.3\$trillion [93.3\$trillion] for $\gamma = 0.89\%$ [$\gamma = 0.65\%$].

Table B.7 provides information about coalitions of eight or nine world regions in the competitive supply-side regime.[75] It is always beneficial to join the coalition since $\Delta \frac{W_i}{n_i} > 0$ holds. Consequently, the grand coalition is stable. In this coalition, global energy welfare amounts to 69.691\$trillion [65.358\$trillion], and each coalition country's per-capita energy welfare net of its fuel firm's profit share amounts to 69.676\$trillion [65.357\$trillion] for $\gamma = 0.89\%$ [$\gamma = 0.65\%$].[76] Thus, Proposition 10 holds for $\gamma = 1.44\%$, $\gamma = 0.89\%$ and $\gamma = 0.65\%$ except that in the competitive demand-side regime the stable coalition gets larger as the natural regeneration rate gets smaller to ensure the ceiling.

In the strategic demand-side regime, no world region has an incentive to join any coalition unless the ceiling would otherwise be violated. For $\gamma = 0.89\%$, no coalition of three or fewer word regions and no coalition without China can adhere the ceiling, whereas each coalition of seven or more world regions with China can adhere the ceiling. Consequently, the grand coalition and the $n = 8$ coalitions are unstable. Furthermore, we find that the $n = 7$ coalitions are unstable, because some world region can leave any of these coalitions without violating the ceiling. The three stable $n = 6$ coalitions are {Chi + Ind + ROW + ROE + Rus + Aus}, {Chi + ROA + EU + ROW + Rus + Aus} and {Chi + ROA + ROW + ROE + Rus + Aus}, the twenty-three stable $n = 5$ coalitions are {Chi + Ind + ROA + ROW + ROE}, {Chi + Ind + ROA + ROW + Rus}, {Chi + Ind + ROA + ROW + Aus}, {Chi + Ind + ROA + ROE + Rus}, {Chi + Ind + ROA + ROE + Aus}, {Chi + Ind + ROA + Rus + Aus}, {Chi + Ind + USA + ROW + ROE}, {Chi + Ind + USA + ROW + Aus}, {Chi + Ind + USA + ROE + Aus}, {Chi + Ind + EU + ROW + ROE}, {Chi + Ind + EU + ROW + Aus}, {Chi + Ind + EU + ROE + Aus}, {Chi + ROA + USA + ROW + ROE}, {Chi + ROA + USA + ROW + Rus}, {Chi + ROA + USA + ROW + Aus}, {Chi + ROA + USA + ROE + Rus}, {Chi + ROA + USA + ROE + Aus}, {Chi + ROA + EU + ROW + ROE}, {Chi + USA + EU + ROW + ROE}, {Chi + USA + EU + ROW + Rus}, {Chi + USA + EU + ROW + Aus}, {Chi + USA + EU + ROE + Rus} and {Chi + USA + EU + ROE + Aus}, and the five stable $n = 4$ coalitions are {Chi + Ind + ROA +

[75] For coalitions of eight world regions, the budget constraint does not bind in the competitive supply-side regime.

[76] Note that the scarcity rent and, thus, the per-capita profit of the fuel firm is smaller in the demand-side regime than in the supply-side regime.

USA}, {Chi + Ind + ROA + EU}, {Chi + Ind + USA + EU}, {Chi + Ind + USA + Rus} and {Chi + ROA + USA + EU}. The smallest stable coalition consists of China, Rest of Asia, Rest of the World, Rest of Europe, Russia and Australia ($n_A = 0.726$), and the largest stable coalition consists of China, India, Rest of Asia and USA ($n_A = 0.806$).

For $\gamma = 0.65\%$, no coalition of four or fewer word regions and no coalition without China can adhere the ceiling, whereas each coalition of eight or more world regions with China can adhere the ceiling. Consequently, the grand coalition is unstable. Furthermore, we find that the $n = 8$ coalitions are unstable, because some world region can leave any of these coalitions without violating the ceiling. The four stable $n = 7$ coalitions are {Chi + Ind + ROA + ROW + ROE + Rus + Aus}, {Chi + Ind + EU + ROW + ROE + Rus + Aus}, {Chi + ROA + EU + ROW + ROE + Rus + Aus} and {Chi + USA + EU + ROW + ROE + Rus + Aus}, the sixteen stable $n = 6$ coalitions are {Chi + Ind + ROA + EU + ROW + ROE}, {Chi + Ind + ROA + EU + ROW + Rus}, {Chi + Ind + ROA + EU + ROW + Aus}, {Chi + Ind + ROA + EU + ROE + Rus}, {Chi + Ind + ROA + EU + ROE + Rus}, {Chi + Ind + ROA + EU + Rus + Aus}, {Chi + Ind + USA + ROW + ROE + Rus}, {Chi + Ind + USA + ROW + Rus + Aus}, {Chi + Ind + USA + ROE + Rus + Aus}, {Chi + ROA + USA + EU + ROW + ROE}, {Chi + ROA + USA + EU + ROW + Rus}, {Chi + ROA + USA + EU + ROW + Aus}, {Chi + ROA + USA + EU + ROE + Rus}, {Chi + ROA + USA + EU + ROE + Aus}, {Chi + ROA + USA + EU + Rus + Aus} and {Chi + ROA + USA + ROW + Rus + Aus}, and the eight stable $n = 8$ coalitions are {Chi + Ind + ROA + USA + EU}, {Chi + Ind + ROA + USA + ROW}, {Chi + Ind + ROA + USA + ROE}, {Chi + Ind + ROA + USA + Rus}, {Chi + Ind + ROA + USA + Aus}, {Chi + Ind + USA + EU + ROW}, {Chi + Ind + USA + EU + ROE} and {Chi + Ind + USA + EU + Rus}. The smallest stable coalition consists of China, India, USA, Rest of Europe, Russia and Australia ($n_A = 0.782$), and the largest stable coalition consists of China, India, Rest of Asia, USA and EU ($n_A = 0.890$).

Table B.8 provides information about coalitions of eight or nine world regions in the strategic supply-side regime.[77] It is always beneficial to join the coalition since $\Delta \frac{W_i}{n_i} > 0$

[77]For coalitions of eight world regions, the budget constraint binds in the strategic supply-side regime.

Table B.8: Large coalitions with the strategic supply-side policy (υ_A, n_A in %; $W|_{\bar{K}=0}, \Delta\frac{W_i}{n_i}$ in \$trillion).

			$\gamma = 0.89\%$		$\gamma = 0.65\%$	
	υ_A	n_A	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$	$W\|_{\bar{K}=0}$	$\Delta\frac{W_i}{n_i}$
WORLD	100	100	69.706	–	65.357	–
WORLD-Aus	86.0	98.5	69.706	22.730	65.357	20.641
WORLD-Rus	84.4	97.3	69.705	22.765	65.356	20.683
WORLD-ROE	94.3	97.0	69.705	22.753	65.356	20.693
WORLD-ROW	93.7	96.2	69.703	22.786	65.355	20.721
WORLD-EU	92.8	91.6	69.692	22.969	65.345	20.874
WORLD-USA	77.6	90.6	69.688	23.022	65.342	20.907
WORLD-ROA	93.4	90.4	69.688	23.015	65.342	20.913
WORLD-Ind	90.8	89.7	69.685	23.043	65.339	20.935
WORLD-Chi	87.0	48.7	69.282	24.388	65.017	22.020

Note: υ_A is the coalition's share of global coal reserves, n_A is the coalition's share of global coal consumption, $W|_{\bar{K}=0}$ is global energy welfare, and $\Delta\frac{W_i}{n_i}$ is the fringe country's increase in per-capita welfare by joining the coalition.

holds. Consequently, the grand coalition is stable. Thus, Propositions 11(i) and (iii) hold for $\gamma = 1.44\%$, $\gamma = 0.89\%$ and $\gamma = 0.65\%$ except that in the strategic demand-side regime the stable coalition gets larger as the natural regeneration rate gets smaller to ensure the ceiling.