Win as a Team or Fail as Individuals: Cooperation and Non-Cooperation in the Climate Tax Game*

Elmar Hillebrand† Marten Hillebrand‡

November 23, 2022

Abstract

This paper studies the strategic interaction of arbitrarily many heterogeneous regions in a macroeconomic growth model of climate change. Regions differ with respect to production technologies, productivity and factor endowments, stocks of fossil fuels, and climate damages. They trade on friction-less international markets for capital and fossil fuels and choose climate tax policies under different scenarios of cooperation and non-cooperation. We derive closed form solutions of optimal policies for both the non-cooperative equilibrium where regions internalize only domestic climate damages and the efficient solution internalizing global climate damages. We extend all these results to cases with partial cooperation where regions form coalitions and choose policies to maximize welfare of their coalition. Finally, we study how transfers can incentivize regions to cooperate and determine a class of optimal transfer schemes under which all regions are better-off under full cooperation relative to the non-cooperative outcome. Numerical simulations based on calibrated parameter values are used to illustrate and quantify our theoretical results under different coalition scenarios.

JEL classification: C63, E61, H21, H23, Q54

Keywords: Climate Change; Multi-region dynamic general equilibrium model; Nash-equilibrium; Cooperative solution; Coalitions; Optimal climate policy.

---

*Acknowledgements. We would like to thank Markus Epp, Marius Jäger, Oliver Landmann, and Klaus Wälde as well as participants of various research seminars including the Tea-time macro seminar at Freiburg University and the Chair Research Day at the University of Mainz for helpful suggestions and comments.

†EEFA Research Institute, Muenster, Germany, email: e.hillebrand@eefa.de

‡Department of Economic Theory, Albert-Ludwigs University Freiburg, Rempartstrasse 10-16, 79085 Freiburg im Breisgau, Germany, email: marten.hillebrand@vwl.uni-freiburg.de (corresponding author)

1 Introduction

Climate change is a global threat requiring a coordinated effort of all nations to significantly and permanently reduce their emissions from burning fossil fuels. Several climate summits accompanied by intense negotiations held over that past years and even decades suggest that there is wide agreement on this necessity. Yet, a political solution that all countries have committed to has not been reached. A major obstacle to implementing climate policies at the global level might be that nations or regions differ substantially along many dimensions such as their state of economic development, reliance on fossil fuels, and the climate damages they are projected to suffer. Therefore, the incentives for implementing climate polices vary considerably across regions. Past climate agreements have also seen the formation of coalitions being an integral part of negotiations. Hence, it is important to understand how the formation of coalitions determines the outcome of the political process and which policies remain feasible.

A theoretical analysis of optimal climate policies and their successful implementation must therefore be based on a framework which incorporates three key features. First, decisions on climate policies are taken by politically autonomous regions acting in their own self-interest rather than for the common good. Second, these regions differ substantially along key economic and other dimensions and, therefore, have different incentives for choosing and implementing a particular climate policy. Third, countries may not be willing to fully cooperate but instead form coalitions of regions with common political interests.

The present paper studies the existence and form of optimal climate policies in a model with heterogeneous regions that incorporates all of the previous features. Our analysis draws on the multi-region framework and results developed in our earlier work Hillebrand & Hillebrand (2019, 2022). In these papers, we focused on optimal climate policies under full cooperation. The present paper extends this approach to a general non-cooperative setup in which the fully cooperative scenario emerges as a special case.

Within this setup, we address the following specific questions. First, which climate policies emerge in a purely non-cooperative setting and how do they differ from the fully cooperative solution? Second, which policies emerge with partial cooperation and the formation of coalitions? Third, how can regions be incentivized to cooperate by means of side payments and how can such transfer policies be characterized?

Answering these and a number of related questions is the general contribution of this paper. Specifically, we derive closed form solutions of optimal climate policies for both the fully non-cooperative and the fully cooperative case and extend these results to scenarios with partial cooperation where regions form coalitions. All these results take a very simple and intuitive form. Third, we provide a complete characterization of transfer schemes which redistribute the gains from cooperation such that each region has an incentive to cooperate. This defines a range of admissible transfers that forms the basis for successful climate negotiations.

Our paper contributes to a large and growing literature studying the climate problem from an economic perspective. Modern representatives of this field adopt the framework of dynamic general equilibrium theory which is the standard approach in macroeconomics.[1] Based on this approach, the paper closest to ours is Hambel, Kraft & Schwartz (2021), henceforth HKS who derive optimal climate policies corresponding to abatement efforts in a fully non-cooperative setup. A main contribution of our work relative to HKS and others is that we allow for regions to trade on international markets for capital and fossil fuels not permitted in HKS. Using the implicit definition of the social cost of carbon from Nordhaus (2014), HKS also obtain a measure of the implied regional emissions taxes. In this paper, we use a different concept of the social cost of carbon derived from an explicitly defined planning problem in line with Golosov et al.(2014) and Hillebrand & Hillebrand (2019). This permits to derive the optimal climate tax in closed form and admits a straightforward economic interpretation.

Our work also contributes to the game theoretic literature on climate change advanced by Harstad (2016, 2012) and Battaglini & Harstad (2016). These papers usually focus on game theoretic aspects and, therefore, employ highly stylized models of both the macroeconomy and the climate system. As a consequence, the cost of climate change and damages are specified directly in a somewhat ad-hoc fashion rather than being derived as an endogenous outcome of the interaction between the economic production sector and the climate system. Relative to this approach, we maintain the full-fledged growth framework of Hillebrand & Hillebrand (2019) featuring an explicit description of the production process and the climate system. Despite this more detailed specification, we retain the virtue of being able to derive analytical results including closed form solutions of optimal climate policies under different levels of cooperation among regions.

A final strand of research studies scenarios of cooperation and non-cooperation based on the RICE model developed in Nordhaus & Yang (1996). Setting aside the conceptual problems facing the RICE model and the derivation of its solution (cf. Denning & Emmerling (2017) or Hillebrand & Hillebrand (2019)), all these studies are confined to purely numerical results and analytical expressions can not be derived.[2] In addition, the RICE-framework severely restricts or even excludes trade among countries.

The paper is organized as follows. Section 2 introduces the model. Section 3 derives the decentralized equilibrium under arbitrary climate policies. Section 4 studies the non-cooperative solution which is compared to the efficient solution under full cooperation in Section 5. Section 6 studies the case with partial cooperation under formation of

[1] Examples are Golosov et al. (2014), Hassler & Krusell (2012), or Gerlagh & Liski (2018).

[2] Models in this category include Bosetto et al. (2003), Eyckmans & Tulkens (2003), Eyckmans & Finus (2003), or Finus et al. (2014).

coalitions. Section 7 presents results from a numerical simulation study. Section 8 concludes. Mathematical proofs and computational details are placed in the appendix.

2 The Model

2.1 World economy

The world economy is divided into $L \geq 1$ autonomous regions, indexed by $\ell \in \mathbb{L} := \{1,\ldots,L\}$. Time evolves in discrete periods $t \in \{0,1,2,\ldots\}$. All variables determined prior to the initial period $t = 0$ are treated as given parameters in the following setup. We express the regional dependence of variables by a superscript ℓ and identify summation of this variable over all regions by a bar superscript. For example, X_t^ℓ will denote fossil fuel consumption of region ℓ in period t such that $\overline{X}_t := \sum_{\ell\in\mathbb{L}} X_t^\ell$ denotes global emissions in period t. The major building blocks of the model are the production sector, the climate model, and the consumption sector which are now explained in detail.

2.2 Production sector

Final output

In each region $\ell \in \mathbb{L}$ a single representative firm produces a homogeneous final output commodity Y_t^ℓ using capital K_t^ℓ and fossil fuels X_t^ℓ as inputs. The production technology is of the general form

$$Y_t^\ell = (1-D_t^\ell)F_t^\ell(K_t^\ell, X_t^\ell). \tag{1}$$

Here, $D_t^\ell \in [0,1]$ denotes endogenous climate damage which is further specified below. Capital input K_t^ℓ is rented in the international capital market at price r_t. Input factor X_t^ℓ subsumes different kinds of fossil fuels such as coal, oil, etc. which are purchased in the global resource market at price v_t. Time-dependence of the production function F_t^ℓ captures both regional population growth as well exogenous technological progress of labor and energy efficiency.

Restrictions on technology

We impose the following standard restrictions on the production function F_t^ℓ in (1). The left-side Inada condition (2) ensures that each factor will be employed at equilibrium.

Assumption 1

The function $F_t^\ell : \mathbb{R}_+^2 \longrightarrow \mathbb{R}_+$ is strictly increasing, strictly concave, and continuously differentiable on $\mathbb{R}_{++}^2$. The partial derivatives satisfy the boundary condition

$$\lim_{z_i \searrow 0} \frac{\partial F_t^\ell(z_1,z_2)}{\partial z_i} = \infty \quad \textit{for both } i = 1,2 \textit{ and all } z = (z_1,z_2) \in \mathbb{R}_{++}^2. \tag{2}$$

Extraction of fossil fuels

Region $\ell \in \mathbb{L}$ possesses an initial stock of fossil fuels $R_0^\ell \geq 0$ which can be extracted at constant unit costs $c_x \geq 0$. Extraction of fossil fuels takes place in a resource sector operated by a single firm which chooses an extraction sequence $(X_t^{\ell,s})_{t\geq 0}$ with $X_t^{\ell,s} \geq 0$ denoting the extraction of fossil fuels in period t. Feasible extraction plans satisfy

$$\sum_{t=0}^{\infty} X_t^{\ell,s} \leq R_0^\ell. \tag{3}$$

Extracted resources $X_t^{\ell,s}$ are supplied to the global resource market in each period t. To avoid trivialities, we impose the initial condition $\sum_{\ell\in\mathbb{L}} R_0^\ell > 0$, i.e., initial world resources of fossil fuels are strictly positive.

2.3 Climate model

Emissions

Emissions of CO_2 are generated by using fossil fuels in production. Measuring fossil fuel inputs X_t^ℓ in (1) directly in units of CO_2, global emission in period t are given by

$$\overline{X}_t := \sum_{\ell\in\mathbb{L}} X_t^\ell. \tag{4}$$

Emissions prior to $t = 0$ are given and we assume that $\overline{X}_{-t} = 0$ for $t > 0$ sufficiently large reflecting the fact that emissions before the industrial age were uniformly zero.

Atmospheric level of carbon

Climate damages in period t are determined by atmospheric carbon concentration S_t relative to the pre-industrial level that depends on past aggregate emissions, i.e.,

$$S_t = \sum_{n=0}^{\infty} \delta_n \overline{X}_{t-n}. \tag{5}$$

The non-negative sequence $(\delta_n)_{n\geq 0}$ in (5) determines the evolution and persistence of emissions in the atmosphere. The specification (5) encompasses various climate models in the literature including Golosov et al. (2014) or Gerlagh & Liski (2018).

Climate damages

Climate damages in region $\ell \in \mathbb{L}$ at time t depend exclusively on carbon concentration S_t given by (5) and are determined by the damage function $D^\ell : \mathbb{R}_+ \longrightarrow [0,1[$,

$$D_t^\ell = D^\ell(S_t) := 1 - \exp\{-\gamma^\ell S_t\}, \quad \gamma^\ell > 0. \tag{6}$$

Regional differences in climate damages are captured by the region-specific parameter γ^ℓ, $\ell \in \mathbb{L}$. The exponential form (6) is also widely used in the literature, cf. again Golosov et al. (2014) or Gerlagh & Liski (2018). Golosov et al. (2014) show that this specification approximates well the damage function of the DICE-model ((Nordhaus & Sztorc 2013)) in the empirically relevant range.

2.4 Consumption sector

Representative consumer

The consumption sector in region $\ell \in \mathbb{L}$ consists of a single representative household who supplies capital $K_t^{\ell,s}$ to the global capital market in each period t. Initial capital $K_0^{\ell,s}$ in period $t = 0$ is taken as given in the decision. In addition, the consumer is entitled to receive all profits from domestic firms and transfers from the government.

Consumer preferences

The household's preferences over non-negative consumption sequences $(C_t^\ell)_{t\geq 0}$ are represented by a standard time-additive utility function

$$U((C_t^\ell)_{t\geq 0}) = \sum_{t=0}^{\infty} \beta^t u(C_t^\ell) \quad \text{where} \quad u(C) = \begin{cases} \frac{C^{1-\sigma}}{1-\sigma} & \text{for} \quad \sigma > 0, \sigma \neq 1 \\ \log(C) & \text{for} \quad \sigma = 1. \end{cases} \tag{7}$$

The discount factor satisfies $0 < \beta < 1$. The previous specification is widely used in models of climate change. It is key for the separability between efficiency and optimality exploited in Hillebrand & Hillebrand (2019) to derive an optimal climate policy.

2.5 Summary of the economy

The economy $\mathscr{E}$ introduced in the previous sections can be summarized by its regional structure, production parameters, climate model, and consumer parameters. Formally,

$$\mathscr{E} = \left\langle \mathbb{L}, \left((F_t^\ell)_{t\geq 0}\right)_{\ell\in\mathbb{L}}, c_x, (\delta_n)_{n\geq 0}, (\gamma^\ell)_{\ell\in\mathbb{L}}, \beta, \sigma \right\rangle. \tag{8}$$

In addition, the initial distribution of capital $(K_0^{\ell,s})_{\ell\in\mathbb{L}}$ and initial stocks of fossil fuels $(R_0^\ell)_{\ell\in\mathbb{L}}$ are given as well as aggregate emissions $(\overline{X}_{-t})_{t>0}$ prior to $t = 0$.

3 Decentralized Equilibrium

The decentralized equilibrium reconciles optimal behavior of consumers and producers in each region with market clearing on regional and global markets. The equilibrium is determined for a given climate policy chosen by each region to be specified next.

3.1 Climate policy

Climate tax

Each region $\ell \in \mathbb{L}$ levies a tax τ_t^ℓ on regional emissions in period t to be paid by the final sector. The tax sequence $(\tau_t^\ell)_{t\geq 0}$ is the first building block of a climate policy.

This sequence may consist of given exogenous numbers or can be generated by a time-invariant rule that depends on endogenous variables.

Transfers
All tax revenue is returned to consumers as a lump sum transfer. As for the distribution of transfers across regions, we distinguish the cooperative and the non-cooperative case.

In the non-cooperative case, each region simply rebates its tax revenue to domestic consumers implying that consumers in region receive a transfer in period t equal to

$$T_t^\ell = \tau_t^\ell X_t^\ell \quad \text{for each } \ell \in \mathbb{L}. \tag{9}$$

In the cooperative case, regions pool their tax revenue and agree on a transfer policy $\theta : \mathbb{L} \longrightarrow \mathbb{R}, \ell \mapsto \theta^\ell$ satisfying $\sum_{\ell \in \mathbb{L}} \theta^\ell = 1$ which determines the transfer share θ^ℓ received by region ℓ. Transfers received by consumers in region ℓ at time t then follow as

$$T_t^\ell = \theta^\ell \sum_{k \in \mathbb{L}} \tau_t^k X_t^k \quad \text{for each } \ell \in \mathbb{L}. \tag{10}$$

The assumption of time-invariant transfers shares imposes no restriction since consumer behavior depends exclusively on lifetime transfers defined below. Any time-dependent distribution of transfers is therefore equivalent to a time-invariant transfer policy in the sense that it implies the same consumption behavior.

3.2 Producer behavior

Final sectors
The representative firm in the final sector chooses non-negative factor inputs in period t to maximize its period-profit, taking climate damages and prices for capital, labor, and fossil fuels as given. The latter includes the tax on emissions. The formal problem determining profits Π_t^ℓ in period t reads:

$$\Pi_t^\ell = \max_{K^\ell, X^\ell \geq 0} \left\{ (1 - D_t^\ell) F_t^\ell (K^\ell, X^\ell) - r_t K^\ell - (v_t + \tau_t^\ell) X^\ell \right\}. \tag{11}$$

Profit maximizing factor demand (K_t^ℓ, X_t^ℓ) solves the standard first order conditions[3]

$$(1 - D_t^\ell) \partial_K F_t^\ell (K_t^\ell, X_t^\ell) = r_t \tag{12a}$$
$$(1 - D_t^\ell) \partial_X F_t^\ell (K_t^\ell, X_t^\ell) = v_t + \tau_t^\ell. \tag{12b}$$

Profits determined by (11) are fully transferred to consumers in region ℓ.

Resource sector
Unlike the final sector, the resource sector solves an intertemporal decision problem

[3]We denote by $\partial_K F(K,X) := \frac{\partial F(K,X)}{\partial K}$ and $\partial_X F(K,X) := \frac{\partial F(K,X)}{\partial X}$ the partial derivatives of a function F.

involving a stream of future profits. To discount these payments to period $t = 0$, we define the time t-discount factor (Arrow-Debreu price) as

$$q_t := \prod_{n=1}^{t} r_n^{-1} \text{ for each } t = 0,1,2,\dots \text{ where } q_0 = 1. \tag{13}$$

The resource firm's discounted profit stream in $t = 0$ generated by an optimal extraction sequence is then determined as

$$\Pi_x^\ell := \max_{(X_t^{\ell,s})_{t\geq 0}} \left\{ \sum_{t=0}^{\infty} q_t (v_t - c_x) X_t^{\ell,s} \,\middle|\, (3) \text{ holds, } X_t^{\ell,s} \geq 0 \text{ for all } t \geq 0 \right\}. \tag{14}$$

Applying standard arguments (cf. Hillebrand & Hillebrand (2019)), the existence of an optimal extraction sequence requires the Hotelling rule

$$v_t - c_x = r_t(v_{t-1} - c_x) \text{ for all } t = 1,2,3,\dots \tag{15}$$

under which maximum profits (11) are given by

$$\Pi_x^\ell = (v_0 - c_x) R_0^\ell. \tag{16}$$

3.3 Consumer behavior

In each period t, the representative consumer in region ℓ receives factor income from supplying capital $K_t^{\ell,s}$ formed in the previous period $t-1$. In addition, the consumer collects profits Π_t^ℓ from the final sector and $\Pi_{x,t}^\ell = (v_t - c_x) X_t^{\ell,s}$ from the domestic resource sector as well as transfers T_t^ℓ from the government. Consumption C_t^ℓ and newly formed capital $K_{t+1}^{\ell,s}$ satisfy the period budget constraint

$$K_{t+1}^{\ell,s} = \Pi_t^\ell + r_t K_t^{\ell,s} + (v_t - c_x) X_t^{\ell,s} + T_t^\ell - C_t^\ell. \tag{17}$$

We interpret $K_{t+1}^{\ell,s}$ as the consumer's net asset position and, therefore, do not exclude negative values. To exclude Poni-schemes, however, we impose the usual condition $\lim_{t\to\infty} q_t K_{t+1}^{\ell,s} \geq 0$. Using this and (16) and defining lifetime transfer income $T^\ell := \sum_{t=0}^{\infty} q_t T_t^\ell$ and the discounted stream of final profits $\Pi^\ell := \sum_{t=0}^{\infty} q_t \Pi_t^\ell$ we can solve (17) forward to obtain the lifetime budget constraint

$$\sum_{t=0}^{\infty} q_t C_t^\ell \leq \Pi^\ell + r_0 K_0^{\ell,s} + (v_0 - c_x) R_0^\ell + T^\ell. \tag{18}$$

The consumer chooses consumption $(C_t^\ell)_{t\geq 0}$ subject to (18) to maximize lifetime utility U defined in (7). Optimality is determined by equality of (18) and the Euler equation

$$C_{t+1}^\ell = C_t^\ell (\beta r_{t+1})^{\frac{1}{\sigma}} \quad \text{for all } t = 0,1,2,\dots \tag{19}$$

3.4 Market clearing and equilibrium

Market clearing

In each period t, capital and fossil fuels are traded on international markets. The market clearing conditions for period t read:

$$\sum_{\ell\in\mathbb{L}} K_t^{\ell} \stackrel{!}{=} \sum_{\ell\in\mathbb{L}} K_t^{\ell,s} \quad \text{and} \quad \sum_{\ell\in\mathbb{L}} X_t^{\ell} \stackrel{!}{=} \sum_{\ell\in\mathbb{L}} X_t^{\ell,s}. \tag{20}$$

Since the Hotelling rule (15) makes resource firms indifferent between the timing of extraction, the extraction sequence $(X_t^{\ell,s})_{t\geq 0}$ will, in general, be indeterminate at equilibrium. Due to (17), the same will be true of regional capital supply $(K_{t+1}^{\ell,s})_{t\geq 0}$. However, regional consumption of fossil fuels $(X_t^{\ell})_{t\geq 0}$ is uniquely determined at equilibrium and satisfies the world resource constraint

$$\sum_{t=0}^{\infty} \sum_{\ell\in\mathbb{L}} X_t^{\ell} = \sum_{t=0}^{\infty} \overline{X}_t \leq \overline{R}_0 := \sum_{\ell\in\mathbb{L}} R_0^{\ell}. \tag{21}$$

Writing profits (11) as $\Pi_t^{\ell} = Y_t^{\ell} - r_t K_t^{\ell} - (v_t + \tau_t^{\ell}) X_t^{\ell}$ and exploiting the definition of transfers (9) or (10) we can sum the consumer's budget constraint (17) over all regions ℓ and combine the result with the market clearing conditions (20) to obtain the evolution of world capital as

$$\sum_{\ell\in\mathbb{L}} K_{t+1}^{\ell} = \sum_{\ell\in\mathbb{L}} Y_t^{\ell} - c_x \sum_{\ell\in\mathbb{L}} X_t^{\ell} - \sum_{\ell\in\mathbb{L}} C_t^{\ell}. \tag{22}$$

Equation (22) can also be interpreted as a market clearing condition for the world commodity market.

Finally, for purposes of a compact definition of equilibrium we combine (4), (5), and (6) to obtain regional climate damages as a function of past regional emissions given by

$$D_t^{\ell} = 1 - \exp\left(-\gamma^{\ell} \sum_{n=0}^{\infty} \delta_n \sum_{\ell\in\mathbb{L}} X_{t-n}^{\ell}\right) \quad \text{for all} \quad \ell\in\mathbb{L}. \tag{23}$$

Definition of equilibrium

We are now in a position to formally define a decentralized equilibrium for this economy.

Definition 1

An equilibrium of $\mathscr{E}$ *consists of climate taxes and transfers* $((\tau_t^{\ell}, T_t^{\ell})_{\ell\in\mathbb{L}})_{t\geq 0}$, *an allocation* $A^* = ((K_t^{\ell*}, X_t^{\ell*}, C_t^{\ell*})_{\ell\in\mathbb{L}})_{t\geq 0}$, *and prices* $P^* = (r_t^*, v_t^*)_{t\geq 0}$ *such that for all* $t = 0, 1, 2, \ldots$*:*

(i) Factor inputs $(K_t^{\ell*}, X_t^{\ell*})$ *solve conditions (12) for given prices* (r_t^*, v_t^*), *taxes* τ_t^{ℓ}, *and damages* D_t^{ℓ} *determined by (23) for each region* $\ell\in\mathbb{L}$.

(ii) Prices P^* *satisfy the Hotelling rule (15) and global fossil fuel consumption is consistent with the world resource constraint (21).*

(iii) *Taxes and transfers and emissions are consistent with (9) or (10).*

(iv) *Regional consumption* $(C_t^{\ell *})_{t\geq 0}$ *satisfies the Euler equation (19) and constraint (18) with equality with discount factors* $(q_t^*)_{t\geq 0}$ *determined by (13).*

(v) *Initial capital* $(K_0^\ell)_{\ell\in\mathbb{L}}$ *satisfies the capital market clearing condition (20) for* $t=0$ *and (22) holds for all* t *with output determined by (1) and damages by (23).*

Properties of equilibrium

The form of the utility function (7) and a frictionless capital market imply a world-consumption distribution that is constant over time. Thus, each region acquires a constant share of world consumption in each period. We state this result formally in the next lemma. The proof is analogous to Hillebrand & Hillebrand (2019) and is omitted.

Lemma 1

Equilibrium consumption of each region $\ell \in \mathbb{L}$ *is a constant fraction* $\mu^\ell \in]0,1[$ *of world consumption* $\overline{C}_t^* := \sum_{\ell\in\mathbb{L}} C_t^{\ell *}$ *in each period* t*, i.e.,*

$$C_t^{\ell *} = \mu^\ell \, \overline{C}_t^* \;\; \text{where} \;\; \mu^\ell = \frac{\Pi^\ell + r_0 K_0^{\ell,s} + (v_0 - c_x) R_0^\ell + T^\ell}{\sum_{k\in\mathbb{L}} \left(\Pi^k + r_0 K_0^{k,s} + (v_0 - c_x) R_0^k + T^k \right)}. \tag{24}$$

4 The Non-Cooperative Equilibrium

This section considers the non-cooperative equilibrium in which each region ℓ sets its climate tax policy $\tau^\ell = (\tau_t^\ell)_{t\geq 0}$ to maximize domestic welfare, taking as given the decisions of other regions. In this non-cooperative scenario, transfers are determined by (9) excluding net transfers between regions. The regionally optimal tax policy can be derived in two steps. In the first step, each region chooses a regionally optimal allocation taking emissions from other regions as well as global prices for capital and fossil fuels as given. The second step then determines regional climate taxes such that the regionally optimal allocation materializes as a decentralized equilibrium.

4.1 Regional planning problem

Constraints

The planner in region ℓ takes local initial stocks of capital $K_0^{\ell,s}$ and fossil fuels $R_0^\ell \geq 0$ as well as aggregate emissions $(\overline{X}_{-t})_{t\geq 1}$ prior to $t=0$ as fixed parameters. Likewise, the sequence $(\overline{X}_t^{-\ell})_{t\geq 0}$ of future emissions $\overline{X}_t^{-\ell} := \sum_{k\in\mathbb{L}\setminus\{\ell\}} X_t^k$ from other regions as well as international prices $P = (r_t, v_t)_{t\geq 0}$ are taken as given. We assume that these prices satisfy the Hotelling rule (15) as they will at equilibrium and define $(q_t)_{t\geq 0}$ by (13).

A crucial deviation from the decentralized solution is that the planner in region ℓ takes into account the impact of its own emissions and those of other regions on local climate damage D_t^ℓ determined by (23). Using the previous notation, we write damages as

$$D_t^\ell = 1 - \exp\left\{-\gamma^\ell \sum_{n=0}^{\infty} \delta_n \left(X_{t-n}^\ell + \overline{X}_{t-n}^{-\ell}\right)\right\} \quad \text{for } \ell \in \mathbb{L}. \tag{25}$$

The regional planner chooses factor inputs to the production technology (1) in each period t. The decision also involves the accumulation of capital and extraction of fossil fuels. However, these variables are not necessarily fully employed in domestic production but can also be traded on global capital and resource markets at the given prices. As before, a superscript 's' signifies variables supplied to these markets. The regional stock of capital then evolves according to the resource condition

$$K_{t+1}^{\ell,s} = (1-D_t^\ell)F_t^\ell(K_t^\ell, X_t^\ell) + r_t\big(K_t^{\ell,s} - K_t^\ell\big) + v_t\big(X_t^{\ell,s} - X_t^\ell\big) - c_x X_t^{\ell,s} - C_t^\ell. \tag{26}$$

It follows from (3) and the Hotelling rule (15) that the extraction sequence $(X_t^{\ell,s})_{t\geq 0}$ chosen by region ℓ generates total discounted revenue given again by (16). Using this and imposing again the No-Ponzi condition $\lim_{T\to\infty} K_{T+1}^{\ell,s} q_T \geq 0$, one can solve (26) forward to obtain the lifetime budget constraint

$$\sum_{t=0}^{\infty} q_t \Big(r_t K_t^\ell + v_t X_t^\ell + C_t^\ell - (1-D_t^\ell)F_t^\ell(K_t^\ell, X_t^\ell)\Big) \leq r_0 K_0^{\ell,s} + (v_0 - c_x)R_0^\ell. \tag{27}$$

This condition somewhat parallels the consumer's condition (18). The main difference is that the planer chooses inputs to final production directly and, more importantly, takes into account the impact of fossil fuel inputs X_t^ℓ on domestic climate damages via (25).

The regional planning problem

Using (25) and (27) the planning problem of region $\ell \in \mathbb{L}$ can now be stated as follows.

$$\max_{(K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}} \left\{ U((C_t^\ell)_{t\geq 0}) \,\middle|\, \text{(25) and (27) hold, } K_t^\ell, X_t^\ell, C_t^\ell \geq 0 \text{ for all } t = 0,1,2,\ldots \right\}. \tag{28}$$

Clearly, the solution to (28) depends on the given emissions $(\overline{X}_t^{-\ell})_{t\geq 0}$ of other regions and prices $P = (r_t, v_t)_{t\geq 0}$. Crucially, the planner is a 'price-taker' on international markets and does not take into account the impact of his decision on the determination of these prices as equilibrium outcomes.

Solution to the problem

Problem (28) is a constrained optimization problem that can be solved by standard Lagrangean methods. The details can be found in Section A.1 in the Appendix. This way, on can derive the following optimality conditions which characterize the solution in addition to the technological and resource constraints. First, consumption satisfies the Euler equation (19) and (27) holds with equality. Second, the marginal product of capital in period t equals its return r_t such that (12a) holds. Finally, resource input in period t earns a marginal product equal to its price v_t plus the total discounted marginal damage in production:

$$(1-D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) = v_t + \sum_{n=0}^{\infty} \beta^n \left(C_{t+n}^\ell / C_t^\ell\right)^{-\sigma} \delta_n \gamma^\ell Y_{t+n}^\ell. \tag{29}$$

Intuitively, the damage-related term in (29) consists of three factors. First, a discount factor $\beta^n \left(C_{t+n}^\ell / C_t^\ell\right)^{-\sigma}$ serving to measure damages in $t+n$ in units of time t consumption. Second, the term δ_n which measures the quantity of emissions at time t that are still in the atmosphere at time $t+n$. Third, the marginal loss in domestic output Y_{t+n}^ℓ at time $t+n$ caused by an additional unit of carbon in the atmosphere. Summation of these factors over all periods $t, t+1, t+2, \ldots$ then gives the total domestic damage generated by one additional unit of emissions in period t.

For later reference, we state the previous result formally in the next lemma.

Lemma 2

Let prices $P = (r_t, v_t)_{t\geq 0}$ and emissions $(\overline{X}_t^{-\ell})_{t\geq 0}$ of other regions be given. If the regional allocation $(K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}$ solves (12a), (19), and (29) with D_t^ℓ determined by (25) for all $t = 0, 1, 2, \ldots$ and satisfies (27) with equality, then it is a solution to the problem (28).

4.2 Non-cooperative solution as a Nash equilibrium

The market maker

In a non-cooperative equilibrium the decisions of all regions $\ell \in \mathbb{L}$ determined as a solution to (28) are mutually compatible in the sense that emissions of all other regions are correctly anticipated and market clearing on global capital and resource markets obtains in each period t. To formalize this idea, we embed the previous structure into the setup of a non-cooperative game and invoke the formal definition of a Nash equilibrium. Since the regional planning problem involves not only the decision variables from other regions, but also the prices of capital and fossil fuels, we introduce an additional player $\ell = 0$ that will be referred to as a 'market-maker'. This player chooses prices $P = (v_t, r_t)_{t\geq 0}$ as his strategy subject to the Hotelling rule (15). Since the Hotelling rule determines all future prices $(v_t)_{t>0}$ from the initial choice v_0 and capital returns and choosing the sequence $(r_t)_{t\geq 1}$ is equivalent to choosing the AD-prices $(q_t)_{t\geq 0}$, we can identify the choice of player $\ell = 0$ with the strategy $(v_0, r_0, (q_t)_{t\geq 0})$.

Decision problem of the market maker

We now set up the market maker's decision problem such that market clearing obtains along the Nash equilibrium defined below. Suppose player $\ell = 0$ bases his choice of prices on the 'mismatch' between demand and supply on international markets determined by the allocations chosen by players. To this end, the resource price v_0 is chosen to maximize the value of excess demand in the resource market.[4] Formally, the price is a solution to the linear problem

$$\max_{v_0}\left\{(v_0-c_x)\left(\sum_{t=0}^{\infty}\sum_{\ell\in\mathbb{L}}X_t^\ell-\overline{R}_0\right)\middle|c_x\le v_0\le v_0^{\max}\right\}. \tag{30a}$$

Here, $v_0^{\max}>0$ is some arbitrary upper bound chosen sufficiently large to exceed any possible equilibrium value v_0^{nc} defined below. Thus, the market maker sets the resource price to its minimal value $v_0=c_x$ if the total demand for resources $\sum_{t=0}^{\infty}\sum_{\ell\in\mathbb{L}}X_t^\ell$ summed over all regions and periods is lower than the available resource stock $\overline{R}_0$. Conversely, if resource demand exceeds this stock, the resource price is set to its maximum value $v_0=v_0^{\max}$. It follows that problem (30a) has an interior solution $c_x<v_0<v_0^{\max}$ (in fact, infinitely many) if and only if the resource constraint (21) holds with equality such that total demand exhausts the global stock.

In a similar fashion, the market maker determines r_0 and the values q_t for $t=0,1,2,\dots$ based on the excess demand for capital in $t=0$ and in the consumption good market in period t, respectively. Formally, these problems read

$$\max_{r_0}\left\{r_0\sum_{\ell\in\mathbb{L}}\left(K_0^\ell-K_0^{\ell,s}\right)\middle|0\le r_0\le r_0^{\max}\right\} \tag{30b}$$

and, for each $t=0,1,2,\dots$

$$\max_{q_t}\left\{q_t\sum_{\ell\in\mathbb{L}}\left(C_t^\ell+c_xX_t^\ell+K_{t+1}^\ell-(1-D_t^\ell)F_t^\ell(K_t^\ell,X_t^\ell)\right)\middle|0\le q_t\le q_t^{\max}\right\}. \tag{30c}$$

Again, the upper bounds $r_0^{\max}$ and $q_t^{\max}$ are chosen sufficiently large and the initial distribution of capital $(K_0^{\ell,s})_{\ell\in\mathbb{L}}$ entering (30b) is given and climate damage D_t^ℓ in (30c) determined by (25) for each ℓ. Linearity of each of the underlying problems then implies that existence of interior solutions to (30b) and (30c) is equivalent to the capital market clearing condition in (20) for $t=0$ and (22) for each $t=0,1,2,\dots$[5]

Definition of non-cooperative equilibrium

Using $\mathbb{L}_0:=\{0\}\cup\mathbb{L}$ as the set of players we can now define a Nash equilibrium as follows.

Definition 2

A non-cooperative (Nash) equilibrium of $\mathscr{E}$ is an allocation $A^{\mathrm{nc}}=((K_t^{\ell,\mathrm{nc}},X_t^{\ell,\mathrm{nc}},C_t^{\ell,\mathrm{nc}})_{\ell\in\mathbb{L}})_{t\ge0}$ and prices $P^{\mathrm{nc}}=(r_t^{\mathrm{nc}},v_t^{\mathrm{nc}})_{t\ge0}$ such that the following holds:

(i) The regional allocation $(K_t^{\ell,\mathrm{nc}},X_t^{\ell,\mathrm{nc}},C_t^{\ell,\mathrm{nc}})_{t\ge0}$ chosen by player $\ell\in\mathbb{L}$ solves the planning problem (28) given prices P^{nc} and emissions $(\overline{X}_t^{-\ell,\mathrm{nc}})_{t\ge0}$ of other regions.

[4]This idea is used in finite-dimensional Walrasian economies to construct a price correspondence of which equilibrium prices obtain as fixed points. Details of this approach can be found, e.g., in Mas-Colell, Whinston & Green (1995, Section 17.C).

[5]Since the choices of r_0, v_0, and $(q_t)_{t\ge0}$ are independent, one could also determined them simultaneously by maximizing the sum of (30a), (30b), and (30c) summed over all $t=0,1,2,\dots$

(ii) *Prices P^{nc} chosen by player $\ell = 0$ satisfy the Hotelling rule (15) and $(v_0^{\mathrm{nc}}, r_0^{\mathrm{nc}})$ and q_t^{nc} defined by (13) are solutions to the decision problems (30a)-(30c) for all $t \geq 0$.*

It is worth noting that (21) may or may not bind along the non-cooperative equilibrium. If it does, fossil fuels carry a scarcity rent implying $v_0 > c_x$ and, by (15), $v_t > c_x$ for all t. On the other hand, if $\sum_{t=0}^{\infty} \sum_{\ell \in \mathbb{L}} X_t^{\ell,\mathrm{nc}} < \overline{R}_0$, the solution to problem (30a) satisfies $v_0 = c_x$ and implies $v_t = c_x$ for all t by (15). In this case, fossil fuels are abundant and there is no scarcity rent.

4.3 Time-consistency of the non-cooperative solution

In this section we show that the non-cooperative equilibrium is time-consistent in the sense that in each future period $t = N \geq 1$, each player if permitted to re-optimize will stick to his original strategy provided everybody else does. In other words, the Nash equilibrium of the non-cooperative game defined above satisfies the one-shot deviation principle, i.e., no player can increase their pay-off by singularly deviating from the equilibrium strategy in any period. Hence, the non-cooperative equilibrium from Definition 2 is time-consistent or also subgame perfect.

To show this formally, define the wealth of the consumer in region $\ell \in \mathbb{L}$ at the beginning of period $t = 0, 1, 2, \dots$ as

$$W_t^\ell := r_t K_t^{\ell,s} + (v_t - c_x) R_t^\ell \tag{31}$$

where R_t^ℓ is the regional resource stock at the beginning of period t determined recursively as

$$R_{t+1}^\ell = R_t^\ell - X_t^{\ell,s} \quad \text{for all } t = 0, 1, 2, \dots \tag{32}$$

Intuitively, wealth W_t^ℓ consist of capital income and the value of the current resource stock net of extraction costs in period t. Note that initial wealth W_0^ℓ appears on the right-hand side of the lifetime budget constraint (27).

Combining definition (31) and (32) with the period-budget constraint (17), the evolution of the wealth sequence $(W_t^\ell)_{t \geq 0}$ is determined recursively by the equation

$$W_{t+1}^\ell = r_{t+1}\left(W_t^\ell + (1 - D_t^\ell) F_t^\ell(K_t^\ell, X_t^\ell) - r_t K_t^\ell - v_t X_t^\ell - C_t^\ell\right) \quad \text{for all } t = 0, 1, 2, \dots \tag{33}$$

with the initial value W_0^ℓ determined by (31) from the given initial values $K_0^{\ell,s}$ and R_0^ℓ. Solving (33) forward one obtains initial wealth at the beginning of time $t = N > 0$ as:

$$q_N W_N^\ell = \sum_{t=0}^{N-1} q_t \left((1 - D_t^\ell) F_t^\ell(K_t^\ell, X_t^\ell) - C_t^\ell - r_t K_t^\ell - v_t X_t^\ell\right) + W_0^\ell. \tag{34}$$

Observe that although the consumer's net capital position $K_t^{\ell,s}$ and regional extraction $X_t^{\ell,s}$ are, in general, indeterminate at equilibrium, the consumer's wealth position is uniquely determined recursively by (34).

Consider now the behavior of the consumer in region $\ell \in \mathbb{L}$ in some period $t = N \geq 1$. Let previous wealth W_{N-1}^{ℓ} and the strategies of all regions $(K_t^{\ell}, X_t^{\ell}, C_t^{\ell})_{0\leq t<N}$ as well as the sequence $(r_t, v_t)_{0\leq t<N}$ chosen by the market maker prior to period $t = N$ be given. In particular, aggregate emissions $(\overline{X}_t)_{t<N}$ prior to period $t = N$ are given. Suppose in period $t = N$ each player $\ell \in \mathbb{L}$ chooses an updated strategy $(C_t^{\ell}, X_t^{\ell}, K_t^{\ell})_{t\geq N}$ to maximize the remaining utility

$$U_N((C_t^{\ell})_{t\geq N}) = \sum_{t=N}^{\infty} \beta^t u(C_t^{\ell}) \tag{35}$$

subject to the updated time N lifetime budget constraint

$$\sum_{t=N}^{\infty} q_{t,N} \left(C_t^{\ell} + r_t K_t^{\ell} + v_t X_t^{\ell} - (1 - D_t^{\ell}) F_t^{\ell}(K_t^{\ell}, X_t^{\ell}) \right) \leq W_N^{\ell} \tag{36}$$

where $q_{t,N} := q_t/q_N = \prod_{n=N+1}^{t} r_n^{-1}$ and W_N^{ℓ} determined by (34) or, equivalently, by (33) setting $t = N - 1$. The time-N re-optimization problem of player $\ell \in \mathbb{L}$ then reads:

$$\max_{(C_t^{\ell}, X_t^{\ell}, K_t^{\ell})_{t\geq N}} \left\{ U_N((C_t^{\ell})_{t\geq N}) | C_t^{\ell}, K_t^{\ell}, X_t^{\ell} \text{ for all } \geq 0, t \geq N, (36) \text{ holds} \right\}. \tag{37}$$

The following lemma describes the properties of solutions to (37). The proof parallels the one of Lemma 2 and is therefore omitted.

Lemma 3

If the sequence $(C_t^{\ell}, X_t^{\ell}, K_t^{\ell})_{t\geq N}$ satisfies (12a), (19), and (29) with D_t^{ℓ} determined by (25) for all $t \geq N$ as well as (36) with equality, then it is a solution to problem (37).

In a similar vein, suppose the market maker $\ell = 0$ is permitted to update his strategy as well by choosing values $(r_t, v_t)_{t\geq N}$ subject to the Hotelling rule (15). Clearly, this choice is equivalent to choosing values v_N and r_N and $q_{t,N}$ for all $t > N$. Suppose these choices are made to maximize the values $(v_N - c_x) \sum_{\ell\in\mathbb{L}} \left(R_N^{\ell} - \sum_{t=N}^{\infty} X_t^{\ell} \right)$, $r_N \left(\sum_{\ell\in\mathbb{L}} K_N^{\ell} - \overline{K}_N^s \right)$, and $q_{t,N} \sum_{\ell\in\mathbb{L}} \left(C_t^{\ell} + r_t K_t^{\ell} + v_t X_t^{\ell} - (1 - D_t^{\ell}) F_t^{\ell}(K_t^{\ell}, X_t^{\ell}) \right)$ for all $t \geq N$, respectively. Here, $\overline{K}_N^s := \sum_{\ell\in\mathbb{L}} \left((1 - D_{N-1}^{\ell}) F_{N-1}^{\ell}(K_{N-1}^{\ell}, X_{N-1}^{\ell}) - C_{N-1}^{\ell} - c_x X_{N-1}^{\ell} \right)$ is world capital supply in period N. Combining this behavior with that of players $\ell \in \mathbb{L}$, we can define a time N Nash equilibrium as a list of continuation strategies $(C_t^{\ell}, X_t^{\ell}, K_t^{\ell})_{t\geq N}$ for each $\ell \in \mathbb{L}$ and $(r_t, v_t)_{t\geq N}$ for $\ell = 0$ which solve the updated decision problems at time N given previous decisions. Now suppose until time $N - 1$, all players followed the strategies prescribed by the non-cooperative equilibrium. Formally, each player $\ell \in \mathbb{L}$ chose $(K_t^{\ell,\text{nc}}, X_t^{\ell,\text{nc}}, C_t^{\ell,\text{nc}})_{0\leq t<N}$ as their strategy and the market maker $\ell = 0$ chose $(r_t^{\text{nc}}, v_t^{\text{nc}})_{0\leq t<N}$. Then, the continuation strategies $(K_t^{\ell,\text{nc}}, X_t^{\ell,\text{nc}}, C_t^{\ell,\text{nc}})_{t\geq N}$ for each player $\ell \in \mathbb{L}$ are clearly feasible, provided the market maker chooses $(r_t^{\text{nc}}, v_t^{\text{nc}})_{t\geq N}$. Moreover, these strategies satisfy the optimality conditions from Lemma 3, implying that each player will find it advantageous to stick to these strategies, provided everybody else does. This establishes the following main result of this section.

Proposition 1
Let $N \geq 0$ be arbitrary and suppose all players $\ell \in \mathbb{L}_0$ followed the strategies of the non-cooperative equilibrium up to time $t = N - 1$. Then, the continuation strategies $(K_t^{\ell,\mathrm{nc}}, X_t^{\ell,\mathrm{nc}}, C_t^{\ell,\mathrm{nc}})_{t \geq N}$ for each player $\ell \in \mathbb{L}$ and $(r_t^{\mathrm{nc}}, v_t^{\mathrm{nc}})_{t \geq N}$ for player $\ell = 0$ constitute a non-cooperative (Nash) equilibrium at stage $t = N$.

We remark that a stronger version of this result would be a Markov-perfect equilibrium where all strategies are generated by time-invariant decision rules defined on a suitable state space. This concept is widely used in game-theoretic studies of the climate problem, see, e.g., (Harstad 2016). In our setting, however, with time-varying production functions (1) due to population growth, technological progress, etc., it is in general not possible to obtain such a time-invariant structure. Therefore, the subgame-perfectness demonstrated in this section appears to be the strongest result possible in our setting.

4.4 Regionally optimal climate policy

Implementing the non-cooperative solution
Returning to the decentralized equilibrium discussed in Section 3, we seek to implement the non-cooperative solution as a decentralized equilibrium. This requires determining the climate tax policy $(\tau_t^\ell)_{t \geq 0}$ for each region ℓ such that the equilibrium allocation and price system satisfy $A^* = A^{\mathrm{nc}}$ and $P^* = P^{\mathrm{nc}}$. Formally, this can be accomplished by ensuring the optimality conditions determining the non-cooperative solution coincide with the equilibrium conditions from Definition 1.

Regionally optimal climate tax
Along the non-cooperative equilibrium, the regional lifetime budget constraint (27) holds with equality. Using the definition of profits in (11) and the form of transfers (9), we obtain

$$(1 - D_t^\ell) F_t^\ell (K_t^\ell, X_t^\ell) - r_t K_t^\ell - v_t X_t^\ell = \Pi_t^\ell + T_t^\ell. \tag{38}$$

Using this in (27) we obtain the consumer's lifetime budget constraint (18) with equality as required. Furthermore, invoking Lemma 2, we see that the Euler equations (19) are automatically satisfied and so is the optimality condition (12a) with respect to capital input. The price sequence P^{nc} satisfies the Hotelling rule (15) by assumption. Finally, comparing the optimality conditions (12b) and (29) wit respect to fossil fuel inputs, we can see that implementing the regionally optimal solution to (28) requires choosing the regional climate tax according to the rule

$$\tau_t^\ell = \sum_{n=0}^{\infty} \beta^n \left(C_{t+n}^\ell / C_t^\ell \right)^{-\sigma} \delta_n \gamma^\ell Y_{t+n}^\ell. \tag{39}$$

Hence, if taxes are determined by the tax rule (39), all conditions determining the decentralized and non-cooperative solution coincide. We state the previous result formally in the following theorem.

Theorem 1
Suppose each region $\ell \in \mathbb{L}$ chooses a climate tax policy $(\tau_t^\ell)_{t\geq 0}$ according to the rule (39) and consumers in region ℓ receive transfers determined by (9). Then, the induced equilibrium allocation and prices satisfy $A^* = A^{\rm nc}$ and $P^* = P^{\rm nc}$.

By Lemma 1, equilibrium consumption satisfies $C_t^\ell = \mu^\ell \overline{C}_t$ permitting to replace the discount factor in (39) by aggregate consumption. This yields (39) in equivalent form

$$\tau_t^\ell = \sum_{n=0}^{\infty} \beta^n \left(\overline{C}_{t+n}/\overline{C}_t\right)^{-\sigma} \delta_n \gamma^\ell Y_{t+n}^\ell. \tag{40}$$

Finally, along a balanced growth path where regional output and consumption grow at constant and identical rate g, (39) reduces to

$$\tau_t^\ell = \gamma^\ell Y_t^\ell \sum_{n=0}^{\infty} \left(\beta(1+g)^{1-\sigma}\right)^n \delta_n. \tag{41}$$

The formula (41) is an extension of the main result in Golosov et al. (2014) who derive the globally optimal climate tax in closed form and show that it can be expressed as a constant fraction of output determined by climate parameters and the discount factor.

5 Equilibrium under Full Cooperation

5.1 The efficient solution

The global planning problem
This section compares the previously derived non-cooperative solution to the globally efficient solution that obtains under full cooperation between regions. Intuitively, the non-cooperative solution is what the world *is doing* while the cooperative solution is what it *should be doing* to combat climate change.
Hillebrand & Hillebrand (2019) show that the efficient solution maximizes utility of a world representative consumer on the set of feasible allocations determined by technological constraints and the climate system. Since the present framework differs from the one used in Hillebrand & Hillebrand (2019), we briefly restate this problem here.
Consider a global planner choosing an aggregate allocation $\overline{A} = ((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}}, \overline{C}_t)_{t\geq 0}$ of production factors for each region along with aggregate global consumption. The term 'aggregate' signified by the bar notation reflects the fact that the allocation $\overline{A}$ specifies only aggregate consumption, but no its distribution across regions. Note however that this aggregation only refers to consumption while production inputs are specified separately for all regions.
The planner takes into account the impact of emissions on climate damage via (23), the global resource constraint (21), and the feasibility constraint

$$\overline{C}_t \leq \sum_{\ell\in\mathbb{L}} \left((1-D_t^\ell)F_t^\ell(K_t^\ell, X_t^\ell) - c_x X_t^\ell - K_{t+1}^\ell\right) \quad \text{for all} \quad t = 0,1,2,\ldots \tag{42}$$

Furthermore, the initial capital allocation $(K_0^\ell)_{\ell\in\mathbb{L}}$ is subject to the initial condition

$$\sum_{\ell\in\mathbb{L}} K_0^\ell \le \overline{K}_0^s := \sum_{\ell\in\mathbb{L}} K_0^{\ell,s}. \tag{43}$$

Based on these constraints, the global planner maximizes utility of a fictitious world representative consumer who consumes aggregate consumption $\overline{C}_t$ in each period t. Formally, the global planning problem can be stated as:

$$\max_{((K_t^\ell,X_t^\ell)_{\ell\in\mathbb{L}},\overline{C}_t)_{t\ge0}} \left\{ U((\overline{C}_t)_{t\ge0}) \middle| (21),(23),(42),(43)\text{ hold, } K_t^\ell, X_t^\ell, \overline{C}_t \ge 0 \text{ for all } t\ge0, \ell\in\mathbb{L} \right\}. \tag{44}$$

We denote the solution to (44) by $\overline{A}^{\text{eff}} = ((K_t^{\ell,\text{eff}}, X_t^{\ell,\text{eff}})_{\ell\in\mathbb{L}}, \overline{C}_t^{\text{eff}})_{t\ge0}$ and refer to it as the *efficient aggregate allocation*. The following lemma characterizes the efficient solution formally. The proof can be found in Section A.2 in Appendix A.

Lemma 4

If the aggregate allocation $\overline{A} = ((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}}, \overline{C}_t)_{t\ge0}$ satisfies the feasibility conditions (42) and (43) with equality for all $t = 0,1,2,\ldots$, the intratemporal efficiency conditions

$$(1-D_t^\ell)\partial_K F_t^\ell(K_t^\ell, X_t^\ell) = (1-D_t^k)\partial_K F_t^k(K_t^k, X_t^k) \tag{45a}$$

$$(1-D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) = (1-D_t^k)\partial_X F_t^k(K_t^k, X_t^k) \tag{45b}$$

for all $k,\ell\in\mathbb{L}$, the intertemporal efficiency conditions

$$\frac{\beta\overline{C}_{t+1}^{-\sigma}}{\overline{C}_t^{-\sigma}}(1-D_{t+1}^\ell)\partial_K F_{t+1}^\ell(K_{t+1}^\ell, X_{t+1}^\ell) = 1 \tag{45c}$$

$$\frac{\beta\overline{C}_{t+1}^{-\sigma}}{\overline{C}_t^{-\sigma}}\left((1-D_{t+1}^\ell)\partial_X F_{t+1}^\ell(K_{t+1}^\ell, X_{t+1}^\ell) - c_x - \hat{\tau}_{t+1}\right) = (1-D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) - c_x - \hat{\tau}_t \tag{45d}$$

for all $\ell\in\mathbb{L}$ with damages determined by (23) and

$$\hat{\tau}_t := \sum_{n=0}^{\infty} \beta^n \left(\overline{C}_{t+n}/\overline{C}_t\right)^{-\sigma} \delta_n \sum_{k\in\mathbb{L}} \gamma^k Y_{t+n}^k \quad \textit{for all } t = 01,2,\ldots \tag{46}$$

as well as the transversality condition $\lim_{T\to\infty}\beta^T\overline{C}_T^{-\sigma}\sum_{\ell\in\mathbb{L}}K_{T+1}^\ell = 0$ and the resource constraint (21), then it is a solution to (44).

Intuitively, the efficient solution equates marginal products across all regions in each period and also intertemporally. Equation (45a) implicitly determines a (shadow) capital return $\hat{r}_t$ which ensures an efficient world capital allocation. This return can be used in (45c) ensuring intertemporally efficient consumption and capital accumulation. Furthermore, condition (45b) allocates fossil fuel usage efficiently to equate marginal products across world regions. This property can be used to define a shadow resource price $\hat{v}_t := (1-D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) - \hat{\tau}_t$ corresponding to the marginal product of the resource in production net of the total discounted damage captured by $\hat{\tau}_t$. With this definition, equation (45d) corresponds to the Hotelling rule (15) ensuring efficient extraction of the resource over time. In particular, if the resource constraint (21) is non-binding at the optimal solution, then $\hat{v}_t = c_x$, as shown in the appendix.

5.2 Implementing the efficient solution

We now explore conditions under which the decentralized allocation A^* is efficient. Formally, the equilibrium allocation is efficient if and only if for all $t = 0,1,2,\dots$

$$K_t^{\ell*} = K_t^{\ell,\text{eff}} \quad \text{and} \quad X_t^{\ell*} = X_t^{\ell,\text{eff}} \quad \text{for all } \ell \in \mathbb{L} \quad \text{and} \quad \sum_{\ell\in\mathbb{L}} C_t^* = \overline{C}_t^{\text{eff}}. \tag{47}$$

Climate policies tax policies under which (47) holds will be termed efficient and denoted by $(\tau_t^{\text{eff}})_{t\geq 0}$. Note that such a policy will necessarily be uniform across all regions. We now show that the efficient climate tax policy takes the following explicit form

$$\tau_t^{\text{eff}} := \sum_{n=0}^{\infty} \beta^n \left(\overline{C}_{t+n}^{\text{eff}} / \overline{C}_t^{\text{eff}}\right)^{-\sigma} \delta_n \sum_{k\in\mathbb{L}} \gamma^k Y_{t+n}^{k,\text{eff}} \quad \text{for all } t = 0,1,2,\dots. \tag{48}$$

A similar result is derived in Hillebrand & Hillebrand (2019) in a slightly different setup. To prove this claim, we must show that the equations from Lemma 4 generating the efficient allocation are satisfied at equilibrium if taxes are chosen based on (48). This result holds independently of how transfers $(T_t^\ell)_{t\geq 0}$ are determined.
The final sector's optimality condition (12a) implies equalization of capital returns (45a). Further, under uniform taxation, optimality condition (12b) implies that marginal products of fossil fuels equalize as required by (45b). Further, the Euler equation (19) can be aggregated and combining it with (12a) to replace the return on capital by its marginal product yields the aggregate Euler equation (45c).[6] Further, solving (12b) for v_t and substituting the result into the Hotelling rule (15) using the form of taxes (48) and replacing again the capital return r_{t+1} by its marginal product based on (12a) gives (45d). At equilibrium, the capital market clearing condition (20) for $t = 0$, the world resource constraint (21), and the commodity market equilibrium condition (22) hold directly. Finally, using the market clearing condition (20) of the capital market individual transversality conditions can be aggregated based on (45c) to obtain the aggregate version in Lemma 4.[7]
We state the previous result formally in the following theorem.

Theorem 2
Suppose all regions choose uniform climate taxes given by (48). Then, the equilibrium allocation $A^* = ((K_t^{\ell*}, X_t^{\ell*}, C_t^{\ell*})_{\ell\in\mathbb{L}})_{t\geq 0}$ is efficient, i.e., (47) holds for all $t = 0,1,2,\dots$

[6] We remark that this aggregation property of the Euler equation is key for the separability result from Hillebrand & Hillebrand (2019) on which the present section is based. It requires the form of preferences (7) and equalization of capital returns via an international capital market.

[7] For this, one can use the (arbitrary but given) sequence of transfers $(T_t^\ell)_{t\geq 0}$ and assume an arbitrary extraction sequence $(X_t^{\ell,s})_{t\geq 0}$ consistent with (3) and satisfying $\sum_{\ell\in\mathbb{L}} X_t^{\ell,s} = \sum_{\ell\in\mathbb{L}} X_t^\ell$ for all t. Then, one can use (17) to recursively obtain the sequence $(K_{t+1}^{\ell,s})$ satisfying the capital market clearing condition in (20) and the transversality condition $\lim_{t\to\infty} K_{t+1}^{\ell,s} \beta^t u'(C_t^{\ell,*})/u'(C_0^{\ell,*}) = \lim_{t\to\infty} K_{t+1}^{\ell,s} \beta^t u'(\overline{C}_t^*)/u'(\overline{C}_0^*) = 0$. Aggregating this condition over all regions using the capital market clearing condition in (20) then gives the result.

The distribution $(\mu^\ell)_{\ell\in\mathbb{L}}$ of aggregate consumption $(\overline{C}^{\text{eff}}_t)_{t\geq 0}$ depends on whether there are transfers between regions. If there is no re-distribution, transfers are determined by (9) and consumption shares follow from equation (24). In a cooperative setup, however, transfers can be used to incentivize regions to implement the globally rather than the regionally optimal tax. This issue will be studied next.

5.3 Incentive-compatible transfer policies

Welfare gain under cooperation
Denote the consumption distribution along the non-cooperative equilibrium by $(\mu^{\ell,\text{nc}})_{\ell\in\mathbb{L}}$ and define aggregate consumption as $\overline{C}^{\text{nc}}_t := \sum_{\ell\in\mathbb{L}} C^{\ell,\text{nc}}_t$. The aggregate allocation $\overline{A}^{\text{nc}} = ((K^{\ell,\text{nc}}_t, X^{\ell,\text{nc}}_t)_{\ell\in\mathbb{L}}, \overline{C}^{\text{nc}}_t)_{t\geq 0}$ associated with the non-cooperative equilibrium satisfies all the constraints in (44). Therefore, the implied total welfare is less than at the efficient equilibrium, i.e.,[8]

$$\overline{U}^{\text{nc}} := U((\overline{C}^{\text{nc}}_t)_{t\geq 0}) < U((\overline{C}^{\text{eff}}_t)_{t\geq 0}) =: \overline{U}^{\text{eff}}. \tag{49}$$

We now explore whether this efficiency gain can be distributed such that each regionis better-off nder efficient taxation relative to the non-cooperative equilibrium.

Pareto-improving consumption shares
To induce a Pareto-improvement under efficient taxation relative to the non-cooperative equilibrium, each region $\ell \in \mathbb{L}$ must receive a consumption share μ^ℓ such that

$$U((\mu^\ell C^{\text{eff}}_t)_{t\geq 0}) \geq U((\mu^{\ell,\text{nc}} C^{\text{nc}}_t)_{t\geq 0}). \tag{50}$$

Exploiting the form of utility (7), we can determine explicit lower bounds μ^ℓ_{crit} for this share such that (50) holds whenever $\mu^\ell \geq \mu^\ell_{\text{crit}}$. These shares are characterized in the following result. The proof can be found in Section A.3 in Appendix A.

Lemma 5
Under efficient taxation of all regions, suppose each region $\ell \in \mathbb{L}$ receives a consumption share

$$\mu^\ell \geq \mu^\ell_{\text{crit}} := \begin{cases} \mu^{\ell,\text{nc}} \left(\overline{U}^{\text{nc}}/\overline{U}^{\text{eff}}\right)^{\frac{1}{1-\sigma}} & \text{if} \quad \sigma > 0, \sigma \neq 1 \\ \mu^{\ell,\text{nc}} \cdot e^{-(1-\beta)(\overline{U}^{\text{eff}} - \overline{U}^{\text{nc}})} & \text{if} \quad \sigma = 1. \end{cases} \tag{51}$$

Then, the induced disaggregated efficient allocation $A = ((K^{\ell,\text{eff}}_t, X^{\ell,\text{eff}}_t, \mu^\ell \overline{C}^{\text{eff}}_t))_{t\geq 0}$ Pareto improves the non-cooperative allocation $A^{\text{nc}} = ((K^{\ell,\text{nc}}_t, X^{\ell,\text{nc}}_t, C^{\ell,\text{nc}}_t))_{t\geq 0}$.

[8] The strict inequality can be inferred from Theorems 1 and 2 by observing that taxes along the non-cooperative equilibrium are determined by (39) rather than by (48). Therefore, $\bar{A}^{\text{nc}}$ cannot be efficient.

Noting that $\mu^{\ell}_{\text{crit}} < \mu^{\ell,\text{nc}}$ for all $\ell \in \mathbb{L}$, condition (51) holds in particular if $\mu^{\ell} = \mu^{\ell,\text{nc}}$, i.e., each region attains the same consumption share as in the non-cooperative equilibrium.

Pareto-improving transfer shares

We now explore how transfers between regions can be designed such that (51) holds. For this purpose, we assume that regions pool their tax revenue and agree on a transfer policy $(\theta^{\ell})_{\ell\in\mathbb{L}}$ determining transfer payments as in (10). Let $\overline{A}^{\text{eff}}$ be the efficient allocation and define the induced equilibrium prices $P^{\text{eff}} = (r_t^{\text{eff}}, v_t^{\text{eff}})_{t\geq 0}$ by setting

$$r_t^{\text{eff}} := (1-D_t^{\ell,\text{eff}})\partial_K F_t^{\ell}(K_t^{\ell,\text{eff}}, X_t^{\ell,\text{eff}}) \quad \text{and} \quad v_t^{\text{eff}} := (1-D_t^{\ell,\text{eff}})\partial_X F_t^{\ell}(K_t^{\ell,\text{eff}}, X_t^{\ell,\text{eff}}) \tag{52}$$

for $t = 0,1,2,\ldots$ with climate damages $D_t^{\ell,\text{eff}}$ determined by (23). Note that the quantities in (52) are well-defined., i.e., independent of the regional index ℓ due to (45a) and (45b) and satisfy the Hotelling rule (15) due to (45d). Denote by $(q_t^{\text{eff}})_{t\geq 0}$ the induced Arrow-Debreu prices defined by (13) and let $\overline{T}_t^{\text{eff}} := \sum_{\ell\in\mathbb{L}} \tau_t^{\text{eff}} X_t^{\ell,\text{eff}}$ denote total tax revenue and $\Pi_t^{\ell,\text{eff}} := Y_t^{\ell,\text{eff}} - r_t^{\text{eff}} K_t^{\ell,\text{eff}} - v_t^{\text{eff}} X_t^{\ell,\text{eff}}$ final profits in region ℓ at time t. As before, write the induced lifetime profit incomes as $\Pi^{\ell,\text{eff}} := \sum_{t=0}^{\infty} q_t^{\text{eff}} \Pi_t^{\ell,\text{eff}}$ and total tax revenue $\overline{T}^{\text{eff}} := \sum_{t=0}^{\infty} q_t^{\text{eff}} \overline{T}_t^{\text{eff}}$. Then, we can use the result from Lemma 1 and condition (24) to establish the following one-to-one relation between consumption shares $(\mu^{\ell})_{\ell\in\mathbb{L}}$ and transfer shares $(\theta^{\ell})_{\ell\in\mathbb{L}}$ along the efficient equilibrium:

$$\mu^{\ell} = \frac{\Pi^{\ell,\text{eff}} + r_0^{\text{eff}} K_0^{\ell,s} + (v_0^{\text{eff}} - c_x) R_0^{\ell} + \theta^{\ell} \overline{T}^{\text{eff}}}{\overline{\Pi}^{\text{eff}} + r_0^{\text{eff}} \overline{K}_0^{s} + (v_0^{\text{eff}} - c_x) \overline{R}_0 + \overline{T}^{\text{eff}}}. \tag{53}$$

Solving (53) for θ^{ℓ} and combining it with the result from Lemma 5 gives the following complete characterization of incentive-compatible transfer policies inducing a Pareto-improvement over the non-cooperative solution.

Theorem 3

Define $(\mu^{\ell,\text{crit}})_{\ell\in\mathbb{L}}$ as in Lemma 5. Suppose each region $\ell \in \mathbb{L}$ chooses efficient taxes determined by (61) and receives a constant share θ^{ℓ} of global tax revenue satisfying

$$\theta^{\ell} \geq \theta^{\ell}_{\text{crit}} := \mu^{\ell,\text{crit}} + \frac{\mu^{\ell,\text{crit}}\left(\overline{\Pi}^{\text{eff}} + r_0^{\text{eff}} \overline{K}_0^{s} + (v_0^{\text{eff}} - c_x)\overline{R}_0\right) - \left(\Pi^{\ell,\text{eff}} + r_0^{\text{eff}} K_0^{\ell,s} + (v_0^{\text{eff}} - c_x) R_0^{\ell}\right)}{\overline{T}^{\text{eff}}}.$$

Then, all regions are better-off relative to the non-cooperative equilibrium.

Observe that the minimal transfer shares defined in Theorem 3 satisfy

$$\sum_{\ell\in\mathbb{L}} \theta^{\ell}_{\text{crit}} = \sum_{\ell\in\mathbb{L}} \mu^{\ell,\text{crit}} - \frac{\left(\sum_{\ell\in\mathbb{L}} \mu^{\ell,\text{crit}} - 1\right)\left(\overline{\Pi}^{\text{eff}} + r_0^{\text{eff}} \overline{K}_0^{s} + (v_0^{\text{eff}} - c_x)\overline{R}_0\right)}{\overline{T}^{\text{eff}}} < 1. \tag{54}$$

Thus, the set of transfer policies $(\theta^{\ell})_{\ell\in\mathbb{L}}$ satisfying the conditions of Theorem 3 is non-empty.

6 Partial Cooperation and Coalitions

6.1 A cooperative setup

The previous sections compared the allocation under full cooperation by all members of $\mathbb{L}$ to the case where there is no cooperation at all and each region solves its own planning problem (28). Clearly, there are many intermediate scenarios where some regions join forces and coordinate their climate policies by forming coalitions.

Coalitions and aggregation

To describe a scenario with coalitions formally, let $\mathscr{P}(\mathbb{L})$ be the power set of $\mathbb{L}$ consisting of all subsets of $\mathbb{L}$. Any non-empty subset $\mathbb{L}' \in \mathscr{P}(\mathbb{L})$ of $\mathbb{L}$ will be referred to as a coalition. The case $\mathbb{L}' = \mathbb{L}$ is called the grand coalition consisting of all regions. Let $\mathbb{L}' \in \mathscr{P}(\mathbb{L})$ be an arbitrary coalition. In what follows, we index variables referring to the coalition as a whole by a superscript $\mathbb{L}'$. For variables referring to individual members, we continue to use the regional index $\ell \in \mathbb{L}'$. For example, $R_0^{\mathbb{L}'} := \sum_{\ell \in \mathbb{L}'} R_0^\ell$ is the initial stock of fossil fuels owned by the coalition $\mathbb{L}'$ while $(X_t^\ell)_{t \geq 0}$ denotes fossil fuel consumption in member region $\ell \in \mathbb{L}'$. Based on this convention, a coalitional variable $x^{\mathbb{L}'}$ obtains as the sum of the list of individual variables $(x^\ell)_{\ell \in \mathbb{L}'}$, i.e., $x^{\mathbb{L}'} = \sum_{\ell \in \mathbb{L}'} x^\ell$.

The coalitional planning problem

Suppose the members of coalition $\mathbb{L}'$ join forces by aggregating their resource stocks and capital endowments. Summing (27) over all regions in $\mathbb{L}'$ one obtains the constraint

$$\sum_{t=0}^{\infty} q_t \left(\sum_{\ell \in \mathbb{L}'} \left((1 - D_t^\ell) F_t^\ell (K_t^\ell, X_t^\ell) - r_t K_t^\ell - v_t X_t^\ell \right) - C_t^{\mathbb{L}'} \right) + r_0 K_0^{\mathbb{L}',s} + (v_0 - c_x) R_0^{\mathbb{L}'} \geq 0. \qquad (55)$$

Defining the complement $-\mathbb{L}' := \mathbb{L} \setminus \mathbb{L}'$, regional climate damages (23) can be written as

$$D_t^\ell = 1 - \exp\left(-\gamma^\ell \sum_{n=0}^{\infty} \delta_n \left(X_{t-n}^{\mathbb{L}'} + X_{t-n}^{-\mathbb{L}'} \right) \right) \quad \text{for } \ell \in \mathbb{L}' \qquad (56)$$

Assume that coalition $\mathbb{L}'$ maximizes the total utility attained by the aggregated consumption stream $(C_t^{\mathbb{L}'})_{\geq 0}$. This approach is based on the separation principle established in Hillebrand & Hillebrand (2019) permitting to first determine an efficient solution by maximizing aggregate utility and then distributing consumption among coalition members based on some suitable weighting scheme. The distribution of consumption among coalition members corresponds to an imputation in the cooperative game defined below.

Based on the previous idea, the coalitional planning problem can be stated as follows:

$$\max_{((K_t^\ell, X_t^\ell)_{\ell \in \mathbb{L}'}, C_t^{\mathbb{L}'})_{t \geq 0}} \left\{ U((C_t^{\mathbb{L}'})_{t \geq 0}) \middle| (55), (56) \text{ hold}, K_t^\ell, X_t^\ell, C_t^{\mathbb{L}'} \geq 0 \text{ for all } t = 0, 1, 2, \ldots \right\}. \qquad (57)$$

As in the regional problem (28), emission $(X_t^{-\mathbb{L}'})_{t \geq 0}$ of all other regions and global prices $P = (r_t, v_t)_{t \geq 0}$ are taken as given in the decision (57). Further, the coalitional planer takes into account the impact of emission of coalition members on damages (56) in each member region $\ell \in \mathbb{L}'$.

Solution to the coalitional planning problem

It is again straightforward to solve the coalitional problem (57) using Lagrangean methods. The proof of the following result parallels the one of Lemma 2 and is omitted.

Lemma 6

Let prices $P = (r_t, v_t)_{t\geq 0}$ and emissions $(X_t^{-\mathbb{L}'})_{t\geq 0}$ of non-coalition members be given. If the allocation $((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}'}, C_t^{\mathbb{L}'})_{t\geq 0}$ solves conditions (12a) and

$$(1-D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) = v_t + \sum_{n=0}^{\infty} \beta^n \left(C_{t+n}^{\mathbb{L}}/C_t^{\mathbb{L}}\right)^{-\sigma} \delta_n \sum_{k\in\mathbb{L}'} \gamma^k Y_{t+n}^k \tag{58}$$

for all $\ell \in \mathbb{L}'$ with D_t^ℓ determined by (56) and the Euler equation

$$C_{t+1}^{\mathbb{L}'} = C_t^{\mathbb{L}'}(\beta r_{t+1})^{\frac{1}{\sigma}} \tag{59}$$

for all $t = 0,1,2,\ldots$ and satisfies (55) with equality, then it is a solution to problem (57).

Note that (58) corrects the marginal product of fossil fuel in production by the discounted future climate damages of all coalition members.

6.2 Coalitional equilibrium

Classes of coalitions

Now let $\mathscr{L} = \{\mathbb{L}_1, \ldots, \mathbb{L}_N\} \subset \mathscr{P}(\mathbb{L})$ be a partition of the set of all regions $\mathbb{L}$ into $1 \leq N \leq L$ coalitions, i.e., $\cup_{n=1}^N \mathbb{L}_n = \mathbb{L}$ and $\mathbb{L}_n \cap \mathbb{L}_m = \emptyset$ for all $n \neq m$ where $n, m \in \{1, \ldots, N\}$. Following Yi (1997), we call $\mathscr{L}$ a *coalition structure*. Suppose the members of each coalition $\mathbb{L}' \in \mathscr{L}$ solve a coalitional planning problem of the form (57). The market maker $\ell = 0$ acts as in the previous section to enforce market clearing. Consistency of coalitional decisions leads to the following definition of a coalitional equilibrium.

Definition 3

Let $\mathscr{L} \subset \mathscr{P}(\mathbb{L})$ be a coalition structure. A coalitional equilibrium is an allocation

$$\bar{A} = \left(\left((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}'}, C_t^{\mathbb{L}'}\right)_{t\geq 0}\right)_{\mathbb{L}'\in\mathscr{L}} \tag{60}$$

and a price system $P = (r_t, v_t)_{t\geq 0}$ such that the following holds:

(i) For each $\mathbb{L}' \in \mathscr{L}$, the allocation $((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}'}, C_t^{\mathbb{L}'})_{t\geq 0}$ solves the coalitional planning problem (57) given prices P and emissions $(X_t^{-\mathbb{L}'})_{t\geq 0}$ of all other regions..

(ii) Prices P chosen by player $\ell = 0$ satisfy the Hotelling rule (15) and (v_0, r_0) and q_t defined by (13) are solutions to the decision problems (30a)-(30c) for all $t \geq 0$.

When we want to emphasize the dependence of coalitional equilibrium on the coalition configuration, we write $\bar{A}^{\mathscr{L}}$ and $P^{\mathscr{L}}$. The bar-notation is used to emphasize that the distribution of consumption within a coalition is not determined and subject to an imputation to be bargained over by coalition members.

6.3 Decentralization of coalitional equilibrium

Based on the optimality condition (58), it is now straightforward to extend the results from Theorem 1 to the following theorem describing the climate policy under which the coalitional allocation is decentralized.

Theorem 4
Let $\mathscr{L} \subset \mathscr{P}(\mathbb{L})$ be a coalition structure. For each coalition $\mathbb{L}' \in \mathscr{L}$, suppose all its members $\ell \in \mathbb{L}'$ set taxes according to the rule

$$\tau_t^\ell = \tau_t^{\mathbb{L}'} := \sum_{n=0}^{\infty} \beta^n \left(\frac{C_{t+n}^{\mathbb{L}'}}{C_t^{\mathbb{L}'}} \right)^{-\sigma} \delta_n \sum_{k \in \mathbb{L}'} \gamma^k Y_{t+n}^k \quad \text{for all } \ell \in \mathbb{L}'. \tag{61}$$

Then, the coalitional equilibrium is decentralized.

On a balanced growth path where aggregate consumption and output in region of the coalition grow at constant rate g, the optimal tax formula (61) reduces again to

$$\tau_t^\ell = \tau_t^{\mathbb{L}'} = \sum_{k \in \mathbb{L}'} \gamma^k Y_t^k \sum_{n=0}^{\infty} \left(\beta(1+g)^{1-\sigma} \right)^n \delta_n \quad \text{for all } \ell \in \mathbb{L}'. \tag{62}$$

Two special coalition structure are of particular interest. The first one is the coalition $\mathscr{L} = \{\{1\}, \ldots, \{L\}\}$ where each region acts individually. In this case, Definition 3 coincides with that of a non-cooperative equilibrium from Definition 2 and so $\bar{A}^{\mathscr{L}} \cong A^{\mathrm{nc}}$.[9]
A second scenario is where $\mathscr{L} = \{\mathbb{L}\}$ corresponding to the grand coalition consisting of all regions. For this case, the tax formula (61) coincides with the efficient solution (48). Therefore, by virtue of Theorem 2 $\bar{A}^{\mathscr{L}} \cong \bar{A}^{\mathrm{eff}}$, i.e., the induced allocation is efficient.
We state these two insights as the following and final main result.

Proposition 2

(i) If $\mathscr{L} = \{\{1\}, \ldots, \{L\}\}$, then $\bar{A}^{\mathscr{L}} \cong A^{\mathrm{nc}}$.

(ii) If $\mathscr{L} = \{\mathbb{L}\}$, then $\bar{A}^{\mathscr{L}} \cong \bar{A}^{\mathrm{eff}}$.

We remark that all results from Section 5.3 on optimal transfer policies designed to induce cooperation among regions carry over to the case with partial cooperation. In fact, such polcies can also be derived in closed form based on the same principles as in the non-cooperative case.

[9]Due to the slightly different arrangement of terms in $\bar{A}^{\mathscr{L}}$ and A^{nc}, we write $\bar{A}^{\mathscr{L}} \cong A^{\mathrm{nc}}$ and $\bar{A}^{\mathscr{L}} \cong \bar{A}^{\mathrm{eff}}$ below to mean that the two allocations are isomorphic rather than directly identical.

7 Quantitative results

7.1 Regional structure and coalition scenarios

This section illustrates and quantifies our theoretical results based on calibrated parameter values chosen to match selected empirical targets. Specifically, we distinguish the non-cooperative and fully cooperative solution as well as two additional scenarios with some key regions not joining the global climate agreement.

Regions
Setting $L = 10$ we distinguish the following region displayed in Table 1. The number of regions is small enough to allow for a compact presentation of the results and large enough to distinguish the main players in past climate agreements.

Table 1: Regions in the simulation model

Region	Label	Index ℓ	Region	Label	Index ℓ
United States	USA	1	India	IND	6
OECD Europe	OEU	2	Russia	RUS	7
Australia & New Zealand	ANZ	3	Brazil	BRA	8
Other High Income	OHI	4	Developing countries	DEV	9
China	CHN	5	Low Income Countries	LIC	10

Time structure
One time period t in our model represents ten years which is a standard choice in the literature. The initial model period $t = 0$ represents the years 2010-2019 and is referred to as the baseline period 2015. Subsequent periods representing years $2020-2029$, $2030-2039$, etc. are referred to by their midpoints $2025, 2035$, etc. Flow variables such as production output or emissions are generally aggregated over the entire period while stocks such as capital or atmospheric carbon usually refer to the beginning of the period.

Coalition scenarios
As for the structure of coalitions, we distinguish four scenarios detailed in the following table. Scenario 1 is the efficient benchmark which is opposed to the completed absence

Table 2: Coalition structures in our simulation

Scenario	Description	Coalition structure
1	Full non-cooperation	$\mathscr{L} = \{\{\ell\} \mid \ell \in \mathbb{L}\} = \{\{1\}, \{2\}, \ldots, \{10\}\}$
2	Full cooperation (grand coalition)	$\mathscr{L} = \{\mathbb{L}\} = \{\{1, \ldots, 10\}\}$
3	Grand coalition except US	$\mathscr{L} = \{\mathbb{L} \setminus \{1\}, \{1\}\}$
4	Grand coalition except China	$\mathscr{L} = \{\mathbb{L} \setminus \{5\}, \{5\}\}$

of cooperation in Scenario 2. Scenarios 3 and 4 correspond to the cases where either the US or China decide not to cooperate. This allows us to evaluate and quantify the relative importance of these players for a successful global climate agreement.

7.2 Functional forms and parameters

Production sector

Similar to Hassler, Krusell & Olovsson (2021) we assume that the production function $F_t^\ell : \mathbb{R}_+^2 \longrightarrow \mathbb{R}_+$ in (1) is of the form

$$F_t^\ell(K,X) = \left[\kappa \left(K^\alpha (Q_{N,t}^\ell N_t^\ell)^{1-\alpha}\right)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\kappa)(Q_{X,t}^\ell X)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \quad 0<\kappa<1, 0<\alpha<1, \varepsilon>0. \tag{63}$$

Here, N_t^ℓ and $Q_{N,t}^\ell$ denote population size and labor efficiency in region ℓ at time t while $Q_{X,t}^\ell$ represents energy efficiency. Parameter ε controls the elasticity of substitution between fossil energy and the capital-labor aggregate.

Variables $Q_{N,t}^\ell$ and $Q_{X,t}^\ell$ capture labor-augmenting and energy-augmenting technical change, respectively. They both grow at constant exogenous rates $g_N \geq 0$ and $g_X \geq 0$ identical for all regions such that

$$Q_{N,t}^\ell = (1+g_N)^t Q_{N,0}^\ell \quad \text{and} \quad Q_{X,t}^\ell = (1+g_X)^t Q_{X,0}^\ell \quad \text{for } t \geq 0. \tag{64}$$

The population size $(N_t^\ell)_{t\geq 0}$ evolves exogenously and becomes constant for $t \geq 19$ corresponding to the year 2200. In our simulations we choose these values consistent with current population sizes and the UN population forecasts for $t = 2100$ and $t = 2200$. Details can be found below and in Section B.2 in Appendix B.

Resource sector

We abstract from the resource scarcity problem studied in Hassler, Krusell & Olovsson (2021) by formally setting $R_0 = \infty$. This seems justified because fossil energy in our model comprises all kinds of fossil fuels including coal which is known to be relatively abundant. This assumption implies a constant resource price $v_t = c_x$.

Climate model

We employ the climate model developed in GHKT which has the advantage of having a forward-recursive structure. The climate state S_t in period t consist of permanent ($S_{1,t}$) and non-permanent ($S_{2,t}$) carbon dioxide which evolve as

$$S_{1,t} = S_{1,t-1} + \phi_L \overline{X}_t \tag{65a}$$

$$S_{2,t} = (1-\phi)S_{2,t-1} + (1-\phi_L)\phi_0 \overline{X}_t. \tag{65b}$$

Setting $\delta_n = \phi_L + (1-\phi_L)\phi_0(1-\phi)^n$ our climate model (5) obtains as a reduced form of (65) with climate state $S_t := S_{1,t} + S_{2,t} - \bar{S}$ where the pre-industrial level is $\bar{S} = 581$ GtC.

7.3 Calibration

Calibration targets

We choose initial productivity parameters $Q^{\ell}_{N,0}$ and $Q^{\ell}_{X,0}$ in (64) to match regional output and emissions displayed in the following table in our baseline period $t = 0$. Formal details on this and how we constructed the data can be found in Sections B.2 and B.3.

Table 3: Calibration targets for baseline period 2010-2019

Variable	Units	USA	OEU	ANZ	OHI	CHN	IND	RUS	BRA	DEV	LIC
$Y_0^{\ell,\text{target}}$	Trn. U.S.$	156.7	188.5	10.6	134.2	134.6	57.1	31.3	28.4	45.8	147.8
$X_0^{\ell,\text{target}}$	Gt C	52.1	37.0	4.1	37.1	78.8	16.2	15.2	3.9	30.8	18.4

Production parameters

The literature contains a broad range of estimates for the elasticity of substitution ε. Papageorgiou et al. (2017) argue that this elasticity 'significantly exceeds unity' and estimate it to be about 2. Hassler et al. (2021) obtain a much smaller value of about 0.02 in their estimation for the U.S. economy. Our choice of $\varepsilon = 0.75$ is somewhat in the middle of these estimates. It implies that energy is a gross complement to other inputs. For $\varepsilon = 1$, $1-\kappa$ is the value of fossil energy relative to GDP. Empirical numbers put this share at about 5%, which is the value we and also Hassler et al. (2021) use, despite the fact that $\varepsilon < 1$ in our simulations. Alternative specifications as low as $\kappa = 0.8$ did not significantly affect our quantitative results reported below. We also follow Hassler et al. (2021) by setting $\alpha = 0.2632$.

Empirically, coal makes up by far the largest share of available stocks of fossil fuels. For this reason, we set extraction costs to 43 U.S. $ per physical ton of the resource which is the value of extraction costst per ton of coal used in GHKT. With a carbon content of 0.5441 tons of C per physical ton of coal, this implies extraction costs of about 79$ per ton of carbon corresponding to the parameter value $c_x = 0.000043/0.5441 = 0.000079$.

Consumer sector

Restricting consumer utility as in (7), we choose $\sigma = 1$ which gives a logarithmic utility function. The annual discount rate is 1.5% which implies a discount factor $\beta = 0.985^{10}$. These values are identical to the ones used by GHKT in their benchmark scenario. Our initial global capital stock is set to $\bar{K}_0 = 0.2$. This value avoids a transitory effect due to capital adjustments in the initial periods.

The exogenous population sequence $(N_t^{\ell})_{t\geq 0}$ in region ℓ is determined based on current population levels and UN Population forecasts for 2050, 2100, 2150, and 2200. Details are provided in Section B.3 in Appendix B.

The growth rates of labor efficiency g_N and energy efficiency g_X are chosen identical to imply an annual growth rate of 1% which is a conservative estimate. This implies $g_N = g_X = 0.105 = 10.5\%$ per decade.

Climate parameters
We follow GHKT by setting $\phi_L = 0.2$, $\phi_0 = 0.393$, and $\phi = 0.0228$ in (65). The initial climate state of permanent ($S_{1,-1}$) and non-permanent ($S_{2,-1}$) atmospheric CO_2 at the beginning of $t = 0$ is chosen consistent with the empirically observed carbon concentration of $S_{2009} = 827$ at the end of year 2009 and $S_{2019} = 878$ in 2019 and global emissions of about 100 Gt C in this decade. This approach is also used in Hillebrand & Hillebrand (2021) to which we refer for details. The implied initial values consistent with these observations and (65) are $S_{1,-1} = 807.7$ and $S_{2,-1} = 19.3$ GtC.
Damage parameters are notoriously hard to measure but tend to be much higher for poorer countries. Our choices in Table 7.3 are based on Hillebrand & Hillebrand (2021) adapted to the more disaggregated regional setting employed here.

Table 4: Climate damage parameters.

Region:	USA	OEU	ANZ	OHI	CHN	IND	RUS	BRA	DEV	LIC
$\gamma^\ell \cdot 10^5$	4.12	2.05	2.05	2.05	4.12	6.22	6.22	6.22	6.22	8.33

Computation
We use a version of the algorithm developed in Hillebrand & Hillebrand (2021). This approach identifies the forward-recursive structure of the model and determines initial consumption $\overline{C}_0$ based on the shooting-principle. The factor allocation in period t is determined as a fixed point of a mapping updating regional outputs and the cost shares of factors in final production. Details can be found in Section B.1 in Appendix B.

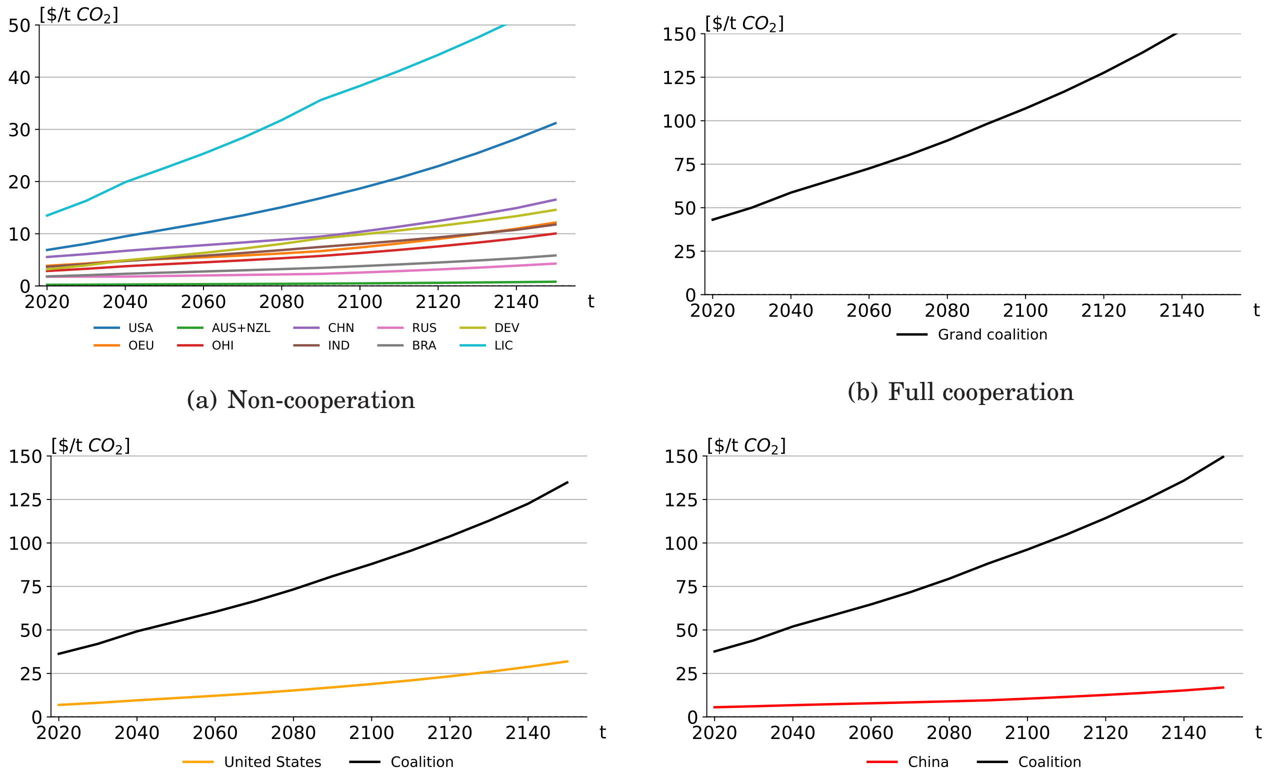
7.4 Results

The following figures depict the evolution of selected variables for the four scenarios described above.

Regional emission taxes
Figure 1 illustrates the evolution of regional emissions taxed for each scenario. The tax values for period $t = 1$ corresponding to the year 2020 are reported in Table 5 below.

Table 5: Climate taxes in 2020 under full, partial, and non-cooperation in \$/t CO_2.

Scenario	USA	OEU	AUS	OHI	CHN	IND	RUS	BRA	DEV	LIC
1	6.88	3.86	0.22	2.86	5.54	3.68	1.77	1.81	3.18	13.48
2	43.11	43.11	43.11	43.11	43.11	43.11	43.11	43.11	43.11	43.07
3	6.90	36.27	36.27	36.27	36.27	36.27	36.27	36.27	36.27	36.27
4	37.63	37.63	37.63	37.63	5.55	37.63	37.63	37.63	37.63	37.63



(a) Non-cooperation

(b) Full cooperation

(c) Grand coalition except US

(d) Grand coalition except China

Figure 1: Regional climate taxes.

In all cases, taxes increase over time due to GDP growth. As one would expect from our tax formulae (39), (48), and (61), non-cooperation leads to taxation much lower compared to all other scenarios. For this reason, the non-cooperative solution represents a lower bound of tax policies in any cooperative scenario. For instance, under non-cooperation, climate taxes in 2020 are less than 10 \$/t CO_2 for all regions except low income countries and range from 0.22 \$/t CO_2 in Australia/New Zealand to 6.89 \$/t CO_2 in the U.S. Low income countries set a slightly higher carbon tax equal to 13.48 \$/t CO_2 in 2020 reflecting higher climate damages in this region internalized by the tax.

Under full cooperation, regions agree to introduce a global carbon tax of 43.07\$/ t CO_2 in 2020. This value is in line with the 34\$ reported in GHKT and also in Hillebrand & Hillebrand (2019) for the year 2010.[10] Our value is slightly higher than the optimal tax (corresponding to the social cost of carbon) obtained from the DICE model which amounts to 36.7 \$/t CO_2 in $t = 2020$.[11] The optimal tax increases again over time re-

[10]The increase is mainly due to GDP growth from 2010 to 2020 and also because we measure GDP in PPP-terms which increases measured GDP notably in less developed countries.

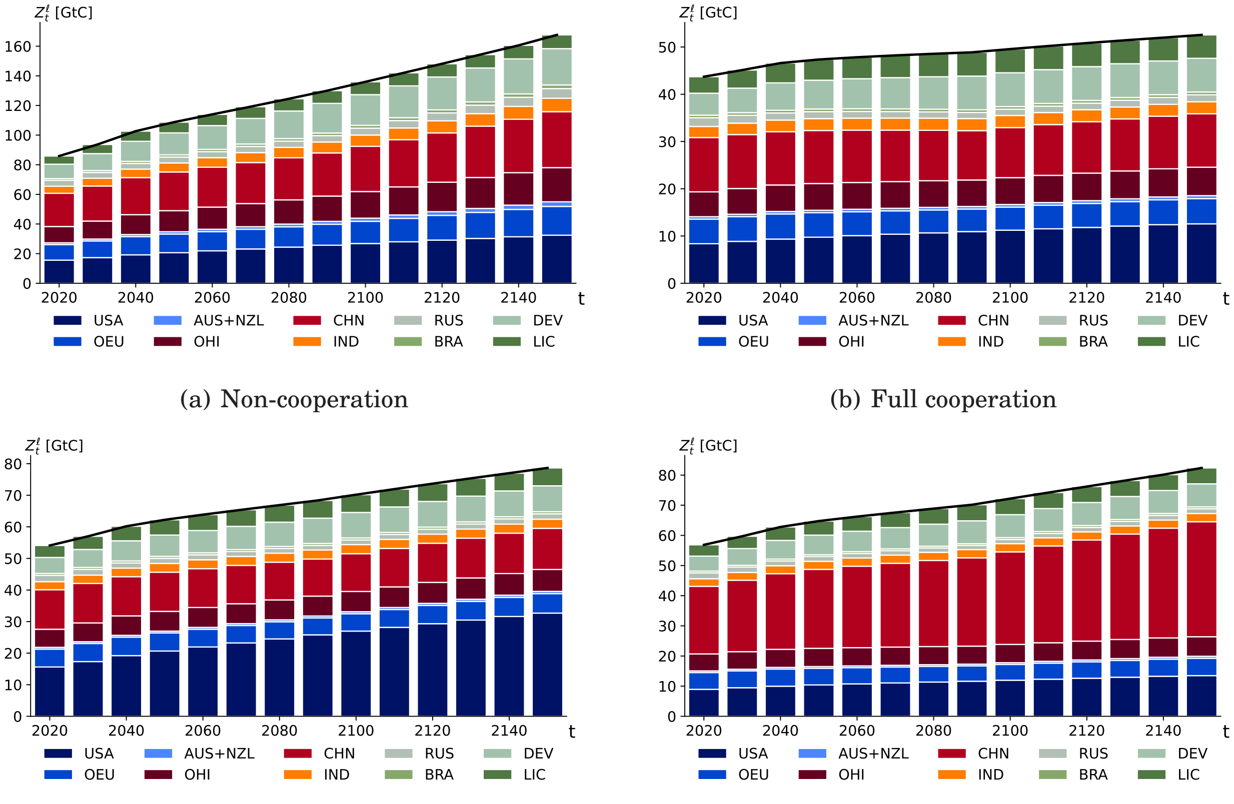
[11]The value is directly taken from the latest spreadsheet version of DICE available online at http://www.econ.yale.edu/~nordhaus/homepage/homepage/DICE2016R-090916ap-v2.xlsm.

flecting the growth trend of GDP in each region to reach a value of 107 \$/t CO_2 in 2100 and of 169 \$/t CO_2 in 2150.

In Scenario 3 where all regions except the U.S. form a coalition, the coalitional tax is 36.23 \$/t CO_2 while the U.S. imposes a much lower tax of merely 6.91 \$/t CO_2. A symmetric picture emerges for Scenario 4 where instead China breaks away from the grand coalition imposing a domestic tax of 5.57 \$/t CO_2 while the coalition sets a much higher tax of 37.59 \$/t CO_2. In both cases, however, the coalitional tax is lower than for the grand coalition because damages in the deviating country are no longer internalized.

Regional emissions

It is clear that any successful climate policy must induce drastic reductions of energy related CO_2-emissions. Figure 2 shows how regional and global emissions evolve over time under all four political scenarios.



(a) Non-cooperation

(b) Full cooperation

(c) Grand coalition except US

(d) Grand coalition except China

Figure 2: Regional emissions.

Introducing the optimal global tax in Scenario 2 leads to a substantial and permanent reduction in emissions. At the global level, emissions in 2020 decline by about 49% relative to Scenario 1 with non-cooperation. In absolute terms, this corresponds to a reduction of more than 42 GtC. This immediate reduction in absolute and relative terms is a little less pronounced but still significant for the two Scenarios 3 and 4 with partial cooperation: A global coalition without the U.S. reduces global CO_2 emissions by 37% corresponding to 31.8 Gt C, a coalition without China by 29.0 Gt C corresponding to a decline of 32%.

At the regional level, there are again sizable differences depending on the degree of fossil fuel dependence and political scenario. Under full cooperation, the most drastic reduction occurs in China where emissions decline by -10.9 GtC (-48.8%) relative to non-cooperation, followed by the U.S. where the decline is -7.3 GtC (-46.4%).

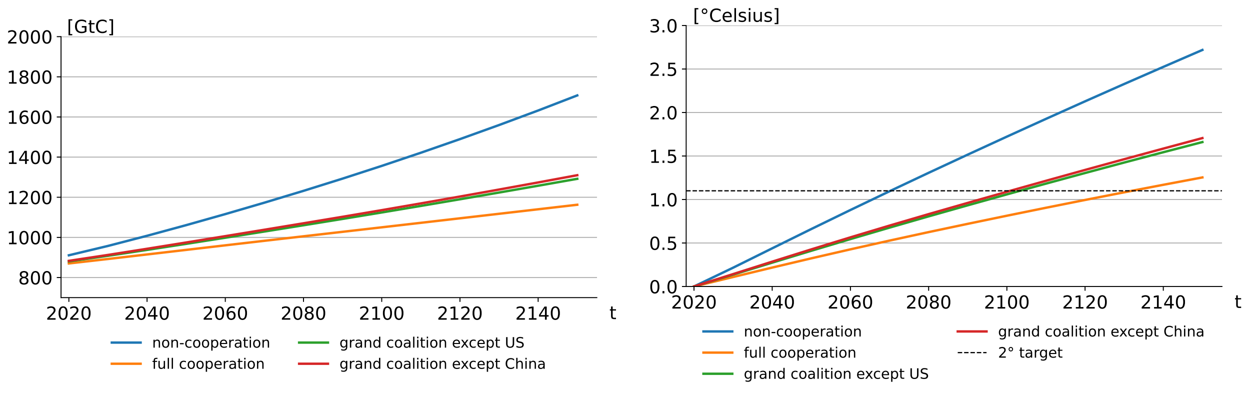
For all three cooperative scenarios, initial reductions are essentially preserved over the entire time window where emissions within the corresponding coalition rise only slightly. This is in stark contrast to the non-cooperative scenario 1 and for the region not part of the world coalition in Scenarios 3 and 4 where emissions continue to grow without bounds due to the exponential increase in fossil fuel consumption. As a consequence, the deviating country acquires an every increasing share of global emissions and becomes the main offender in both Scenarios 3 and 4.

Climate and temperature

The main variable representing global warming and the success of climate policies is the change in global temperature determined by the so-called Arrhenius relation:

$$TEMP_t = 3\log(S_t/581)/\log 2. \tag{66}$$

Figure 3 shows the change in global temperature relative to the baseline period (right) for our four political scenarios. An important benchmark is the 2°-target set by the 2015 Paris Agreement (cf. UNFCCC (2015)) which limits the increase in global temperature to less than two degrees until 2100 relative to the pre-industrial level. Data from NASA (2018) shows that global temperature in 2015 already exceeded the pre-industrial level by 0.9 °C. For this reason, the two-degree target represented by the dashed line in Figure 3 corresponds to an increases of 1.1 °C relative to 2015.



(a) CO_2-concentration

(b) Global temperature

Figure 3: Atmospheric carbon concentration and global temperature.

The main insight here is that under full cooperation global temperature increases by 1.014 °C until 2100 which is therefore in line with the two-degree target. By contrast, the two-degree target is exceeded as early as 2070 under non-cooperation and exactly in 2100 under both political scenarios of partial cooperation and increases exponentially thereafter.

8 Conclusions

The present paper provides closed-form solutions of optimal climate polices under different scenarios of cooperation between regions. These results provide the basis for understanding the quantitative effects of climate policy for a given coalition structure. They also hold the key for designing the redistribution scheme discussed in this paper under which all countries have an incentive to cooperate and participate in a global climate agreement to implement the efficient level of taxation.

One goal of future research is to better understand the formation process of coalitions and their stability. A recent study of this type can be found in Vosooghi, Arvaniti & van der Ploeg (2022) who determine the equilibrium coalition structure endogenously. However, their setup does not allow for trade between countries on international markets. Thus, it would be interesting to extend their results to a setting with regional trade assumed in our paper.

A second line of research would be to make technological change endogenous and directed as do Acemoglu et al.(2012) and Hassler et al. (2021). Understanding the impact of regional technological change and its spill-over to other regions as well as the scope for direct technology transfers across regions is likely to be a key factor in solving the climate problem. Such an innovation policy could complement the incentives set by tax policies studied in this paper.

A Mathematical results and proofs

A.1 Proof of Lemma 2

The boundary behavior (2) of each F_t^ℓ and utility function (7) ensure that any solution to (28) satisfies $K_t^\ell > 0$, $X_t^\ell > 0$, and $C_t^\ell > 0$. Thus, we can dispense with non-negativity constraints and use (27) as the single constraint. Define the Lagrangean function

$$\mathscr{L}\Big((K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}, \lambda\Big) := \sum_{t=0}^{\infty} \beta^t u(C_t^\ell) + \lambda\Bigg(\sum_{t=0}^{\infty} q_t \Big(e^{-\gamma^\ell \sum_{n=0}^{\infty}\delta_n (X_{t-n}^\ell + X_{t-n}^{-\ell})} F_t^\ell(K_t^\ell, X_t^\ell) \\ - r_t K_t^\ell - v_t X_t^\ell - C_t^\ell\Big) + r_0 K_0^{\ell,s} + (v_0 - c_x) R_0^\ell\Bigg).$$

For each $t = 0,1,2,\ldots$ the derivatives with respect to consumption and capital read:

$$\frac{\partial \mathscr{L}\left((K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}, \lambda\right)}{\partial C_t^\ell} = \beta^t u'(C_t^\ell) - \lambda q_t \overset{!}{=} 0 \tag{67a}$$

$$\frac{\partial \mathscr{L}\left((K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}, \lambda\right)}{\partial K_t^\ell} = \lambda q_t \left(e^{-\gamma^\ell \sum_{n=0}^\infty \delta_n (X_{t-n}^\ell + X_{t-n}^{-\ell})} \partial_K F_t^\ell(K_t^\ell, X_t^\ell) - r_t\right) \overset{!}{=} 0 \tag{67b}$$

Solving (67a) gives $\lambda = \beta^t u'(C_t^\ell)/q_t$ for all t which implies the Euler equation (19). Moreover, $\lambda > 0$ implies that (27) holds with equality by virtue of the Kuhn-Tucker theorem. Equation (67b) implies (12a). Finally, the partial derivative with respect to fossil fuels read:

$$\frac{\partial \mathscr{L}\left((K_t^\ell, X_t^\ell, C_t^\ell)_{t\geq 0}, \lambda\right)}{\partial X_t^\ell} = \lambda q_t \left((1 - D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) - v_t - \sum_{n=0}^\infty \frac{q_{t+n}}{q_t} \gamma^\ell \delta_n Y_{t+n}^\ell\right) \overset{!}{=} 0 \tag{68}$$

with D_t^ℓ defined as in (23) and Y_t^ℓ as in (1). Rearranging (68) using that by virtue of (13) and (19) $q_{t+n}/q_t = \beta^n (C_{t+n}/C_t)^{-\sigma}$ implies (29). ■

A.2 Proof of Lemma 4

To derive the optimality conditions (45) we adopt a standard Lagrangean approach. The boundary behavior (2) of each F_t^ℓ and of the utility function (7) ensures that any solution to (44) satisfies $K_t^\ell > 0$, $X_t^\ell > 0$, and $\overline{C}_t > 0$ for all t and ℓ. Thus, we can dispense with non-negativity constraints. Define the Lagrangean function

$$\mathscr{L}\left(((K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}}, \overline{C}_t)_{t\geq 0}, (\lambda_t)_{t\geq 0}, \mu_X, \mu_K\right) := \sum_{t=0}^\infty \beta^t u(\overline{C}_t) + \mu_X \left(\overline{R}_0 - \sum_{t=0}^\infty \sum_{\ell\in\mathbb{L}} X_t^\ell\right)$$
$$+ \sum_{t=0}^\infty \lambda_t \left(\sum_{\ell\in\mathbb{L}} e^{-\gamma^\ell \sum_{n=0}^\infty \delta_n \sum_{k\in\mathbb{L}} X_{t-n}^k} F_t^\ell(K_t^\ell, X_t^\ell) - \overline{C}_t - \sum_{\ell\in\mathbb{L}} K_{t+1}^\ell - c_x \sum_{\ell\in\mathbb{L}} X_t^\ell\right) + \mu_K \sum_{\ell\in\mathbb{L}} \left(K_0^{\ell,s} - K_0^\ell\right).$$

For each $t = 0,1,2,\ldots$ and $\ell \in \mathbb{L}$ the first order conditions read :

$$\frac{\partial \mathscr{L}(-)}{\partial \overline{C}_t} = \beta^t u'(\overline{C}_t) - \lambda_t \overset{!}{=} 0 \quad \text{for all } t = 0,1,2,\ldots \tag{69a}$$

$$\frac{\partial \mathscr{L}(-)}{\partial K_0^\ell} = \lambda_0 (1 - D_0^\ell)\partial_K F_0^\ell(K_0^\ell, X_0^\ell) - \mu_K \overset{!}{=} 0 \quad \text{for all } \ell \in \mathbb{L} \tag{69b}$$

$$\frac{\partial \mathscr{L}(-)}{\partial K_t^\ell} = \lambda_t (1 - D_t^\ell)\partial_K F_t^\ell(K_t^\ell, X_t^\ell) - \lambda_{t-1} \overset{!}{=} 0 \quad \text{for all } \ell \in \mathbb{L} \text{ and } t = 1,2,3,\ldots \tag{69c}$$

$$\frac{\partial \mathscr{L}(-)}{\partial X_t^\ell} = -\mu_X + \lambda_t \left((1 - D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) - c_x\right)$$
$$- \sum_{n=0}^\infty \lambda_{t+n} \delta_n \sum_{k\in\mathbb{L}} \gamma^k Y_{t+n}^k \overset{!}{=} 0 \quad \text{for all } \ell \in \mathbb{L} \text{ and } t = 0,1,2,\ldots \tag{69d}$$

Solving (69a) gives

$$\lambda_t = \beta^t u'(\overline{C}_t) = \beta^t \overline{C}_t^{-\sigma} \quad \text{for all} \quad t = 0,1,2,\ldots \tag{70}$$

Using (70) in (69c) gives the Euler equation (45c). Since the term $\beta\overline{C}_{t+1}^{-\sigma}/\overline{C}_t^{-\sigma}$ in (45c) is independent of ℓ, this equation and (69b) imply (45a). Furthermore, solving (69d) using (70) and defining $\hat{\tau}_t$ as in (46) gives

$$\mu_X = \lambda_t \left((1 - D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) - c_x - \hat{\tau}_t \right) \quad \text{for all } \ell \in \mathbb{L} \text{ and } t = 0,1,2,\ldots \tag{71}$$

Since the l.h.s. in (71) and the term $\hat{\tau}_t$ are both independent of the regional index ℓ, this implies (45b). Further, since the l.h.s. in (71) is also independent of time, combining it with (70) gives (45d). Equations (69b) and (70) imply $\mu_K > 0$ and $\lambda_t > 0$ such that (42) and (43) hold with equality by means of the Kuhn-Tucker theorem. The transversality condition ensures that consumption does not implode.
Remark: The term $\hat{x}_t := \mu_X/\lambda_t$ plays the role of a scarcity rent which by (71) evolves as

$$\hat{x}_{t+1} = \mu_X/\lambda_{t+1} = \frac{\lambda_t}{\lambda_{t+1}}\hat{x}_t = \frac{\overline{C}_t^{-\sigma}}{\beta\overline{C}_{t+1}^{-\sigma}}\hat{x}_t \quad \text{for all } t = 0,1,2,\ldots \tag{72}$$

This corresponds to the variable $v_t - c_x$ along the decentralized equilibrium. The initial value $\hat{x}_0 = \mu_X/\lambda_0$ ensures that either $\sum_{t=0}^{\infty}\sum_{\ell\in\mathbb{L}} X_t^\ell = \overline{R}_0$ or $\sum_{t=0}^{\infty}\sum_{\ell\in\mathbb{L}} X_t^\ell < \overline{R}_0$ in which case $\mu_X = 0$ by the Kuhn-Tucker Theorem and $\hat{x}_0 = 0$. In the latter case, by (71)

$$(1 - D_t^\ell)\partial_X F_t^\ell(K_t^\ell, X_t^\ell) = c_x + \hat{\tau}_t \quad \text{for all } \ell \in \mathbb{L} \text{ and } t = 0,1,2,\ldots \tag{73}$$

This corresponds to the case $v_t = c_x$ along the decentralized solution. ■

A.3 Proof of Lemma 5

The argument depends on parameter σ in (7). We distinguish the following cases.
First, assume that $\sigma \neq 1$ in (7). Then, U is homogenous of degree $1-\sigma > 0$ permitting to write the utilities in (50) as

$$U((\mu^\ell \overline{C}_t^{\text{eff}})_{t\geq 0}) = (\mu^\ell)^{1-\sigma}\overline{U}^{\text{eff}} \quad \text{and} \quad U((\mu^{\ell,\text{nc}}\overline{C}_t^{\text{nc}})_{t\geq 0}) = (\mu^{\ell,\text{nc}})^{1-\sigma}\overline{U}^{\text{nc}}. \tag{74}$$

Suppose $\sigma < 1$. Then, U is positive-valued and solving (74) for μ^ℓ using $\overline{U}^{\text{eff}} > 0$ gives the following condition for the consumption share of region ℓ:

$$\mu^\ell \geq \mu^\ell_{\text{crit}} := \mu^{\ell,\text{nc}} \left(\overline{U}^{\text{nc}}/\overline{U}^{\text{eff}} \right)^{\frac{1}{1-\sigma}}. \tag{75}$$

Since $0 < \overline{U}^{\text{nc}}/\overline{U}^{\text{eff}} < 1$ and $\sigma < 1$, we have $\mu^\ell_{\text{crit}} < \mu^{\ell,\text{nc}}$ as claimed.
Suppose $\sigma > 1$. Then, U is negative-valued and solving (74) for μ^ℓ yields again the same

condition (75). Since now $\overline{U}^{\text{nc}}/\overline{U}^{\text{eff}} > 1$, $\sigma > 1$ implies $\mu^{\ell}_{\text{crit}} < \mu^{\ell,\text{nc}}$ again.
As a second case, assume that $\sigma = 1$ implying logarithmic period utility in (7). Then,

$$U((\mu^{\ell}\overline{C}^{\text{eff}}_{t})_{t\geq 0}) = \frac{\log(\mu^{\ell})}{1-\beta} + \overline{U}^{\text{eff}} \quad \text{and} \quad U((\mu^{\ell,\text{nc}}\overline{C}^{\text{nc}}_{t})_{t\geq 0}) = \frac{\log(\mu^{\ell,\text{nc}})}{1-\beta} + \overline{U}^{\text{nc}}. \tag{76}$$

Using this result in (74) and solving again for μ^{ℓ} gives

$$\mu^{\ell} \geq \mu^{\ell}_{\text{crit}} := \mu^{\ell,\text{nc}} \cdot e^{-(1-\beta)(\overline{U}^{\text{eff}}-\overline{U}^{\text{nc}})}. \tag{77}$$

Again, $\overline{U}^{\text{eff}} > \overline{U}^{\text{nc}}$ implies that $\mu^{\ell}_{\text{crit}} < \mu^{\ell,\text{nc}}$ also in this case. ■

B Details on the simulations

In this section we report additional details on our computation and calibration strategy.

B.1 Computational details

Equilibrium conditions for period t
Consider an arbitrary period $t \geq 0$. Let aggregate capital supply $\overline{K}^{s}_{t}$, exogenous population and productivity variables $(N^{\ell}_{t}, Q^{\ell}_{K,t}, Q^{\ell}_{X,t})_{\ell\in\mathbb{L}}$, and the climate state $(S_{1,t-1}, S_{2,t-1})$ from the previous period be given. Based on the approximation formula (62), suppose taxes in region $\ell \in \mathbb{L}$ are determined by

$$\tau^{\ell}_{t} = \sum_{k\in\mathbb{L}_{\ell}} \gamma^{k} Y^{k}_{t} \sum_{n=0}^{\infty} \beta^{n}\delta_{n} \tag{78}$$

where $\mathbb{L}_{\ell} \subset \mathbb{L}$ is the coalition that region ℓ as a member of.[12] Using the CES-form of production (63) and defining the cost shares

$$\eta^{\ell}_{K,t} := \frac{\kappa\left((K^{\ell}_{t})^{\alpha}(Q^{\ell}_{N,t}N^{\ell}_{t})^{1-\alpha}\right)^{\frac{\varepsilon-1}{\varepsilon}}}{\left(F^{\ell}_{t}(K^{\ell}_{t},X^{\ell}_{t})\right)^{\frac{\varepsilon-1}{\varepsilon}}} \quad \text{and} \quad \eta^{\ell}_{X,t} := \frac{(1-\kappa)(Q^{\ell}_{X,t}X^{\ell}_{t})^{\frac{\varepsilon-1}{\varepsilon}}}{\left(F^{\ell}_{t}(K^{\ell}_{t},X^{\ell}_{t})\right)^{\frac{\varepsilon-1}{\varepsilon}}} = 1-\eta^{\ell}_{K,t} \tag{79}$$

one can write final sector's optimality conditions (12) as

$$r_{t} = \alpha\eta^{\ell}_{K,t}Y^{\ell}_{t}/K^{\ell}_{t} \quad \text{and} \quad c_{x} + \sum_{k\in\mathbb{L}_{\ell}} \gamma^{k}Y^{k}_{t} \sum_{n=0}^{\infty} \beta^{n}\delta_{n} = \eta^{\ell}_{X,t}Y^{\ell}_{t}/X^{\ell}_{t}, \quad \ell\in\mathbb{L}. \tag{80}$$

[12]In particular, $\mathbb{L}_{\ell} = \{\ell\}$ in the non-cooperative and $\mathbb{L}_{\ell} = \mathbb{L}$ in the fully cooperative case. In our simulations, we compared the approximated value (78) ex-post to the true solution (61) to show that the approximation is excellent.

Moreover, using (65) and (4) in (6) permits to write final output in region ℓ as

$$Y_t^\ell = \exp\left(-\gamma^\ell\left(S_{1,t-1} + (1-\phi)S_{2,t-1} + (\phi_L + (1-\phi_L)\phi_0)\sum_{k\in\mathbb{L}} X_t^k - \bar{S}\right)\right) F_t^\ell(K_t^\ell, X_t^\ell). \quad (81)$$

The temporary equilibrium problem is to determine the factor allocation $(K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}}$ consistent with optimal producer behavior (80) and the market clearing condition

$$\sum_{\ell\in\mathbb{L}} K_t^\ell = \overline{K}_t^s \quad (82)$$

with outputs $(Y_t^\ell)_{\ell\in\mathbb{L}}$ determined by (81) and cost shares as in (79). Rearranging the first condition in (80) and summing over all regions using (82) gives

$$r_t = \alpha \sum_{\ell\in\mathbb{L}} \eta_{K,t}^\ell Y_t^\ell / K_t^s. \quad (83)$$

Using (83) in (80) and re-arranging both conditions, exploiting that $\eta_{X,t}^\ell = 1 - \eta_{K,t}^\ell$ gives

$$K_t^\ell = \frac{\eta_{K,t}^\ell Y_t^\ell}{\sum_{h\in\mathbb{L}} \eta_{K,t}^h Y_t^h} \cdot K_t^s \quad \text{and} \quad X_t^\ell = \frac{(1-\eta_{K,t}^\ell) Y_t^\ell}{c_x + \sum_{k\in\mathbb{L}_\ell} \gamma^k Y_t^k \sum_{n=0}^\infty \beta^n \delta_n}. \quad (84)$$

Next we show how the temporary equilibrium problem can be computed numerically.

Computing the factor allocation in period t.

Let arbitrary values $\hat{H} := (\hat{Y}_t^\ell, \hat{\eta}_{K,t}^\ell)_{\ell\in\mathbb{L}} \in \mathbb{H} := (\mathbb{R}_{++} \times [0,1[)^L$ be given. Using $\hat{H}$ in (84) determines the implied factor allocation $\hat{G} := (\hat{K}_t^\ell, \hat{X}_t^\ell)_{\ell\in\mathbb{L}} \in \mathbb{G} := \mathbb{R}_{++}^{2L}$. This step defines a first mapping $\Phi_G : \mathbb{H} \to \mathbb{G}$, $\hat{H} \mapsto \Phi_G(\hat{H}) := \hat{G}$. Substituting the values $\hat{G}$ back into (79) and (81) yields the updated values $\tilde{H} := (\tilde{Y}_t^\ell, \tilde{\eta}_{K,t}^\ell)_{\ell\in\mathbb{L}} \in \mathbb{H}$ defining a second mapping $\Phi_H : \mathbb{G} \to \mathbb{H}$, $\hat{G} \mapsto \Phi_H(\hat{G}) := \tilde{H}$. The composition $\Phi := \Phi_H \circ \Phi_G : \mathbb{H} \to \mathbb{H}, \hat{H} \mapsto \Phi(\hat{H}) := \tilde{H}$ is a self-map on $\mathbb{H}$ and the equilibrium solution $H := (Y_t^\ell, \eta_{K,t}^\ell)_{\ell\in\mathbb{L}}$ is a fixed point of Φ. It turned out that Φ is globally asymptotically stable such that simply iterating Φ forward, starting with an arbitrary guess H_0 yields the equilibrium solution $\lim_{n\to\infty} \Phi^n(H_0) = H = (Y_t^\ell, \eta_{K,t}^\ell)_{\ell\in\mathbb{L}}$. The implied factor allocation then obtains as $G = (K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}} = \Phi_G(H)$.

Equilibrium state dynamics

The endogenous state variable in period t is $\xi_t = (\overline{K}_{t+1}^s, \overline{C}_t, S_{1,t}, S_{2,t})$. To uncover the forward-recursive structure of equilibrium, let the previous state ξ_{t-1} in period t be given. Using $\overline{K}_t^s$ and $(S_{1,t-1}, S_{2,t-1})$ along with the exogenous variables $(N_t^\ell, Q_{K,t}^\ell, Q_{X,t}^\ell)_{\ell\in\mathbb{L}}$ we can determined the factor allocation $(K_t^\ell, X_t^\ell)_{\ell\in\mathbb{L}}$ and regional outputs $(Y_t^\ell)_{\ell\in\mathbb{L}}$ and the auxiliary variables $(\eta_{K,t}^\ell, \eta_{X,t}^\ell)_{\ell\in\mathbb{L}}$ as described in the previous step. Using the implied aggregate emissions $\overline{X}_t := \sum_{\ell\in\mathbb{L}} X_t^\ell$ in (65) determines the new climate state $(S_{1,t}, S_{2,t})$. Computing the implied capital return r_t as in (83), current aggregate consumption follows from the Euler equation $\overline{C}_t = \beta r_t \overline{C}_{t-1}$. Aggregate capital supply $\overline{K}_{t+1}^s$ then follows from the resource constraint (22), completing the determination of ξ_t.

Given initial world capital $\overline{K}_0^s$ and the climate state $(S_{1,-1}, S_{2,-1})$, we determine initial aggregate consumption $\overline{C}_{-1}$ (or, equivalently, $\overline{C}_0 = (\beta r_0)^{1/\sigma}\overline{C}_{-1}$) such that the transversality condition $\lim_{t\to\infty} \beta^t \overline{K}_{t+1}^s / \overline{C}_t = 0$ holds, i.e., consumption neither explodes nor implodes. It turned out that there is always a unique value C_0 for which this is the case.

B.2 Calibration details

Determining productivity parameters

Let the initial climate state $(S_{1,-1}, S_{2,-1})$ and initial capital supply $\overline{K}_0^s$ in $t=0$ be given. For each region $\ell \in \mathbb{L}$, set output and emissions to the target levels defined in Table 3, i.e, $Y_0^\ell = Y_0^{\ell,\text{target}}$ and $X_0^\ell = X_0^{\ell,\text{target}}$. Consider the non-cooperative scenario where each region sets taxes based on the approximation (41). Then, initial regional taxes compute as $\tau_0^\ell = \gamma^\ell Y_0^\ell \sum_{n=0}^{\infty} \beta^n \delta_n$. Using this result and the given values of regional output Y_0^ℓ and emissions X_0^ℓ one can explicitly solve the second condition in (84) to obtain the values $\eta_{X,0}^\ell = 1 - \eta_{K,0}^\ell$ for all $\ell \in \mathbb{L}$. This also determines $\eta_{K,0}^\ell$ for each ℓ, which can be used in (84) together with the given output levels $(Y_0^\ell)_{\ell\in\mathbb{L}}$ to infer the initial capital allocation $(K_0^\ell)_{\ell\in\mathbb{L}}$. Further, one obtains from (81) the values $F_0^\ell(K_0^\ell, X_0^\ell)$ as

$$F_0^\ell(K_0^\ell, X_0^\ell) = Y_0^\ell \cdot \exp\left(\gamma^\ell \left(S_{1,-1} + (1-\phi)S_{2,-1} + (\phi_L + (1-\phi_L)\phi_0)\sum_{k\in\mathbb{L}} X_0^k - \bar{S}\right)\right). \tag{85}$$

Using the values $\eta_{X,0}^\ell$, X_0^ℓ, and $F_0^\ell(K_0^\ell, X_0^\ell)$ one can solve the second condition in (79) explicitly for $Q_{X,0}^\ell$. Analogously, one can use $\eta_{K,0}^\ell$, K_0^ℓ, $F_0^\ell(K_0^\ell, X_0^\ell)$ and the given value N_0^ℓ to solve the first condition in (79) explicitly for $Q_{N,0}^\ell$.

B.3 Data

Regional GDP and emissions

We use annual regional PPP-adjusted GDP data expressed in current international \$ from the World Bank (2020). Data on regional emissions were obtained from the IEA Energy Balances (2021) expressed in Gt of CO_2 and converted to GtC using a conversion factor of 12/44. All values are aggregated over different countries based on our regional distinction and years 2010-2019. .

Population data

We used current population sizes and predictions from the 'Population Projection 2300' provided by the United Nations (2004) displayed in Table 6. Population sizes for 2010-2019 are averaged to obtain the value used in $t = 0$. To obtain population forecasts for each period $t = 1, \ldots, 19$ we linearly interpolate between forecasts for 2050, 2100, 2150, and 2200 from Table 6. After $t = 19$ corresponding to the year 2200, the population in all regions is assumed to be constant.

Table 6: Regional population for baseline period 2010-2019 and in future periods

Variable	Description	USA	OEU	ANZ	OHI	CHN	IND	RUS	BRA	DEV	LIC
N_0^ℓ	Pop. 2015 [mio.]	319.33	558.20	28.28	277.45	1374.64	1301.92	143.82	203.53	2518.25	578.43
N_4^ℓ	Pop. 2050 [mio.]	408.70	583.74	30.07	330.84	1395.18	1531.44	101.46	33.14	3474.70	863.53
N_9^ℓ	Pop. 2100 [mio.]	437.16	477.91	28.83	301.89	1181.50	1458.36	79.54	212.45	3919.31	951.91
N_{14}^ℓ	Pop. 2150 [mio.]	452.75	480.79	29.21	292.52	1149.12	1308.19	83.08	202.21	3580.55	860.90
N_{19}^ℓ	Pop. 2200 [mio.]	470.05	499.47	30.36	298.26	1200.73	1304.53	86.74	208.83	3537.94	848.62

References

ACEMOGLU, D., P. AGHION, L. BURSZTYN & D. HEMOUS (2012): “The Environment and Directed Technical Change”, *American Economic Review*, 102(1), 131–166.

BATTAGLINI, M. & B. HARSTAD (2016): “Participation and Duration of Environmental Agreements”, *Journal of Political Economy*, 124(1), 160–204.

BOSELLO, F., B. BUCHNER & C. CARRARO (2003): “Equity, Development, and Climate Change Control”, *Journal of the European Economic Association*, 1, 601–611.

DENNING, F. & J. EMMERLING (2017): “A Note on Optima with Negishi Weights”, Working paper.

EYCKMANS, J. & M. FINUS (2003): “Coalition formation in a global warming game : how the design of protocols affects the success of environmental treaty-making”, LIDAM Discussion Papers CORE 2003088, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

EYCKMANS, J. & H. TULKENS (2003): “Simulating coalitionally stable burden sharing agreements for the climate change problem”, *Resource and Energy Economics*, 25, 299–327.

FINUS, M., B. RUNDSHAGEN & J. EYCKMANS (2014): “Simulating a sequential coalition formation process for the climate change problem: first come, but second served?”, *Annals of Operations Research*, 220(1), 5–23.

GERLAGH, R. & M. LISKI (2018): “Consistent climate policies”, *Journal of the European Economic Association*, 16(1), 1–44.

GOLOSOV, M., J. HASSLER, P. KRUSELL & A. TSYVINSKI (2014): “Optimal Taxes on Fossil Fuel in General Equilibrium”, *Econometrica*, 82(1), 41–88.

HAMBEL, C., H. KRAFT & E. SCHWARTZ (2021): “The social cost of carbon in a non-cooperative world”, *Journal of International Economics*, 131, 103490.

HARSTAD, B. (2012): “Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations”, *Review of Economic Studies*, 79, 1527–1557.

——— (2016): "The Dynamics of Climate Agreements", *Journal of the European Economic Association*, 14(3), 719–752.

HASSLER, J. & P. KRUSELL (2012): "Economics And Climate Change: Integrated Assessment In A Multi-Region World", *Journal of the European Economic Association*, 10(5), 974–1000.

HASSLER, J., P. KRUSELL & C. OLOVSSON (2021): "Directed Technical Change as a Response to Natural Resource Scarcity", *Jounral of Political Economcy*, 129(11), 3039–3072.

HILLEBRAND, E. & M. HILLEBRAND (2019): "Optimal Climate Policies in a Dynamic Multi-Country Equilibrium Model", *Journal of Economic Theory*, 179, 200–239.

——— (2021): "Who Pays the Bill? Climate Change, Taxes, and Transfers in a Multi-Region Growth Model", Working paper, University of Freiburg.

INTERNATIONAL ENERGY AGENCY (2021): "World Energy Balances 2021", .

MAS-COLELL, A., M. WHINSTON & J. GREEN (1995): *Microeconomic Theory*. Oxford University Press, New York a.o.

NASA (2018): "Global Land-Ocean Temperature Index", Discussion paper, NASA Goddard Institute for Space Studies.

NORDHAUS, W. & P. SZTORC (2013): "DICE 2013R: Introduction and User's Manual", Discussion paper, Website: dicemodel.net.

NORDHAUS, W. & Z. YANG (1996): "A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies", *American Economic Review*, 86, 741–765.

NORDHAUS, W. D. (2014): "Estimates of the Social Cost of Carbon: Concepts and Results from the DICE-2013R Model and Alternative Approaches", *Journal of the Association of Environmental and Resource Economists*, 1, 273–312.

PAPAGEORGIOU, C., M. SAAM & P. SCHULTE (2017): "Substitution between Clean and Dirty Energy Inputs: A Macroeconomic Perspective", *Review of Economics and Statistics*, 99(2), 281–290.

UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE (2015): "Paris Agreement", https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

VOSOOGHI, S., M. ARVANITI & F. VAN DER PLOEG (2022): "Self-Enforcing Climate Coalitions for Farsighted Countries: Integrated Analysis of Heterogeneous Countries", CESifo working paper, CESifo.

World Bank (2020): "World Development Indicators (WDI)", https://datacatalog.worldbank.org/dataset/world-development-indicators.

Yi, S.-S. (1997): "Stable Coalition Structures with Externalities", *Games and Economic Behavior*, 20, 201–237.